ISSN 1847-3946

mipro 2023

organizer upro



46th

ICT and Electronics Convention

May 22 - 26, 2023 Opatija, Croatia

Lampadem tradere



mipro - innovative promotional partnership

mipro proceedings

My profession.
My organization.
My IEEE.



Discover the benefits of IEEE membership.

Join a community of more than 365,000 innovators in over 150 countries. IEEE is the world's largest technical society, providing members with access to the latest technical information and research, global networking and career opportunities, and exclusive discounts on education and insurance products.

Join today
www.ieee.org/join

IEEE

MIPRO 2023

46th ICT and Electronics Convention

May 22 – 26, 2023
Opatija, Croatia

Proceedings

Conferences:

Human-Computer Interaction /HCI

Smart Industries and Digital Ecosystems /SIDE

Microelectronics, Electronics and Electronic Technology /MEET

Optoelectronics and Photonics /OPHO

Data Science and Biomedical Engineering /DS-BE

Telecommunications & Information /CTI

Computers in Education /CE

Distributed Computing and Cyber-Physical Systems /DC-CPS

Artificial Intelligence Systems /AIS

Information Systems Security /ISS

Business Intelligence Systems /miproBIS

Digital Economy and Digital Society /DE-DS

Information and Communication Technology Law /ICTLAW

Engineering Education /EE

Software and Systems Engineering /SSE

MIPRO Junior - Student Papers /SP

Edited by:
Karolj Skala

International Program Committee

Karolj Skala, General Chair (Croatia),
Lejla Abazi-Bexheti (North Macedonia),
Enis Afgan (United States),
Saša Aksentijević (Croatia),
Slaviša Aleksić (Germany),
Slavko Amon (Slovenia),
Krešo Antonović (Croatia),
Zoran Aralica (Croatia),
Michael E. Auer (Austria),
Viktor Avbelj (Slovenia),
Dubravko Babić (Croatia),
Snježana Babić (Croatia),
Dinko Bačić (United States),
Darinko Bago (Croatia),
Tadej Bajd (Slovenia),
Ante Bakić (Croatia),
Hrvoje Balen (Croatia),
Marko Banek (Croatia),
Mirta Baranović (Croatia),
Bartosz Bebel (Poland),
Nina Begičević Ređep (Croatia),
Ladjel Bellatreche (France),
Boris Blumenschein (Croatia),
Damir Boras (Croatia),
Darko Bosnar (Croatia),
Adrian Boukalov (Belgium),
Lovro Božičević (Croatia),
Ricardo Branco (Portugal),
Laura Brandimarte (United States),
Mario Brčić (Croatia),
Karla Brkić (Croatia),
Ljiljana Brkić (Croatia),
Marian Bubak (Poland),
Andrea Budin (Croatia),
Željko Butković (Croatia),
Domenico Caputo (Italy),
Željka Car (Croatia),
Jesús Carretero Pérez (Spain),
Constantinos K. Coursaris (Canada),
Bojan Cukic (United States),
Alfredo Cuzzocrea (Italy),
Duško Čakara (Croatia),
Stipo Čelar (Croatia),
Marina Čičin-Šain (Croatia),
Dragan Čišić (Croatia),
Marko Čupić (Croatia),
Davor Davidović (Croatia),
Vlado Delić (Serbia),
Matjaž Depolli (Slovenia),
Saša Dešić (Croatia),
Dražen Dragičević (Croatia),
Marin Karuza (Croatia),
Ivan Kaštelan (Serbia),
Zlatko Katalenić (Slovenia),
Ana Katalinić Mucalo (Croatia),
Tihomir Katulić (Croatia),
Marko Đurasević (Croatia),
Todd Eavis (Canada),
Maurizio Ferrari (Italy),
Tiziana Ferrari (Netherlands),
Renato Filjar (Croatia),
Matteo Francia (Italy),
Andreja Gajović (Croatia),
Tihana Galinac Grbac (Croatia),
Enrico Gallinucci (Italy),
Dragan Gamberger (Croatia),
Gordan Gledec (Croatia),
Matteo Golfarelli (Italy),
Stjepan Golubić (Croatia),
Miran Gosta (Croatia),
Vera Gradišnik (Croatia),
Simeon Grazio (Croatia),
Andrej Grgurić (Croatia),
Stjepan Groš (Croatia),
Matija Gulić (Croatia),
Nina Gumzej (Croatia),
Marjan Gusev (North Macedonia),
Jaak Henno (Estonia),
Željko Hocenski (Croatia),
Tatjana Holjevac (Croatia),
Marko Horvat (Croatia),
Vlasta Hudek (Croatia),
Darko Huljenić (Croatia),
Robert Inkret (Croatia),
Ivo Ipšić (Croatia),
Mile Ivanda (Croatia),
Marina Ivašić-Kos (Croatia),
Hannu Jaakkola (Finland),
Tomislav Jagušt (Croatia),
Domagoj Jakobović (Croatia),
Vojko Jazbinšek (Slovenia),
Leonardo Jelenković (Croatia),
Jeffrey L. Jenkins (United States),
Bojan Jerbić (Croatia),
Dragan Jevtić (Croatia),
Andrej Jokić (Croatia),
Alan Jović (Croatia),
Alen Jugović (Croatia),
Admela Jukan (Germany),
Nenad Jukić (United States),
Oliver Jukić (Croatia),
Irena Jurdana (Croatia),
Ozren Jureković (Croatia),
Marko Jurić (Croatia),
Đani Juričić (Slovenia),
Nikola Kadoić (Croatia),
Jurij Matija Kališnik (Slovenia),
Mirko Poljak (Croatia),
Tonka Poplas Susič (Slovenia),
Aleksandra Rashkovska (Slovenia),
Robert Repnik (Slovenia),
Libuša Révészová (Slovakia),

Tonimir Kišasondi (Croatia),
Zalika Klemenc-Ketiš (Slovenia),
Tihomir Knežević (Croatia),
Dragi Kocev (Slovenia),
Mario Konecki (Croatia),
Marko Koričić (Croatia),
Gregor Kosec (Slovenia),
Igor Kotenko (Russia),
Božidar Kovačić (Croatia),
Zdenko Kovačić (Croatia),
Miklos Kozlovszky (Hungary),
Danica Kragić Jensfelt (Sweden),
Dieter Kranzlmüller (Germany),
Marjan Krašna (Slovenia),
Đorđe Krivokapić (Serbia),
Lene Krøl Andersen (Denmark),
Ashwinee Kumar (India/Belgium),
Nadica Kunštek (Croatia),
Benjamin Kušen (Croatia),
Marko Lacković (Croatia),
Erich Leitgeb (Austria),
Jadran Lenarčič (Slovenia),
Hrvoje Lisičar (Croatia),
Dražen Lučić (Croatia),
Duško Lukač (Germany),
Igor Ljubi (Croatia),
Ruizhe Ma (United States),
Goran Marković (Croatia),
Leslie Martinich (United States),
Romana Matanovac Vučković (Croatia),
Ludek Matyska (Czech Republic),
Mladen Mauher (Croatia),
Igor Mekterović (Croatia),
Željka Mihajlović (Croatia),
Branko Mikac (Croatia),
Anđelko Milardović (Croatia),
Hrvoje Mlinarić (Croatia),
Mladen Mrvelj (Croatia),
Gorana Mudronja (Croatia),
Neeta Nain (India),
Jadranko F. Novak (Croatia),
Dario Ogrizović (Croatia),
Vladimir Olujić (Croatia),
Predrag Pale (Croatia),
Dana Paľová (Slovakia),
Panče Panov (Slovenia),
Branimir Pejčinović (United States),
Nedjeljko Perić (Croatia),
Ana Perić Hadžić (Croatia),
Dana Petcu (Romania),
Juraj Petrović (Croatia),
Duc Truong Pham (UK),
Stjepan Picek (Netherlands),
Damir Pintar (Croatia),
Vincenzo Piuri (Italy),
Mihaela Vranić (Croatia),
Miroslav Vrankić (Croatia),
Neven Vrček (Croatia),
Boris Vrdoljak (Croatia),
Yingwei Wang (Canada),
Slobodan Ribarić (Croatia),
Dubravko Sabolić (Croatia),
Ioan Sacala (Romania),
Davor Salamon (Croatia),
Saonee Sarker (Sweden),
Christoph Schneider (Spain),
Jörg Schulze (Germany),
Bruno Siciliano (Italy),
Sandro Skansi (Croatia),
Zoran Skočir (Croatia),
Tadej Slapnik (Slovenia),
Mladen Sokele (Croatia),
Ana Sović Kržić (Croatia),
Matija Srbić (Croatia),
Vlado Sruk (Croatia),
Hrvoje Stančić (Croatia),
Uroš Janez Stanič (Slovenia),
Mario Stipčević (Croatia),
Vjeran Strahonja (Croatia),
Tomislav Suligoj (Croatia),
Aleksandar Szabo (Croatia),
Marija Šćulac Domac (Croatia),
Dina Šimunić (Croatia),
Frano Škopljanac-Mačina (Croatia),
Dejan Škvorc (Croatia),
Zorislav Šojat (Croatia),
Andreja Špernjak (Slovenia),
Vitomir Štruc (Slovenia),
Marija Šutina (Croatia),
Marko Švaco (Croatia),
Velimir Švedek (Croatia),
Darko Švelec (Croatia),
Viktor Švigelj (Slovenia),
Zheng-Hua Tan (Denmark),
Nikola Tanković (Croatia),
Jason Thatcher (United States),
Antonio Teixeira (Portugal),
Edvard Tijan (Croatia),
A Min Tjoa (Austria),
Ivan Tomašić (Sweden),
Mladen Tomić (Croatia),
Roman Trobec (Slovenia),
Joe S. Valacich (United States),
Mladen Varga (Croatia),
Lucija Vejmelka (Croatia),
Domen Verdnik (Croatia),
Matjaž Veselko (Slovenia),
Linda Vicković (Croatia),
Ivan Vidaković (Croatia),
Marijana Vidas-Bubanja (Serbia),
Marko Vidović (Croatia),
Slavko Vidović (Croatia),
Davor Vinko (Croatia),
Ernest Vlačić (Croatia),
Goran Vojković (Croatia),
Mario Weber (Croatia),
Markus Weinmann (Germany),
David Wilson (United States),
Roman Wyrzykowski (Poland),
Kristijan Zimmer (Croatia)

organized by

MIPRO Croatian Society

technical cosponsorship

IEEE Region 8
IEEE Croatia Section
IEEE Croatia Section Computer Chapter
IEEE Croatia Section Electron Devices/Solid-State Circuits Joint Chapter
IEEE Croatia Section Education Chapter
IEEE Croatia Section Communications Chapter
CroSI Association

under the auspices of

Ministry of Science and Education of the Republic of Croatia
Ministry of the Sea, Transport and Infrastructure of the Republic of Croatia
Ministry of Economy and Sustainable development of the Republic of Croatia
Ministry of Regional Development and EU Funds of the Republic of Croatia
Ministry of Foreign and European Affairs of the Republic of Croatia
Ministry of Tourism and Sport of the Republic of Croatia
Ministry of Agriculture of the Republic of Croatia
Central State Office for the Development of Digital Society
Primorje-Gorski kotar County
City of Rijeka
City of Opatija
Croatian Regulatory Authority for Network Industries - HAKOM
Croatian Power Exchange – CROPEX

patrons

University of Zagreb
University of Rijeka
Juraj Dobrila University of Pula
Ruđer Bošković Institute, Zagreb
Jožef Stefan Institute, Ljubljana, Slovenia
University of Zagreb, Faculty of Electrical Engineering and Computing
University of Zagreb, Faculty of Organization and Informatics, Varaždin
University of Rijeka, Faculty of Maritime Studies
University of Rijeka, Faculty of Engineering
University of Rijeka, Faculty of Economics and Business
Josip Juraj Strossmayer University of Osijek, Faculty of Education
Zagreb University of Applied Sciences
Croatian Academy of Engineering – HATZ
Cybernetics Society, Rijeka
Ministry of the Sea, Transport and Infrastructure of the Republic of Croatia
Croatian Regulatory Authority for Network Industries - HAKOM
Ericsson Nikola Tesla, Zagreb
T - Croatian Telecom, Zagreb
Končar - Electrical Industry, Zagreb
Hrvatska elektroprivreda, Zagreb

sponsors

Hrvatska elektroprivreda, Zagreb
Končar - Electrical Industry, Zagreb
Telemach, Zagreb
Storm Group, Zagreb
InfoDom, Zagreb
dSPACE Engineering, Zagreb
Ericsson Nikola Tesla, Zagreb
Altpro, Zagreb
Adnet, Zagreb
Siemens Energy, Zagreb
Mjerne Tehnologije, Zagreb

All papers are published in their original form

For Publisher:

Karolj Skala

Publisher:

Croatian Society for Information, Communication and
Electronic Technology – MIPRO
Office: Jadranski trg 1/II, HR-51000 Rijeka, Croatia
Phone/Fax: (+385) 51 423 984

Printed by:

GRAFIK, Rijeka

ISSN 1847-3946

Copyright © 2023 by MIPRO

All rights reserved. No part of this book may be reproduced in any form, nor may be stored in a retrieval system or transmitted in any form, without written permission from the publisher.

CONTENTS

LIST OF PAPER REVIEWERS

LIST OF AUTHORS

FOREWORD

HUMAN-COMPUTER INTERACTION

PAPERS

Measuring State and Trait Attention Control Using Mouse Movements 5
D. Wilson, J.L. Jenkins, J.S. Valacich

Getting What You Paid For: Assessing Participant Experience Parameters for Amazon Mechanical Turk (MTurk) Workers on Survey Response Quality 11
M. Kumar, D. Kim, P.A. Weisgarber, J.S. Valacich, J.L. Jenkins

Detecting Goal-Oriented vs. Browsing Users Through Behavior Analysis 17
J.L. Jenkins, A. Denison, J.S. Valacich, D. Wilson

Physiological and Socio-Behavioral Determinants of Viral Video User Engagement 23
D. Bačić, C.A. Gilstrap, N. Jukić

Exploring Pie Charts and Part-To-Whole Alternatives: Eye-tracking Approach 29
D. Bačić, A. Krbanjević, N. Jukić

Parental Trust in Automated Detection of Cyberpredators 34
L. Brandimarte

Mobile Technologies and Live Streaming Commerce: A Systematic Review and Lexical Analysis 40
C.A. Gilstrap, C.M. Gilstrap

Dehumanized Avatars: Unethical Behavior in the Metaverse 49
C. Castagna, S. Demir, M. Weinmann

Tangible Data Exploration: Creating Card Games for Sensemaking 54
A. Wolff, A. Knutas

Analytics Dashboards and User Behavior: Evidence from GitHub 60
J. Schilpp, F. Pethig, H. Hoehle

Using Hobbies as Proxy for Gamification Player Types 66
S. Rantanen, A. Knutas, J. Kasurinen

An Eye-Tracking Solution Using Consumer Grade Webcams for Potential Concussion Diagnosis and Evaluation 71
N. Caporusso, G. Sanders, B. Thaman, E. Hall

Do I Trust My Medical Professional on a Virtual Consultation? Biosensors, Human Computer Interaction and Trust-Building 77
B. Ye, C. Milewicz, S. Mujumdar, M. Khayum

The Influence of the Internet and Social Networks on Behavioral Problems of Primary and Secondary School Students in Croatia 80
N. Flego, L. Josipović Deranja, T. Orehovački

Cloud Accounting as a Factor in Protection from Cyber Attacks and Theft of Accounting Data 86
V. Vinšalek Stipić, M. Vičić

Principles for Designing for Data Exploration for a Non-Expert Audience
N. Tylosky, A. Knutas, A. Wolff 92

Stranded Away: Implementation and User Experience Evaluation of an Indie Platformer Game Developed Using Unity Engine 96
L. Blašković, A. Žužić, T. Orehovački

Analyzing the Impact of Digitality and AI on HCI: Four Dimensions of Intelligent Digitality That Afford Innovative Designs 102
D. Te'eni

Evaluating the Impact of Different User Engagement Elements on Software Applications 107
M. Hajarian, P. Diaz, I. Aedo

Interaktivna potpuno imerzivna virtualna stvarnost računalne simulacije brodske strojarnice 113
D. Ogrizović

SMART INDUSTRIES AND DIGITAL ECOSYSTEMS

PAPERS

Digitalization of Innovation Networks: Theoretical and Empirical Issues 121
V. Omelyanenko, I. Pidorycheva, H. Shevtsova, O. Omelianenko, N. Rudenko, O. Yurchenko

Sensor System for Real-time Water Quality Monitoring 126
E. Simonoska, D. Capeska Bogatinoska, I. Dimitrievski, R. Malekian

Proposal of Prediction Model for Smart Agriculture Based on IoT Sensor Data 132
J. Fondaj, M. Hamiti, S. Krrabaj, B. Selimi, E. Idrizi

BIM Based Information Management in Renewable Energy Projects 138
M. Ljuban, M. Curavić, L. Budin, I. Duilo, M. Delimar

Case Management Model Design for Obtaining Documentation Using FieldWork 4 RES Application in Renewable Energy Sources Projects 145
Z. Tolvajčić, T. Ugrina, M. Curavić, I. Duilo, L. Budin, M. Delimar

Influence of Personal Recommendations towards Electricity Consumption on Consumer Behavior 150
J. Héjjová, J. Bucko

Advanced and Smart Maintenance of Induction Motors Based on an Algorithm for Motor Faults Recognition 158
S. Tvorić, T. Đuran, V. Jančić, D. Sente

Analytics Use Cases for Landside Traffic Optimization in the Catchment Area of the Airport: Case Study of Zagreb Airport 165
M. Tica, I. Štimac, K. Vidović, S. Vojvodić, I. Stipanović

MICROELECTRONICS, ELECTRONICS AND ELECTRONIC TECHNOLOGY

PAPERS

Back-Hopping in Ultra-Scaled MRAM Cells 175
M. Bendra, S. Fiorentini, J. Ender, R.L. de Orio, T. Hadámek, N.P. Jørstad, B. Pruckner, S. Selberherr, W. Goes, V. Sverdlov

Tunneling Attenuation and Leakage Current in MoS2 Nanoribbon MOSFETs 179
I. Prevarić, M. Matić, M. Poljak

Transport Properties and Device Performance of Quasi-One-Dimensional MoS2 FETs 184
M. Matić, M. Poljak

Bandgap Narrowing in Silicene Nanoribbons with Metal Edge Contacts 189
M. Poljak, M. Matić

Determining Graphene and Substrate Quality from the Coupled Hall Mobility Measurements and Theoretical Modeling 195
K. Japec, M. Matić, R. Lukose, M. Lisker, M. Lukosius, M. Poljak

Stress-dependent MOSFET Model for Use in Circuit Simulations 200
A. Žamboki, L. Gočan, J. Mikulić, N. Bako, G. Schatzberger, T. Marković, A. Barić

Low-Power Frequency-Locked Loop Circuit with Static Frequency Offset Cancellation 206
L. Gočan, A. Žamboki, N. Bako, J. Mikulić, G. Schatzberger, T. Marković, A. Barić

Low-Power CMOS Frequency Comparator 212
M. Kováč, M. Potočný, D. Arbet, R. Ondica, R. Ravasz, V. Stopjaková

Measurement System for Characterization of a Resistor Array in 180-nm CMOS Technology 218
I. Franković, F. Mikić, J. Mikulić, N. Bako, G. Schatzberger, A. Barić

Performance Analysis of 1-MHz Voltage-Controlled Ring Oscillator Designed in 180-nm CMOS Technology for Phase-Locked Loop 224
A. Traživuk, A. Barić

Analysis of Energy Consumption for SPI and I2C Communications in Ultra-Low Power Embedded Systems 229
J. Zidar, I. Aleksi, T. Matić

Genetic Algorithm Based Optimization of Circular Planar Coil Geometry with Homogeneous Magnetic Field Distribution 234
D. Bilandžija, D. Vinko, L. Filipović

Analysis of the Effect of Differential-Mode and Common-Mode Impedance Matching on the Common-Mode Rejection Ratio of a Differential Electro-Optical Voltage Probe 239
H. Štimac, A. Barić

Measurement Setup for Characterizing Immunity of Integrated Circuits to Pulsed Electric Fields Using the IC Stripline 245
C. Cetin, M. Magerl, C. Stockreiter

Characterization of Ground Bounce and Conducted Emissions of Integrated Isolated DC-DC Converter 250
A. Križanić, R. Blečić, A. Broznić, A. Barić

Characterization of Class-E Resonant Converter Operating in MHz-Range 255
I. Krois, A. Barić

Design and Characterization of the Output Network of a High-Current Buck Converter 258
K. Šolaja, R. Blečić, A. Broznić, A. Barić

Developing a Model of Buck Converter for EMI Filter Optimization by Circuit Simulations 263
J. Baća, R. Blečić, A. Broznić, A. Barić

Implementation of Voltage Regulation in a Spread-Spectrum-Clocked Buck Converter 269
J. Kundrata, A. Barić

Uninterruptible Power Supply System for Railway Infrastructure with Power Factor Correction and Immunity to Energy Strike 275
A. Hećimović, L. Patrlj, V. Šunde, Ž. Ban

Reduction of Current Harmonics in Data Center Power System Using Single-Phase Shunt Active Power Filter 281
V. Zeleničić, M. Miletić, K. Raič Raguž, D. Sumina, I. Erceg

OPTOELECTRONICS AND PHOTONICS

INVITED PAPER

Polarisation Based Entanglement Distribution Quantum Networking 291
M.J. Clark, R. Wang, S. Bahrani, M. Peranić, O. Alia, M. Lončarić, Ž. Samec, A. Radman, M. Stipčević, R. Nejabati, D. Simeonidou, J. Rarity, S.K. Joshi

PAPERS

Monolithic Receiver with Nine Single-Photon Avalanche Diodes 297
A. Kuttner, S. Marschner, B. Steindl, K. Schneider-Hornstein, M. Hofbauer, H. Zimmermann

Occupancy Determination by Backscattered Visible Light Sensing 302
O. El Ouahabi, C. Fragner, A.P. Weiss, E. Leitgeb

An Experimental Investigation of Whispering Gallery Mode Microsphere Resonators for Sensing, Lasing, and Nonlinear Optical Applications 308
S. Thekke Thalakkal, D. Ristić, D. Zhivotkov, G. Nunzi Conti, M. Ivanda

DATA SCIENCE AND BIOMEDICAL ENGINEERING

PAPERS

DATA SCIENCE

Missing Values Interpolation in PurpleAir Sensor Data Based on a Correlation with Neighboring Locations Using KNIME Analytics Platform 319
S. Omanović, A. Midžić, Z. Avdagić, D. Pozderac, A. Toroman

Spectral Analysis, Agglomerative, Mean Shift and Affinity Propagation Algorithms, Use on the Content from Social Media for Low-Resource Languages 324
M.H. Hoti, J. Ajdari, X. Zenuni, M. Hamiti

Interactive Redescription Set Mining and Exploration 331
I. Kozjak, M. Mihelčić

Fake News Detection: A Comprehensive Survey 337
E. Tata, J. Ajdari, N. Besimi

Graphical User Interface to Perform Glacier Simulations with PISM 343
M. Urbanč, M. Depolli

Computational Performance Aspects of CROCO-BFM Coupling 349
F. Strniša, M. Vodopivec, G. Kosec

OpenMP Offloading and OpenACC Programming Model Approach for Object-Oriented Plasma Device Algorithms 354
E. Krishnasamy, I. Vasileska, L. Kos, P. Bouvry

Benchmark DPC++ Code and Performance Portability on Heterogeneous Architectures 359
N. Mijić, D. Davidović

A View into the Future of Academic Publishing 366
Z. Šojat, K. Skala

ROBOTICS TECHNOLOGIES AND APPLICATIONS

Biped Robot Walking Based on Deep Reinforcement Learning 375
T. Tadić, P. Ćurković

Human Intention Recognition in Collaborative Environments Using RGB-D Camera 381
J. Marić, L. Petrović, I. Marković

ROS Framework for Distributed Control of Networks of Dynamical Systems 387
M. Rossi, A. Jokić

BIOMEDICAL ENGINEERING

Heart Rate Variability Monitoring with Savvy ECG Sensor during Dental Surgery 397
I. Tomašić, A. Prkić, A. Lesin, D. Kalibović Govorko, N. Tomašić, I. Medvedec Mikić

Optimizing Heartbeat Classification Using Bézier Interpolation 403
M. Gusev, N. Petrovski, L. Tonkovikj

Fetal Monitoring: Multi-Channel Fetal ECG Denoising Based on Artificial Intelligence Approach 409
D.D. Taralunga

ECG Compression Based on Successive Differences 413
M. Gusev, M. Jovanov, A. Shekerov, G. Temelkov

Smart Glasses for Gait Analysis of Parkinson's Disease Patients 419
I. Kiprijanovska, F. Panchevski, S. Stankoski, M. Gjoreski, J. Archer, J. Broulidakis, I. Mavridou, B. Hayes, T. Guerreiro, C. Nduka, H. Gjoreski

Attention-based U-net: Joint Segmentation of Layers and Fluids from Retinal OCT Images 425
M. Melinščak

Recognizing Activities of Daily Living Using Multi-sensor Smart Glasses 431
S. Stankoski, B. Sazdov, J. Broulidakis, I. Kiprijanovska, B. Sofronievski, S. Cox, M. Gjoreski, J. Archer, C. Nduka, H. Gjoreski

Reconstruction of Short Genomic Sequences with Graph Convolutional Networks 437
L. Vrček, X. Bresson, T. Laurent, M. Schmitz, M. Šikić

Fortuna Detects Novel Splicing in Drosophila scRNASeq Data 444
B. Borozan, L. Borozan, D. Ševerdija, D. Matijević, S. Canzar

Using De Novo Metagenome Assembly for Improved Metagenomic Classification 450
J. Lipovac, K. Križanović

Hyperparameter Optimization of Graph Neural Networks for mRNA Degradation Prediction 457
V. Vodilovska, S. Gievska, I. Ivanoska

Examination of Different Representations of Proteins Using Protein Ray-based Descriptor and Deep Learning Models 463
G. Mirceva, A. Naumoski, A. Kulakov

Enabling Hyperparameter-Tuning of AI Models for Healthcare Using the CoE RAISE Unique AI Framework for HPC 469
M. Riedel, C. Barakat, S. Fritsch, M. Aach, J. Busch, A. Lintermann, A. Schuppert, S. Brynjólfsson, H. Neukirchen, M. Book

TELECOMMUNICATIONS & INFORMATION

PAPERS

NEW INFORMATION AND COMMUNICATION TECHNOLOGIES IN PRACTICE

Smart Work Order 481
Z. Bosić, P. Gusić, M. Zubčić, M. Marković, T. Žitnik

Bringing Children Closer to Science and the Universe with New Technology 488
P. Lahovsky, A. Ljubek, K. Mišura, M. Žilak

An Overview of 3D Holographic Visualization Technologies and Their Applications in Education 494
A. Kešelj, K. Žubrinić, M. Miličević, M. Kuzman

COMMUNICATION PROTOCOLS AND SYSTEM PERFORMANCE

An Analysis of the TR069 (CWMP) Protocol 503
I. Bašičević

Implementation of Automated Testing Solution for Voice Services 509
M. Ivančić, N. Štokić, T. Žilić

RFID Inventory Management System Sampling Optimization Based on Zebra Android Framework 515
I. Benke, I. Heđi, E. Ciriković

DATA MANAGEMENT

Study of Data Transfer Nodes Infrastructure in Enabling Big Data Movement between Research and Cloud Storage Networks 523
D. Regvart, M. Mikuc, V. Dakić

Using Low Power Wide Area Networks to Provide out of Band Management for Data Centers 528
J. Redžepagić, Z. Morić, D. Regvart

Establishing the Guidelines for Using Refurbished Hardware in Creating New Data Centers to Lower the Amount of E-Waste 532
J. Redžepagić, V. Dakić, Z. Morić

INFORMATION QUALITY AND SECURITY

The Challenges to Information Security in Croatian Chamber of Economy during COVID-19 Pandemic 539
D. Lučić

Do Croatian Citizens Want the Digitalization of Elections? Differences in Socio-Demographic Characteristics 544
V. Pavlović Vinogradac, B. Cerin, D. Vanđelić, A. Prišćan

PAPA for IoT - Role of Data Accuracy in IoT Deployment and Data-centric Decision-Making 548
M. Vermanen, J. Naskali, V. Harkke, J. Koskinen

DATA PROCESSING AND SIGNAL TRANSMISSION

User Input Search - Custom Motion Estimation Algorithm Optimized for UAVs 557
J. Benjak, D. Hofman

FOGO - An Optimized Fog and Edge Computing Method for VANETs 562
R. Ćorluka, I. Matoš, D. Damjanović, J. Balen

On Signal Attenuation in 5G Higher Frequencies Transmissions with Different Area Scenario 568
T. Iliev, I. Stoyanov, I. Penkov, G. Mihaylov, A. Fazylova, G. Balbayev, I. Beloev, S. Kosunalp

Outage Probability Determining for Wireless Systems in the Presence of Beaulieu-Xie Fading and Co-channel Interference Rayleigh Modeled 574
D. Krstić, S. Suljović, D.S. Gurjar, S. Yadav

COMPUTERS IN EDUCATION

PAPERS

Applying TAM (Technology Acceptance Model) to Predict Effective ICT Use of Preschool Teachers during the COVID-19 Pandemic 583
T. Velki, M. Miočić

Data Acquisition and Corpus Creation for Phishing Detection 589
I. Dunđer, S. Seljan, M. Odak

Information Extraction from Security-Related Datasets 595
S. Seljan, N. Tolj, I. Dunđer

A Gamified Learning Scenario for Introducing Pupils to the Scientific Method through Experimentation and Programming 601
D. Kager, A. Marinšek

The Reality of Digital Teaching of History in Croatia 607
M. Hajdarović

Increasing Students' Motivation and Improving Outcome by Changing the Course Examination from Test to Project on a University Course 611
D. Bele, D. Kučak, Đ. Pašić

Model for Assessing E-Learning Courses Considering Multiple Visual and Technical Indicators 616
M. Bankovska, D. Borissova, K. Rasheva-Yordanova

Teaching Computer Security in a Secondary School 622
J. Mottl

What Do We Know about Learning – Conversations with chatGPT 628
J. Henno, H. Jaakkola, J. Mäkelä

A Treasure Trove at Your Fingertips: Analysis of Contents from Europeana with Regard to Their Copyright Status 634
M. Duić

QR Codes as an Educational Tool for Implementing the BYOD Approach in Physics Lessons 640
M. Drushlyak, Y. Sabadosh, P. Mulesa, E. Diemientiev, A. Yurchenko, O. Semenikhina

Windows Admin GUI Model for Learning PowerShell Commands 646
D. Tuličić, D. Delija, G. Sirovatka, M. Mrkoci

Education of Teachers at Higher Education towards Changing Education Expectations 652
D. Paľová, I. Korobaničová

Simulation of Corporate Work-life as a Part of Higher Education 658
D.Paľová, M. Vejačka

Potential Implementation of Augmented Reality Technology in Education 664
M. Brizar, D. Kažović

Challenges of Adoption of Cloud Computing Solutions in Higher Education: Case Study Republic of Kosovo 669
L. Shala Riza, J. Ajdari, M. Hamiti

Augmenting Student Education Using the RealityScan Application for Generating 3D Content 675
V.V. Kozov, B. Ivanova

Implementation of a Facilitative Approach to Teaching Mathematical Disciplines to Future Mathematics Teachers by Means of the MAPLE Package 681
I.V. Shyshenko, Y.O. Chkana, O.V. Martynenko, O.M. Udovychenko, I.I. Stotskyi, O. Semenikhina

Digital Technologies in the Activities of Physical Culture and Sports Specialists: Features of Professional Training 687
P. Rybalko, O. Shukatka, S. Lazorenko, V. Kyselov, N. Skachedub, Y. Kozeruk Paryshkura

Exploring the Pedagogical Use of AI-Powered Chatbots Educational Perceptions and Practices 692
E. Rakovac Bekeš, V. Galzina

Automatic Evaluation of Student Software Solutions in a Virtualized Environment 698
M. Fabijanić, G. Đambić, B. Skračić, M. Kolarić

Digital Literacy of Students at the Faculty of Humanities and Social Sciences Zagreb: An Empirical Study 704
R. Vrana

Effective Educational Ukrainian Practices of the Formation of Media Literacy 710
Y. Rudenko, S. Ahadzhanova, K. Ahadzhanov-Honsales, O. Bieliaieva, A. Korovai, O. Semenikhina

The Formation of Infomedia Literacy of Students in a Media Tournament 716
M. Yachmenyk, I. Kharchenko, O. Semenog, N. Kyrylenko, M. Ostroha, S. Bohoslavskyi, O. Semenikhina

Digital Technologies in Teaching Physics: An Analysis of Existing Practices 722
A. Yurchenko, Y. Khvorostina, V. Shamonia, M. Soroka, O. Semenikhina

Fostering Critical and Computational Thinking in the Field of Primary and Secondary Education in non-STEM Subjects by Using Data Sets and Applications 728
I. Vlahović, I. Ogrizek Biškupić

Gamification in Education: Building an Escape Room Using VR Technologies 734
A. Staneva, T. Ivanova, K. Rasheva-Yordanova, D. Borissova

The Ubiquity of Mathematics in the Practical Basics of Computing and Programming 740
M. Bednjanec, D. Bednjanec, B. Kuhar

From Classroom to Online Environment – The Comparison Analysis of the E-Learning Standards before and during the COVID-19 Pandemic 744
K. Aleksić-Maslać, P. Vranešić, B. Sinković

Validating a Model of Smart Service System, Supporting Teachers to Create Educational Maze Video Games 749
A. Antonova

STREAM Education - Potential for Engaging Students in Generating Innovative Green Ideas and Development of Transversal Skills 755
K. Dyulgerova, D. Atanasova, M. Milanova

When User Experience Design Meets LMS Desktop Users: How to Mitigate Mobile-Oriented Interface? 761
M. Banek Zorica, J. Klindžić

Survey of Recent Research Topics on Effective Use of Communication Tools in Higher Education 766
M. Nurminen, H.-M. Järvinen, J. Viteli, M. Saari, P. Rantanen

Does Daily Use of Digital Technologies Influence the Reading and Information Literacy of 15-year-old Students? 772
V. Lang, A. Špernjak, A. Šorgo

Self-Assessment of Pre-service English Teachers' Digital Competence 778
T.K. Harakchiyska

Examining the Relationship Between Students' Social Media Usage, Smartphone Checking During Lessons/Studying, and Academic Achievement (GPA) 784
M. Kućar, V. Vidaček Hainš, R. Kovačić

Face-to-Face and Online Communication and Etiquette in Secondary Education 790
E. Martinčević, V. Vidaček-Hainš

Non-Formal Education: Dynamics of Changes in the Popularity of Mass Open Online Courses 795
N. Dehtiarova, O. Zhmud, V. Makarova, O. Hontar, M. Zakharevych

Comparison of Pandemic and Post-Pandemic Education at Higher Education Computer Science Course 800
D. Paľová, J. Štofa, M. Vejačka

AI Comics as Art: Scientific Analysis of the Multimedia Content of AI Comics in Education 806
K. Bedi

Teenagers' Gaming Practices and Their Performance in STEM Courses 810
M. Homen, V. Juričić

Education via ICT Due to Corona Crisis – Our View on Pros and Cons 816
L. Révészová

Comparative Analysis of Student Assessment during the COVID-19 Crisis 821
A. Pongrac Pavlina, K. Pavlina

ICT Support for Successful Face-to-Face and Online Teaching Process 825
T. Alajbeg, M. Sokele, S. Morić, F. Brkić

Programmable Questions in Edgar 831
I. Mekterović, Lj. Brkić, V. Krstić

Inclusion of Digital Competence in Educational Physics Study in Croatia and Slovenia 837
F.M. Grgurin, R. Repnik

The Influence of the COVID 19 to the Future of Learning 843
M. Krašna, T. Bratina

How to Measure Co-Creation in the Digital Environment of Higher Education? 847
T. Mijač, M. Jadrić, M. Ćukušić

Classification of Text, Image and Audio Messages Used for Cyberbullying on Social Medias 853
E. Idrizi, M. Hamiti

International Cooperation for Digital Innovations in Primary Schools 859
B. Denys, B. Klimczuk

The Role of Informatics Textbooks in Primary School in Croatia 863
A. Glavaš, D. Marković

The Use of Smartphones and Tablets for Teaching Biology at a Distance: A Comparison between Rural and Urban Schools 869
V. Lang, A. Šorgo

Using VIDI-X Microcomputers in High School and College Education 873
B. Balon, A. Skendžić, M. Simić

How Well Students Perceive Their Understanding of Logic Programming Course Content? 880
S. Lovrenčić, V. Sekovanić

Prototyp of an Optical - Electrical Measuring System for the Determination of the Gravitational Constant for Teaching Purposes 885
C. Ungermanns, W. Werth, A. Schöndorfer, C. Madritsch

Computers in Education 889
H. Jaakkola, J. Henno, J. Mäkelä

ICT and Changes in Education and Professional Activity of Journalists 896
M. Galić

Physics-Electronics Practical Work Model Accessible and Manipulable Remotely via the Web 901
L. Yade, A.D. Gueye, P.L.T. Sow, A. Ndiaye

How to Use E-Learning App in Education and Presentation of Erasmus Project 906
M. Sertić, T. Ćosić, M. Kerstner

Educational Processor with Single-cycle Instructions SVEU16 911
S. Ribić, K. Hodžić

Student Perspectives on Source-Code Plagiarism: Case Study of Three Programming Courses 917
M. Novak

Utilizing Robotics in Maths in Primary Education: A Case Study 923
A. Goltsiou, C. Sofianopoulou

Web Application for Learning Mathematics 928
N. Stojkovikj, L.K. Lazarova, M. Ananieva

Teachers' Expectations for Digivisio 2030 – A Joint Higher Education Digital Transformation Project in Finland 933
J. Multisilta, T. Mattila, K. Lempinen

Digital Culture and Students' Awareness of Their Participation in This Phenomenon 938
L. Bandić, P. Perić, T. Babić

Students' Perception of Influencer Marketing on Instagram 944
D. Blažević, T. Babić

Tolerance for Disagreement in Communication via Instagram 950
D. Njegovan, B. Frigan, T. Babić

Using Drones in Teaching Computer Science 955
P. Voštinár

A Critical Analysis of Students' Cheating in Online Assessment in Higher Education: Post-COVID-19 Issues and Challenges Related to Conversational Artificial Intelligence 961
G. Bubaš, A. Čižmešija

The Development of Smart E-Portfolio Generator System for University Students: An SDLC Approach 967
A. Shuhaiber, A. Alkuwaiti, M. Alremeithi, H. Almenhali

A Smart Desk: A Smart Solution for Young Students 975
A. Shuhaiber, R. Hiyasat, A. Abuelsamen, N. Ghanem, M. Alhammadi, A. Alghafri

Pair Programming Education Aided by ChatGPT 981
B. Banić, M. Konecki, M. Konecki

Evaluating Learners' Reactions to an Online Training Programme: A Case Study of Maldives National University 986
I. Balaban, J. Badari, I. Ružić

The Impact of Short and Very Short Videos on the Effectiveness of Teaching and the Principles of Their Development 992
J. Žufić, S. Bodiš

Teacher's Motivation for Applying the Unplugged MEMA Method for Early Programming Teaching 996
S. Babić, M. Čičin-Šain

Formativno vrednovanje u online okruženju nastave matematike 1000
R. Soldo, D. Kurtić

Modeliranje linearnom funkcijom u osnovnoj školi 1005
I. Nađ

Rješavanje jednadžbi u osnovnoj školi uz Geogebru 1009
I. Nađ

Industrijski IT certifikati iz područja računalnih mreža 1013
D. Valenčić, M. Miholjek, A. Skendžić

Projekt GAMMA: Igranje ili učenje? 1021
K. Brleković, D. Ivanović-Ižaković

Čarobni svijet snjegovića knjigoljubaca – izrada digitalne slikovnice 1025
S. Barbarić

Stavovi i percepcije učenika i studenata prema plagiranju 1029
M. Mirković, D. Možnik

Medijski tehničari – nositelji projekata za mlade 1035
Lj. Ille, R. Fic

Digitalni alati u inkluzivnom odgoju i obrazovanju 1039
J. Bistrović, T. Pavičić Zajec, T. Ređep

Suvremene metode poučavanja robotike i umjetne inteligencije 1045
I. Ružić

Motivacija nastavnika za usvajanje znanja vezanih uz nove tehnologije 1049
A. Tihomirović, S. Šišić

Primjena e-učenja u visokoškolskoj nastavi: pristupi instrukcijskom dizajnu, pedagoški scenariji, modeli hibridnog poučavanja i razine kompleksnosti u primjeni 1054
G. Bubaš

Umjetna inteligencija kao izvor potpore nastavnicima u pripremi i izvođenju nastave 1060
K. Maček Blažeka

Knjižnica u Glideu 1066
K. Udina

Digitalni alati i filozofija uma: Može li stroj razmišljati? 1070
M. Kerstner, M. Sertić

eTwinning – što kažu učenici? 1075
I. Naranđa

DISTRIBUTED COMPUTING AND CYBER-PHYSICAL SYSTEMS

PAPERS

Comparing the Use of Custom-built and Commercial Off-the-shelf Data Gathering Devices in IoT Systems 1083
M. Saari, J. Harjamäki, M. Nurminen, P. Rantanen

Challenges of Combining Open and Commercial Data Sources in Visitor Mobility Estimations 1088
J. Grönman, P. Rantanen, P. Sillberg, T. Pohjola, T. Jönkkäri

An Overview of Diffusion Models for Text Generation 1094
H. Čeović, M. Šilić, G. Delač, K. Vladimir

An Overview of State-of-the-Art Solutions for Scene Text Detection 1100
M. Džida, D. Vukadin, M. Šilić, G. Delač, K. Vladimir

A Transfer Learning Method for Hate Speech Detection 1106
E. Šmuc, G. Delač, M. Šilić, K. Vladimir

Detection of Shilling Attacks on Collaborative Filtering Recommender Systems by Combining Multiple Random Forest Models 1112
V. Grozdanić, K. Vladimir, G. Delač, M. Šilić

High-Performance Serverless Request Generator: Capable of Generating a Hundred Thousand Requests Per Second 1117
D. Mileski, M. Gusev

Future Dynamic Computing and User Profile Execution Considering User Migration in Social Media Contexts 1123
A. Abiola Periola, A.A. Alonge, K.A. Ogudo

Brief Introduction to Active Noise Control 1130
D. Miljković

Active Noise Control Using Waveform Synthesis with Improved Convergence and Tracking 1136
D. Miljković

Higher-Level Experimental Prototype of a Control Device for Dynamic Beverage-Cooling Process 1141
T. Špoljarić, B. Vuletić Komljen, Lj. Cvitaš, G. Malčić, J. Matuško, M. Miletić

Artificial Intelligence-based Predictive Fuel Blending Control for Flare Gas Mitigation 1147
J. Kir Hromatko, Š. Ileš, R. Huljev, V. Vučković

Program Solution for Regulation and Monitoring of Evaporator Section in Individual Quick-Freezing Process 1153
T. Špoljarić, D. Gadže, I. Šulekić

Electrical Profile of User with Photovoltaic System 1159
I. Petrović, D. Koprivanac, E. Ciriković

Analiza rada fotonaponske elektrane radi optimizacije potrošnje energenata 1163
D. Koprivanac, I. Petrović, I. Heđi

ARTIFICIAL INTELLIGENCE SYSTEMS

PAPERS

ARTIFICIAL INTELLIGENCE THEORY

Binary Dynamic Models of Structural Synthesis of Programs 1173
G.A. Oparin, V.G. Bogdanova, A.A. Pashinin

The Concept of Cognition as Categorization in the Development of New Metaheuristics and Algorithms Inspired by Nature 1179
D. Tuličić, N. Ivković

Speeding up the Solving of Logical Equivalence Checking Problems with Disjunctive Diagrams 1184
V. Kondratiev, I. Otpuschennikov, A. Semenov

Empirical Analysis of the RC2 MaxSAT Algorithm 1190
S. Kochemazov, V. Kondratiev, I. Gribanova

Challenges in Collective Intelligence: A Survey 1196
K. Poje, M. Brčić, M. Kovač, D. Krleža

The Study of the Target Set Selection Problem under Deterministic Linear Threshold Model Using Evolutionary Algorithms 1202
M. Smirnov, S. Kochemazov, A. Semenov

Neuroevolution for the Sustainable Evolution of Neural Networks 1208
E. Otović, J. Lerga, D. Kalafatović, G. Mauša

IMAGE AND VIDEO PROCESSING AND ANALYSIS

Influence of Quality of Pixel Level Annotations on Text Detection Performance in Natural Images 1217
I. Dorkić, M. Brisinello, R. Grbić, M. Herceg

Effects of Applying Identified Road Lane Lines on Vehicle Autopilot Model Driving Performance 1222
V. Diklić, F. Matković, K. Pardon

Estimating a Nonradial Vignetting Shape 1228
D. Potoč, D. Petrinović

The Impact of Image Processing on Perceptual Hash Values 1235
M. Ferenčak, P. Grd, I. Tomičić

Usporedba funkcija gubitka za semantičku segmentaciju objekata u prometu 1241
G. Oreški, S. Aničić

NATURAL LANGUAGE PROCESSING

The Role of Knowledge Management in Transition to Industry 5.0 1249
S. Lovrenčić

Learning Translation Model to Translate Croatian Dialects to Modern Croatian Language 1256
B. Penkova, M. Mitreska, K. Ristov, K. Mishev, M. Simjanoska

Sentiment of the Tweets on Russo-Ukrainian War: The Social Network Analysis 1262
A. Poleksić, S. Martinčić-Ipšić

Compressing Sentence Representation with Maximum Coding Rate Reduction 1269
D. Ševerdija, T. Prusina, A. Jovanović, L. Borozan, J. Maltar, D. Matijević

DistilBERT and RoBERTa Models for Identification of Fake News 1275
A. Kitanovski, M. Toshevska, G. Mirceva

Analysis of Pro-Russian Tweets during Russian Invasion of Ukraine 1280
J. Katalinić

Syllable and Morpheme Segmentation of Macedonian Language 1286
M. Mitreska, K. Zdravkova

Multilingual Named Entity Recognition Solution for Optimizing Parcel Delivery in Online Commerce: Identifying Person and Organization Names 1292
M. Pajas, A. Radovan, I. Ogrizek Biškupić

Augmentacija podataka za klasifikaciju kratkih tekstova 1298
I. Hrga

MACHINE LEARNING APPLICATIONS

Using Generated LIBS Data as a Base for Neural Network Architecture Development 1307
J. Lehtonen, T. Aaltonen

Imitation Learning for Financial Applications 1312
S. Goluža, T. Bauman, T. Kovačević, Z. Kostanjčar

The Comparison of Different Feature Extraction Methods in Musical Instrument Classification 1318
N. Rodin, D. Pinčić, K. Lenac, D. Sušanj

OTHER TOPICS

Driven by Artificial Intelligence (AI) – Improving Operational Efficiency and Competitiveness in Business 1327
K. Šiber Makar

A Survey on Usage of Multimedia Databases for Emotion Elicitation: A Quantitative Report on How Content Diversity Can Improve Performance 1333
M. Horvat, P. Jerčić

AI Chatbot for Job Interview 1340
N. Boudjani, V. Colas, C. Joubert, D. Ben Amor

Consumer Class Side Scanning Sonar Dataset for Human Detection 1346
T. Aaltonen

Network Model of Multiagent Communication of Traffic Inspection for Supervision and Control of Passenger Transportation in Road and City Traffic 1352
A. Đukić, R. Bjelošević, M. Stojčić, M.K. Banjanin

Closed-loop Artificial Pancreas Development: A Review 1358
S. Glumčević, Z. Mašetić, B. Viteškić

INFORMATION SYSTEMS SECURITY

PAPERS

The Forensic Significance of Indexing Applications on the Windows Operating System 1369
I. Špoljarić, D. Delija, G. Sirovatka

Implementation of Biometric Verification of a Fingerprint Whose Image is Taken from a Glass Surface 1373
M. Knezović, M. Žagar, D. Delija, G. Sirovatka, D. Možnik

The Latest Developments and Future Perspectives of Artificial Intelligence Systems for In-Vehicle Communication Intrusion Detection 1379
G. Marvin

Deep Learning within the Web Application Security Scope – Literature Review 1385
M. Kaniški, J. Dobša, D. Kermek

Social Engineering Aspects of Email Phishing: An Overview and Taxonomy 1391
I. Tomičić

Data Leaks to Third Parties in Web Services for Vulnerable Groups 1398
R. Carlsson, S. Rauti, T. Heino

Privacy in Popular Children's Mobile Applications: A Network Traffic Analysis 1403
R. Carlsson, S. Rauti, S. Laato, T. Heino, V. Leppänen

Analysis of DMARC Implementation in Republic of Croatia 1409
D. Pranić

Towards Improving Online Security Awareness Skills with Phishing and Spoofing Labs 1415
A. Kerr, T. Hynninen

Secure, Accessible, Virtual Voting Infrastructure (SAVVI): Reducing Barriers for Disabled and Overseas Voters 1420
T. Selker, J.M. Pelletier

JavaScript Library Version Detection 1430
V. Pagon, B. Skendrović, I. Kovačević, S. Groš

Detecting JavaScript Libraries Using Identifiers and Hashes 1436
S. Lončarević, B. Skendrović, I. Kovačević, S. Groš

Detection and Analysis of Obfuscated and Minified JavaScript in the Croatian Web Space 1442
T. Dujmović, B. Skendrović, I. Kovačević, S. Groš

Forensic Implication of a Cyber-Enabled Fraud Taking Advantage of an Offline Adversary-in-the-Middle (AiTM) Attack 1448
D.O. Lawal, D.W. Gresty, D.E. Gan, T.C. Durojaiye

BUSINESS INTELLIGENCE SYSTEMS

PAPERS

Metrics for Estimating Accuracy, Reliability, and Bias in Peer Assessment 1459
M. Fertalj, Lj. Brkić, I. Mekterović

Exploring the Relationship between Indoor Playrooms and Population in Skopje 1465
A. Naumoski, G. Mirceva, K. Mitreski

Predicting Customer Behavior in Support Channels Using Machine Learning 1470
M. Begović, E. Avdagić-Golub, B. Memić, A. Kosovac

Modeling Tennis Matches Using Monte Carlo Simulations Incorporating Dynamic Parameters 1476
J. Krčadinac, E.M. Maruševec, L. Jerković, I. Kovač, J. Zloić, A. Šarčević, M. Vranić

Data Warehouse-Based Analytical System in Private Higher Education Institution 1482
M. Fabijanić, D. Ružak, A. Novosel, T. Hlupić

Opinion Mining on E-commerce Live Broadcast of Agricultural Products Based on WLDA-Apriori Model 1488
B. Wang, Y. Wang, Z. Wang

CroStats — Visualization of Population in Croatia 1494
V. Perković, F. Belac, K. Boroš, M. Kos, L. Humski, D. Pintar

A New Approach to Semantic Parsing of Metonymic Phrases by a Business Intelligence System 1500
V.A. Fomichov

A Common Pentest Output Schema for Business Intelligence System Ingestion 1506
S. Sharma, J.M. Pelletier, B. Stackpole

Partial SQL Query Assessment 1512
M. Fabijanić, I. Mekterović

DIGITAL ECONOMY AND DIGITAL SOCIETY

PAPERS

Risk Management in Development of Croatian Maritime Cargo Single Window System 1523
S. Aksentijević, E. Tijan, K. Nikolozo, A. Perić Hadžić

E-Accounting System as a Service (E-ASaaS): A Conceptual Overview 1528
N. Vlahović

A Preliminary Survey into the Use of Business Process Tools in Croatian Companies 1534
Lj. Milanović Glavan

The Importance of ITIL4 Adoption for IT Service Management in Insurance Companies 1541
A. Ćerimagić Hasibović, A. Tanović, A. Granulo

Classifying Relevant Causes of Employee Absenteeism Using Machine Learning 1547
I. Fosić, A. Živković, I. Fosić

Application of Blockchain Technology and NFTs in a Museum Environment 1553
M. Bilogrivić, H. Stublić

Digital Strategies of Service Sector Enterprises: Promising B2B Decisions 1559
O. Omelianenko, V. Omelyanenko, O. Kudrina, I. Zihunova, D. Kostyrko, S. Lytvynenko

Factoring Economic Biases Out of Tokenomics 1564
T. Baldo, M. Migliardi

Adaptation of European Enterprises to COVID-19 Pandemic: Cluster Analysis Findings 1570
M. Pejić Bach, B. Jaković, I. Jajić

The 5G-supported Unmanned Aerial Vehicles for Emergency Cases Response 1576
Z. Paladin, Ž. Lukšić, N. Kapidani, M. Montagud, M. Fernández-Dasí, S. Srinidhi, T. Wöllert, G. Boustras

Implementation and Cost-Effectiveness of Blockchain Technology within the IP Ecosystems 1582
P. Karanikić

Factors Affecting the Use of Online Recommendation Systems in E-Commerce in Croatia 1587
V. Šebek, S. Antolović, C. Kramarić

Co-creating Community-level Indicators: Involving Communities in the Digitalization of Energy for Empowering Energy Citizenship 1592
A. Kumar, A. Wolff, B. Naqvi

Exploring Regional Innovation Dynamics Using Directed Acyclic Graphs 1598
F. Molinari, D. Čišić, B. Kovačić

Methodological Framework for the Comparative Evaluation of the Emission Footprint of Virtual Remote Work and Travel by Autonomous Electric Vehicle to the Workplace 1605
M. Berković, A. Omerhodžić, A. Džananović

Internet of Behavior – The Transformation of Customer Relationship Management in Logistics 1611
A. Agatić, E. Tijan, S. Aksentijević, A. Pucihar

Auto-generated Plans and Decisions as an Instrument of Crisis Management 1616
D. Labaš, M. Galić, M. Pejić Bach

The Legal Status of Digital Nomads 1622
H. Arbutina, Z. Šinković, L. Pribisalić

Comparative Analysis of Machine Learning Algorithms on Data Sets of Different Characteristics for Digital Transformation 1628
D. Oreški, I. Pihir, D. Višnjić

Clustering of Croatian Software Development Companies Based on Their Financial Indicators 1634
I. Konecki, D. Oreški, N. Kadoić

Artificial Intelligence in Healthcare Services – Regulation, Implementation and Future Challenges 1640
M. Boban, M. Klarić

INFORMATION AND COMMUNICATION TECHNOLOGY LAW

PAPERS

Comparative Analysis of the AI Regulation of the EU, US and China from a Privacy Perspective 1651
V. Gábor Rádi

Algorithmic Contracts in European and United States Contract Law – A Comparative Legal Analysis 1657
K. Biczysko-Pudełko

Medical Data in the Digital Era - Legal Challenges Related to Providing Information Security, Applying GDPR and Respecting the Professional Secrecy 1662
K. Świtała

Technical Solutions Supporting the Online RTBF in the CJEU and ECHR Jurisprudence 1672
N. Gumzej

Legal Assessment of the National Cybersecurity System in Poland in the Light of the New Developments in the NIS2 Directive 1679
A. Besiekierska

Online Audiovisual Content, Video Sharing Platforms and Regulation under DSA and AVMSD 1683
H. Lisičar, T. Katulić, M. Jurić

Interplay of Gatekeepers' Obligations and Consumer Rights under the Digital Markets Act 1689
I. Kanceljak

Analysis of IoT Devices Security for Household Applications 1695
M. Milenković, R. Nikolić, S. Čelan, P. Petrošanec

Privacy Notice Informativeness: in a Search for Benchmark 1701
M. Alić

DeShame Croatia: Student Reactions and Consequences of Online Sexual Harassment in High Schools 1706
L.Vejmelka, R. Matković, M. Rajter, T. Ramljak

Capacities of Western Balkan Economies (and Their Public Sectors) to Respond to Ransomware Attacks 1713
Đ. Krivokapić, A. Nikolić, I. Živković

Status of Open Data (Sub)Ecosystem in Croatia: National Open Data Portal 1720
A. Musa, P. Đurman

Information Security among SMEs in Hungary - An Overview 1726
N. Mike, E. Krén, T. Kecskeméti

Anti-corruption Measures in the New Croatian Public Administration Office Management Regulation 1731
G. Vojković, T. Katulić

Privacy Paradox and Generation Z 1737
M. Alić, L. Sopić

Zračna luka kao osnova za uvođenje inovativnih tehnoloških rješenja u postupku kreiranja pametnog grada zračne luke 1743
M. Milenković

ENGINEERING EDUCATION

PAPERS

Development of Renewable Energy Course for Electrical Engineering Program 1753
F.R. Shahroury, I. Abuishmais, H.H. Ahmad

Hybrid Agile Approach in Software Engineering Education – A Case Study 1758
M. Kaluža, S. Čandrlić, M. Ašenbrener Katić

Computer-supported Education of the Reserve Officers of the Armed Forces 1764
M. Čolić

Design, Development and Control of a Ball-on-Beam Control System Using Industrial Equipment 1770
N. Muškinja

Using the Matlab Programming Environment in the Course of Student Training in the Electrical Machines Discipline 1776
V.S. Ruseva, A.H. Krasteva

Exploring Pre-scoring Clustering for Short Answer Grading 1782
L. Petricioli, K. Skračić, J. Petrović, P. Pale

Identifying Higher Software Engineering Education's Design-Reality Gaps in Rural India 1787
P. Spiesberger, L. Bürstmayr, R. Vallon, T. Grechenig

Understanding the Teamwork Challenges of Software Engineering Students 1793
L. Dorić, N. Luburić, J. Slivka, A. Kovačević

Stimulating (Open) Data Literacy at the Basis of Society: Approaches for Active Learning and Teaching to Young Children 1799
I. Bosnić, A. Kuveždić Divjak, B. van Loenen

The Relationship between Extracurricular Activities and Student Achievement in the Affective Domain: The Case of a Vocational Electrical Engineering School 1805
D. Delač, D. Purković

Academic Motivation of Sophomore and Junior Electrical Engineering Students 1810
A. Gero, B. Catz

Teaching and Assessing Ethics in a Specialized Professional Skills Course 1814
B. Pejčinović

Levels and Differences in Professional Identity in Engineering Undergraduates 1819
V. Putarek, J. Petrović

Virtual and Augmented Reality in Mechanical Engineering Education 1827
G. Ivanova, A. Ivanov, L. Zdravkov

Video Creation as a Catalyst of Value Change in Engineering Education 1833
R. Habash, N.M. Abdulkadir

Ensuring Global Perspectives within Reading Lists to Increase Students' Engagement 1839
G.J. Collins

Virtual Electric Machines Laboratory, Requirements and Practical Realization 1845
P.J. van Duijsen, D.C. Zuidervliet

Hierarchical Approach in Modeling and Simulation of Power Electronics for Education 1851
P.J. van Duijsen, D.C. Zuidervliet

Teaching Introductory Parallel Programming Using Two-Player Online Games 1857
S. Manoharan, X. Ye

Curriculum Development for Socially Responsible and Sustainably Acting Engineers 1862
M. Derda, M. Wedel, M. Albrecht

Income Prospects of Engineering Graduates in Croatia 1868
A. Tecilazić

Vertikalna i horizontalna usklađenost inženjerskog obrazovanja u Hrvatskoj: rezultati istraživanja EUROGRADUATE 1873
A. Tecilazić, L. Mršić

SOFTWARE AND SYSTEMS ENGINEERING

PAPERS

Role, Importance and Significance of Software Quality 1883
F. Témolé, D. Atanasova

Identification of Code Properties that Support Code Smell Analysis 1889
S. Prokić, N. Luburić, J. Slivka, A. Kovačević

Migrating to a Microservice Architecture: Benefits and Challenges 1895
S. Salii, J. Ajdari, X. Zenuni

Optimization and Parallelization of Object-Relational Mappers 1903
F. Jovanov, V. Zdraveski, M. Gusev, M. Kostoska

A Survey of End-to-End Congestion Mechanisms in the Field of IoT 1909
D. Fonović, S. Sovilj, N. Tanković

Security and Privacy Concerns Associated with the Internet of Things (IoT) and the Role of Adapting Blockchain and Machine Learning - A Systematic Literature Review 1915
E. Leka, E. Hoxha, G. Rexha

Applying Balanced Scorecard in Software Process Improvement: A Case Study of Small Software Organization 1922
B. Karahodža, E. Avdagić-Golub, A. Čolaković

Application of Hybrid Project Management Methodology in Development of Software Systems 1928
S. Urlić, Ž. Car

MIPRO JUNIOR - STUDENT PAPERS

PAPERS

Analyzing Sentiment of Reddit Posts for the Russia-Ukraine War 1939
A. Krivičić, S. Martinčić-Ipšić

Classification of Crimes Using Machine Learning Techniques for National Crime Data 1945
M. Trpchevska, A. Dedinec

Optimizing Vision Transformer Performance with Customizable Parameters 1951
E. Ibrahimović

Cardano - What Is It and How to Start Working with It 1957
S. Barac, I. Botički, G. Perković, V. Radošević, I. Terzić

Keeping Drivers Alert: A Solution for Monitoring Driver Attention in Assisted-Driving Vehicles 1963
R. Jutreša, P. Peer, Ž. Emeršič, J. Kim

Teaching Programming and Microcontrollers with an Arduino Remote Laboratory Application 1968
A. Bukovac, E. Pleše, U. Maravić, P. Petrović, T. Jagušt

Awareness of Croatian Citizens about the Advantages and Disadvantages of IoT Devices and Their Safety 1972
P. Petrošanec, S. Čelan, R. Nikolić, M. Milenković

Device for Testing the Windings of Electrical Machines 1978
M. Lonić, V. Zuppa Bakša, G. Vujisić

Electrical Scheme Digitization Using Deep Learning Methods 1984
M. Putak, V. Zuppa Bakša, A. Bednjanec

Body Fat Percentage Calculation: A Linear Regression Model on Croatian Tennis Players Morphology 1990
M. Vasilj, V. Vučetić, M. Sukreški, D. Bojanjac

Visualization of Three-Dimensional Ultrasound Data 1995
H. Hrženjak, Ž. Mihajlović

Istraživanje izazova i problema suradnje dizajnera i razvojnih programera tijekom procesa razvoja digitalnog proizvoda 2001
A. Ilić, M. Čarapina, V. Uglješić

LIST OF PAPER REVIEWERS

Aaltonen, T. (Finland)
Abiola Periola, A. (South Africa)
Aksentijević, S. (Croatia)
Andročec, D. (Croatia)
Antić, T. (Croatia)
Antolić, Ž. (Croatia)
Ašenbrener Katić, M. (Croatia)
Avbelj, V. (Slovenia)
Babić, J. (Croatia)
Babić, S. (Croatia)
Babić, T. (Croatia)
Bačić, D. (United States)
Bagić Babac, M. (Croatia)
Balaban, I. (Croatia)
Balković, M. (Croatia)
Banek, M. (Croatia)
Banek Zorica, M. (Croatia)
Basoglu, A. (United States)
Benčić, F.M. (Croatia)
Blašković, B. (Croatia)
Blečić, R. (Croatia)
Boban, M. (Croatia)
Bonefačić, D. (Croatia)
Bosnić, I. (Croatia)
Bračun, D. (Slovenia)
Brandimarte, L. (United States)
Bratina, T. (Slovenia)
Brčić, M. (Croatia)
Brkić, K. (Croatia)
Brkić, Lj. (Croatia)
Bronzin, T. (Croatia)
Bubaš, G. (Croatia)
Byrd, M. (United States)
Campbell, D. (United States)
Car, S. (Croatia)
Car, Ž. (Croatia)
Craig, K. (United States)
Crnokić, B. (Bosnia and Herzegovina)
Čavrak, I. (Croatia)
Čeke, D. (Bosnia and Herzegovina)
Čelar, S. (Croatia)
Čerkez, N. (Croatia)
Čičin-Šain, M. (Croatia)
Črnjarić-Žic, N. (Croatia)
Čupić, M. (Croatia)
Ćurko, K. (Croatia)
Darin, T. (Brazil)
Davidović, D. (Croatia)
Delač, G. (Croatia)
Delimar, M. (Croatia)
Depolli, M. (Slovenia)
Dunđer, I. (Croatia)
Đurasević, M. (Croatia)
Eckhardt, A. (Austria)
Etinger, D. (Croatia)
Frid, N. (Croatia)
Friganović, K. (Croatia)
Galinac Grbac, T. (Croatia)
Gernhardt, D. (Croatia)
Glavaš, J. (Croatia)
Glumčević, S. (Bosnia and Herzegovina)
Golfarelli, M. (Italy)
Golub, M. (Croatia)
Golubić, M. (Croatia)
Golubić, S. (Croatia)
Gradišnik, M. (Slovenia)
Gradišnik, V. (Croatia)
Grbić, R. (Croatia)
Grd, P. (Croatia)
Grgić, K. (Croatia)
Grgurina, I. (Croatia)
Grimes, M. (United States)
Grönman, J. (Finland)
Gül, B.K. (Turkey)
Gulić, M. (Croatia)
Gumzej, N. (Croatia)
Hashim, M. (United States)
Henno, J. (Estonia)
Hoić-Božić, N. (Croatia)
Holenko Dlab, M. (Croatia)
Horvat, M. (Croatia)
Horvat, N. (Croatia)
Hrkać, T. (Croatia)
Hrvatić, N. (Croatia)
Hudek, V. (Croatia)
Huljenić, D. (Croatia)
Humski, L. (Croatia)
Ipšić, I. (Croatia)
Ivanjko, E. (Croatia)
Ivašić-Kos, M. (Croatia)
Ivković, N. (Croatia)
Ivšac Pavliša, J. (Croatia)
Jaakkola, H. (Finland)
Jagušt, T. (Croatia)
Jakobović, D. (Croatia)
Jakšić, D. (Croatia)
Jančec, L. (Croatia)
Janeković, I. (Australia)
Jardas, M. (Croatia)
Jazbinšek, V. (Slovenia)
Jelić, L. (Croatia)
Jevtić, D. (Croatia)
Ježić, G. (Croatia)
Jovanović, V. (Netherlands)
Jovičić, E. (Croatia)
Jović, A. (Croatia)
Jukić, N. (United States)
Jurdana, I. (Croatia)
Jurić, M. (Croatia)
Jurković, I. (Croatia)
Kadoić, N. (Croatia)
Kainrath, K. (Austria)

Kalafatić, Z. (Croatia)
Kalpić, D. (Croatia)
Kanceljak, I. (Croatia)
Karadimce, A. (North Macedonia)
Karan, M. (Croatia)
Karanikić, P. (Croatia)
Karuza, M. (Croatia)
Kasurinen, J. (Finland)
Kaštelan, I. (Serbia)
Katulić, A. (Croatia)
Katulić, F. (Croatia)
Katulić, T. (Croatia)
Keto, H. (Finland)
Kim, D. (United States)
Kocev, D. (Slovenia)
Kocijan, K. (Croatia)
Komadina, A. (Croatia)
Konecki, M. (Croatia)
Konecki, M. (Croatia)
Koričić, M. (Croatia)
Kostanjčar, Z. (Croatia)
Kota, M. (Croatia)
Kovač, P. (Croatia)
Kovačević, I. (Croatia)
Kovačić, A. (Croatia)
Kovačić, B. (Croatia)
Kralj, N. (Croatia)
Krašna, M. (Slovenia)
Krivić, P. (Croatia)
Krois, I. (Croatia)
Kumar, M. (United States)
Kunda, I. (Croatia)
Kunštek, N. (Croatia)
Kužina, V. (Croatia)
Lisičar, H. (Croatia)
Lončarić, I. (Croatia)
Lovrenčić, R. (Croatia)
Lučić, D. (Croatia)
Lukač, D. (Germany)
Ljubić, S. (Croatia)
Magdalenić, I. (Croatia)
Mäkinen, T. (Finland)
Mandarić, K. (Croatia)
Mandić, L. (Croatia)
Mandić, T. (Croatia)
Margušić, D. (Croatia)
Marković, I. (Croatia)
Markuš, N. (Croatia)
Marović, M. (Croatia)
Martinčić-Ipšić, S. (Croatia)
Matijević, D. (Croatia)
Mauša, G. (Croatia)
Medved Rogina, B. (Croatia)
Mekovec, R. (Croatia)
Mekterović, I. (Croatia)
Merdjanovska, E. (Croatia)
Mesec, P. (Croatia)
Mihajlović, Ž. (Croatia)
Mihoković, L. (Croatia)
Mihova, P. (Bulgaria)
Mikac, B. (Croatia)
Mikuc, M. (Croatia)
Mikulić, J. (Croatia)
Milanović, I. (Serbia)
Milašinović, B. (Croatia)
Milenković, M. (Croatia)
Miličević, K. (Croatia)
Milišić, P. (Croatia)
Mirceva, G. (North Macedonia)
Mlinarić, D. (Croatia)
Možek, M. (Slovenia)
Mrčela, L. (Croatia)
Mudronja, G. (Croatia)
Multisilta, J. (Finland)
Nain, N. (India)
Nuić, H. (Croatia)
Nurminen, M. (Finland)
Ogrizović, D. (Croatia)
Oreški, G. (Croatia)
Palestri, P. (Italy)
Paľová, D. (Slovakia)
Panov, P. (Slovenia)
Pavlin-Bernardić, N. (Croatia)
Pečar-Ilić, J. (Croatia)
Pejčić, T. (Croatia)
Pejčinović, B. (United States)
Pejić Bach, M. (Croatia)
Pentland, S. (United States)
Perković, T. (Croatia)
Pesek, I. (Slovenia)
Petricioli, L. (Croatia)
Petrović, J. (Croatia)
Pihir, I. (Croatia)
Pintar, D. (Croatia)
Pleština, V. (Croatia)
Poljak, M. (Croatia)
Popović, S. (Croatia)
Potoč, D. (Croatia)
Protrka, N. (Croatia)
Pushkar, L. (Croatia)
Putarek, V. (Croatia)
Radmilović, M. (Serbia)
Radovan, A. (Croatia)
Rafajac, O. (Croatia)
Randić, M. (Croatia)
Rantanen, P. (Finland)
Rashkovska Koceva, A. (Slovenia)
Rašan, I. (Croatia)
Repnik, R. (Slovenia)
Reveszova, L. (Slovakia)
Ribić, S. (Bosnia and Herzegovina)
Rogelj, P. (Slovenia)
Saari, M. (Finland)
Schatten, M. (Croatia)
Schneider, C. (Spain)
Schuetzler, R. (United States)
Semenikhina, O. (Ukraine)
Sillberg, P. (Finland)
Simidjievski, N. (Slovenia)
Skala, K. (Croatia)

Skansi, S.	(Croatia)
Skendžić, A.	(Croatia)
Skračić, K.	(Croatia)
Somova, E.	(Bulgaria)
Sović Kržić, A.	(Croatia)
Sovilj, S.	(Croatia)
Stančin, I.	(Croatia)
Stančin, K.	(Croatia)
Stemberger, W.	(Croatia)
Stipčević, M.	(Croatia)
Stipetić, S.	(Croatia)
Stojčić, M.	(Bosnia and Herzegovina)
Storjak, I.	(Croatia)
Strniša, F.	(Slovenia)
Strugar, I.	(Croatia)
Suligoj, T.	(Croatia)
Sužnjević, M.	(Croatia)
Šekoranja, B.	(Croatia)
Ševerdija, D.	(Croatia)
Šiber Makar, K.	(Croatia)
Šilić, M.	(Croatia)
Šimić, D.	(Croatia)
Škopljanac-Mačina, F.	(Croatia)
Škorić, J.	(Serbia)
Škvorc, D.	(Croatia)
Šoić, R.	(Croatia)
Šojat, Z.	(Croatia)
Štengl, B.	(Croatia)
Štruc, V.	(Slovenia)
Šuligoj, F.	(Croatia)
Tafro, A.	(Croatia)
Tanković, N.	(Croatia)
Telišman Prtenjak, M.	(Croatia)
Tijan, E.	(Croatia)
Tomasović, Ž.	(Croatia)
Tomašić, I.	(Sweden)
Tomičić, I.	(Croatia)
Tomičić Furjan, M.	(Croatia)
Tomičić-Pupek, K.	(Croatia)
Tomić, D.	(Croatia)
Tomić, M.	(Croatia)
Trobec, R.	(Slovenia)
Tuličić, D.	(Croatia)
Urem, F.	(Croatia)
Valacich, J.S.	(United States)
Vejačka, M.	(Slovakia)
Vejmelka, L.	(Croatia)
Velčić, I.	(Croatia)
Vidaček Hainš, V.	(Croatia)
Vidaković, J.	(Croatia)
Vinko, D.	(Croatia)
Vladimir, K.	(Croatia)
Vlahinić, S.	(Croatia)
Vlahović, N.	(Croatia)
Vojković, G.	(Croatia)
Voštinár, P.	(Slovakia)
Vrabič, R.	(Slovenia)
Vranić, M.	(Croatia)
Vražić, M.	(Croatia)
Vrček, N.	(Croatia)
Vrdoljak, B.	(Croatia)
Vrhovski, Z.	(Croatia)
Vukić, Đ.	(Croatia)
Vukmirović, S.	(Croatia)
Vuković, M.	(Croatia)
Vušak, E.	(Croatia)
Wang, K.	(United States)
Weber, D.	(United States)
Weisgarber, P.	(United States)
Wolff, A.	(Finland)
Wu, X.	(Germany)
Wyrzykowski, R.	(Poland)
Yueh, R.	(United States)
Zaharija, G.	(Croatia)
Zakarija, I.	(Croatia)
Zdravkova, K.	(North Macedonia)
Žajdela Hrustek, N.	(Croatia)
Žufić, J.	(Croatia)

AUTHOR INDEX

Aach, M. 469
Aaltonen, T. 1307, 1346
Abdulkadir, N.M. 1833
Abiola Periola, A. 1123
Abuelsamen, A. 975
Abuishmais, I. 1753
Aedo, I. 107
Agatić, A. 1611
Ahadzhanova, S. 710
Ahadzhanov-Honsales, K. 710
Ahmad, H.H. 1753
Ajdari, J. 324, 337, 669, 1895
Aksentijević, S. 1523, 1611
Alajbeg, T. 825
Albrecht, M. 1862
Aleksi, I. 229
Aleksić-Maslać, K. 744
Alghafri, A. 975
Alhammadi, M. 975
Alia, O. 291
Alić, M. 1701, 1737
Alkuwaiti, A. 967
Almenhali, H. 967
Alonge, A.A. 1123
Alremeithi, M. 967
Ananieva, M. 928
Aničić, S. 1241
Antolović, S. 1587
Antonova, A. 749
Arbet, D. 212
Arbutina, H. 1622
Archer, J. 419, 431
Ašenbrener Katić, M. 1758
Atanasova, D. 755, 1883
Avdagić, Z. 319
Avdagić-Golub, E. 1470, 1922
Babić, S. 996
Babić, T. 938, 944, 950
Bačić, D. 23, 29
Baća, J. 263
Badari, J. 986
Bahrani, S. 291
Bako, N. 200, 206, 218
Balaban, I. 986
Balbayev, G. 568
Baldo, T. 1564
Balen, J. 562
Balon, B. 873
Ban, Ž. 275
Bandić, L. 938
Banek Zorica, M. 761
Banić, B. 981
Bankovska, M. 616
Banjanin, M.K. 1352
Barac, S. 1957
Barakat, C. 469
Barbarić, S. 1025
Barić, A. 200, 206, 218, 224, 239, 250, 255, 258, 263, 269
Bašičević, I. 503
Bauman, T. 1312
Bedi, K. 806
Bednjanec, A. 1984
Bednjanec, D. 740
Bednjanec, M. 740
Begović, M. 1470
Belac, F. 1494
Bele, D. 611
Beloev, I. 568
Ben Amor, D. 1340
Bendra, M. 175
Benke, I. 515
Benjak, J. 557
Berković, M. 1605
Besiekierska, A. 1679
Besimi, N. 337
Biczysko-Pudełko, K. 1657
Bieliaieva, O. 710
Bilandžija, D. 234
Bilogrivić, M. 1553
Bistrović, J. 1039
Bjelošević, R. 1352
Blašković, L. 96
Blažević, D. 944
Blečić, R. 250, 258, 263
Boban, M. 1640
Bodiš, S. 992
Bogdanova, V.G. 1173
Bohoslavskyi, S. 716
Bojanjac, D. 1990
Book, M. 469
Borissova, D. 616, 734
Boroš, K. 1494
Borozan, B. 444
Borozan, L. 444, 1269
Bosić, Z. 481
Bosnić, I. 1799
Botički, I. 1957
Boudjani, N. 1340
Boustras, G. 1576
Bouvry, P. 354
Brandimarte, L. 34
Bratina, T. 843
Brčić, M. 1196
Bresson, X. 437
Brisinello, M. 1217
Brizar, M. 664
Brkić, F. 825
Brkić, Lj. 831, 1459
Brleković, K. 1021
Broulidakis, J. 419, 431
Broznić, A. 250, 258, 263
Brynjólfsson, S. 469
Bubaš, G. 961, 1054
Bucko, J. 150
Budin, L. 138, 145
Bukovac, A. 1968
Bürstmayr, L. 1787
Busch, J. 469

Canzar, S. 444
Capeska Bogatinoska, D. 126
Caporusso, N. 71
Car, Ž. 1928
Carlsson, R. 1398, 1403
Castagna, C. 49
Catz, B. 1810
Cerin, B. 544
Cetin, C. 245
Chkana, Y.O. 681
Ciriković, E. 515, 1159
Clark, M.J. 291
Colas, V. 1340
Collins, G.J. 1839
Cox, S. 431
Curavić, M. 138, 145
Cvitaš, Lj. 1141
Čandrlić, S. 1758
Čarapina, M. 2001
Čelan, S. 1695, 1972
Čeović, H. 1094
Čičin-Šain, M. 996
Čišić, D. 1598
Čižmešija, A. 961
Čolaković, A. 1922
Čolić, M. 1764
Ćerimagić Hasibović, A. 1541
Ćorluka, R. 562
Ćosić, T. 906
Ćukušić, M. 847
Ćurković, P. 375
Dakić, V. 523, 532
Damjanović, D. 562
Davidović, D. 359
de Orio, R.L. 175
Dedinec, A. 1945
Dehtiarova, N. 795
Delač, D. 1805
Delač, G. 1094, 1100, 1106, 1112
Delija, D. 646, 1369, 1373
Delimar, M. 138, 145
Demir, S. 49
Denison, A. 17
Denys, B. 859
Depolli, M. 343
Derda, M. 1862
Diaz, P. 107
Diemientiev, E. 640
Diklić, V. 1222
Dimitrievski, I. 126
Dobša, J. 1385
Dorić, L. 1793
Dorkić, I. 1217
Drushlyak, M. 640
Duić, M. 634
Duilo, I. 138, 145
Dujmović, T. 1442
Dunđer, I. 589, 595
Durojaiye, T.C. 1448
Dyulgerova, K. 755
Džananović, A. 1605
Džida, M. 1100
Đambić, G. 698
Đukić, A. 1352
Đuran, T. 158
Đurman, P. 1720
El Ouahabi, O. 302
Emeršič, Ž. 1963
Ender, J. 175
Erceg, I. 281
Fabijanić, M. 698, 1482, 1512
Fazylova, A. 568
Ferenčak, M. 1235
Fernández-Dasí, M. 1576
Fertalj, M. 1459
Fic, R. 1035
Filipović, L. 234
Fiorentini, S. 175
Flego, N. 80
Fomichov, V.A. 1500
Fondaj, J. 132
Fonović, D. 1909
Fosić, I. 1547
Fosić, I. 1547
Fragner, C. 302
Franković, I. 218
Frigan, B. 950
Fritsch, S. 469
Gábor Rádi, V. 1651
Gadže, D. 1153
Galić, M. 896, 1616
Galzina, V. 692
Gan, D.E. 1448
Gero, A. 1810
Ghanem, N. 975
Gievska, S. 457
Gilstrap, C.A. 23, 40
Gilstrap, C.M. 40
Gjoreski, H. 419, 431
Gjoreski, M. 419, 431
Glavaš, A. 863
Glumčević, S. 1358
Gočan, L. 200, 206
Goes, W. 175
Goltsiou, A. 923
Goluža, S. 1312
Granulo, A. 1541
Grbić, R. 1217
Grd, P. 1235
Grechenig, T. 1787
Gresty, D.W. 1448
Grgurin, F.M. 837
Gribanova, I. 1190
Grönman, J. 1088
Groš, S. 1430, 1436, 1442
Grozdanić, V. 1112
Guerreiro, T. 419
Gueye, A.D. 901
Gumzej, N. 1672
Gurjar, D.S. 574
Gusev, M. 403, 413, 1117, 1903
Gusić, P. 481
Habash, R. 1833
Hadámek, T. 175
Hajarian, M. 107
Hajdarović, M. 607

Hall, E. 71
Hamiti, M. 132, 324, 669, 853
Harakchiyska, T.K. 778
Harjamäki, J. 1083
Harkke, V. 548
Hayes, B. 419
Hećimović, A. 275
Heđi, I. 515, 1163
Heino, T. 1398, 1403
Héjjová, J. 150
Henno, J. 628, 889
Herceg, M. 1217
Hiyasat, R. 975
Hlupić, T. 1482
Hodžić, K. 911
Hoehle, H. 60
Hofbauer, M. 297
Hofman, D. 557
Homen, M. 810
Hontar, O. 795
Horvat, M. 1333
Hoti, M.H. 324
Hoxha, E. 1915
Hrga, I. 1298
Hrženjak, H. 1995
Huljev, R. 1147
Humski, L. 1494
Hynninen, T. 1415
Ibrahimović, E. 1951
Idrizi, E. 132, 853
Ileš, Š. 1147
Ilić, A. 2001
Iliev, T. 568
Ille, Lj. 1035
Ivančić, M. 509
Ivanda, M. 308
Ivanoska, I. 457
Ivanov, A. 1827
Ivanova, B. 675
Ivanova, G. 1827
Ivanova, T. 734
Ivanović-Ižaković, D. 1021
Ivković, N. 1179
Jaakkola, H. 628, 889
Jadrić, M. 847
Jagušt, T. 1968
Jajić, I. 1570
Jaković, B. 1570
Jančić, V. 158
Japec, K. 195
Järvinen, H.-M. 766
Jenkins, J.L. 5, 11, 17
Jerčić, P. 1333
Jerković, L. 1476
Jokić, A. 387
Jönkkäri, T. 1088
Jørstad, N.P. 175
Joshi, S.K. 291
Josipović Deranja, L. 80
Joubert, C. 1340
Jovanov, F. 1903
Jovanov, M. 413
Jovanović, A. 1269
Jukić, N. 23, 29
Juričić, V. 810
Jurić, M. 1683
Jutreša, R. 1963
Kadoić, N. 1634
Kager, D. 601
Kalafatović, D. 1208
Kalibović Govorko, D. 397
Kaluža, M. 1758
Kanceljak, I. 1689
Kaniški, M. 1385
Kapidani, N. 1576
Karahodža, B. 1922
Karanikić, P. 1582
Kasurinen, J. 66
Katalinić, J. 1280
Katulić, T. 1683, 1731
Kažović, D. 664
Kecskeméti, T. 1726
Kermek, D. 1385
Kerr, A. 1415
Kerstner, M. 906, 1070
Kešelj, A. 494
Kharchenko, I. 716
Khayum, M. 77
Khvorostina, Y. 722
Kim, D. 11
Kim, J. 1963
Kiprijanovska, I. 419, 431
Kir Hromatko, J. 1147
Kitanovski, A. 1275
Klarić, M. 1640
Klimczuk, B. 859
Klindžić, J. 761
Knezović, M. 1373
Knutas, A. 54, 66, 92
Kochemazov, S. 1190, 1202
Kolarić, M. 698
Kondratiev, V. 1184, 1190
Konecki, I. 1634
Konecki, M. 981
Konecki, M. 981
Koprivanac, D. 1159, 1163
Korobaničová, I. 652
Korovai, A. 710
Kos, L. 354
Kos, M. 1494
Kosec, G. 349
Koskinen, J. 548
Kosovac, A. 1470
Kostanjčar, Z. 1312
Kostoska, M. 1903
Kostyrko, D. 1559
Kosunalp, S. 568
Kovač, I. 1476
Kovač, M. 1196
Kováč, M. 212
Kovačević, A. 1793, 1889
Kovačević, I. 1430, 1436, 1442
Kovačević, T. 1312
Kovačić, B. 1598
Kovačić, R. 784
Kozeruk Paryshkura, Y. 687

Kozjak, I. 331
Kozov, V.V. 675
Kramarić, C. 1587
Krasteva, A.H. 1776
Krašna, M. 843
Krbanjević, A. 29
Krčadinac, J. 1476
Krén, E. 1726
Krishnasamy, E. 354
Krivičić, A. 1939
Krivokapić, Đ. 1713
Križanić, A. 250
Križanović, K. 450
Krleža, D. 1196
Krois, I. 255
Krrabaj, S. 132
Krstić, D. 574
Krstić, V. 831
Kučak, D. 611
Kućar, M. 784
Kudrina, O. 1559
Kuhar, B. 740
Kulakov, A. 463
Kumar, A. 1592
Kumar, M. 11
Kundrata, J. 269
Kurtić, D. 1000
Kuttner, A. 297
Kuveždić Divjak, A. 1799
Kuzman, M. 494
Kyrylenko, N. 716
Kyselov, V. 687
Laato, S. 1403
Labaš, D. 1616
Lahovsky, P. 488
Lang, V. 772, 869
Laurent, T. 437
Lawal, D.O. 1448
Lazarova, L.K. 928
Lazorenko, S. 687
Lehtonen, J. 1307
Leitgeb, E. 302
Leka, E. 1915
Lempinen, K. 933
Lenac, K. 1318
Leppänen, V. 1403
Lerga, J. 1208
Lesin, A. 397
Lintermann, A. 469
Lipovac, J. 450
Lisičar, H. 1683
Lisker, M. 195
Lončarević, S. 1436
Lončarić, M. 291
Lonić, M. 1978
Lovrenčić, S. 880, 1249
Luburić, N. 1793, 1889
Lučić, D. 539
Lukose, R. 195
Lukosius, M. 195
Lukšić, Ž. 1576
Lytvynenko, S. 1559
Ljuban, M. 138
Ljubek, A. 488
Maček Blažeka, K. 1060
Madritsch, C. 885
Magerl, M. 245
Makarova, V. 795
Mäkelä, J. 628, 889
Malčić, G. 1141
Malekian, R. 126
Maltar, J. 1269
Manoharan, S. 1857
Maravić, U. 1968
Marić, J. 381
Marinšek, A. 601
Marković, D. 863
Marković, I. 381
Marković, M. 481
Marković, T. 200, 206
Marschner, S. 297
Martinčević, E. 790
Martinčić-Ipšić, S. 1262, 1939
Martynenko, O.V. 681
Maruševec, E.M. 1476
Marvin, G. 1379
Mašetić, Z. 1358
Matić, M. 179, 184, 189, 195
Matić, T. 229
Matijević, D. 444, 1269
Matković, F. 1222
Matković, R. 1706
Matoš, I. 562
Mattila, T. 933
Matuško, J. 1141
Mauša, G. 1208
Mavridou, I. 419
Medvedec Mikić, I. 397
Mekterović, I. 831, 1459, 1512
Melinščak, M. 425
Memić, B. 1470
Midžić, A. 319
Migliardi, M. 1564
Mihajlović, Ž. 1995
Mihaylov, G. 568
Mihelčić, M. 331
Miholjek, M. 1013
Mijač, T. 847
Mijić, N. 359
Mike, N. 1726
Mikić, F. 218
Mikuc, M. 523
Mikulić, J. 200, 206, 218
Milanova, M. 755
Milanović Glavan, Lj. 1534
Milenković, M. 1695, 1743, 1972
Mileski, D. 1117
Miletić, M. 281
Miletić, M. 1141
Milewicz, C. 77
Miličević, M. 494
Miljković, D. 1130, 1136
Miočić, M. 583
Mirceva, G. 463, 1275, 1465
Mirković, M. 1029
Mishev, K. 1256

Mišura, K. 488
Mitreska, M. 1256, 1286
Mitreski, K. 1465
Molinari, F. 1598
Montagud, M. 1576
Morić, S. 825
Morić, Z. 528, 532
Mottl, J. 622
Možnik, D. 1029, 1373
Mrkoci, M. 646
Mršić, L. 1873
Mujumdar, S. 77
Mulesa, P. 640
Multisilta, J. 933
Musa, A. 1720
Muškinja, N. 1770
Nađ, I. 1005, 1009
Naqvi, B. 1592
Naranđa, I. 1075
Naskali, J. 548
Naumoski, A. 463, 1465
Ndiaye, A. 901
Nduka, C. 419, 431
Nejabati, R. 291
Neukirchen, H. 469
Nikolić, A. 1713
Nikolić, R. 1695, 1972
Nikolozo, K. 1523
Novak, M. 917
Novosel, A. 1482
Nunzi Conti, G. 308
Nurminen, M. 766, 1083
Njegovan, D. 950
Odak, M. 589
Ogrizek Biškupić, I. 728, 1292
Ogrizović, D. 113
Ogudo, K.A. 1123
Omanović, S. 319
Omelianenko, O. 121, 1559
Omelyanenko, V. 121, 1559
Omerhodžić, A. 1605
Ondica, R. 212
Oparin, G.A. 1173
Orehovački, T. 80, 96
Oreški, D. 1628, 1634
Oreški, G. 1241
Ostroha, M. 716
Otović, E. 1208
Otpuschennikov, I. 1184
Pagon, V. 1430
Pajas, M. 1292
Paladin, Z. 1576
Pale, P. 1782
Paľová, D. 652, 658, 800
Panchevski, F. 419
Pardon, K. 1222
Pashinin, A.A. 1173
Pašić, Đ. 611
Patrlj, L. 275
Pavičić Zajec, T. 1039
Pavlina, K. 821
Pavlović Vinogradac, V. 544
Peer, P. 1963
Pejčinović, B. 1814
Pejić Bach, M. 1570, 1616
Pelletier, J.M. 1420, 1506
Penkov, I. 568
Penkova, B. 1256
Peranić, M. 291
Perić Hadžić, A. 1523
Perić, P. 938
Perković, G. 1957
Perković, V. 1494
Pethig, F. 60
Petricioli, L. 1782
Petrinović, D. 1228
Petrošanec, P. 1695, 1972
Petrović, I. 1159, 1163
Petrović, J. 1782, 1819
Petrović, L. 381
Petrović, P. 1968
Petrovski, N. 403
Pidorycheva, I. 121
Pihir, I. 1628
Pinčić, D. 1318
Pintar, D. 1494
Pleše, E. 1968
Pohjola, T. 1088
Poje, K. 1196
Poleksić, A. 1262
Poljak, M. 179, 184, 189, 195
Pongrac Pavlina, A. 821
Potoč, D. 1228
Potočný, M. 212
Pozderac, D. 319
Pranić, D. 1409
Prevarić, I. 179
Pribisalić, L. 1622
Prišćan, A. 544
Prkić, A. 397
Prokić, S. 1889
Pruckner, B. 175
Prusina, T. 1269
Pucihar, A. 1611
Purković, D. 1805
Putak, M. 1984
Putarek, V. 1819
Radman, A. 291
Radošević, V. 1957
Radovan, A. 1292
Raič Raguž, K. 281
Rajter, M. 1706
Rakovac Bekeš, E. 692
Ramljak, T. 1706
Rantanen, P. 766, 1083, 1088
Rantanen, S. 66
Rarity, J. 291
Rasheva-Yordanova, K. 616, 734
Rauti, S. 1398, 1403
Ravasz, R. 212
Redžepagić, J. 528, 532
Ređep, T. 1039
Regvart, D. 523, 528
Repnik, R. 837
Révészová, L. 816
Rexha, G. 1915

Ribić, S. 911
Riedel, M. 469
Ristić, D. 308
Ristov, K. 1256
Rodin, N. 1318
Rossi, M. 387
Rudenko, N. 121
Rudenko, Y. 710
Ruseva, V.S. 1776
Ružak, D. 1482
Ružić, I. 986, 1045
Rybalko, P. 687
Saari, M. 766, 1083
Sabadosh, Y. 640
Salii, S. 1895
Samec, Ž. 291
Sanders, G. 71
Sazdov, B. 431
Schatzberger, G. 200, 206, 218
Schilpp, J. 60
Schmitz, M. 437
Schneider-Hornstein, K. 297
Schöndorfer, A. 885
Schuppert, A. 469
Sekovanić, V. 880
Selberherr, S. 175
Selimi, B. 132
Selker, T. 1420
Seljan, S. 589, 595
Semenikhina, O. 640, 681, 710, 716, 722
Semenog, O. 716
Semenov, A. 1184, 1202
Sente, D. 158
Sertić, M. 906, 1070
Shahroury, F.R. 1753
Shala Riza, L. 669
Shamonia, V. 722
Sharma, S. 1506
Shekerov, A. 413
Shevtsova, H. 121
Shuhaiber, A. 967, 975
Shukatka, O. 687
Shyshenko, I.V. 681
Sillberg, P. 1088
Simeonidou, D. 291
Simić, M. 873
Simjanoska, M. 1256
Simonoska, E. 126
Sinković, B. 744
Sirovatka, G. 646, 1369, 1373
Skachedub, N. 687
Skala, K. 366
Skendrović, B. 1430, 1436, 1442
Skendžić, A. 873, 1013
Skračić, B. 698
Skračić, K. 1782
Slivka, J. 1793, 1889
Smirnov, M. 1202
Sofianopoulou, C. 923
Sofronievski, B. 431
Sokele, M. 825
Soldo, R. 1000
Sopić, L. 1737
Soroka, M. 722
Sovilj, S. 1909
Sow, P.L.T. 901
Spiesberger, P. 1787
Srinidhi, S. 1576
Stackpole, B. 1506
Staneva, A. 734
Stankoski, S. 419, 431
Steindl, B. 297
Stipanović, I. 165
Stipčević, M. 291
Stockreiter, C. 245
Stojčić, M. 1352
Stojkovikj, N. 928
Stopjaková, V. 212
Stotskyi, I.I. 681
Stoyanov, I. 568
Strniša, F. 349
Stublić, H. 1553
Sukreški, M. 1990
Suljović, S. 574
Sumina, D. 281
Sušanj, D. 1318
Sverdlov, V. 175
Świtała, K. 1662
Šarčević, A. 1476
Šebek, V. 1587
Ševerdija, D. 444, 1269
Šiber Makar, K. 1327
Šikić, M. 437
Šilić, M. 1094, 1100, 1106, 1112
Šinković, Z. 1622
Šišić, S. 1049
Šmuc, E. 1106
Šojat, Z. 366
Šolaja, K. 258
Šorgo, A. 772, 869
Špernjak, A. 772
Špoljarić, I. 1369
Špoljarić, T. 1141, 1153
Štimac, H. 239
Štimac, I. 165
Štofa, J. 800
Štokić, N. 509
Šulekić, I. 1153
Šunde, V. 275
Tadić, T. 375
Tanković, N. 1909
Tanović, A. 1541
Taralunga, D.D. 409
Tata, E. 337
Te'eni, D. 102
Tecilazić, A. 1868, 1873
Temelkov, G. 413
Témolé, F. 1883
Terzić, I. 1957
Thaman, B. 71
Thekke Thalakkal, S. 308
Tica, M. 165
Tihomirović, A. 1049
Tijan, E. 1523, 1611
Tolvajčić, Z. 145
Tolj, N. 595

Tomašić, I. 397
Tomašić, N. 397
Tomičić, I. 1235, 1391
Tonkovikj, L. 403
Toroman, A. 319
Toshevska, M. 1275
Traživuk, A. 224
Trpchevska, M. 1945
Tuličić, D. 646, 1179
Tvorić, S. 158
Tylosky, N. 92
Udina, K. 1066
Udovychenko, O.M. 681
Uglješić, V. 2001
Ugrina, T. 145
Ungermanns, C. 885
Urbanč, M. 343
Urlić, S. 1928
Valacich, J.S. 5, 11, 17
Valenčić, D. 1013
Vallon, R. 1787
van Duijsen, P.J. 1845, 1851
van Loenen, B. 1799
Vanđelić, D. 544
Vasileska, I. 354
Vasilj, M. 1990
Vejačka, M. 658, 800
Vejmelka, L. 1706
Velki, T. 583
Vermanen, M. 548
Vičić, M. 86
Vidaček Hainš, V. 784, 790
Vidović, K. 165
Vinko, D. 234
Vinšalek Stipić, V. 86
Višnjić, D. 1628
Viteli, J. 766
Viteškić, B. 1358
Vladimir, K. 1094, 1100, 1106, 1112
Vlahović, I. 728
Vlahović, N. 1528
Vodilovska, V. 457
Vodopivec, M. 349
Vojković, G. 1731
Vojvodić, S. 165
Voštinár, P. 955
Vrana, R. 704
Vranešić, P. 744
Vranić, M. 1476
Vrček, L. 437
Vučetić, V. 1990
Vučković, V. 1147
Vujisić, G. 1978
Vukadin, D. 1100
Vuletić Komljen, B. 1141
Wang, B. 1488
Wang, R. 291
Wang, Y. 1488
Wang, Z. 1488
Wedel, M. 1862
Weinmann, M. 49
Weisgarber, P.A. 11
Weiss, A.P. 302
Werth, W. 885
Wilson, D. 5, 17
Wolff, A. 54, 92, 1592
Wöllert, T. 1576
Yachmenyk, M. 716
Yadav, S. 574
Yade, L. 901
Ye, B. 77
Ye, X. 1857
Yurchenko, A. 640, 722
Yurchenko, O. 121
Zakharevych, M. 795
Zdraveski, V. 1903
Zdravkov, L. 1827
Zdravkova, K. 1286
Zeleničić, V. 281
Zenuni, X. 324, 1895
Zhivotkov, D. 308
Zhmud, O. 795
Zidar, J. 229
Zihunova, I. 1559
Zimmermann, H. 297
Zloić, J. 1476
Zubčić, M. 481
Zuidervliet, D.C. 1845, 1851
Zuppa Bakša, V. 1978, 1984
Žagar, M. 1373
Žamboki, A. 200, 206
Žilak, M. 488
Žilić, T. 509
Žitnik, T. 481
Živković, A. 1547
Živković, I. 1713
Žubrinić, K. 494
Žufić, J. 992
Žužić, A. 96

FOREWORD

The *46th ICT and Electronics Convention MIPRO 2023* was held from May 22nd to May 26th, 2023 in Opatija, Croatia. The Convention consisted of fifteen conferences under the titles: *Human-Computer Interaction (HCI), Smart Industries and Digital Ecosystems (SIDE), Microelectronics, Electronics and Electronic Technology (MEET), Optoelectronics and Photonics (OPHO), Data Science and Biomedical Engineering (DS-BE), Telecommunications & Information (CTI), Computers in Education (CE), Distributed Computing and Cyber-Physical Systems (DC-CPS), Artificial Intelligence Systems (AIS), Information Systems Security (ISS), Business Intelligence Systems (miproBIS), Digital Economy and Digital Society (DE-DS), Information and Communication Technology Law (ICTLAW), Engineering Education (EE)* and *Software and Systems Engineering (SSE)*. An additional conference was dedicated to the works of students: *MIPRO Junior-Student Papers (SP)*.

This comprehensive Book of Proceedings includes all positively reviewed and presented papers from each of these conferences. All the papers were reviewed by an International review board and the list of reviewers is contained in the Book of Proceedings. These papers were written by the authors from universities, scientific institutions, educational institutions, industry and other participants.

The convention was organized by the Croatian ICT Society MIPRO with the help of numerous patrons and sponsors to whom we owe our sincere thanks. We especially single out our general sponsor Hrvatska elektroprivreda, golden sponsor Končar - Electrical Industry along with silver sponsor Telemach and bronze sponsors Storm Group, Ericsson Nikola Tesla, dSPACE and Infodom.

To all who helped organizing the 46th ICT and Electronics Convention *MIPRO 2023* as well as editing of this Book of Proceedings we extend our heartfelt thanks.

Academician Karolj Skala, PhD
International Program Committee
General Chair

HCI

International Conference on

HUMAN-COMPUTER INTERACTION

Steering Committee

Chairs: Dragan Čišić, University of Rijeka, Croatia
Joe Valacich, University of Arizona, Tucson, United States
Jeffrey Jenkins, Brigham Young University, Provo, United States

Members: Dinko Bačić, Loyola University Chicago, United States
Laura Brandimarte, University of Arizona, Tucson, United States
Constantinos K. Coursaris, HEC Montréal, Canada
Nenad Jukić, Loyola University Chicago, United States
Božidar Kovačić, University of Rijeka, Croatia
Dario Ogrizović, University of Rijeka, Croatia
Saonee Sarker, Lund University, Sweden
Christoph Schneider, IESE Business School, Barcelona, Spain
Chee-Wee Tan, Copenhagen Business School, Denmark
Jason Thatcher, Temple University, United States
Markus Weinmann, University of Cologne, Germany
David Wilson, Brigham Young University Marriott, Highland, United States

PAPERS

Measuring State and Trait Attention Control Using Mouse Movements

D. Wilson*, J. L. Jenkins*, J. S. Valacich**
* Brigham Young University, Provo, UT, USA
** University of Arizona, Tucson, AZ, USA
davidwilson@byu.edu

***Abstract* - Attention is a scarce resource in an increasingly distracting world. Although computing devices might well be blamed for much of the distraction we experience, these devices also afford an opportunity for feedback and improvement, as seen in a variety of contexts. This paper provides initial evidence that mouse movement data can be used to measure users' attention control. Attention control refers to the ability to maintain top-down attention to goal-relevant stimuli and resist the distracting influence of irrelevant stimuli competing for attention. During a controlled experiment with conditions of varying levels of distraction, we captured participants' mouse movements to determine whether mouse movement data can be used to detect both state and trait attention control. We discuss various applications of this approach as a feedback mechanism in situations where maintaining focus is an important determinant of system success, such as in online learning or training scenarios.**

Keywords – attention, distraction, mouse tracking, behavior analysis, human-computer interaction

I. INTRODUCTION

Attention is a scarce resource in the modern economy [1], becoming scarcer all the time as devices, feeds, and notifications compete for attention in an increasingly crowded technological landscape [2]. Among other effects, the proliferation of these technologies has had a general impact on many users' ability to manage and focus their attention [3], driven mainly by a near-constant barrage of distractions [4].

Even without an array of distractions vying for attention, our species' inability to sustain focused attention is well documented across a broad range of stimuli and environments [5]. Maintaining focus on a particular goal is effortful, requiring executive functions that consume cognitive resources such as working memory [6]. When those resources are depleted—because of distractions or from frequent task switching—cognitive performance suffers [7, 8].

Although technology certainly plays a role in this attention problem, it may also be part of the solution. Research has demonstrated that real-time feedback about whether one is becoming distracted can help focus a wandering mind [9] and attention feedback systems have been studied in a variety of contexts, from distracted driving [10] to the workplace [11, 12].

A key drawback of existing attention feedback systems is their reliance on dedicated sensors to detect and monitor attention or focus. These include eye tracking [12, 13], functional magnetic resonance imaging (fMRI) [9, 14], or electroencephalogram (EEG) devices [15]. Although effective in providing real-time monitoring of attention, these technologies may not be practical for more general use, particularly in everyday web contexts.

This paper describes an alternative approach to real-time attention monitoring using mouse movements and describes the results of an initial test of the approach in a web-based setting. Based on the concept of attention control (AC) [16] and drawing on dual-process theory [17], we hypothesize that mouse movement data can be used to measure trait AC (i.e., an individual's general propensity to maintain attention during goal-directed tasks) and to detect changes in state AC (i.e., to determine when an individual has become distracted during a given task).

In summary, the research question that guides this research is:

Can mouse-cursor movements be used to measure both trait and state attention control during a goal-directed task?

We address this research question by first summarizing prior research on attention feedback systems and related applications of mouse movement data. We then use dual-process theory and the response activation model to explain how and why mouse cursor movements can serve as a proxy for attention control. We summarize a controlled experiment that evaluates our hypotheses and conclude with a discussion of the results of our study and implications for future research.

II. PREVIOUS RESEARCH

Dual-process theories divide mental processing into two broad systems or types [18]: one that is reflexive, automatic, impulsive, and habit-driven, and another that is rational, logical, controlled, and cognitive [19]. Higher cognition—referred to as *executive control* [20]—results from the second system directing attention to the task at hand and, when necessary, intervening to override irrelevant distractions that activate the lower-level, automatic system [21, 22].

Based on this dual-process view, many attention monitoring systems measure neurological activity, watching for signs that mental processing has shifted from the higher-level system to the reflexive, automatic system, indicating that the individual has become distracted. For example, EEG devices can detect sustained attention and provide real-time feedback to the wearer, significantly increasing focused attention [15]. Similar EEG monitoring and feedback systems have shown promise in improving outcomes in real-world scenarios, including therapeutic settings [23] and in the workplace [11].

Another application of attention monitoring and feedback uses fMRI to provide real-time feedback that indicates when someone has lost focus—a phenomenon known as mind wandering [24]. Researchers monitored brain-imaging data in real time and intervened when lapses in attention were detected, improving the focus and performance of study participants who received this "neurofeedback" [14].

Although these monitoring techniques provide relatively direct measurement of attention and focus, they are not practical for many everyday situations, despite the increasing availability of more wearable EEG sensors [e.g., 25]. By contrast, mouse movement data is ubiquitous in desktop computing situations and requires no specialized hardware. Mouse data has been used to infer attention in web-based settings [e.g., 26, 27], though much of this prior work has used mousing data as a way to measure and improve search engine results [e.g., 28, 29] or has been exploratory in nature without incorporating a theoretical lens to understand the mechanisms through which attention and distraction affect mouse movements [30, 31].

However, prior research has demonstrated that mouse movement data also reveals dynamic cognitive processes [32, 33], based on the notion that cognitive demands compromise motor function. Thus, mouse movements can detect deception [34], frustration [35], or fraud [36], using methods compatible with real-time monitoring that can detect state changes [37]. As dual-process theory argues, maintaining attention during goal-directed behavior is a process that evolves over time, whether because of a wind-wandering loss of focus or because of distractions. Thus, the present research explores whether mouse movements could be used to detect differences in attention, as informed by the cognitive resource perspective detailed in the next section.

III. Theory and Hypotheses

Dual-process theory explains that goal-directed behavior is guided by the intentional application of attention to a given task [20]. This process of *attention control*—defined as maintaining focus on goal-relevant information, particularly in the face of distraction [16]—requires cognitive effort and is compromised when cognitive resources are strained [38-40].

Individuals differ in their capacity for sustaining the cognitive resources AC demands [41]. Those who measure higher in this individual trait are less susceptible to distraction [42] and tend to perform better in, for example, academic settings [43]. Individuals can also improve their trait AC through practice or training, becoming less susceptible to distractions and better able to sustain focus over time [44, 45].

Distinct from trait AC, state AC is a measure of an individual's ability to maintain attention during a given task and can be influenced by various aspects that influence the cognitive demands of the task. For example, a perceived threat [46] or other negative emotion [38] will decrease state AC, and distractions or other sources of cognitive burden during a task will temporarily deplete an individual's AC [40, 47]. As prior research has shown that trait and state AC are distinct concepts [45], we treat them separately in this work, exploring whether mouse movement data can serve as a proxy measure for one's trait AC (see H1) as well as measure differences in state AC resulting from experimentally induced distraction (see H2).

Dual-process theory is used to explain the underlying mechanisms linking attention and distraction to motor movements, most notably in the Response Activation Model (RAM) [48]. The RAM argues that as a person interacts with the surrounding environment, stimuli with actionable potential are evaluated and selectively inhibited as movement toward goal-relevant stimuli is ultimately accomplished. Thus, the attention and movement systems are intimately linked [48]—perhaps resulting from their co-evolutionary roots [49]—and monitoring hand movements can reliably reveal many different attention-related processes [50].

A growing body of prior mousing research has demonstrated that cognitive demands result in slower, more deviant, inefficient mouse movements, whether because of negative emotions [35], deception [34], or various forms of cognitive conflict [33]. Following this prior work and the dual-process RAM [48], we argue that a mouse user's ability to maintain focus—both in general and in the face of temporary distractions—can be inferred from that user's mouse movement patterns. Specifically, we hypothesize that higher AC will be characterized by faster, more efficient (i.e., with shorter distances), while low AC will be associated with slow, distracted, inefficient mousing patterns. These assumptions are incorporated for trait and state AC separately in the following hypotheses:

H1: Mouse movements of individuals with low trait AC will be (a) slower and (b) cover more distance than individuals with high trait AC.

H2: Mouse movements of individuals in a distracted state will be (a) slower and (b) cover more distance than individuals who are not in a distracted state.

IV. Methodology

We tested our hypotheses using a within-subjects experiment in which participants completed a simple mouse-clicking task on a website created by the research team.

A. Procedure

The task was adapted from a common experimental task in the attention research literature in which

participants perform a visual search to identify a single target among several distractors [51]. A typical mousing–attention study uses highly controlled tasks with a single movement path from a starting location to a target [52]. We expanded and adapted the task to allow for more extensive mousing behavior that more closely approximates real-world web environments. Specifically, each trial asked the participant to click on a series of four numbered targets in sequential order. While the participant clicked the targets, the interface either displayed the numbered targets alone (high attention control) or also displayed four distractor targets and a distractor image to the right of the target area (low attention control). The distractor targets were visually identical to the numbered targets except that they contained letters. (The interfaces for the low AC and high AC treatments are shown in Figures 1 and 2.) The order of the conditions was randomly counterbalanced across the participants.

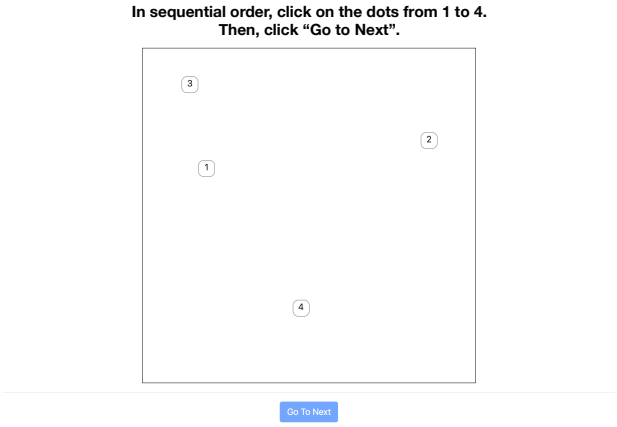


Figure 1. High Attention Control Task.

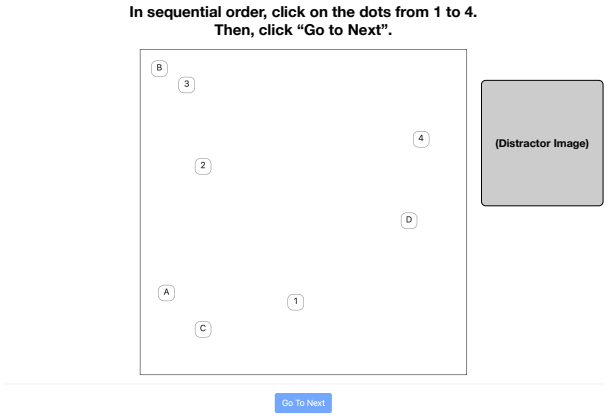


Figure 2. Low Attention Control Task.

After completing the clicking task, participants were presented with a concluding survey. As a manipulation check, the survey displayed screenshots of the interfaces used in each condition and asked the participants to rate them in terms of how distracting the interface was. The post-survey also collected the *attentional control scale* (ACS) [53], a widely used survey measure of trait AC, as well as demographic information.

B. *Participants*

A total of 223 students from a large, private university participated in the study in exchange for 0.25% extra credit applied to a participating management course of their choice. Approximately 49% of the participants were male, 85% were from the USA, and the average age was 21.8.

C. *Mouse Movement Data*

The web interface contained embedded JavaScript that captured mouse movement data during each trial. The library captured the x- and y-location of mouse movements, along with timestamps, which were then sent to a server to calculate speed and distance features.

We calculated the total distance for each trial by summing the Euclidean distances between each pair of x-y positions a_i and a_{i+1}, given by the following equation: $d(a_i, a_{i+1}) = \sqrt{(a_i^{(x)} - a_{i+1}^{(x)})^2 + (a_i^{(y)} - a_{i+1}^{(y)})^2}$. This translates to a total distance of $D = \sum_{i=1}^{n-1} d(a_i, a_{i+1})$ between the recorded points a_1, a_2, …, a_n. Speed was then calculated as a function of cursor distance D and trial time t, and is thus measured in pixels per second: $v = D / t$.

D. *Analysis*

We first used the manipulation check data to ensure that our experimental design reduced state AC as intended. A Wilcoxon rank sum test indicated that participants ranked the interface from the low AC condition significantly higher in terms of how distracting the interface was during the task.

To test whether our two mouse movement measures could be used to effectively differentiate between different levels of trait AC, we used the ACS survey items. The ACS contains two subscale scores that represent different aspects of trait AC—*focusing* and *shifting* [54]. Thus, we first performed a basic factor analysis to evaluate the factor structure of the items, eliminating items that loaded poorly on their intended factor or that significantly cross-loaded on both factors. The remaining items were then averaged to derive a single, continuous measure of trait AC that could be included in subsequent analyses.

We first tested the first-order correlations among trait AC and our measures of speed and distance. The results revealed significant correlation between speed and distance ($r = .65$), but nonsignificant correlations close to zero between trait AC and speed ($r = .04$), and trait AC and distance ($r = .03$). Two additional tests for a possible relationship between trait AC and the two mouse movement measures were performed to further confirm the lack of results. First, to eliminate possible noise from the distractors used in the low AC condition, we performed the same correlation analysis using only mousing data from each participant's high AC trial. Second, we categorized each participant's trait AC score into one of four quartiles, comparing those individuals in the bottom quartile with those in the top quartile. Both of

these additional tests revealed no significant relationship between trait AC and the mousing measures (i.e., H1 was not supported). Thus, the trait AC measures were excluded from subsequent analysis. (These nonsignificant results are discussed further in the discussion section.)

Next, we tested whether the mouse movement data could effectively differentiate between state AC, comparing each participant's mousing data from the two experimental conditions in a mixed-effects linear regression analysis. Mixed-effects models use random effects to account for individual differences in repeated observations [55]—including differences in natural computing ability or, importantly, trait AC that may not be well captured by the ACS measures—and fixed effects to model the treatment effects—in our case, the high versus low AC manipulation. (Table 1 provides the descriptive statistics for the two conditions.)

We specified one model with mousing speed as the dependent variable and a second with mousing distance as the dependent variable, allowing a random intercept for each participant and specifying the AC manipulation as a fixed effect. The regression model results are summarized in Table 2. Both models revealed significant effects resulting from the AC manipulation, providing initial evidence that mouse movements may be used as a proxy measurement for state AC. Thus, H2 was supported. The implications of our results and plans for future follow-up work are detailed in the next section.

TABLE I. MEAN (SD) OF MOUSING MEASURES BY CONDITION

Trial Condition	*Speed (px/s)*	*Distance (px)*
High AC (no distractions)	460 (130)	3,180 (926)
Low AC (with distractions)	427 (135)	3,550 (1,086)

TABLE II. MIXED-EFFECTS MODEL RESULTS

Fixed Effect	*Estimate*	*Std. Err.*	*df*	*t-value*
DV: Speed (px/s)				
(Intercept)	459.6	8.9	331.9	51.8^{***}
Low AC	-32.3	8.1	222	-4.0^{***}
DV: Distance (px)				
(Intercept)	3179.7	67.6	381.2	41.1^{***}
Low AC	370.8	73.7	222	5.0^{***}

V. DISCUSSION AND FUTURE RESEARCH

The goal of this research was to investigate whether mouse movement data can be used as a proxy measure for the attention (or, conversely, distraction) of the user in a web-based setting. The dual-process view explains how attention is managed by two different systems, with distractions adding additional cognitive burden during goal-directed tasks [20]. Prior research has demonstrated that dynamic feedback about individuals' level of sustained attention can significantly improve their focus and other related outcomes [14, 15], but cumbersome brain monitoring technologies are not well suited to web-based administration. Based on this same dual-process view of attention, we explored whether mousing behaviors could be used to identify both trait and state AC. Using a controlled, within-subjects experiment, we captured mousing behaviors during one high-AC and one low-AC task, hypothesizing that the depleted AC caused by visual distractions would cause systematic differences in mousing speed and distance. We also measured trait AC using the ACS [53], allowing us to test whether those same mouse movement features could reliably reveal individual differences in trait AC.

The analysis revealed no significant relationships between trait AC and the mouse movement features. There are several plausible explanations for these nonsignificant results. First, this preliminary study design was a first attempt at investigating a potential link between trait AC and mousing behaviors and should not be considered comprehensive evidence for or against an observable relationship between the two. Prior research has demonstrated that individual differences in working memory capacity—closely related to trait attention control [41]—influence one's susceptibility to distraction, observable in hand movements such as while using a handheld touchscreen pen [56]. The link between these individual differences and attention-related performance may also only apply to longer tasks [57] or in situations of real-world pressure like when an audience is watching [58]. Our preliminary experimental design is not likely to create those nuanced effects, and a more refined study design may produce the conditions necessary to observe the hypothesized effects.

Another likely explanation for the nonsignificant trait AC results is the scale chosen to measure the construct. Although the ACS [53] is arguably the most widely used survey measure of trait AC, recent research has provided convincing evidence that the ACS may not correlate well with objective measures of AC [59, 60]. Thus, using a trait AC score derived from the ACS may be an inaccurate way to measure the construct and test the relationship. An improved study design could use a more objective measure of trait AC to afford a more reliable test of the hypothesize relationships between trait AC and mouse movements.

However, our analysis provided evidence in support of H2, namely, that mouse movement data show promise as a method for identifying if a web user has become distracted. The state AC manipulation (in which various distractors were displayed among the goal-relevant targets) produced significant fixed effects in both the mousing speed and mousing distance models. This provides initial evidence that mouse movement data may be used in a web-based setting to detect when a user has become distracted. These findings have important implications for future research.

First, state AC can be influenced by various internal and external factors, including stress, fatigue, and distraction. Having a way to measure the momentary AC in real-time could therefore provide insight into these

cognitive states, which could be used to inform the design and delivery of related interventions.

Relatedly, research has shown that feedback interventions that are tailored to individuals' cognitive states can be more effective than interventions that are not tailored [61]. Thus, a measurement technique such as that suggested by our results could be used in a range of different customized feedback loops to enhance outcomes in situations where user attention is crucial—most notably, perhaps, in the context of online learning.

A final potential application for our preliminary findings further investigation is in helping users develop metacognitive skills, such as self-awareness and self-regulation. Metacognitive skills are important for learning and cognitive performance, as they allow individuals to monitor and adjust their cognitive processes to meet the demands of the task [24]. Systems that provide real-time information on users' attention states can help users develop metacognitive skills by increasing their awareness of their own cognitive processes.

Despite these promising findings, the preliminary study reported here leaves room for further refinement and improvement. The mousing measures evaluated constitute a high-level representation of what is likely a complex cognitive process. Alternative explanations for the observed effects—such as frustration or task complexity—need to be ruled out. Moreover, expanding the approach to test the method's validity during more complex and varied attention-related tasks will be required before a real-time attention feedback system based on this technology could be used reliably. Despite these limitations, however, the results reported here provide initial evidence that mouse movement data may contain signals that, with proper refinement and isolation, may prove valuable in detecting and promoting more focused attention.

VI. Conclusion

This research applied mouse movement data to a unique but promising area of potential application—monitoring users' attention in real-time. Based on dual-process theory and aligned with attention feedback systems in offline settings, we developed hypotheses predicting how mouse movements will differ based on differences in attention control. A controlled experiment provided initial evidence that attention control impacts a user's mousing behavior while also suggesting several promising opportunities for further development and refinement of the theory and approach used. With further exploration and additional empirical evidence, using this approach to monitor and respond in real time to users' cognitive states could produce significant value in web-based contexts where attention is a crucial resource relevant to system success.

VII. References

[1] J. Falkinger, "Limited attention as a scarce resource in information-rich economies," The Economic Journal, vol. 118, no. 532, pp. 1596-1620, 2008.

[2] A. Goldfarb, "Stop Letting Modern Distractions Steal Your Attention," in New York Times, ed, 2019.

[3] J. Firth et al., "The "online brain": how the Internet may be changing our cognition," World Psychiatry, vol. 18, no. 2, pp. 119-129, 2019.

[4] L. D. Rosen, L. M. Carrier, and N. A. Cheever, "Facebook and texting made me do it: Media-induced task-switching while studying," Comput Hum Behav, vol. 29, no. 3, pp. 948-958, 2013.

[5] M. M. Chun, J. D. Golomb, and N. B. Turk-Browne, "A taxonomy of external and internal attention," Annu Rev Psychol, vol. 62, pp. 73-101, 2011.

[6] A. D. Baddeley, "Is working memory still working?," Am Psychol, vol. 56, no. 11, pp. 851–864, 2001, doi: 10.1037/0003-066x.56.11.851.

[7] S. Leroy, "Why is it so hard to do my work? The challenge of attention residue when switching between work tasks," Organizational Behavior and Human Decision Processes, vol. 109, no. 2, pp. 168-181, 2009.

[8] J. Graydon and M. W. Eysenck, "Distraction and cognitive performance," Eur J Cogn Psychol, vol. 1, no. 2, pp. 161-179, 1989.

[9] E. Awh and E. K. Vogel, "Attention: feedback focuses a wandering mind," Nature neuroscience, vol. 18, no. 3, pp. 327-328, 2015.

[10] B. Donmez, L. N. Boyle, and J. D. Lee, "Safety implications of providing real-time feedback to distracted drivers," Accident Analysis & Prevention, vol. 39, no. 3, pp. 581-590, 2007.

[11] P. Mijović et al., "Towards continuous and real-time attention monitoring at work: reaction time versus brain response," Ergonomics, vol. 60, no. 2, pp. 241-254, 2017.

[12] M. Langner, P. Toreini, and A. Maedche, "EyeMeet: A Joint Attention Support System for Remote Meetings," 2022, pp. 1-7.

[13] M. Langner, P. Toreini, and A. Maedche, "Information Systems and Neuroscience, NeuroIS Retreat 2020," Lect Notes Information Syst Organisation, pp. 266-275, 2020, doi: 10.1007/978-3-030-60073-0_31.

[14] M. T. Debettencourt, J. D. Cohen, R. F. Lee, K. A. Norman, and N. B. Turk-Browne, "Closed-loop training of attention with real-time brain imaging," Nature neuroscience, vol. 18, no. 3, pp. 470-475, 2015.

[15] G. Tuckute, S. T. Hansen, T. W. Kjaer, and L. K. Hansen, "Real-Time decoding of attentional states using closed-loop EEG neurofeedback," Neural Computation, vol. 33, no. 4, pp. 967-1004, 2021.

[16] C. Draheim, R. Pak, A. A. Draheim, and R. W. Engle, "The role of attention control in complex real-world tasks," Psychon B Rev, vol. 29, no. 4, pp. 1143-1197, 2022.

[17] M. I. Posner, "Orienting of Attention," Q J Exp Psychol, vol. 32, no. 1, pp. 3-25, 1980, doi: 10.1080/00335558008248231.

[18] B. Gawronski and L. A. Creighton, "Dual process theories," The Oxford handbook of social cognition., pp. 282-312, 2013.

[19] J. S. B. T. Evans and K. E. Stanovich, "Dual-Process Theories of Higher Cognition," Perspect Psychol Sci, vol. 8, no. 3, pp. 223-241, 2013, doi: 10.1177/1745691612460685 PMID - 26172965.

[20] M. I. Posner and G. J. DiGirolamo, "Cognitive Neuroscience: Origins and Promise," Psychol Bull, vol. 126, no. 6, pp. 873–889, 2000, doi: 10.1037/0033-2909.126.6.873.

[21] D. A. Norman and T. Shallice, "Attention to Action: Willed and Automatic Control of Behavior," in Consciousness and Self-Regulation, R. Davidson, R. Schwartz, and D. Shapiro Eds. New York, NY: Plenum Press, 1986, pp. 1–18.

[22] R. Desimone and J. Duncan, "Neural Mechanisms of Selective Visual Attention," Annu Rev Neurosci, vol. 18, no. 1, pp. 193–222, 1995, doi: 10.1146/annurev.ne.18.030195.001205.

[23] S. Aliakbaryhosseinabadi, D. Farina, and N. Mrachacz-Kersting, "Real-time neurofeedback is effective in reducing diversion of attention from a motor task in healthy individuals and patients with amyotrophic lateral sclerosis," Journal of Neural Engineering, vol. 17, no. 3, p. 036017, 2020.

[24] J. W. Schooler, J. Smallwood, K. Christoff, T. C. Handy, E. D. Reichle, and M. A. Sayette, "Meta-awareness, perceptual decoupling and the wandering mind," Trends Cogn Sci, vol. 15, no. 7, pp. 319-326, 2011.

[25] M. T. Knierim, C. Zimny, and J. Seitz, "Monitoring Cognitive Load with 3D-Printed EEG Headphones," 2022.

[26] M. C. Chen, J. R. Anderson, and M. H. Sohn, "What can a mouse cursor tell us more? Correlation of eye/mouse movements on web

browsing," presented at the CHI '01 Extended Abstracts on Human Factors in Computing Systems, Seattle, WA, 2001.
[27] J. Hurtienne, M. Landeck, S. Ludwig, and D. Löffler, "Beyond Eye Tracking Analogies: Cursor Trajectories as Subtle Cues to Detect Distracting UI Elements," presented at the CHI '14 Extended Abstracts on Human Factors in Computing Systems, Toronto, ON, Canada, 2014.
[28] M. Speicher, S. Nuck, L. Wesemann, A. Both, and M. Gaedke, "From TMR to turtle: predicting result relevance from mouse cursor interactions in web search," Journal of Web Engineering, pp. 386-413, 2015.
[29] Y. Chen, Y. Liu, M. Zhang, and S. Ma, "User satisfaction prediction with mouse movement information in heterogeneous search environment," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 11, pp. 2470-2483, 2017.
[30] V. Navalpakkam and E. Churchill, "Mouse Tracking: Measuring and Predicting Users' Experience of Web-based Content," presented at the CHI '12 Extended Abstracts on Human Factors in Computing Systems, Austin, TX, 2012.
[31] K. Tzafilkou and N. Protogeros, "Mouse behavioral patterns and keystroke dynamics in End-User Development: What can they tell us about users' behavioral attributes?," Comput Hum Behav, vol. 83, no. 1, pp. 288-305, 2018, doi: 10.1016/j.chb.2018.02.012.
[32] M. J. Spivey, M. Grosjean, and G. Knoblich, "Continuous attraction toward phonological competitors," Proc National Acad Sci, vol. 102, no. 29, pp. 10393–10398, 2005, doi: 10.1073/pnas.0503903102.
[33] J. B. Freeman, "Doing Psychological Science by Hand," Curr Dir Psychol Sci, vol. 27, no. 5, pp. 315–323, 2018, doi: 10.1177/0963721417746793.
[34] J. L. Jenkins, J. G. Proudfoot, J. S. Valacich, G. M. Grimes, and J. F. Nunamaker, Jr., "Sleight of Hand: Identifying Concealed Information by Monitoring Mouse-Cursor Movements," J Assoc Inf Syst, vol. 20, no. 1, pp. 1–32, 2019, doi: 10.17705/1jais.00527.
[35] M. Hibbeln, J. L. Jenkins, C. Schneider, J. S. Valacich, and M. Weinmann, "How Is Your User Feeling? Inferring Emotion Through Human-Computer interaction Devices," MIS Quarterly, vol. 41, no. 1, pp. 1–21, 2017, doi: 10.25300/misq/2017/41.1.01.
[36] M. Weinmann, J. S. Valacich, C. Schneider, J. L. Jenkins, and M. Hibbeln, "The Path Of The Righteous: Using Trace Data To Understand Fraud Decisions In Real Time," MIS Quarterly, vol. 46, no. 4, pp. 2317–2336, 2022, doi: 10.25300/misq/2022/17038.
[37] L. Yang and S.-F. Qin, "A Review of Emotion Recognition Methods From Keystroke, Mouse, and Touchscreen Dynamics," Ieee Access, vol. 9, pp. 162197-162213, 2021.
[38] M. W. Eysenck, N. Derakshan, R. Santos, and M. G. Calvo, "Anxiety and Cognitive Performance: Attentional Control Theory," Emotion, vol. 7, no. 2, pp. 336–353, 2007, doi: 10.1037/1528-3542.7.2.336.
[39] J. Dang, F. Björklund, and M. Bäckström, "Self‐control depletion impairs goal maintenance: A meta‐analysis," Scand J Psychol, vol. 58, no. 4, pp. 284‐293, 2017, doi: 10.1111/sjop.12371.
[40] S. Forster and N. Lavie, "Failures to ignore entirely irrelevant distractors: the role of load," Journal of Experimental Psychology: Applied, vol. 14, no. 1, p. 73, 2008.
[41] L. F. Barrett, M. M. Tugade, and R. W. Engle, "Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind," Psychol Bull, vol. 130, no. 4, pp. 553-573, 2004, doi: 10.1037/0033-2909.130.4.553 PMID - 15250813.
[42] N. Unsworth and R. W. Engle, "The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory," Psychological review, vol. 114, no. 1, p. 104, 2007.
[43] M. M. McClelland, C. E. Cameron, C. M. Connor, C. L. Farris, A. M. Jewkes, and F. J. Morrison, "Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills," Developmental psychology, vol. 43, no. 4, p. 947, 2007.
[44] X. Luo, L. Zhang, and J. Wang, "The Benefits of Working Memory Capacity on Attentional Control under Pressure," Front Psychol, vol. 8, p. 1105, 2017, doi: 10.3389/fpsyg.2017.01105 PMID - 28740472.
[45] B. Chin et al., "Mindfulness Interventions Improve Momentary and Trait Measures of Attentional Control: Evidence From a Randomized Controlled Trial," J Exp Psychology Gen, vol. 150, no. 4, pp. 686-699, 2021, doi: 10.1037/xge0000969 PMID - 32969686.
[46] A. J. Kim and B. A. Anderson, "Threat Reduces Value-Driven but Not Salience-Driven Attentional Capture," Emotion, vol. 20, no. 5, pp. 874–889, 2020, doi: 10.1037/emo0000599.
[47] K. S. Moore, C. B. Porter, and D. H. Weissman, "Made you look! Consciously perceived, irrelevant instructional cues can hijack the attentional network," Neuroimage, vol. 46, no. 1, pp. 270–279, 2009, doi: 10.1016/j.neuroimage.2009.01.042.
[48] T. N. Welsh and D. Elliott, "Movement trajectories in the presence of a distracting stimulus: Evidence for a response activation model of selective reaching," Q J Exp Psychology Sect, vol. 57, no. 6, pp. 1031–1057, 2004, doi: 10.1080/02724980343000666.
[49] M. L. Anderson, "Neural reuse: A fundamental organizational principle of the brain," Behavioral and brain sciences, vol. 33, no. 4, pp. 245-266, 2010.
[50] D. G. Gozli and J. Pratt, "Seeing while acting: hand movements can modulate attentional capture by motion onset," Atten Percept Psychophys, vol. 73, no. 8, pp. 2448-2456, 2011, doi: 10.3758/s13414-011-0203-x PMID - 21870205.
[51] H.-C. Nothdurft, "Salience and target selection in visual search," Vis Cogn, vol. 14, no. 4-8, pp. 514–542, 2006, doi: 10.1080/13506280500194162.
[52] M. A. Dieciuc, H. M. Maranges, and W. R. Boot, "Trait self-control does not predict attentional control: Evidence from a novel attention capture paradigm," Plos One, vol. 14, no. 12, pp. 1–15, 2019, doi: 10.1371/journal.pone.0224882 PMID - 31830063.
[53] D. Derryberry and M. A. Reed, "Anxiety-related attentional biases and their regulation by attentional control," Journal of abnormal psychology, vol. 111, no. 2, p. 225, 2002.
[54] M. R. Judah, D. M. Grant, A. C. Mills, and W. V. Lechner, "Factor structure and validation of the attentional control scale," Cognition & emotion, vol. 28, no. 3, pp. 433-451, 2014.
[55] J. C. Pinheiro and D. M. Bates, Mixed-effects models in S and S-PLUS. New York, NY: Springer-Verlag, 2000.
[56] D. Domkin, P. Sörqvist, and H. O. Richter, "Distraction of eye-hand coordination varies with working memory capacity," J Motor Behav, vol. 45, no. 1, pp. 79-83, 2013.
[57] D. van Son, A. Angelidis, M. A. Hagenaars, W. van der Does, and P. Putman, "Early and late dot‐probe attentional bias to mild and high threat pictures: Relations with EEG theta/beta ratio, self‐reported trait attentional control, and trait anxiety," Psychophysiology, vol. 55, no. 12, p. e13274, 2018.
[58] P. Furley and G. Wood, "Working Memory, Attentional Control, and Expertise in Sports: A Review of Current Literature and Directions for Future Research," J Appl Res Mem Cogn, vol. 5, no. 4, pp. 415-425, 2016, doi: 10.1016/j.jarmac.2016.05.001.
[59] P. J. F. Clarke and J. Todd, "Lessons unlearned: A conceptual review and meta-analysis of the relationship between the Attention Control Scale and Objective Attention Control," Cognition and Emotion, vol. 35, no. 8, pp. 1447-1459, 2021.
[60] J. Todd, L. Notebaert, and P. J. F. Clarke, "The association between self-report and behavioural measure of attentional control: Evidence of no relationship between ACS scores and antisaccade performance," Pers Indiv Differ, vol. 184, p. 111168, 2022.
[61] K. C. Ewing, S. H. Fairclough, and K. Gilleade, "Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop," Frontiers in human neuroscience, vol. 10, p. 223, 2016.

Getting What You Paid For: Assessing Participant Experience Parameters for Amazon Mechanical Turk (MTurk) Workers on Survey Response Quality

M. Kumar*, D. Kim*, P.A. Weisgarber*, J.S. Valacich* and J.L. Jenkins**
* University of Arizona/Management Information Systems Department, Tucson, AZ, USA
** Brigham Young University/Information Systems Department, Provo, UT, USA
manasvik@arizona.edu

***Abstract* – MTurk is a powerful and widely used method for completing online Human Intelligence Tasks (HITs) such as website testing or completing psychological surveys. Researchers typically use various platform-provided work experience parameters to choose qualified and unqualified participants. Conventional wisdom suggests that participants with greater experience and historically high acceptance rates (i.e., higher-rated) will generate better quality data than their peers with lower experience and low acceptance rates (i.e., lower-rated). We examine the limits of this assumption by comparing responses and engagement behaviors between higher-rated, experienced (HE) participants and lower-rated, inexperienced (LI) participants while answering online surveys. We administered an online survey where participants first answered questions related to their MTurk account profile and then answered questions related to their personality. LI participants provide more inaccurate responses when answering factual questions (i.e., higher error rates) compared to HE participants. They also exhibit lower engagement behaviors when answering personality-related questions, resulting in marginally lower reliability scores for survey constructs. We are encouraged by the systematic differences we observed between the two populations and urge researchers to consider our findings before selecting optimal work experience parameters.**

Keywords – MTurk, Online Surveys, Data Quality

I. INTRODUCTION

Online surveys are a popular method utilized by researchers to collect data. While other alternatives exist [1], Amazon Mechanical Turk (MTurk) remains the most popular crowdsourcing platform that enables requesters (researchers and practitioners) to post several Human Intelligence Tasks (HITs) that workers can complete. Reference [2] suggests that compared to other traditional sample sources (i.e., university students), MTurk remains the most widely used platform due to its vast participant pool, convenient data collection, economical cost, and flexible research design options. Despite its popularity, MTurk faces several challenges due to its small, active subset of workers who complete most of the HITs available. It is estimated that nearly 80% of all HITs are completed by 20% of its participant pool [3]. Therefore, there are concerns that the platform is ineffective as the few workers who complete most of the tasks are familiar with the traditional techniques used to ensure data quality and can successfully circumvent them [4].

To combat these participant pool selection challenges, MTurk provides certain workers with the qualification of a "Master Worker": a qualification that is provided after the workers submit high-quality work for a sustained period. While the explicit details of how a worker becomes a "Master Worker" are not shared, they are often more expensive and do not provide significantly better data than other workers [5]. MTurk also allows its requesters to screen participants for its HITs based on their HIT acceptance rate, number of HITs completed or other criteria to ensure that adequate data quality requirements may be satisfied, and expenses are decided accordingly. In this study, we address where this separation line might be as we adjust the various qualification criteria to determine when higher-rated, experienced (HE) participants generate better quality data compared to lower-rated, inexperienced (LI) participants. While other studies have investigated the quality of data generated from less commonly used sections of MTurk's participant pool using survey responses alone [6], our study utilizes both response and engagement behavior to compare and evaluate the quality of data generated by participants of vastly different HIT acceptance and experience levels. By unobtrusively capturing paradata [7] (i.e., data such as mouse movements, clicks, timestamps etc., that is generated during the response generation process), we utilize fine-grained movement-based metrics to identify potentially problematic response generation behaviors when a participant answers a survey question. Thus, by using varied data sources, we provide a holistic comparison of the differences between data generated by HE and LI participants.

To demonstrate the efficacy of our approach, we administered a survey where HE and LI participants completed a two-part online survey where they answered factual questions (i.e., questions about their account) and traditional self-report survey questions (i.e., questions about their personality). Our analysis indicates that there are distinctive differences between HE and LI participants both in terms of their responses and their engagement

behaviors. While answering factual questions, HE participants typically have lower error rates compared to their LI peers. They also spend more time generating the response to these questions. While answering traditional survey questions, LI participants generate responses that have marginally lower reliability scores and they often engage in extreme response generation behavior by requiring greater response times. Our results suggest that HE participants generate marginally better quality data than LI participants, suggesting that conventional wisdom on utilizing experience and acceptance rates to select participants is justified. The rest of the paper is organized as follows. First, we will briefly review the concerns regarding data quality in online surveys and how paradata may be used to address them. Next, we briefly describe the various aspects of the methodological approach. We will then present the results from our study, followed by a brief discussion and the conclusion.

II. BACKGROUND

Data collected using online surveys may contain responses that are "invalid" (i.e., does not represent the true response value for a given question) and therefore poor quality [8]. A family of techniques commonly used to identify such responses involves the use of attention check questions. These questions require the participant to select a specific response (e.g., "Please select Moderately Inaccurate for this item."), as discussed in [9]. As mentioned earlier, one of the biggest concerns associated with MTurk participants is the small percentage of active workers who complete a large proportion of HITs. As a result of these numerous interactions, these workers become skilled in identifying attention check questions easily [10], thereby negating one of the most popular approaches used by researchers to identify inattentive participants who generate such invalid responses. Therefore, current traditional approaches may not be adequate to identify inattentive participants, especially if they may have answered similar questions earlier.

An alternative approach to identifying responses generated by inattentive participants involves collecting paradata. To illustrate the effectiveness of this approach, consider the three hypothetical responses to a survey question depicted in Figure 1. In Parts a, b, and c, the participants chose to select the same option: "Agree". However, the time taken to generate the responses (i.e., response time) varies greatly. Without using paradata, it appears that all three participants generated the same response, and it is impossible to distinguish between them. In contrast, by observing the response times, it is now possible to comment on the response generation process of each participant and therefore on the true "accuracy" of each response. Anecdotally, it has been observed that minimum amount of time to generate a meaningful response is 2 seconds [9]. This contrasts with the time taken by participant A, who takes around 1 second to generate a response. It is more likely that this participant didn't pay adequate attention while generating the response for this question, and therefore, this response is more likely to be invalid. Note that this response is also characterized by decisive, quick movements with less deviations. Participant B takes around 3 seconds to

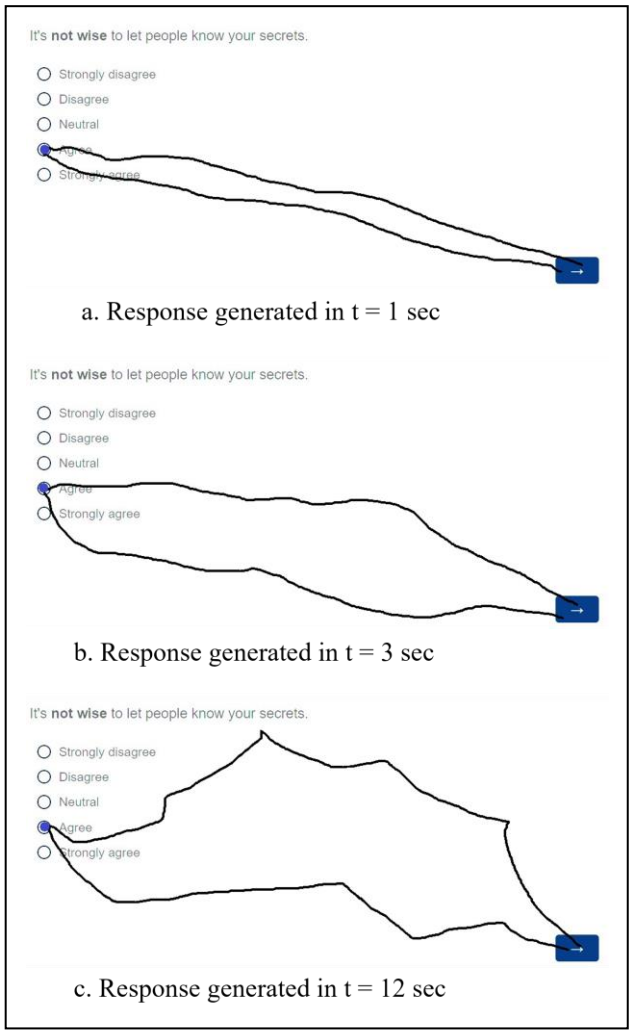


Figure 1: Representation of responses generated by three hypothetical participants – t = Response times with mouse movement navigation.

generate a response. This is close to the median response time taken by other participants to answer this question. As the participant had sufficient time to read the question and the movement is characterized by adequate, precise, deliberation; it is more likely that the response generated is a good quality response. Participant C takes 12 seconds to generate their response. This is significantly longer than the time taken by most participants to answer this question. Compared to previous participants, the movement here is characterized by greater deviations, potential pauses, and indecisive movements. It is therefore more likely that the response generated is invalid, as the participant may have been distracted or multitasking. Therefore, by utilizing fine-grained mouse cursor movement, it is possible to derive metrics (i.e., response time) that can be used to understand engagement behavior at an item level. These metrics may also be used to identify particularly poor-quality responses, or bad data, in a manner that is completely unobtrusive and can be performed after data collection has been completed.

Some of the major challenges of MTurk research include inattention, misrepresentation, selection bias, non-naivete and social-desirability bias [2]. As illustrated earlier, paradata may be used to improve data quality by capturing engagement behaviors during response

generation. Engagement behavior metrics have been used to examine a variety of phenomena, including multitasking [11], social desirability bias [12], distractions [13], adequacy of outlier definitions [14], and impact of question formats on response quality [15]. In particular, there is comprehensive research that examines the impact of using various response time-based metrics to improve response quality [16]. In this study, we utilize both survey responses and engagement behaviors to compare data quality of responses generated by HE and LI participants.

III. METHODOLOGY

We designed a two-part study to compare responses and engagement behaviors between HE and LI participants. In the first part, we asked participants to select information about their MTurk account, including information about the number of HITs they had completed and their HIT approval rating. In the second part of the study, they self-report scores about their personality dimensions of psychopathy and sadism. Data collection occurred in two phases. In the first phase, we collected data from LI participants while in the second phase, we collected data from HE participants. Therefore, both groups of participants (HE and LI participants) answered factual questions (i.e., questions about their account) and traditional self-report survey questions (i.e., questions about their personality). We also captured fine-grained mouse cursor data while the participants generated their responses. Thus, the data collected from this study allows us to compare both responses and engagement behaviors between HE and LI participants.

A. Pilot Studies

Prior to conducting the primary study, we conducted a series of pilot studies to determine optimal thresholds for experience and HIT acceptance rates in MTurk participants. As we were interested in comparing workers who hadn't achieved "Master" status, we considered different thresholds to adequately represent typical experience and acceptance rates of other "regular" workers. Reference [5] found that the average number of HITs completed by a regular worker is around 19,791 while the average approval rate is around 97.9%. We concluded that a LI participant should have a lower threshold compared to the average participant; therefore we selected appropriate thresholds for experience (i.e., 5,000 and 10,000) and acceptance rates (i.e., 90 and 95). We then conducted several pilot studies to determine the availability of participants satisfying these threshold buckets. We found that to collect a typical number of participants (n=100), it is advisable to select a threshold of 10,000 for HITs completed and 90 to 95 for acceptance rate. Therefore, we determined that participants who completed lesser than 10,000 HITs and have an HIT acceptance rate between 90 to 95 are LI participants. Consequently, participants who completed greater than 10,000 HITs and have an HIT acceptance rate of 95 and above are HE participants.

B. Survey Design

In the first part of the two-part survey, participants were required to provide information about their MTurk account. This section includes information about their HIT approval rating (categorized as 90 or less, 90 to 95, and 95 or more) and the number of HITs they had completed (categorized as 0 to 10,000 and 10,001 or more), in accordance with the thresholds determined in the pilot studies. The participants were requested to answer these questions "as best as they can", and that these responses may be used to "recruit them in future studies." We also included an attention question ("22 times 22 is:") to identify inattentive participants. Only participants who successfully answered the attention check question could participate in the second part of the study. In the second part of the study, participants were required to provide self-report responses to the Dark-Tetrad Personality questionnaire, as adapted from [17]. Items for all measures utilized a 5-point Likert Scale ranging from "Strongly Disagree" to "Strongly Agree."

C. Participants

We recruited participants from Amazon Mechanical Turk in two phases. In the first phase, participants who completed less than 10,000 HITs and had a HIT approval rating lower than 95% (LI participants) were recruited for the study. In the second phase, participants who completed greater than 10,000 HITs and had a HIT approval rating higher than 95% were recruited (HE participants). The study consisted of two parts, and participants were paid $0.05 for completing the first part and $0.95 for completing the second part of the study. Only data from participants who completed both parts of the study were used in future analyses. Of the 200 participants recruited for the study (100 for each phase), 84 participants completed both parts of the study in the first phase, while 91 participants completed both parts of the study in the second phase.

D. Engagement Data Tracking

We utilized the Qualtrics survey system to host the survey used for data collection. A custom JavaScript library that captures mouse-cursor movements (e.g., x and y coordinates of the cursor, timestamps, etc.) and other behavioral data (e.g., mouse clicks, interaction with HTML elements, etc.) was embedded in the survey to track engagement behavior. The data captured by this library was sent to a web service for storage and further processing. Through in-house python code developed by the research team, we create specialized metrics to understand engagement behaviors of participants at a question level. In this study, we utilize two metrics: response time and response switches. Response time may be understood as the time taken by the participant to generate a response for a survey question. Response switches indicate the number of intermediate selections made prior to selecting a final response. These metrics have previously been shown to identify poor quality responses in online surveys [18]. For example, excessive response switches could indicate indecision, lack of comprehension, or the presence of response bias. Consider Figure 2, which presents two responses to the question of HIT approval ratings used in this study. In Part a, the participant answers the question directly without any intermediate selections. However, Part b shows the response generated by the participant, who has a response

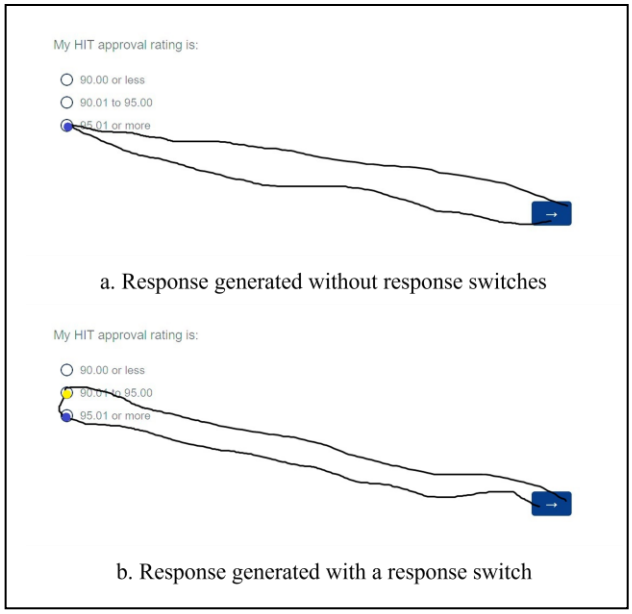


Figure 2: Representation of responses generated by two hypothetical participants – Response Switches with mouse movement navigation.

switch. It is likely that this participant chose to select a lower approval rating before finally deciding to select the alternative. As mentioned earlier, this response behavior could indicate indecision, lack of comprehension, or (as is more likely in this case) the presence of a social response bias. Similarly, while response times have been used as a proxy for engagement, extreme response generation behaviors: either too fast or too slow, may not be ideal. To determine these engagement behavioral "outliers", we adopted a procedure similar to that utilized in [18]. We first compute the average response time utilized by an individual to generate a response for questions pertaining to a particular construct. We then compute the median (M) and median absolute deviation (MAD) times of all average response times computed for all individuals. If an individual's average response time is greater than M + 3*MAD, then the participant spent an unusual amount of time answering questions pertaining to the construct and is therefore considered an engagement behavioral outlier.

IV. Analysis and Results

In this section, we provide results for the various analyses conducted to compare responses and engagement behaviors between HE and LI participants. As mentioned earlier, both groups of participants provided responses to factual questions (i.e., questions about their account) and traditional self-report survey questions (i.e., questions about their personality). In addition, we also collected fine-grained mouse movement data while these responses were generated using a JavaScript library. Previous studies have shown that relevant metrics derived from such data may be used to identify responses of poor data quality as they provide information about participant engagement behaviors [19]. In this study, we answer the following questions:

- How do responses obtained by HE and LI participants differ for factual questions?
- How do responses obtained by HE and LI participants differ for traditional survey questions?
- How does engagement behavior exhibited by HE and LI participants differ for factual questions?
- How does engagement behavior exhibited by HE and LI participants differ for traditional survey questions?

A. *Response Data*

We compared the survey responses received from both HE and LI participants for different question types (i.e., factual questions and traditional self-report survey questions). In factual questions, we asked participants to select the number of HITs they completed and their HIT approval rate "as best as they can." As we know the actual values of these parameters, we determined the error rate of their responses. For the HITs completed question, 5% of HE participants provided erroneous answers while 8% of LI participants provided erroneous answers. For the HIT approval rate question, the difference is more apparent. 49% of LI participants provided erroneous answers for this question, while around 6% of HE participants provided erroneous answers. Thus, for both factual questions, HE participants provided less erroneous answers compared to LI participants.

We also compared the survey responses of HE and LI participants to more traditional survey questions (i.e., the Dark Tetrad Questionnaire). One approach to evaluate the quality of survey data is through reliability, with coefficients greater than 0.7 being the recommended standard for MTurk data [20]. For our study, we evaluate the Cronbach's alpha value for the constructs of psychopathy and sadism for data from HE and LI separately. For psychopathy, the Cronbach's alpha obtained from analyzing HE and LI participants' survey responses is 0.931 and 0.886, respectively. Similarly, the Cronbach's alpha obtained by analyzing the responses from HE and LI participants for sadism is 0.89 and 0.843, respectively. In either case, the values of Cronbach's alpha are greater for HE participants than LI participants, while both are above the recommended threshold.

B. *Engagement Data*

We compared engagement behaviors through metrics generated from fine-grained mouse movement data as the participants generated their responses to both factual questions and traditional self-report survey questions. For the factual questions, we created several question-level metrics and analyzed for differences between HE and LI participants. We found that when participants are asked to report the number of HITs they've completed, the HE participants have greater response times compared to their LI peers ($t(114)=-2.87$, $p<0.01$). Note that HE participants generated fewer erroneous responses on this question compared to the LI participants. Similarly, we found that when participants are asked to report their HIT approval rating, LI participants have a significantly greater number of response switches compared to their HE peers ($t(119)=2.68$, $p<0.01$). It is important to note that HE participants generated significantly fewer erroneous responses compared to LI participants. The implications of these results are discussed in the next section.

TABLE I. SUMMARIZED RESULTS OF ALL ANALYSES PERFORMED

Type size	Response		Engagement Behavior	
	HE Participant	***LI Participant***	***HE Participant***	***LI Participant***
HIT Completion Question	Error (5%)	Error (8%)	Greater Response Time	Less Response Time
HIT Approval Question	Error (6%)	Error (49%)	Fewer Response Switches	More Response Switches
Survey Questions	Cronbach's Alpha (0.931 and 0.89)	Cronbach's Alpha (0.886 and 0.843)	Less proportion of slow-moving outliers	Greater proportion of slow-moving outliers

We analyzed the difference in engagement behaviors for the traditional self-report survey questions by identifying engagement behavioral outliers, as discussed in the Methodology section. For questions pertaining to psychopathy construct, we found that the participants who exhibited extreme engagement behaviors (greater response times) are disproportionately LI participants (52% compared to a baseline of 48%). We observe similar results when analyzing extreme engagement behaviors for questions pertaining to sadism construct as well (53% compared to a baseline of 48%). Thus, a greater proportion of engagement behavioral outliers were found to be LI participants for both constructs. Table 1. summarizes the results obtained for all analyses performed.

V. DISCUSSION

While comparing differences in responses between HE and LI participants, the results obtained are intuitive. For the factual questions, we observed that HE participants produced less erroneous responses compared to LI participants. This is expected as HE participants are more careful and aware of important questions in the survey compared to LI participants. It is also worth mentioning that the error rate of LI participants on the HIT approval rate question (49%) is significantly greater than the error rates among others. This is in part due to the instruction provided which indicated that their responses may be used for future recruitment. It is possible that the LI participants are aware of the importance researchers place on HIT approval rates in their selection process and decided to be deceptive to ensure future selection. Comparing differences in responses to the traditional survey questions revealed similarly intuitive results. The Cronbach's alpha value in all cases is above the recommended threshold, as expected for a widely used survey. While the Cronbach's alpha values are greater for responses generated by HE participants, the differences are not large enough to warrant further analysis.

Comparing differences in engagement behaviors between HE and LI participants provides several interesting results, especially when viewed in conjunction with differences in responses. For the factual question concerning number of HITs completed, we observed that HE participants have significantly greater response times than LI participants. Note that for this question, the error rate among HE participants is lower. This is intuitive, as response times are often seen as a proxy for engagement, and it is likely that HE participants spent more time on this question to ensure that they answered it correctly. For the other factual question concerning the HIT approval rate, we observed that LI participants have significantly greater number of response switches compared to HI participants. They also have a significantly large error rate on this question. As mentioned earlier, a large number of response switches indicates indecision, improper comprehension or the presence of response bias. In this study, it was clearly indicated that participants should answer the questions "as best as they can" and that their responses may be used for future recruitment. It is more likely that the LI participants engaged in social desirability bias and selected options that inflated their approval ratings.

Comparing engagement behaviors between HE and LI participants while answering traditional survey questions revealed that LI participants are likely to have engaged in extreme engagement behaviors (i.e., have unusually greater response times). As mentioned earlier, participants who engage in extreme response generation behaviors are found to generate poor quality responses. Note that for these questions, the Cronbach's alpha (while above the recommended value) is lower for responses generated by LI participants. It is typically assumed that relatively inexperienced "amateurs" are easily distracted from their work [10], and these distractions could result in excessive response times among these participants.

In addition to highlighting the response and behavioral differences between HE and LI participants while answering survey questions, we also demonstrate the effectiveness of fine-grained mouse movement-based metrics to identify poor quality responses. While several studies have identified and acknowledged concerns with MTurk's participant pool [21], MTurk (and other related crowdsourcing services) provides researchers with access to several features that can mitigate common concerns [22]. By integrating a custom JavaScript code to collect data and developing useful metrics to understand engagement behavior, we provide researchers with an alternative approach to identifying poor quality responses that goes beyond existing practices. The suggested approach may also be used on other crowdsourcing services or traditional recruitment approaches that lack prescreening and other related features provided by MTurk to maintain data quality standards.

A. Limitations

In this study, participants were required to answer factual questions about their MTurk account. They were instructed to answer these questions "as best as they can", and that their responses may be used for future recruitments. While these instructions were intended to ensure correctness of their response, the LI participants may have intentionally provided deceptive responses to improve their odds of recruitment. This wasn't possible for the HE participants as they couldn't provide a more

deceptive, inaccurate answer. In the future, we advise researchers to utilize factual questions that do not provide any group of participants unintended incentives to provide inaccurate responses.

In this study, we compared responses and response generation behaviors between HE and LI participants. We could not collect data for higher-rated, inexperienced participants and lower-rated, experienced participants as these groups of participants did not have easily available workers. It is, however, possible that by utilizing a different threshold or multiple thresholds (i.e., experienced > 100,000 HITs and inexperienced < 10,000 HITs), we may obtain a set of available workers whose data may be compared. We acknowledge that these thresholds will continue to evolve as MTurk grows and advise researchers to utilize the latest available data to select appropriate thresholds.

VI. CONCLUSION

In this study, we examined whether traditionally higher-rated, experienced participants generate better quality data in online surveys compared to lower-rated, inexperienced participants. Both HE and LI participants were required to answer a survey containing factual and typical survey questions. We analyzed differences between both groups of participants not only through their responses but also using their engagement behaviors, as captured by fine-grained mouse movement-based metrics. Our analysis revealed that HE participants provide less erroneous responses to factual questions, likely because they spend more time on these questions and do not exhibit potentially deceptive behavior. We also found that LI participants provide slightly less reliable (but acceptable) responses to traditional survey questions and are more likely to engage in extreme engagement behaviors.

REFERENCES

[1] E. Peer, L. Brandimarte, S. Samat, and A. Acquisti, "Beyond the Turk: Alternative platforms for crowdsourcing behavioral research," in Journal of Experimental Social Psychology, 70, 2017, pp. 153-163.

[2] H. Aguinis, I. Villamor, and R. S. Ramani, "MTurk research: Review and recommendations," in Journal of Management 47, no. 4, 2021, pp. 823-837.

[3] K. Fort, G. Adda, and K. B. Cohen, "Amazon Mechanical Turk: Gold mine or coal mine?," in Computational Linguistics, 2011, pp. 413-420.

[4] A. W. Meade, and S. B. Craig, "Identifying careless responses in survey data," in Psychological methods 17, no. 3, 2012, pp. 437.

[5] E. Loepp, and J. T. Kelly, "Distinction without a difference? An assessment of MTurk Worker types," in Research & Politics 7, no. 1 , 2020, 2053168019901185.

[6] J. Robinson, C. Rosenzweig, A. J. Moss, and L. Litman, "Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool," in PloS one 14, no. 12 , 2019, e0226394.

[7] F. Kreuter, Improving surveys with paradata: Analytic uses of process information. John Wiley & Sons, 2013.

[8] P.G. Curran, "Methods for the detection of carelessly invalid responses in survey data," in Journal of Experimental Social Psychology, 66, 2016, pp. 4-19.

[9] J.L. Huang, P. G. Curran, J. Keeney, E. M. Poposki, and R. P. DeShon, "Detecting and deterring insufficient effort responding to surveys," in Journal of Business and Psychology 27 , 2012, pp. 99-114.

[10] E. Peer, J. Vosgerau, and A. Acquisti. "Reputation as a sufficient condition for data quality on Amazon Mechanical Turk," in Behavior research methods 46 ,2014, pp. 1023-1031.

[11] L. Zwarun, and A. Hall. "What's going on? Age, distraction, and multitasking during online survey taking," in Computers in human behavior 41 , 2014, pp. 236-244.

[12] J.L. Jenkins, J. S. Valacich, and P. Williams, "Human-computer interaction movement indicators of response biases in online surveys," in International Conference on Information Systems, 2018, Seoul, Korea.

[13] A. Wenz, "Do distractions during web survey completion affect data quality? Findings from a laboratory experiment," in Social Science Computer Review 39, 2021, pp.148-161.

[14] J.K. Höhne, and S. Schlosser, "Investigating the adequacy of response time outlier definitions in computer-based web surveys using paradata," in Social Science Computer Review, 36, 2018, pp. 369-378.

[15] J.K. Höhne, S. Schlosser, and D. Krebs, "Investigating cognitive effort and response quality of question formats in web surveys using paradata," in Field Methods 29, 2017, pp. 365-382.

[16] M. Matjašič, V. Vehovar, and K.L. Manfreda, "Web survey paradata on response time outliers: A systematic literature review," in Advances in Methodology and Statistics, 15, 2018, pp. 23-41.

[17] D.L. Paulhus, "Toward a taxonomy of dark personalities," in Current Directions in Psychological Science, 23, no. 6 , 2014, pp. 421-426.

[18] M. Kumar, D. Kim, , J. S. Valacich, J. L. Jenkins, and A. Dennis, "Improving the Quality of Survey Data: Using Answering Behavior as an Alternative Method for Detecting Biased Respondents," 2021, SIGHCI 2021 Proceedings, 13.

[19] M. Kumar, D. Kim, , J. S. Valacich, and J. L. Jenkins, "Too Fast? Too Slow? A Novel Approach for Identifying Extreme Response Behavior in Online Surveys," 2022, SIGHCI 2022 Proceedings, 14.

[20] C.J. Holden, T. Dennie, and A. D. Hicks, "Assessing the reliability of the M5-120 on Amazon's Mechanical Turk," in Computers in Human Behavior 29, no. 4 , 2013, pp. 1749-1754.

[21] D. Hauser, G. Paolacci, and J. Chandler, "Common concerns with MTurk as a participant pool: Evidence and solutions," in F. R. Kardes, P. M. Herr, and N. Schwarz (Eds.), Handbook of research methods in consumer psychology, 2019, pp. 319–337

[22] J. Chandler, P. Mueller, and G. Paolacci, "Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers," Behavior research methods, 46, 2014, pp. 112-130.

Detecting Goal-Oriented vs. Browsing Users Through Behavior Analysis

J. L. Jenkins*, A. Denison*, J. S. Valacich**, D. Wilson*
* Brigham Young University, Provo, UT, USA
** University of Arizona, Tucson, AZ, USA
jjenkins@byu.edu

Abstract - **Understanding user personas in e-commerce is important for promoting successful online interactions. Two common personas include goal-oriented and browsing users. A goal-oriented user has the intention of completing a specific task as efficiently as possible (e.g., purchasing a product). A browsing user explores for information that will ultimately determine the next objective (e.g., to purchase or look elsewhere). A website that customizes content to a goal-oriented versus browsing user will improve the user experience and ultimately maximize conversion. In this research, we provide a methodology for differentiating between goal-oriented and browsing users by monitoring users' behavior on the website. We conducted a study where participants were randomly assigned to either a goal-oriented task to find a product or told to simply browse the website. Based on the study's results, we discuss suggestions to assist future human-computer interaction (HCI) researchers on how to design behavior-monitoring studies to accurately portray goal-oriented versus browsing users. We also provide insights into the need for considering the motivation of Amazon's Mechanical Turk workers to appropriately utilize them as a sample population in future behavior-monitoring studies.**

Keywords - e-commerce persona, mouse tracking, behavior analysis, goal-oriented users, browsing users, human-computer interaction

I. INTRODUCTION

Enabling effective online interactions is essential for business. In the United States, nearly 30% of all business is done online [1]. Consumers frequently expect to be able to interact with companies online. Yet, e-commerce conversion rates (i.e., purchasing) are abysmally small, being less than 3% on average [2]. Businesses spend billions of dollars searching for ways to improve their online sales and better meet consumers' needs [3].

One successful technique to improve conversion is to understand user personas and customize the online experience to match user personas. A persona refers to "archetypical users whose goals and characteristics represent the needs of a larger group of users" [4]. Examples of personas include goal directed users—a user that has the intention of completing a specific task as efficiently as possible—and browsing users—a user who is exploring for information that will ultimately determine the next objective (e.g., to purchase or look elsewhere). Understanding personas can help business build empathy with their users and provide direction on how to optimize the users experience, ultimately increasing user satisfaction and thus conversion.

Despite the importance of understanding user personas, detecting user personas is very difficult. It often requires the use of historical data gained by tracking a user across the web through third-party cookies. However, third party cookies are highly criticized due to privacy concerns, and users and companies alike are beginning to block them [5]. In addition, regulations, such as GDPR, ban the use of third-party cookies without the user's consent [6].

We propose that one way to detect user personas without the negative side effects of third-party cookies is to analyze how a user is browsing a website in real time. This can be done by recording and analyzing the users' mouse-cursor movements and determine if the user is goal-directed or browsing a website. Mouse cursor movements have been shown to provide valuable insight into users' attentional, emotional, and cognitive states online [e.g., 7, 8]. We extend this research to explore whether mouse-cursor movements can differentiate among users' personas. In summary, we address the following research questions:

Can mouse-cursor movements be used to differentiate between goal-directed and browsing users online?

We address this research question by first summarizing relevant previous research. We then draw on attentional selection theory and response activation model to explain how mouse movement speed, x-flips, and y-flips differentiate between goal-directed and browsing users. We summarize a controlled experiment in an e-commerce setting to test our hypotheses. We then end by discussing the observations from our study and suggest implications for future research.

II. PREVIOUS RESEARCH

Research has shown that monitoring a person's behavior on a website (e.g., the computer mouse, touchscreen, keyboard, etc.) can reveal valuable insights into the user. For example, eye fixations and mouse movements are correlated. Namely, monitoring the computer mouse can show where users' attention is, thus we can better gage what the user is viewing or doing based on the mouse location [9]. Earlier research shows that considering mouse-cursor activity can provide insight

into what users find relevant in a search result and predicts where they will click [10].

On a more sophisticated level, mouse cursor tracking can provide valuable information into the cognitive and emotional state of the user. For example, in a credibility setting, mouse-cursor movements have been used to detect fraud [11], compliance [12], and even deception [13]. In an ecommerce setting, monitoring the computer mouse can provide insight into whether a user is frustrated or experiencing heightened cognitive demand [14]. For example, Hibbiln and colleagues [7] examined how negative emotions influence mouse movements via their burden on working memory. They confirmed in three studies in an online shopping context that mouse-cursor movements and emotion are indeed correlated.

Our research builds on extant literature to examine whether monitoring mouse-cursor movements can be used to differentiate between users who are goal-directed and users who are browsing a website. Previous research has shown that there are eye tracking fixation differences between users who are goal-directed and those who are viewing a website for recreation [15]. While valuable for understanding how attention differs for those user personas, the use of eye tracking is limited in normal e-commerce situations. Hence, there is a need to create theoretically sounds indicators of goal-directed and browsing users that can be used in the "wild" – i.e., normal use situations. In the following section, we describe how mouse movement speed, x-flips, and y-flips are theoretically sound candidates of indicators to differentiate between goal-directed versus browsing users in a natural e-commerce environment.

III. Theory and Hypotheses

To develop our hypotheses, we draw on attentional selection theory. Namely, there are three different strategies that derive where someone allocates attention: goal-driven selection, stimulus-driven selection, and history-driven selection [16]. A goal-driven selection strategy suggests that a person is looking for something specifically and filters out other stimuli that are not relevant to finding that target. For example, if someone is searching for a specific pair of pants on a clothing website, the person will not devote attention to or be distracted by shirts on the website.

Conversely, in a stimulus-driven selection strategy, a person is not specifically looking for one target. Rather, the person is observing the environment and various stimuli compete for attention. The most salient stimuli will ultimately capture the attention. Salience is determined both by presentation (e.g., bold items, large items, color, etc.) and also preference (e.g., you like cars more than clothes, so the cars catch your attention).

Finally, the history-driven selection strategy posits that people allocate attention based on their history with the environment. For example, people ignore an item in a room because through past experience, they know it is not relevant to their current situation.

For the context of this study, we focus specifically on the first two selection processes—goal-driven and stimulus-driven strategies—as they are relevant for differentiating between goal-directed users, and users who are browsing information on an unfamiliar website. It is worth noting, however, that the history-driven selection process is very relevant if someone already has experience with a website. Importantly, how one allocates attention will influence how one navigates a website with a computer mouse, touchscreen, or other input device. Namely, when something draws a person's attention, it increases the likelihood of moving to that object.

This phenomenon is explained through the Response Activation Model, which describes how competing movement responses influence hand movements. The theory explains "attention and action are intimately linked" [17]. All targets that capture a user's attention will prime specific movement responses. To prime a behavior refers to subconsciously programming a movement response (transmitting nerve impulses to the hand and arm muscles) to move toward the stimulus (i.e., a link, image, etc.). The biological reason for priming movements is to enable faster reaction times. It also helps the person focus attention on the target element. For example, if various items catch a users' attention on a page as possible destinations (i.e., because they find the stimuli interesting), the brain will subconsciously and simultaneously prime movements to the various items and people will likely move, at least somewhat, toward the items that catch their attention (see Figure 1 for an example).

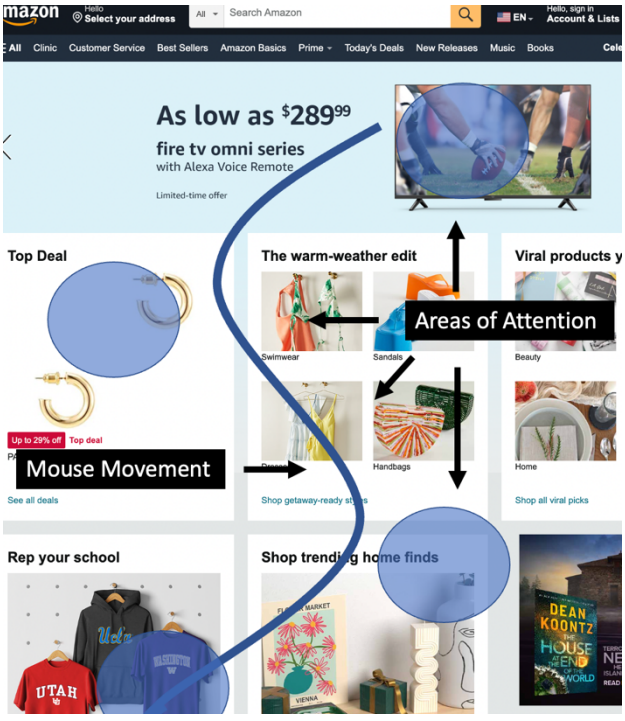


Figure 1. Example of how items that catch a users attention may influence mouse movements.

Combining the stimulus-driven strategy of attentional selection theory and the response activation model, we predict that goal-directed users will have faster mouse movement speed than browsing users. When a user is

browsing in a shopping scenario, they are motivated to explore different information and not immediately make a purchase. Browsing is done for two primary reasons: 1) to research a category of items or 2) for leisurely pleasure [18]. Typically, when people are doing something leisurely, they are not in a rush, they move at a somewhat consistent rate taking in the world as they go. In terms of attentional selection theory, they are in a stimulus-driven selection mode. They take time to allow stimuli to catch their attention, and often move towards those stimuli to focus on them. This increase attention to various stimuli results in the user moving slower as it consumes cognitive resources that could otherwise be used to quickly progress through a task [7]. Conversely, in a goal driven setting, the user utilizes a goal-driven selection strategy, ignoring irrelevant items to their task, and devoting more resources to quickly completing the task.

To illustrate the relation between speed, attention, and motivation, imagine two people walking down the street. One is out for pleasure, window shopping as they go. This person tends to look in the windows as they walk, going at a slow to average pace to devote attention to various items. The other person is out shopping for a very specific item. They walk quickly and with purpose looking in the windows only for the item that might meet their needs, and not devoting attention to other items. This allows them to more quickly accomplish their task.

In summary, when a user is browsing and various items capture attention, the user will overall move the mouse slower. However, when a person is goal directed, the user will move the mouse more quickly ignoring other stimuli to more efficiently accomplish the task.

H1. Goal-directed users will have faster speed than browsing users.

We also predict that users who are goal directed will have more deviation in their movements versus those uses who are browsing. When users browse, they are more likely to utilize a stimulus-driven selection strategy to determine where they devote their attention. Various items of potential interest may catch their attention. As previously explained, when these items catch their attention, the response activation model predicts that they are more likely to move their hand (and in a computer context, the computer mouse) towards the items, resulting in a less direct path.

To illustrate this, again consider the two in-person shoppers. This time, imagine them in a store instead of just window shopping. Let's assume that person A has a strict shopping list. Since she has specific goals, she will walk directly towards the items she needs. This means she will most likely take the shortest and quickest path to those items, assuming she knows where the items are. Compare this with shopper B who we assume isn't sure exactly what he is wanting from the store. He will spend more time wandering around the store and taking indirect routes before arriving at an item they think they want.

Translating these shopping habits to a digital world, one way to measure less direct movements is through x- and y-flips. An x-flip refers to a change in direction on the x-axis and a y-flip refers to a change in direction on the y-axis. Goal directed users are not likely to change directions as much, as they are trying to efficiently complete the task without being distracted by other stimuli. However, a browsing user is likely to switch directions more often as the user's attention, and thus movements, move more frequently among items. In summary, we predict:

H2. Goal-directed users will have fewer x-flips than browsing users.

H3. Goal-directed users will have fewer y-flips than browsing users.

IV. METHODOLOGY

To test our hypotheses, we designed a between-subject experiment where half of the participants were assigned to a goal-directed task and half of the participants were assigned to a browsing task on the same online store interface.

The study was conducted online. First, a survey was administered that collected basic demographics on participants. Afterwards, participants were randomly assigned to the goal-directed treatment or a browsing treatment on a storefront created by the research team to ensure no one had previous experience with the store (see Figure 2)

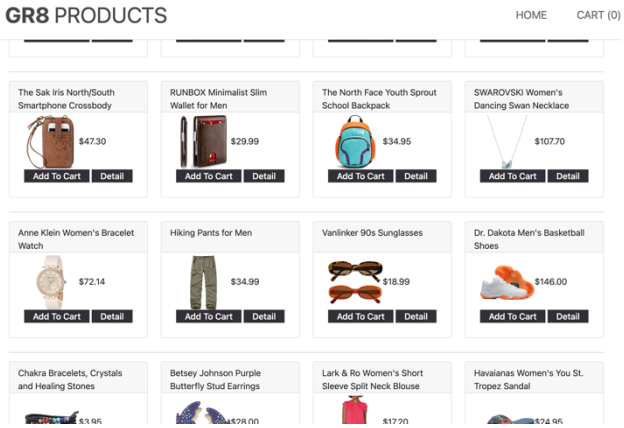


Figure 2. Example of how items that catch a users attention may influence mouse movements.

A. Goal-driven treatment

In the goal driven treatment, participants were told:

You are interviewing for a new job. To prepare, you need to find nice business clothes to wear that fit the company's requirements found below:

- Navy or black bottoms

- A button up top

- Dark colored dress shoes

Visit the website below and find clothes that meet the requirements. Take a screenshot of your final shopping cart and upload it.

B. Browsing treatment

In the browsing treatment, participants were told:

Pretend it is your birthday. Hypothetically, someone gives you the option of $50 in cash or you can get 3 items off the website below. Browse the website for a bit and put your decision below.

If you choose three items, take a screenshot of your shopping cart for later.

C. Data capture

Participants then visited the website to complete their task. This website contained embedded JavaScript that captured behavioral data as the user interacted with the website. The library captured the x- and y-location of mouse movements, along with timestamps. It then sends the data to the server to calculate speed, x-flips, and y-flips.

To calculate speed, we first calculated the cursor distance for each participant using the Euclidean distance between two x/y positions a_i and a_{i+1}: $d(a_i, a_{i+1}) = \sqrt{(a_i^{(x)} - a_{i+1}^{(x)})^2 + (a_i^{(y)} - a_{i+1}^{(y)})^2}$, leading to a total distance of $D = \sum_{i=1}^{n-1} d(a_i, a_{i+1})$ between the recorded points a_1, a_2, …, a_n. Cursor speed was then calculated as a function of cursor distance D and movement time t during the task, measured in pixels per millisecond: $v = D / t$.

X- and y-flips were calculated through logic that recorded which way the user was moving on the x- and y-axis. Then, if the direction of movement changed, the number of x- or y-flips were incremented depending on whether the change occurred on the x-axis, y-axis, or both. Figure 3 – 6 demonstrates scenarios of no flips, an x-flip, a y-flip, and a combined x- and y-flip.

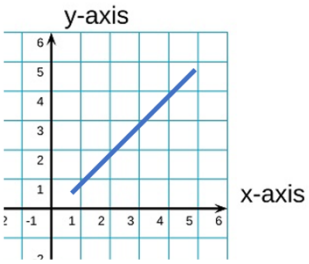


Figure 3. No flip

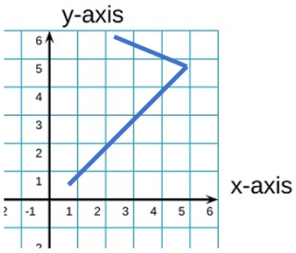


Figure 4. X-flip

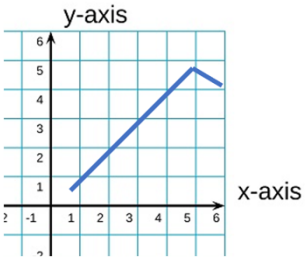


Figure 5. Y-flip

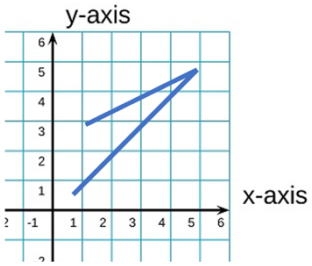


Figure 6. Both x- and y-flip

We also recorded which treatment all participants were in so we could examine differences in speed and x- and y-flips per treatment.

D. Participants

One-hundred-sixty-six participants were recruited from Amazon's Mechanical Turk to participate in the study. They were paid $2 for an approximately 10-minute task, which is equivalent to a $12 hourly rate. Forty-three percent of participants were between the age of 35 – 44; the second most common age group was 25-34 (31%). Sixty-one percent of the participants were male. The most common ethnicities were White (61%), Asian (32%), and African American (6%). Participants were limited to those who had a computer mouse. In some cases, technical difficulties made it impossible to calculate statistics and those participants were therefore excluded as denoted in the analysis.

E. Analysis

We conducted Welch two sample t-tests to test our hypotheses. First, we examined whether speed was different between users in the goal-directed task and the browsing task. The mean speed in the browsing group was 1.99 (sd = 1.59) pixels / ms, and the mean for the goal directed group was 1.36 (sd = 1.36) pixels / ms. The difference was significant: t = 3.30 (104.90), $p < .001$. Hence, the people in the browsing condition moved significantly faster than people in the goal-directed condition. H1 was thus not supported; although it was significant, it was significant in the opposite direction as hypothesized.

Next, we examined whether x- and y-flips were different between users in the goal-directed tasks and the browsing task. The mean x-flips and y-flips in the browsing group was 21.04 (sd = 18.27) and 26.11 (sd = 24.51) respectively, and the means for the goal directed group were 29.98 (sd = 26.50) and 37.32 (sd = 35.73) respectively. Both the differences for x-flips [t (143.79) = -2.52, p < .01] and y-flips [t (143.40) = -2.34, p < .05] were significant. Hence, H2 and H3 were not supported; although they were significant, they were significant in the opposite direction as hypothesized. The results are summarized in Table 1.

TABLE I. SUMMARY OF RESULTS

Hypothesis	*Browsing mean (standard deviation)*	*Goal Driven mean (standard deviation)*	*Result*
H1. Goal-directed users will have faster speed than browsing users	1.99 (1.59) pixels / ms	1.36 (0.64) pixels / ms	Opposite Effect Observed***
H2. Goal-directed users will have fewer x-flips than browsing users	21.04 (18.27)	29.98 (26.50)	Opposite Effect Observed **
H3. Goal-directed users will have fewer y-flips than browsing users	26.11 (24.51)	37.32 (35.73)	Opposite Effect Observed *

* p < .05, ** p < .01, p < .001

V. DISCUSSION AND FUTURE RESEARCH

In our study, we examined the influence of a goal-directed task on movement speed, x-flips, and y-flips compared to a browsing task. We predicted that one who is goal directed would move faster to achieve the goal and deviate less in terms of x- and y- flips. To test our hypotheses, we created an experiment where participants recruited from Amazon Mechanical Turk were randomly assigned to a shopping task where they were asked to find a specific product (goal-directed task) or explore the website to find products that might be interesting to them (browsing task). Contrary to our predictions, participants in the browsing task had less deviation and moved slower (H1, H2, and H3 not supported; significant opposite effect observed).

Upon further evaluation of the experimental design, we are hesitant to conclude that indeed a person in a browsing state would normally move faster and have less deviation than a person in a goal driven state. Rather, we believe the difference is due to the sample population: Amazon Mechanical Turk. Namely, despite our manipulations, it is probable that the majority of Mechanical Turk Workers were driven by a high goal-directed motive: to adequately complete the task as quickly as possible to be paid so they can move on to the next task. Thus, we propose that perhaps both groups could be in a goal-directed state, but the participants who receive the goal-directed manipulation had more requirements to complete the task (find a specific product) that required more careful evaluation, slowing down speed and creating more deviation in their search task. Whereas people in the browsing manipulation did not have requirements to find any product specifically, so they were able to put in much more minimal effort and could complete the task with lesser involvement. This is evident in that participants in the browsing group completed the task in 20 seconds less than people in the goal-directed tasks (71,284 ms compared to 93,702 ms on average).

We believe that our observations potentially have important implications for the use of Amazon Mechanical Turk as research subjects [19-21]. One must consider how the goal-directed motivation of workers to complete tasks efficiently to move onto the next task interacts with the manipulations in the experiment. There are several reasons why an Amazon Mechanical Turk worker's motivation to complete a task efficiently differs from a normal population's motivation as shown in Table II.

TABLE II. POSSIBLE DIFFERENCES BETWEEN AN AMAZON MECHANICAL TURK WORKER AND OTHER POPULATIONS

Category	*Amazon Mechanical Turk Worker*	*Non-Crowd Sourced Worker*
Compensation	Daily compensation is determined by how many tasks can be completed	Daily compensation is limited to this task; there is likely not another immediate task to complete
Task Experience	Has experience in completing tasks efficiently.	Doing an experimental task is novel.
Motivation	More universely motivated to complete task efficiently to earn more money, resulting in a systematic pattern.	A variety of motivations that are randomly distributed across participants, resulting in random noise.

In the case of our experiment, not considering the motivation may have unintentionally changed our manipulations. Instead of having a goal-directed and browsing manipulation, we more likely had a goal-directed with specific requirements manipulation and a goal-directed with minimal requirements manipulation (Figure 7)

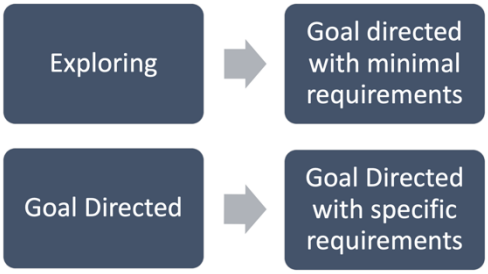


Figure 7. How a Amazon Mechanical Turk sample population influenced the intended manipulations of this study

Hence, we recommend that researchers ask the following question before selecting Amazon Mechanical Turk as the sample population:

Does the motivation of the sample population matter?

In many cases, it does not. For example, it likely matters less when participants are selecting their preference for different products, doing classification tasks, or providing their opinion on a variety of matters. However, it likely does matter when the researchers are trying to manipulate motivation, because the results of the study may not generalize to random populations that do not have a systematically higher motivation to complete the task efficiently to increase their daily compensation. It is also questionable whether Amazon Mechanical Turk workers are appropriate for tasks where the research is attempting to generalize elements of timing to other populations (movement speed, classification time, navigation patterns, etc.).

VI. Conclusion

This paper explores how to analyze mouse-cursor movements to categorize browsing and goal-directed users in an online setting. Drawing on attentional selection theory and response activation model, we predict how mouse-movement speed, x-flips, and y-flips will differentiate between users who are browsing a website versus goal-directed users. We conducted an experiment where participants were randomly assigned to either a goal-oriented task or a browsing task in a mock online store. Our hypotheses were not supported; rather, we found a significant opposite effect. However, we speculate that the results were largely affected by the systematic motivation of our selected sample population – Amazon's Mechanical Turk. We discuss learnings on when it is appropriate to use Amazon's Mechanical Turk, and implications for designing future studies to classify user personas.

References

[1] IBISworld. "Percentage of business conducted online." (accessed February 6, 2023).

[2] BigCommerce. "30 ecommerce conversion rate optimization steps to help boost sales." bigcommerce.com/articles/ecommerce/conversion-rate-optimization/ (accessed.

[3] The Business Research Company, "Search engine optimization services global market report," 2021.

[4] P. Faller. "https://xd.adobe.com/ideas/process/user-research/putting-personas-to-work-in-ux-design/." Adobe. (accessed February 6, 2023).

[5] N. Statt. "Apple updates safari's anti-tracking tech with full third-party cookie blocking." The Verge. https://www.theverge.com/2020/3/24/21192830/apple-safari-intelligent-tracking-privacy-full-third-party-cookie-blocking (accessed February 6, 2023).

[6] C. Dilmegani. "Third Party Cookies in 2023: What Will Their Ban Change?" https://research.aimultiple.com/third-party-cookie/#:~:text=After%20Safari%20and%20Firefox%2C%20Google%20will%20stop%20Chrome%E2%80%99s,such%20as%20the%20General%20Data%20Protection%20Regulation%20%28GDPR%29 (accessed February 6, 2023).

[7] M. Hibbeln, J. L. Jenkins, C. Schneider, J. S. Valacich, and M. Weinmann, "How is your user feeling? Inferring emotion through human-computer interaction devices," Mis Quart, vol. 41, no. 1, pp. 1–21, 2017, doi: 10.25300/misq/2017/41.1.01.

[8] P. Boi, G. Fenu, L. D. Spano, and V. Vargiu, "Reconstructing user's attention on the web through mouse movements and perception-based content identification," ACM Transactions on Applied Perception (TAP), vol. 13, no. 3, pp. 1-21, 2016.

[9] H. P. Lu and P. Y. J. Su, "Factors affecting purchase intention on mobile shopping web sites," Internet Research, vol. 19, no. 4, pp. 442-458, 2009.

[10] J. Huang, R. W. White, G. Buscher, and K. Wang, "Improving searcher models using mouse cursor activity," in international ACM SIGIR Conference on Research and Development in Information Retrieval, 2012, pp. 195-204.

[11] M. Weinmann, J. S. Valacich, C. Schneider, J. L. Jenkins, and M. Hibbeln, "The path of the righteous: Using trace data to understand fraud decisions In real time," Mis Quart, vol. 46, no. 4, pp. 2317–2336, 2022, doi: 10.25300/misq/2022/17038.

[12] J. L. Jenkins, J. S. Valacich, A. F. Zimbelman, and M. F. Zimbelman, "Detecting noncompliant behavior in organizations: How online survey responses and behaviors reveal risk," Journal of Management Information Systems, vol. 38, no. 3, pp. 704-731, 2021.

[13] J. L. Jenkins, J. G. Proudfoot, J. S. Valacich, G. M. Grimes, and J. F. Nunamaker, Jr., "Sleight of hand: Identifying concealed information by monitoring mouse-cursor movements," J Assoc Inf Syst, vol. 20, no. 1, pp. 1–32, 2019, doi: 10.17705/1jais.00527.

[14] G. M. Grimes and J. Valacich, "Mind over mouse: The effect of cognitive load on mouse movement behavior," presented at the International Conference on Information Systems, Fort Worth, TX, 2015, 9.

[15] Y. M. Hwang and K. C. Lee, "Using eye tracking to explore consumers' visual behavior according to their shopping motivation in mobile environments," Cyberpsychology, Behavior, and Social Networking, vol. 20, no. 7, pp. 442-447, 2017.

[16] J. Theeuwes, "Goal-driven, stimulus-driven, and history-driven selection," Current Opinion in Psychology, vol. 29, pp. 97-101, 2019.

[17] T. N. Welsh, "The relationship between attentional capture and deviations in movement trajectories in a selective reaching task," Acta Psychol, vol. 137, no. 3, pp. 300–308, 2011, doi: 10.1016/j.actpsy.2011.03.011.

[18] P. H. Bloch and M. L. Richins, "Shopping without purchase: An investigation of consumer browsing behavior," ACR North American Advances, 1983.

[19] P. B. Lowry, J. D'Arcy, B. Hammer, and G. D. Moody, "'Cargo Cult' Science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, Including mechanical turk and online panels," The Journal of Strategic Information Systems, vol. 25, no. 3, pp. 232-240, 2016.

[20] G. Paolacci, J. Chandler, and P. G. Ipeirotis, "Running experiments on Amazon Mechanical Turk," Judgment and Decision making, vol. 5, no. 5, pp. 411-419, 2010.

[21] J. H. Cheung, D. K. Burns, R. R. Sinclair, and M. Sliter, "Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations," Journal of Business and Psychology, vol. 32, pp. 347-361, 2017.

Physiological and Socio-Behavioral Determinants of Viral Video User Engagement

D. Bačić*, C. Gilstrap ** and N. Jukić*
* Loyola University Chicago, Chicago, USA
**University of Southern Indiana, Evansville, USA
dbacic@luc.edu, cgilstrap1@usi.edu, njukic@luc.edu

***Abstract* - The use of social media platforms like TikTok, Facebook, and YouTube to consume videos has seen significant growth in recent years. Some videos can generate high levels of engagement - views, likes, shares, or comments - at a rapid pace, making them "viral." However, understanding what causes these videos to become viral or predicting virality is still largely a mystery. This study aimed to uncover if biometrics-based emotion and arousal data, such as facial muscle and skin conductance data, could predict user engagement and contribute to video virality. The experiment used 64 participants, who watched 13 videos and had their facial expressions and galvanic skin response (GSR) data recorded throughout the viewing experience. The study then used an XGBoost classifier and 42 collected features, including physiological data and socio-behavioral responses, to predict user engagement (willingness to like, share, or comment) with over 80% accuracy. The results indicate that a combination of facial expression, GSR, and socio-behavioral data can accurately distinguish between high and low user engagement without needing to ask viewers anything about the videos or analyze video content. This study elevates the role of viewers' physiological and subconscious responses to video content across the viewing experience.**

Keywords - Video virality; Emotions; Facial expression; Galvanic skin response; XGBoost; Predictive modeling

I. INTRODUCTION

Video sharing has experienced tremendous growth over the previous decade [1]. Industry marketers and researchers alike have recognized that some videos garner high viewer engagement. Knowledge of how these videos become viral remains limited and elusive. Not surprisingly, the task of predicting which videos will become viral is even more challenging [2]. The present paper accepts this challenge by exploring the use of biometric-based user emotion data in conjunction with more advanced machine learning-based analytics to predict individuals' willingness to engage with videos on social media.

Social media platforms such as TikTok, Facebook, and YouTube offer social interaction, information, news, and entertainment via video modalities for large segments of the population [1]. Achieving 'viral' status has become the holy grail of digital marketing as well as a status sought by social video content creators. However, achieving virality is hard, rare, and valuable [3].

Researchers know emotions to be essential in the context of video virality and viewer engagement, but little progress has been made to convert emotional content-to-virality knowledge into actionable, predictive models. Recent advances in emotion detection across human-computer interaction research streams generally [4] and biometrics specifically based upon facial muscle movement, skin conductance, and other observations [5] offer new insights into users' affective states, including those of video viewers. New biometric methods, based upon Facial Action Coding System (FACS) [6-7] and skin conductance [8], assess emotions and emotion-related arousal at an instinctive and subconscious level during video consumption rather than in moments or metrics after video consumption thus increasing the likelihood video virality can be predicted.

The present study intends to generate a video virality predictive model based upon biometric data captured during the whole of the video viewing activity. To do so, the following research question is posed: Can users' physiological manifestation of emotions captured through facial expressions and skin conductance during video viewing be used to predict viewers' video engagement successfully?

II. BACKGROUND

A. Video Virality

A viral video is one viewed and shared expeditiously across popular social platforms. Common to all virality definitions is that a viral video takes advantage of the network effect [9], where the term 'viral' is a metaphoric reference to a contagious virus easily spreading quickly from one host to another. Emotion-focused research found that expressing emotional connection and emotional generosity in social media may prompt viral video sharing [10], while videos using emotion-eliciting strategies were more likely to be shared [3]. The relationship between video virality messages and emotions was also explored through the lens of consumers' emotional reactions to videos [11].

Video virality is of interest to computer science and information systems disciplines, especially in predictive modeling. For instance, Tik Tok dance video virality prediction was proposed using a multi-modal framework that integrates skeletal, holistic appearance, facial and scenic cues [12]. Others used models leveraging Support Vector Regression [13] and neural networks [14] to predict the future popularity of YouTube videos.

Problematic for virality studies however, previous research has relied upon two processes which may lead to incomplete data or incomplete analyses: non-subconscious data collection and self-reporting. To respond to both of these issues, the present study forges new paths for virality research by including subconscious data collected from video viewers during video viewing with no reliance upon self-reporting after the viewing experience.

B. *Biometric Sensors: Facial Muscle Movement and Galvanic Skin Response*

More recent advances in biometrics, based on Facial Action Coding System (FACS), provide new methods of assessing emotions at an instinctive and subconscious level. FACS, a method designed to help classify human fascial muscle movement, is used by facial expression researchers and human coders to identify and understand facial action units (AU) or basic facial movement building blocks from which more complex facial expressions can be understood [17]. More recently, tools automating facial expression analysis have emerged. Consequently, we have witnessed an increase in research and practical application of automated facial expression tools over the last three decades.

Galvanic Skin Response (GSR), also referred to as electrodermal activity (EDA), Skin Conductance Response (SCR), or Psychogalvanic Reflex (PGR), is a measure of the variation of conductivity of human skin due to sweat secretion [8]. GSR research has found that the increase in SCR (or "GSR peak") is linked to an increase in emotional arousal [8]. SCR, as a response to a stimulus extant usually within 1-5 seconds following the stimulus, is called Event-Related SCR [18], and it has been deployed by researchers to detect correlations between the arousal level of emotions [19], [20]. Although GSR data cannot provide the direction of the emotion it measures by itself, simultaneous use of a GSR sensor and automated FACS is a particularly effective way of understanding both the valance and the arousal associated with an emotional response [21].

Extant literature suggests that a viewer's emotions trigger participatory behavior as an essential element to video virality. Building on this notion, the present experiment was designed and deployed to collect rich facial expression consisting of over 20 emotional channels/AUs, seven basic emotions, valence, and visual engagement and attention variables, as well as GSR (peaks per minute) datasets to predict participatory behavior via users' willingness to like, share, and/or comment upon the videos they view.

III. Experiment And Data Collection

A. *Subjects*

Sixty-four undergraduate students from a public Midwestern university participated in data collection. About 34% (22) of participants identified as females. 97% (62) of subjects were between 19 and 22 with an average age of 20.2 years old, and most considered themselves dependents. About 72% (46) of participants were employed part-time (<40 hrs. per week). Virtually all participants identified as "white" (63) and never married (62). Most participants recently shared, commented, or liked video content on social media (during the day of the experiment or in the last seven days), with 'like' being the most frequent engagement activity.

B. *Experiment Set-up and Data Collection*

Subjects watched 13 videos in random order (See Appendix, Table 3 for a complete list). During video viewing, each participant's face and reactions were recorded using an HD camera and collected using the iMotions biometric platform through its Affectiva AFFDEX algorithm and iMotions' Facial Expression module. Similarly, GSR data was collected using the Shimmer GSR+ sensor and iMotions' GSR module. A Shimmer GSR+ device was placed on participants two fingers of one hand to monitor skin conductivity between two reusable electrodes (Figure 1).

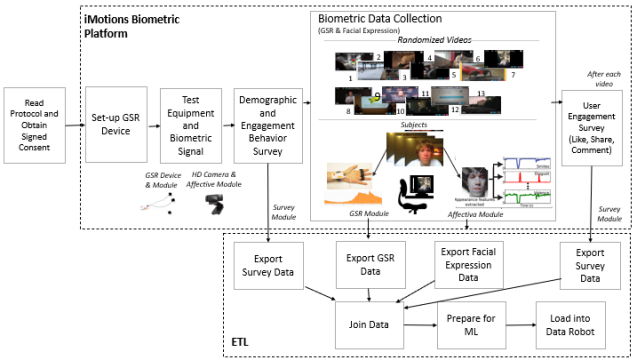


Figure 1: Experimental Setup & Data Collection

Before biometric data collection, each participant completed a survey to collect socio - demographic and past video engagement behavioral data. The answers were recorded and labeled as socio-behavioral features of 'Last Time Shared' (LAST_ENG1), 'Last Time Commented On' (LAST_ENG2), 'Last Time Liked' (LAST_ENG3) (for each question, the answer options were 'Today', 'Last 7 days', 'Last 30 Days', 'Last 360 Days', and 'Never), Income, Gender, Education & Employment status.

Each participant's video recording was post-processed using Affectiva's AI algorithm (AFFDEX) to capture twenty emotional channels, three valance categories (positive, neutral, and negative valance), two engagement metrics (engagement and attention time), and seven basic emotions (anger, contempt, disgust, fear, joy, surprise, and sadness) from participants' facials expressions. Counts for each biometric metric collected were excluded so that the input file only includes % of the time for each variable. This was done to eliminate bias due to differences in video lengths. Given the exploratory nature of the present research, and a relatively low number of features and data records (fast computation time), this study selected to leave all 42 features for model building (See Appendix, Table 2 for a complete list).

After each video, participants were asked to rate how likely they were to like, share, and comment on the video (1-7 scale). A prediction target feature was calculated (AVG_USER _ENGAGEMENT), using the average of Like, Share, Comment variable scores and then discretized

into High (above mean) and Low (below mean) as the prediction target.

IV. MODELING

A. Predictive Algorithm and Model Selection

Once data was loaded into DataRabot, this study used a number of classifiers and 77 resulting models to find the best model. Cross-entropy loss, or Log Los, as a measure of the inaccuracy of predicted probabilities, was used to rank the models, an appropriate measure given the binary classification problem (High (1)/Low (0)). All models were created using 80% of data for training & validation and 20% as a holdout to test the model. K-folds (5 folds) were used in all models The initial list of models was ranked using validation results based on 64% of the data used in training the models (in training and validation folds, The top eight models were kept on the list and used in cross-validation (across five folds). The top model (again, based on Log loss performance), using eXtreme Gradient Boosted Trees (XGBoost) Classifier survived and was used against the hold-out or 20% of data (167 records) that our model was blind to when the best model was selected.

XGBoost is a very efficient, parallel version of the Gradient Boosting Machines (GBM) classifier [22] that has been heavily optimized for faster runtimes and higher predictive accuracy [23]. GBMs are an advanced algorithm (supervised) for fitting highly accurate predictive models. GBMs generalize Freund and Schapire's AdaBoost algorithm that handles arbitrary loss functions. XGBoost has been applied successfully across contexts and has become a widely popular tool among Kaggle competitors [24] and Data Scientists.

B. Model Evaluation

To assess the performance of the XGBoost-based model, this study followed standard performance measures: (i) True Positive (TP): The number of high user engagement videos correctly predicted as resulting in high user engagement. (ii) False Positive (FP): The number of low user engagement videos wrongly predicted as resulting in high user engagement, (iii) True Negative (TN): The number of low user engagement videos correctly predicted as resulting in low user engagement, and (iv) False Negative (FN): The number of high user engagement videos wrongly predicted as resulting in low user engagement (Figure 2).

n=167

		Predicted: Low (0)	Predicted: High (1)	
Actual	Low (0)	TN 57	FP 28	85
	High (1)	FN 12	TP 70	92
		69	98	

Figure 2: Confusion Matrix

Next, the study applied confusion matrix terms to calculate standard performance measures to evaluate the model (Table 1) using a specific threshold. The Accuracy measure suggests that the candidate model is correct 76.05% of the time or misclassification rate of 23.95% (Error Rate). Sensitivity (or Recall) indicates that when in actuality users indicated high engagement, the candidate model predicted it correctly 85.37% of the time, while incorrectly 14.63% (False Negative Rate). On the other hand, when users indicated low engagement, the candidate model predicted it correctly 67.06% (Specificity) and incorrectly 32.94% (Fallout) of time. Precision measure suggests that when the model predicts "high" user engagement, it is 71.43% correct, and incorrect 28.57% (False Discovery Rate). The F1 Score is an overall measure of a model's accuracy that combines precision and recall. In our case, our model reported the F1 Score of 77.78%.

TABLE 1: MODEL PERFORMANCE

AUC	0.8086	F1 Score	0.7778
Accuracy	0.7605	Error Rate	0.2395
Sensitivity	0.8537	False Neg. Rate	0.1463
Specificity	0.6706	Fallout	0.3294
Precision	0.7143	False Disc. Rate	0.2857

This study evaluated the model's ROC curve by focusing on Area Under Curve (AUC), or the area under the ROC curve (Figure 3). The candidate model's AUC of 0.8086 suggests an 80.9% chance that the model will correctly distinguish viewers' rating as a highly engaging video from a low engaging video. AUC of 0.5 suggests no discrimination, 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding [25]. Based on this AUC heuristic, the candidate model is considered excellent.

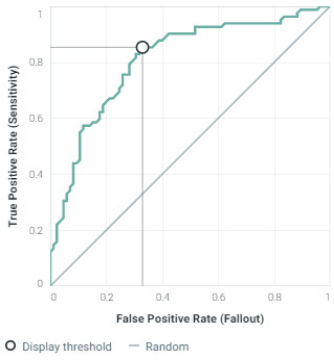


Figure 3: ROC

C. Model Interpretation

The present study used feature impact to interpret how each key feature is driving the model's decision.

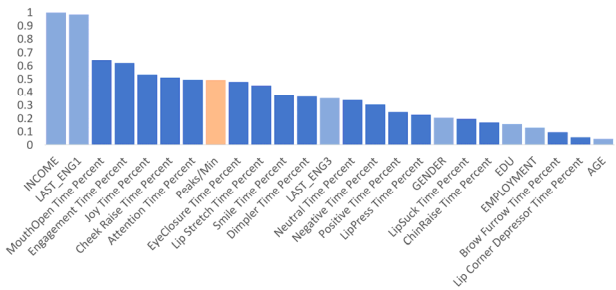


Figure 4: Feature Impact - Top 25 Features

Figure 4 displays the top 25 features, with 'INCOME' and 'LAST_ENG1' (Last Time Shared) 'having the largest impact from the group of socio-demographic and behavioral features. 'Mouth Open Time Percent', 'Cheek Raise Time Percent', and 'Eye Closure' features have the highest impact from the group of emotional channel or action unit features. 'Engagement Time Percent', 'Joy Time Percent', and 'Attention Time Percent' features have the highest impact from the group of features representing a group of combined action units based on FACS. 'Peaks/Min', the only GSR-collected feature, also appears to impact the model's classification decision substantially. Other features not included in Figure 3 have a normalized impact value of less than 0.04.

V. DISCUSSION

This study explored the possibility of using biometrics-based emotion and arousal data (fascial muscle and skin conductance data) to predict user engagement as an essential ingredient to video virality. The AUC evaluation of the candidate model suggests more than 80 % chance that a mix of facial expression, GSR, and socio-behavioral data will correctly distinguish between high and low user engagement.

This study's findings encourage the use of biometric data in predicting and understating factors influencing video virality. The present study finds that physiological measures offer nuance beyond basic emotions and valance by focusing on granular facial muscle movements to extract action units. Furthermore, the concept of arousal is not limited to the verbal description of emotion intensity but can be captured through a more objective measure of skin conductance.

Next, the present results emphasize the value of existing research investigating the emotional components of user experience relative to viewer engagement and virality [26]. In the process of interpreting the present model, our analysis confirmed the role of viewers' emotional responses such as joy, positive valance, and smile action unit. The predictive model also confirmed the role of arousal ('Peaks/Min') in user engagement, further validating the role of emotional arousal.

Focusing on analytics and systems, Information Systems, Computer Science, and Decision Science research communities can utilize the resulting data to continue improving video virality predictive models while further exploring video engagement from data and systems perspectives. Facial expression data, in particular, is valuable for modeling vis-a-vis machine learning as it provides information with various levels of granularity incorporating actions units and combinations of those units to capture emotions, valence, attention, and engagement. These information-rich features minimize the need for significant data engineering and further feature extractions that can add levels of complexity to interpret advanced algorithm results.

A. Practical Implication

From a practical perspective, the current paper is commercially contextualized, given that video strategy is an increasingly important component of marketing strategy and that companies and digital platforms are competing for the finite resource of viewers' attention. Video platforms offer significant entertainment and commercial value predicated upon eliciting user engagement. Given the ubiquity of videos and the potential value of viral videos in customer engagement, entertainment, marketing, and overall business strategy, improvements in (i) understanding factors that influence user engagement, (ii) ability to predict engagement, and (iii) ability to integrate physiological data into existing video creation and selection have the aforementioned clear and practical implications.

The present study edifies video creators' intent on virality-strategy as to what emotional responses to target (joy, positive valence, smile) when creating video content. Similarly, when multiple video versions are created, video makers can use a combination of social and behavioral-related and biometric features identified herein and use them with advanced classifiers to decide which video to be released. Combining this predictive model's accuracy with the known economic impact of video virality provides evidence that time and resources investment can yield a return on investment. However, investments should not completely replace existing processes inclusive of user engagement/video domain knowledge but, rather, should augment and enhance it.

B. Limitations

The present study has several limitations. First, the characteristics of the participants might have influenced the results. Second, the selection of videos and various video elements can provoke unique reactions from participants. Third, this study selected only two types of biometric data - facial expression and GSR – as the source of emotion and arousal data. Fourth, this study should be evaluated through the aforementioned research goal and exploratory analysis and predictive modeling scope.

VI. CONCLUSIONS

The present study's ability to create a predictive model represents the first attempt to explore the predictive use of this biometric data type in the context of video virality. In the process, the study (i) introduced the potential of facial expression and GSR data, (ii) confirmed the role of emotional response during video viewing as well as the resulting tendency of an audience to engage with viewed videos, (iii) elevated the need to integrate both socio-behavioral and physiological data, and (iv) provided evidence of the value of advanced analytics in the context of biometric data and video virality.

REFERENCES

[1] M. L. Khan, "Social media engagement: What motivates user participation and consumption on YouTube?," Comput Human Behav, vol. 66, pp. 236–247, 2017, doi: https://doi.org/10.1016/j.chb.2016.09.024.

[2] K. Motoki, S. Suzuki, R. Kawashima, and M. Sugiura, "A Combination of Self-Reported Data and Social-Related Neural Measures Forecasts Viral Marketing Success on Social Media," Journal of Interactive Marketing, vol. 52, pp. 99–117, 2020, doi: https://doi.org/10.1016/j.intmar.2020.06.003.

[3] E. Akpinar and J. Berger, “Valuable virality,” Journal of Marketing Research, vol. 54, no. 2, pp. 318–330, 2017, doi: 10.1509/jmr.13.0350.

[4] K. Sailunaz, M. Dhaliwal, J. Rokne, and R. Alhajj, “Emotion detection from text and speech: A survey,” Soc Netw Anal Min, vol. 8, no. 1, Dec. 2018, doi: 10.1007/s13278-018-0505-2.

[5] A. De Keyser, Y. Bart, X. Gu, S. Q. Liu, S. G. Robinson, and P. K. Kannan, “Opportunities and challenges of using biometrics for business: Developing a research agenda,” J Bus Res, vol. 136, pp. 52–62, Nov. 2021, doi: 10.1016/j.jbusres.2021.07.028.

[6] P. Ekman, W. V Friesen, and J. C. Hager, “Facial action coding system: The manual: Research Nexus devision of Network Information Research Corporation,” Salt Lake City, UT, 2002.

[7] C. H. Hjortsjö, Man’s Face and Mimic Language. Studen litteratur, 1969.

[8] M. Benedek and C. Kaernbach, “A continuous measure of phasic electrodermal activity.,” J Neurosci Methods, vol. 190, no. 1, pp. 80–91, Jun. 2010, doi: 10.1016/j.jneumeth.2010.04.028.

[9] T. Broxton, Y. Interian, J. Vaver, and M. Wattenhofer, “Catching a Viral Video,” in 2010 IEEE International Conference on Data Mining Workshops, 2010, pp. 296–304. doi: 10.1109/ICDMW.2010.160.

[10] A. Nikolinakou and K. W. King, “Viral video ads: Emotional triggers and social media virality,” Psychol Mark, vol. 35, no. 10, pp. 715–726, 2018, doi: 10.1002/mar.21129.

[11] E. Botha and M. Reyneke, “To share or not to share: the role of content and emotion in viral marketing,” J Public Aff, vol. 13, no. 2, pp. 160–171, May 2013, doi: https://doi.org/10.1002/pa.1471.

[12] J. Wang et al., “Will You Ever Become Popular? Learning to Predict Virality of Dance Clips,” CoRR, vol. abs/2111.0, 2021.

[13] T. Trzcinski and P. Rokita, “Predicting Popularity of Online Videos Using Support Vector Regression,” IEEE Trans Multimedia, vol. 19, no. 11, pp. 2561–2570, 2017, doi: 10.1109/TMM.2017.2695439.

[14] W. Stokowiec, T. Trzcinski, K. Wolk, K. Marasek, and P. Rokita, “Shallow reading with Deep Learning: Predicting popularity of online content using only its title,” CoRR, vol. abs/1707.0, 2017.

[15] P. Ekman, W. V Friesen, and J. C. Hager, “Facial action coding system: The manual: Research Nexus devision of Network Information Research Corporation,” Salt Lake City, UT, 2002.

[16] C. H. Hjortsjö, Man’s Face and Mimic Language. Studen litteratur, 1969.

[17] W. Li, F. Abtahi, and Z. Zhu, “Action unit detection with region adaptation, multi-labeling learning and optimal temporal fusing,” Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 6766–6775, 2017, doi: 10.1109/CVPR.2017.716.

[18] M. E. Dawson, Schell, AM, and D. L. Filion, “The electrodermal system,” in Handbook of psychophysiology, B. G. Cacioppo JT, Tassinary LG, Ed. Cambridge: University Press, 2007, pp. 159–81.

[19] A. Nakasone, H. Prendinger, and M. Ishizuka, “Emotion Recognition from Electromyography and Skin Conductance,” in The Fifth International Workshop on Biosignal Interpretation (BSI-05), 2005, pp. 209–2012.

[20] W. Boucsein, Electrodermal activity. Springer Science & Business Media, 2012.

[21] D. Bačić, “Understanding Business Dashboard Design User Impact: Triangulation Approach Using Eye-tracking, Facial Expression, Galvanic Skin Response and EEG Sensors,” in AMCIS 2017, 2017.

[22] R. P. Sheridan, W. M. Wang, A. Liaw, J. Ma, and E. M. Gifford, “Extreme Gradient Boosting as a Method for Quantitative Structure-Activity Relationships,” J Chem Inf Model, vol. 56, no. 12, pp. 2353–2360, 2016, doi: 10.1021/acs.jcim.6b00591.

[23] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, vol. 13-17-Augu, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.

[24] R. Bekkerman, “The Present and the Future of the KDD Cup Competition: an Outsider’s Perspective,” 2015. https://www.linkedin.com/pulse/present-future-kdd-cup-competition-outsiders-ron-bekkerman/

[25] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, Applied logistic regression, vol. 398. John Wiley & Sons, 2013.

[26] A. Dobele, A. Lindgreen, M. Beverland, J. Vanhamme, and R. van Wijk, “Why pass on viral messages? Because they connect emotionally,” Bus Horiz, vol. 50, no. 4, pp. 291–304, 2007, doi: https://doi.org/10.1016/j.bushor.2007.01.004.

APPENDIX

TABLE 2: FEATURE LIST

#	Feature	Feature Type	#	Feature	Feature Type
1	GENDER	Socio-Dem.	22	Brow Furrow Time Percent	AU/Em. Channel
2	AGE	Socio-Dem.	23	Brow Raise Time Percent	AU/Em. Channel
3	EDU	Socio-Dem.	24	Lip Corner Depressor T. %	AU/Em. Channel
4	MARITUAL_ST	Socio-Dem.	25	Smile Time Percent	AU/Em. Channel
5	EMPLOYMENT	Socio-Dem.	26	InnerBrowRaise T. Percent	AU/Em. Channel
6	INCOME	Socio-Dem.	27	EyeClosure Time Percent	AU/Em. Channel
7	LAST_ENG1	Behavioral	28	NoseWrinkle Time Percent	AU/Em. Channel
8	LAST_ENG2	Behavioral	29	UpperLipRaise Time %	AU/Em. Channel
9	LAST_ENG3	Behavioral	30	LipSuck Time Percent	AU/Em. Channel
10	Engagement Time %	Engagement	31	LipPress Time Percent	AU/Em. Channel
11	Attention Time Percent	Attention	32	MouthOpen Time Percent	AU/Em. Channel
12	Anger Time Percent	Basic Emotion	33	ChinRaise Time Percent	AU/Em. Channel
13	Sadness Time Percent	Basic Emotion	34	Smirk Time Percent	AU/Em. Channel
14	Disgust Time Percent	Basic Emotion	35	LipPucker Time Percent	AU/Em. Channel
15	Joy Time Percent	Basic Emotion	36	Cheek Raise Time Percent	AU/Em. Channel
16	Surprise Time Percent	Basic Emotion	37	Dimpler Time Percent	AU/Em. Channel
17	Fear Time Percent	Basic Emotion	38	Eye Widen Time Percent	AU/Em. Channel
18	Contempt Time %	Basic Emotion	39	Lid Tighten Time Percent	AU/Em. Channel
19	Positive Time Percent	Valence	40	Lip Stretch Time Percent	AU/Em. Channel
20	Negative Time Percent	Valence	41	Jaw Drop Time Percent	AU/Em. Channel
21	Neutral Time Percent	Valence	42	Peaks/Min	GSR/Arousal

TABLE 3: VIDEO SUMMARY

Video	Interesting/ Suprising	Amusing/ Funny	Annoying/ Boring	Emotional/ Intense	Length	Views
1 - 0.38 second rubric's cube solve https://youtu.be/nt00QzKuNVY	X				0:30	29,348,799
2-Amtrack snow-mo collision https://youtu.be/V-Cb9x70gYQ	X				0:43	31,102,210
3-Dogs favorite toy https://youtu.be/l7op92W7voE		X			0:59	53,855,299
4-Maddie waiting for the beat https://youtu.be/REmpPC9gPq0		X			1:11	13,152,888
5-My cat mumbles https://youtu.be/5ZRMV--H0AA			X		0:53	5,792,368
6-Robot beats - 1 am not a robot https://youtu.be/fsF7enQY8uI	X				0:30	29,296,624
7-Unreal rescue in baton https://youtu.be/UbiPT5VMo8E				X	1:46	4,416,589
8-Chimp at zoo throws poo https://youtu.be/0veiTgUQLKw		X			0:18	4,215,179
9-DISC profile theory https://youtu.be/0veiTgUQLKw			X		1:15	<100
10-The e-LEMON-ators https://youtu.be/M-D8DDjkcRk			X		1:46	<100
11-Accounting circle https://youtu.be/H9axyByA_rw			X		1:31	<100
12-Budweiser brewed by veterans https://www.ispot.tv/ad/wIao/budweiser-folds-of-honor-brewed-by-vets-for-vets				X	0:28	Not released
13-Budweiser 2017 super bowl https://youtu.be/7ZmlRtpzwos				X	1:01	1,863,872

Exploring Pie Charts and Part-To-Whole Alternatives: Eye-tracking Approach

D. Bačić, A Krbanjevic and N. Jukić
Loyola University Chicago, Chicago, USA
dbacic@luc.edu, akrbanjevic@luc.edu, njukic@luc.edu

***Abstract* - In the field of business information visualization (BIV), there are numerous options for displaying data. One popular yet debated choice is the use of pie charts, as their effectiveness and efficiency have been called into question. Despite their widespread use, there is limited empirical research on the effectiveness and efficiency of pie charts compared to other display types. Additionally, existing research primarily focuses on measuring user impact through accuracy and time, neglecting the impact on users' physiological attention and cognitive resources. This study aims to fill this gap by comparing the use of pie charts to seven alternative data representation types in a part-to-whole task. The study employed a randomized, within-subject experiment with 21 participants, utilizing eye-tracking technology to evaluate user performance (accuracy and time) and cognitive effort (eye fixation). The results showed that pie charts were more accurate than bar charts but required more cognitive effort than stacked bars and treemaps. Donut charts required the most cognitive effort among all tested data representations. The study highlighted that time to complete the task may not be the best indicator of user experience and accuracy while elevating the importance of accounting for users' cognitive resources.**

Keywords - pie chart; cognitive effort; eye tracking; data visualization; pie chart alternatives; part-to-whole task

I. INTRODUCTION

Given the relative ease of chart creation using commercial business intelligence tools, visual representations such as pie charts, bar charts, treemaps, and bubble charts are gaining popularity amongst data analysts and other business practitioners. Continued practitioner (especially their executive ranks) and academic fascination with pie chart data representation when displaying solutions to part-to-whole problems is a driving motivation for the present research. There is a lack of conclusive empirical data focused on understanding overall pie chart effectiveness, efficiency, and comparison to alternative display types. Similarly, the available research is almost exclusively focused on measuring user impact through accuracy and time and is not addressing users' physiological attentional resources and cognition. We address this gap by conducting an experiment comparing pie charts to seven alternatives and evaluating user impact relative to performance (accuracy, time) and cognitive effort using eye-tracking technology.

Despite their popularity, the use of pie charts has become heavily criticized [1], [2], in several foundational [3] and recent studies favoring the alternative forms of displays [4]–[7]. For proponents of pie charts, there is a relative consensus that pie charts are best when used for part-to-whole relationships, with the slice size relating to other parts or the whole [7]. However, the proliferation of pie chart misuse (3D display option, large number of pie slices, poor labeling, and color choices) and general lack of understanding of how users 'read' pie charts [8] has resulted in two vocal camps and a vigorous debate centered on the pie chart's appropriateness and user impact.

This current study aims to impact the ongoing discussion and debate about pie charts. More specifically, this study seeks to explore a more thorough evaluation of pie charts by comparing them to alternative options. This evaluation will focus on the impact of pie charts and other part-to-whole alternatives on users and will consider various aspects, including visual attention, cognitive effort, judgment accuracy, and speed of judgment. In the process, the study will utilize the latest eye-tracking technology.

The study will begin by briefly addressing the relevant literature related to pie charts and alternative options, as well as cognitive effort measured through eye-tracking. This will provide the necessary context and background information to understand the purpose and significance of the experiment. Following the literature review, the study will provide an overview of the experiment procedures and setup. The next stage of the study will analyze and present the collected experimental data. The results will be discussed and analyzed, taking into account the findings' implications and limitations and providing suggestions for future research.

II. LITERATURE REVIEW

The first instance of a pie chart was in William Playfair's 1801 Statistical Breviary [9], which defined the pie chart as a simple information graphic whose principal purpose is to show the relationship of a part to the whole [10]. The academic critiques of pie charts eventually led to their growing popularity within business dashboards, flashy presentations, and popular media [2], [11]. A study pitted pie charts against bar charts and ultimately found that pie charts outperformed stacked bar charts but not all types of bar charts [3]; another set of different studies found that pie charts were useful and effective in part-to-whole comparisons [12] for quick comparison calculations. In the context of alternative charts, a study found that treemaps performed worse than pie charts in accuracy, while a pie chart took longer to read and that slice size itself may impact user performance irrespective of display type [4].

Although academic criticism of pie charts continued [4], [7], [8], it is clear that an ongoing debate has prompted a necessity to investigate further user impact when assessing pie charts and their alternatives. While there has

been some progress made in our understanding of pie chart [13] and its alternatives usage through observing user performance [4]–[7], [14] little progress has been made in measuring users' visual attention and cognitive effort, a primary contribution of the present study.

As part-to-whole data visualization literature evaluated user performance through accuracy and time, suggestions have been made that cognitive processes and cognitive effort may be promising avenues to clarify existing findings and inconclusive results [2], [4]. Cognitive effort can be defined as cognitive resources (perceptions, memory, and judgment) needed to complete a task [15]. There are several ways to measure cognitive effort, physiological [16] and perceptual [17]. Technological advances have been made in more recent times, and relevant research found numerous ways to measure cognitive effort physiologically through fMRI, EEG, and eye tracking [15], [16], [18]–[21].

Eye-tracking has been established as the primary physiological sensing technology used to analyze visual stimuli, as it provides valuable insight into a user's attention and effort while they are reading and scanning displayed information [20], [21]. This is achieved by capturing metrics based on eye fixation, saccades, blinking, and pupil dilation [18]. Beyond psychology, eye tracking has been widely adopted in various fields to measure visual attention distribution and assess the impact of user interactions. Salient to this research, studies have found that visualizations with improper use of colors can lead to increased cognitive effort and longer decision-making times [22]. Additionally, eye tracking has been utilized to examine conventional dashboard design guidelines [23], cognitive effort as the cognitive fit mechanism using tables and graphs [21], and to evaluate the impact of pie chart visual properties (arc, angle, area) [13]. Fixation-based eye tracking measures (fixation is the stabilization of the eye on an object), such as fixation count and duration, continue to be the preferred method for evaluating cognitive processes and effort [16], [21], [13].

III. Experiment

We conducted an experiment to evaluate users' performance and cognitive effort when evaluating pie charts and alternative charts. In our experiment, we measured user performance through accuracy, speed, and effort. Demographic data (education level, age, gender identity, and major) and task answers were collected through a survey module integrated into the iMotions biometric platform. Eye tracking data was collected using the SmartEye AI-X EyeTracker (0.5-degree accuracy, 60hz sample rate, 0.01-degree precision). All images were presented on a 22", 1920 x 1080 resolution monitor, and collected data were stored and analyzed using iMotions software (version 9.1). The default iMotions I-VT filter setting was used to process gaze information, including minimum fixation duration of 60ms.

A. Experimental Design

We conducted a within-subject experiment where chart type, chart order, and part size (2%, 3%, 5%, 9%, 16%, 19%, 62%, 67%) were fully randomized to avoid bias due to the learning effect. Subjects were asked to estimate the size of the part relative to the whole (100%) for a category.

A randomization algorithm was created to present each subject with all eight chart types and all part sizes (in total). For example, subject 1 was presented with eight experimental conditions (part size in parenthesis): pie chart (participant is shown 1 of the 8 pie charts, for example, the one with 16%-part size), horizontal bar (5%), vertical bar (2%), lollipop chart (3%), bubble chart (19%), stacked bar (9%), treemap (67%), and donut chart (62%). The combinations for subject 2 will then be randomized so that they were also given one instance of each part size combined with different charts yet still exposed to each chart type (also randomized) (Figure 1).

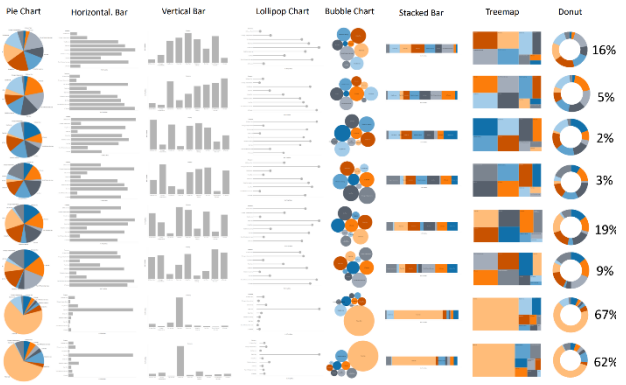


Figure 1: Fully randomized within subject design

B. Participants

Students from a large private midwestern university participated and were recruited from two business undergraduate classes (Data Visualization, Introduction to Information Systems). Students were rewarded with a small number of class bonus points for participating. There were 26 participants, 50% males, all undergraduates, mostly business majors, 77% with junior or senior class standing, and an average age of 20.53 years (Table 1).

TABLE 1: PARTICIPANTS

Gender				**Age**		
Male	13	50.0%		18	1	3.8%
Female	12	46.2%		19	1	3.8%
Non-Binary	1	3.8%		20	10	38.5%
				21	11	42.3%
Class				22	3	11.5%
Freshman	2					
Sophomore	14			**Major**		
Junior	6			Business	25	96.2%
Senior	4			Non-business	1	3.8%

C. Experimental Conditions

As mentioned, subjects were given a baseline pie chart and seven alternative charts (Figure 2) while controlling for the number of categories (each chart displayed eight).

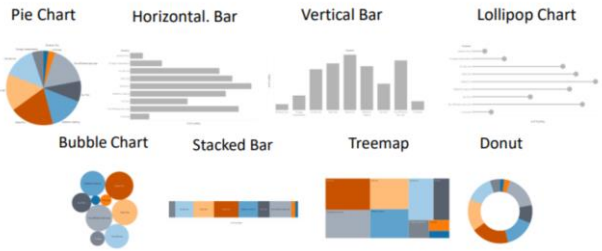


Figure 2: Experimental conditions

We implemented standard best practices found in chart design, such as using a color-blind palette. For each chart, except the donut chart, the categories were directly labeled on the chart itself. The color was used in addition to text to identify categories in all cases except in horizontal bar, vertical bar, and lollipop charts to preserve graphical best practices and mimic what practitioners experience when designing these types of charts.

D. Variables

Cognitive effort and visual attention were measured through fixation duration (average length of all fixations) and count (the number of eye fixations). Accuracy was measured as the subjects' answer nominal % difference (error) relative to the correct answer (part size). Time was measured in milliseconds.

IV. Data Analysis

Data from five participants were excluded from the analysis due to poor eye-tracking calibration and misunderstanding of the task. After removing invalid data points from five respondents, the remaining data from the 21 participants (43% males), all undergraduates, average age= 20.6 yrs (SD=0.18)) was used for analysis.

Before evaluating performance across experimental conditions (chart types), participant performance is compared across all possible task solutions for all chart types combined (Table 2).

TABLE 2: ACCURACY ACROSS PART SIZES

Task Solution (a)	Avg. Solution (b)	Error (Nom) (b-a)	Error Size % (b-a)/a
2%	6.6%	4.6%	232%
3%	9.3%	6.3%	211%
5%	14.2%	9.2%	185%
9%	18.9%	9.9%	110%
16%	27.6%	11.6%	72%
19%	28.8%	9.8%	52%
62%	68.8%	6.8%	11%
67%	71.6%	4.6%	7%

Recall that task solutions (for example, in the case of a pie chart, the solution is slice size % relative to total pie) were varied and randomized to minimize the potential bias if using constant/single task solution. This approach eliminates the learning bias and aids in the generalizability of the results. The average error size % ranges from 7% (for tasks with 67% as the correct solution) to near or over 200% (for tasks with 2%, 3%, and 5% as correct solutions, which was to be expected).

In addition to the error size, users' performance data included time, fixation count, and fixation duration. Given the within-subject design, and since we wanted to specifically compare the user performance using a typical 2D pie chart to each alternative, a pairwise t-test comparing the respective means was conducted (t-test, paired two sample for means). Its findings are summarized in Table 3 (Accuracy and Time) and Table 4 (Fixation count and duration).

In case of accuracy, a significant difference was found in four pairs; between pie (M=4.643, SD=4.090) and horizontal bar (M=16.238, SD= 18.3356); t(20)=-3.1195, p=0.0054, pie and vertical bar (M=15.119, SD= 14.734); t(20)=-3.2431, p=0.0041, pie and lollipop (M=13.667, SD=14.2945); t(20)=-3.1051, p=0.0056, and pie and bubble (M=8.00 , SD=5.992); t(20)=-2.1871, p=0.0408.

TABLE 3: ACCURACY AND TIME

Chart Type	Accuracy (Nom. Error)		Time (Milliseconds)	
	Mean	SD	Mean	SD
Pie Chart	4.643	4.090	14192.3	8094.3
Horizontal Bar	16.238	18.3356	17801.1	6142.21
	t(20)=-3.1195, **p=0.0054**		t(20)=-1.9218, **p=0.069**	
Vertical Bar	15.119	14.734	12715.4	5967.0
	t(20)=-3.2341, **p=0.0041**		t(20)=0.8635, p=0.3981	
Lollipop	13.667	14.294	1773.1	11315.4
	t(20)=-3.1051, **p=0.0056**		t(20)=-1.3099, p=0.2051	
Bubble Chart	8.00	5.992	13761.2	8122.8
	t(20)=-2.1871 **p=0.0408**		t(20)=0.2192 p=0.8287	
Stacked Bar	4.667	4.3050	13117	5865.6
	t(20)=-0.0198, p=0.9844		t(20)=0.6330, p=0.5339	
Treemap	5.857	6.7474	14569.3	5551.97
	t(20)=-0.8172, p=0.4234		t(20)=-0.1998, p=0.8436	
Donut Chart	5.5238	4.366	13361.7	7252.17
	t(20)=-0.7941, p=0.4365		t(20)=0.5352, p=0.5979	

In the case of time, a significant difference was found in one pair; between pie (M=14192.3, SD=8094.33) and horizontal bar (M=17801.1, SD=6142.21); t(20)=(-1.9218), p=0.069[1].

In the case of fixation count, a significant difference was found in three pairs; between pie (M=37.6, SD=18); and horizontal bar (M=51.81, SD=17.64); t(20)=-2.7510, p=0.0123; pie and lollipop (M=51.8, SD=31.4); t(20)=-1.8604, p=0.0776; and pie and donut (M=31.381, SD=14.2658); t(20)=1.7324, p=0.0986. For the fixation duration, a significant difference was found in five pairs; between pie (M=286.921, SD=67.48) and horizontal bar (M=243.707, SD=41.907); t(20)=3.841, p=0.0010; pie and vertical bar (M=237.404, SD=49.686); t(20)=4.31, p=0.0003, pie and lollipop (M=234.853, SD=39.30); t(20)=4.293, p=0.0004, pie and treemap (M=246.744, SD=56.93); t(20)=2.894, p=0.01, and pie and donut chart (M=326.574 , SD=90.588); t(20)=-3.206, p=0.004.[2]

[1] To minimize the potential for Type 1 error results could be evaluated α level of 0.014 (=0.1/7 – seven tests). At this level, the difference in accuracy between a pie chart and bubble chart, and in time between a pie chart and vertical bar becomes insignificant.

[2] When evaluating the results at the α level of 0.014 (to control for Type 1 error), the findings remain unchanged for fixation duration, while for fixation count the only the difference between pie and horizontal chart is significant.

TABLE 4: FIXATION COUNT AND DURATION

Chart Type	**Fixation Count**		**Fixation Duration**	
	Mean	SD	Mean	SD
Pie Chart	37.6	18	286.921	67.48
Horizontal Bar	51.81	17.64	243.707	41.907
	t(20)=-2.7510, **p=0.0123**		t(20)=3.841, **p=0.0010**	
Vertical Bar	38.143	16.593	237.404	49.686
	t(20)=-0.1042, p=0.9181		t(20)=4.31, **p=0.0003**	
Lollipop	51.8	31.4	234.853	39.30
	t(20)=-1.8604, **p=0.0776**		t(20)=4.293, **p=0.0004**	
Bubble Chart	37.238	19.071	275.665	48.073
	t(20)=0.0738, p=0.942		t(20)=0.908, p=0.375	
Stacked Bar	36.905	16.577	266.184	60.208
	t(20)=0.1471, p=0.885		t(20)=1.484, p=0.1534	
Treemap	42.5	15.2	246.744	56.93
	t(20)=-0.9708, p=0.3432		t(20)=2.894, **p=0.01**	
Donut Chart	31.381	13.793	326.574	90.588
	t(20)=1.7324, **p=0.0986**		t(20)=-3.206, **p=0.004**	

V. DISCUSSION

Our goal was to understand the relative performance of selected alternatives compared to commonly used pie charts. Unlike existing research that primarily focused on the evaluation of accuracy and speed, we also focused on capturing users' visual attention and cognitive effort by measuring their eye fixation count and duration. In the process, we discovered several important findings.

First, regarding accuracy (estimation error), pie charts outperformed alternatives focusing on data encoding through length, including a horizontal bar, vertical bar, and lollipop. This confirms that data encoding through length is ineffective for part-to-whole tasks. A likely explanation is that bar charts provide a challenging way to estimate the total compared to area-based charts. Interestingly, the pie chart also outperformed the bubble chart (by almost 100%), a data representation that encodes data through the size of the area of the circle. Participants did not perform better when using other data representations that rely on data encoding through the area (stacked bar, donut chart, and treemap) than when using a traditional 2D pie chart. When it comes to time, the only significant difference measured when comparing the pie chart to a horizontal bar. Therefore, time may not be the best way to understand users' efforts or cognitive processes that lead to differences in accuracy.

Second, our findings on physiological indicators of attention and effort provide a more nuanced understanding of user impact. More specifically, although there was no difference in accuracy and time, our results indicate that the pie chart requires a longer fixation duration than the stacked bar and treemap. Hence, this is the first study to provide biometric evidence that pie charts require more effort to achieve the same level of accuracy and speed. Alternatively, we found that the donut chart requires more effort than all tested alternatives, including the pie chart. Our findings suggest that the recent increase in popularity of donut charts may come at the expense of users' cognitive effort and its use should be carefully considered in practice.

Third, our experiment suggests that any experiments using part-to-whole analysis must control the potential impact of part size. We found that small data parts (slices/areas/segments) generate more significant accuracy errors when compared to larger data parts. Although not surprising, we provide empirical evidence of the bias and caution others in generalizing any findings if not controlling for a potential effect of a data part. We successfully controlled for part size by randomizing the data part in addition to randomizing chart type and chart type order.

VI. LIMITATIONS AND FUTURE RESEARCH

Future research should consider other ways in which eye-tracking data can provide insights through gaze heatmaps (Figure 3), fixation sequences for viewing patterns, areas of interest for understanding revisits, and potential confusion and uncertainty. Similarly, saccadic movement-based metrics such as peak amplitude count, peak velocity, and pupillometry can provide a more nuanced understanding of user impact. One way to deploy this technology in future research is to focus on part-size user implications. In our exploratory post-hoc analysis of gaze data (heat map replays), we found initial indications of users potentially deploying different visual strategies when evaluating pie slices due to their size (small vs. large) or position (relative to easy to estimate 0, 45, 90, 135, and 180 angles). This potential research stream further amplifies the need for care in experiment design (controls) regarding data/part size.

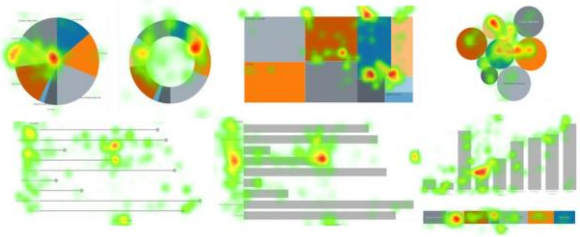


Figure 3: Gaze heat map

While our findings offer practical implications, they should be generalized only to the part-to-whole task type used in the present experiment. Future research should consider the use of pie charts in other task types. For example, tasks to assess the overall distribution of categories and understand the dominant/minority groups. Another research avenue may be analyzing business insights from pie charts and their alternatives.

Although a fully randomized and within-subject design allows for detecting significant relationships, future research should focus on replicating our findings with more participants. The present study, with only 21 participants, allows for Type II error (false negatives), therefore, the results should be interpreted cautiously. The participants were all undergraduate students from a single university, which may limit the generalizability of the results to other populations. Future research should consider expanding the population's size, age, occupation, and geographic distribution.

The experiment was conducted in a controlled lab setting, which may not accurately reflect how individuals use charts in real-world scenarios. While fixation duration and fixation count can provide insights into visual attention, they may not fully capture cognitive effort. Other measures such as subjective ratings or physiological (pupillometry, EEG, fMRI) measures could be used to supplement the fixation data.

Lastly, we used charts with eight categories. Future research should focus on understanding the potential impact of the number of categories (e.g., the number of pie slices) on user performance.

VII. CONCLUSION

In conclusion, the present study provides evidence that eye-tracking, particularly fixation duration, can add valuable insights regarding user impact when using pie charts and other data representations in solving part-to-whole tasks. First, although pie charts outperform length-based charts (vertical and horizontal bar charts and lollipop chart) when it comes to accuracy, this comes at the expense of more significant effort (fixation duration). On the other hand, while yielding the same accuracy, pie charts require more effort than alternative area-based charts (treemap and stacked bar). Second, in part-to-whole tasks, time is a poor predictor of accuracy and effort. Third, donut charts required the most cognitive effort among all tested data representations. Lastly, users tend to be biased toward overestimating the part size. However, the overestimation is not constant, suggesting a need for care when designing experiments focused on part-of-whole tasks.

REFERENCES

[1] E. R. Tufte, "The visual display of information." Cheshire, Ct.: Graphics Press, 1983.

[2] C. N. Knaflic, *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons, 2015.

[3] W. S. Cleveland and R. McGill, "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *J Am Stat Assoc*, vol. 79, no. 387, pp. 531–554, 1984.

[4] R. Kosara, "The Impact of Distribution and Chart Type on Part-to-Whole Comparisons.," in *EuroVis (Short Papers)*, 2019, pp. 7–11.

[5] V. T. Nguyen, K. Jung, and V. Gupta, "Examining data visualization pitfalls in scientific publications," *Vis Comput Ind Biomed Art*, vol. 4, no. 1, pp. 1–15, 2021.

[6] F. E. Sandnes, A. Flønes, W.-T. Kao, P. Harrington, and M. Issa, "Searching for extreme portions in distributions: A comparison of pie and bar charts," in *Cooperative Design, Visualization, and Engineering: 17th International Conference, CDVE 2020, Bangkok, Thailand, October 25–28, 2020, Proceedings 17*, 2020, pp. 342–351.

[7] H. Siirtola, "The cost of pie charts," in *2019 23rd International Conference Information Visualisation (IV)*, 2019, pp. 151–156.

[8] D. Skau and R. Kosara, "Arcs, angles, or areas: Individual data encodings in pie and donut charts," in *Computer Graphics Forum*, 2016, vol. 35, no. 3, pp. 121–130.

[9] W. Playfair, *The statistical breviary*. Wallis, 1801.

[10] I. Spence and S. Lewandowsky, "Displaying proportions and percentages," *Appl Cogn Psychol*, vol. 5, no. 1, pp. 61–77, 1991.

[11] S. Few, "Data visualization for human perception," *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 2014.

[12] D. Simkin and R. Hastie, "An information-processing analysis of graph perception," *J Am Stat Assoc*, vol. 82, no. 398, pp. 454–465, 1987.

[13] A. Krbanjevic and D. Bačić, "Exploring Pie Chart Visual Properties and Design Elements: Eye-Tracking Approach," in *AMCIS 2022 Proceedings*, Aug. 2002.

[14] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman, "The science of visual data communication: What works," *Psychological Science in the public interest*, vol. 22, no. 3, pp. 110–161, 2021.

[15] E. Cooper-Martin, "Measures of cognitive effort," *Mark Lett*, vol. 5, pp. 43–56, 1994.

[16] M. Shojaeizadeh, S. Djamasbi, R. C. Paffenroth, and A. C. Trapp, "Detecting task demand via an eye tracking machine learning system," *Decis Support Syst*, vol. 116, pp. 91–101, 2019.

[17] Dinko. Bačić and R. Henry, "Task-Representation fit's impact on cognitive effort in the context of decision timeliness and accuracy: A cognitive fit perspective," *AIS Transactions on Human-Computer Interaction*, vol. 10, no. 3, pp. 164–187, 2018.

[18] M. K. Eckstein, B. Guerra-Carrillo, A. T. M. Singley, and S. A. Bunge, "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?," *Dev Cogn Neurosci*, vol. 25, pp. 69–91, 2017.

[19] K. Kurzhals, B. Fisher, M. Burch, and D. Weiskopf, "Eye tracking evaluation of visual analytics," *Inf Vis*, vol. 15, no. 4, pp. 340–358, 2016.

[20] D. Bačić, "Biometrics and business information visualization: Research review, agenda and opportunities," in *HCI in Business, Government, and Organizations: 5th International Conference, HCIBGO 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings 5*, 2018, pp. 671–686.

[21] D. Bačić and R. Henry, "Advancing our understanding and assessment of cognitive effort in the cognitive fit theory and data visualization context: Eye tracking-based approach," *Decis Support Syst*, vol. 163, p. 113862, 2022.

[22] P. Bera, "How colors in business dashboards affect users' decision making," *Commun ACM*, vol. 59, no. 4, pp. 50–57, 2016.

[23] D. Peebles and P. C.-H. Cheng, "Modeling the effect of task and graphical representation on response latency in a graph reading task," *Hum Factors*, vol. 45, no. 1, pp. 28–46, 2003.

Parental Trust in Automated Detection of Cyberpredators

L. Brandimarte
Eller College of Management, University of Arizona, Tucson AZ, USA
lbrandimarte@arizona.edu

***Abstract* - Online cyberpredators are a serious threat against children who are increasingly using social media and messaging systems to interact with strangers. At the same time, monitoring children's online activity is challenging for parents, due to the numerous platforms and ways a child can access them. Automated approaches could help detect dangerous conversations, but their adoption may prove difficult due to algorithm aversion – the tendency of humans to place less trust in recommendations by automated systems as compared to the judgment of other humans. Three online experiments investigate whether and under what conditions parents are willing to adopt automated systems for detection of cyberpredators, and when they are willing to trade potentially sensitive information about their children's online interaction, as well as individual control and agency. Factors tested for an effect on parental trust include accuracy of predictions by humans versus the algorithm, and storage and management of data in the cloud or local client. Implications for researchers and developers of automated systems of cybercrime detection are presented and discussed.**

Keywords – algorithmic aversion; cyberpredators; experiments.

I. INTRODUCTION

Online predators (cyberpredators, or criminals that target minors in the online world) are a serious threat against children who increasingly use social networking and messaging services to interact with strangers. For instance, one study found that one in nine teens received unwanted online solicitations [1]. Parents are advised to monitor their children's use of social media, but this is extremely difficult in practice, given the variety of networking services and access methods that a child can choose from.

Automated approaches which employ, for instance, natural language processing methods could be of tremendous benefit, but the adoption of such approaches might be hindered by a general lack of faith in algorithmic decision making. On the one hand, an automation system would relieve the user of a tedious task, analyzing a child's online communications and detecting cyberpredation – a task of critical importance. On the other hand, adoption of an automation tool to substitute the work traditionally performed manually by parents requires significant trust from the users.

Computer automation is well used and trusted for many applications. A user would never question a computer's ability in a domain such as arithmetic or spell checking, in which functionality is precisely defined and correctness is clear. Instead, trust can be a problem in the use of systems whose correctness criteria are more complicated, such as adaptive agents [2], recommender systems [3], and autonomous driving [4]. In these domains, there are many possible solutions to a problem and some form of *judgement* must be used to choose the best solution. For example, a self-driving car may need to choose how best to avoid a slow pedestrian by either speeding up to clear the intersection before the pedestrian enters, or by slowing down and wait for the pedestrian to cross. Both solutions represent a trade-off in travel time versus safety, and judgement is required to make the decision.

The tendency of humans to place less trust in the judgement of automated systems is often referred to as *algorithm aversion* [5, 6]. In addition to a generalized distrust of automated decision-making, several other factors may impact the ability of a user to trust a tool implementing a cyberpredator detection approach. For instance, algorithmic bias against minority groups, based on identifiable speech patterns, could result in false negatives, which would leave children vulnerable, as well as false positives, which would lead to unfair accusations against innocent individuals. Therefore, perceived accuracy of the algorithm is crucial. Data privacy is also critical, because a cyberpredator detection approach requires access to private, potentially sensitive communications, and a parent must rely on the system to keep those communications in confidence.

In this paper, we explore the effect of algorithm aversion on parents' willingness to adopt an automated detection tool for cyberpredators, the role played by accuracy of the tool, as well as some of the mechanisms which, according to existing work, may lie behind algorithm aversion. We run three experimental studies and gather evidence that algorithm aversion may not be as widespread as existing literature suggests, and that, instead, it may be a function of the study design and the specific context of interest.

A. Related work

Much recent literature has focused on how people perceive algorithmic decision making, and the factors influencing the propensity of people to follow or reject suggestions coming from an algorithm. There seems to be much heterogeneity in people's attitudes towards such suggestions, depending on the specific context under consideration. In journalism, there seems to be a bias against humans. An experiment in South Korea [7] showed that, when an article's author is described as a

human, both the general public and journalists rate it of lower quality as compared to the same exact article but described as authored by an algorithm. One wonders if this result would replicate in other countries and contexts, or if, on the other hand, it is specific to a particular culture, historical moment, or specific context.

In another experiment using forecasting tasks (predicting MBA student performance based on admission data; [5]), people showed a bias against algorithms. The mechanism behind this bias was that people observing a human and an algorithm making wrong predictions tend to lose faith in the algorithm faster than losing faith in the human. The same set of authors found a way to counter such algorithm aversion in a follow-up paper [8], in which they show that, when given some level of control (to modify the algorithm's prediction, thus simply using the algorithm as a tool or a starting point), people are more willing to accept the recommendation by the algorithm.

Another set of experiments [9] comparing medical suggestions provided by human or artificial intelligence also show a bias against the algorithm. This bias appears to be caused by the perception that algorithms are not able to take into account the uniqueness of an individual. The bias therefore disappears if the algorithmically generated suggestion is described as personalized or directed to someone other than the self.

In a series of three experiments, we tested whether automated recognition of cyberpredators is trusted by adults. There are at least two levels of trust that need to be taken into consideration. First of all, we consider **Trust in Correctness**, trust intended as belief that the system's prediction is accurate or, at least, more likely to be accurate than the adult's own prediction. Second, we consider **Trust in Privacy**, trust towards the system which needs to be given access to potentially sensitive data in order to perform the assessment. Are people willing to trade sensitive information as well as individual control and agency (for the detection of online dangers) when it comes to identifying cyberpredators? And under what circumstances?

B. Trustworthiness of Algorithmic Detection

What does it mean for a system to be trusted? In information security, trust by people is not necessarily a quality that one wants to achieve in a system. Many distributed systems, such as blockchain for instance, make of the lack of required trust by users toward a central authority a major strength. Similarly, one of the main advantages of end-to-end encryption in communication is that users do not need to trust the service provider in order to feel secure and protected when exchanging messages or materials. Therefore, a good rule of thumb in information security seems to be that one should not want to trust a system [10] because trust is not necessary – the system "just works."

However, not all relationships can be fully decentralized, and whenever there is a service provider that acts as a centralized authority in the relationship, there always exists a form of vulnerability [11] that makes trust a necessary condition to the acceptance and to the lasting of that relationship. The question then becomes: What makes average users (not necessarily experts) trust that a centralized system "just works?" What convinces them to believe in it and ultimately use it?

Different disciplines have provided a variety of definitions of trust. In machine learning, a model is trustworthy if it is secure and resilient to attacks by adversaries [12]. In human-computer interaction, researchers developed a definition of online trust as an evolution of its offline counterpart. In the real world, "trust is the social capital that can create cooperation and coordination." In the cyber-world, trust becomes "an attitude of confident expectation in an online situation of risk that one's vulnerabilities will not be exploited" [13]. Economists define trust in a similar fashion as "a willingness to bet that another person will reciprocate a risky move (at a cost to themselves)" [14], and often measure it via the trust game [15]. This is a game in which one player is given a sum of money and a choice to either keep it to him/herself or to pass (part of) it to the other player. The rules of the game are that as soon as the money is passed to the other player, it triples, and the other player can either keep the tripled sum all to him/herself or reward the trust of the first player and split it. A similar definition is used in the management information systems literature, where "Emotional trust is the extent to which one feels secure and psychologically comfortable about depending on the trustee" [16]. Behavioral decision scientists have looked specifically at trust towards algorithms/computers in the medical context, and have defined it simply as "trust in the quality of the recommendation" [17]. Similarly, in marketing, trust "is the perceived competence (i.e. credibility) and benevolence of a target" [18], a definition that some have also used in reference to relationships with robots [19].

Our interpretation of trust encompasses all these definitions and, ultimately, refers to the preconditions for a cyberpredator detection algorithm to be adopted by parents. As briefly noted above, we consider two such preconditions: trust in correctness and trust in privacy. Notice that these are not necessarily two disjunct concepts. Recent work on willingness to download contact tracing apps in order to reduce the spread of SARS-CoV-2 shows that people are willing to trade-off some of their privacy if they perceive the technology to be accurate [20].

Trust in correctness is operationalized in our studies as the described accuracy of the algorithm (experimentally manipulated in Study 2) and as perceived efficacy of the machine used to perform the necessary computations. Trust in privacy is studied by manipulating the location in which the computation occurs (either on one's machine, requiring no data transfer, or on the cloud, thus requiring data transfer via the internet; see Study 3) and through a series of measures that cover various kinds of concerns that existing literature has identified as important in privacy choices. Detailed description of the studies is reported in the following sections. Overall, the three experimental studies investigated the prevalence of algorithm aversion among parents in the context of automated detection of cyberpredators (Study 1); the role

of relative accuracy of algorithmic versus human detection of cyberpredators (Study 2); the role of privacy concerns associated with control and relinquishing agency in decisions concerning one's children (all studies), and with sensitive data transfers (Study 3). Results indicate that algorithm aversion may not be as prevalent as earlier work suggested, but perhaps a result of elicitation methods used in experimental procedures or of specific contexts.

II. STUDY 1

A. Methods

Study 1 was a basic test of algorithmic aversion. Participants (all parents recruited via the Amazon Mechanical Turk platform for a 10-minute survey and paid $0.5) were told to imagine they had a chance to have their child's online conversations over the last 4 weeks assessed for free to ensure such conversations did not involve an online predator. Participants were reassured that conversations were anonymized and that the child's identity remained confidential. The study only manipulated (between-subjects) the entity performing the assessment: either an algorithm (algo condition) or a criminologist (human condition) trained to detect predatory conversations. Participants were told that the algorithm/criminologist had historically identified cyberpredators correctly 82-85% of the times, so accuracy was the same across conditions. Participants were then asked to indicate on a scale from 1 (Definitely distrust) to 7 (Definitely trust) how much they trusted or distrusted the algorithm/criminologist to make an accurate assessment. A preference for the criminologist would provide evidence of algorithmic aversion.

Participants were then asked a series of questions which were the same across all three studies. The first one measured interest in having the child's online conversations over the last 4 weeks assessed, and it was followed by an open-ended request to explain the reasons for the choice. Then, in order to measure whether algorithmic aversion was a function of personal involvement (it is possible that in general people trust algorithms, but when it comes to the safety of their own children their emotions and preference for a sense of personal agency dominate their choices), a question asked whose judgement *other parents* should trust more when it comes to detecting predatory conversations – a Human Criminologist or an Algorithm. Answers were on a scale from 1 (Definitely criminologist) to 7 (Definitely algorithm), so a lower value indicated an expectation of algorithmic aversion by other parents.

The next set of questions measured frequency of use of technology by the child, expected level of safety while the child is online, participants' familiarity with privacy/ security concepts and technology in general, and participants' use of Privacy Enhancing Technologies (PETs).

To test for algorithmic aversion in a context different from cyberpredators, participants were then told that artificial intelligence technologies are used in many different fields, including healthcare. For instance, they are being used to improve diagnoses for certain types of cancer. They were then asked to consider a scenario in which a family member – including their child – needed a potential cancer diagnosis, and were asked how likely they would be, on a scale from 1 (Extremely unlikely) to 4 (Extremely likely), to seek out a doctor that uses artificial intelligence for cancer diagnoses.

To better explore the mechanism behind algorithmic aversion, the next series of questions asked, on a scale from 1 (Strongly agree) to 4 (Strongly disagree), whether participants agreed that: AI can perform decision-making just as well as humans; Humans are more reliable for makings decisions rather than AI; A human doctor knows them better than any machine/computer ever could; they expected humans to make mistakes, but machines/ computers should be infallible; they were more willing to accept recommendations made by AI than humans so long as they can modify it, even in the slightest way; and Humans more quickly learn from their mistakes and adapt their beliefs faster than machines.

The study ended with a set of standard demographic questions (age, race, gender, country of residence, occupation, education, salary range).

B. Results

We recruited 582 participants. After screening out those who did not complete the study or failed the attention check question (they answered affirmatively to the question of whether they were born before the year 1920), the usable sample consisted of 423 participants (59% male, M_{age} = 34.84, SD_{age} = 9.11). They showed marginal algorithmic aversion, as the average level of trust towards the criminologist (M_{human} = 5.27, SD_{human} = 1.08) was higher than trust towards the algorithm (M_{algo} = 5.13, SD_{algo} = 1.15, 1-tailed t-test $p < 0.1$). Interestingly, algorithmic aversion disappeared when participants imagined other parents were facing a similar decision: the average answer was right in the middle of the scale (M = 4.38, indicating indifference between algorithm and criminologist) and there was no difference across conditions (t-test p-value = 0.49). There was also no evidence of algorithmic aversion in the healthcare scenario, where the median response (and the mode) was 3 (Somewhat likely to seek out a doctor who uses AI for cancer diagnosis) and there was no difference across conditions.

The majority (73%) of participants expressed interest in having their child's conversations assessed, and there was no difference across conditions (z-test $p > 0.1$). Questions on potential mechanisms were not discriminant, as all distributions showed a vast majority of participants agreeing (either somewhat or strongly) with each statement.

III. STUDY 2

A. Methods

Study 2 tested whether algorithmic aversion disappears when the relative accuracy of algorithms versus humans is manipulated. Identical to Study 1, participants (again, all parents recruited via the Amazon

Experiments were funded by the MIS Department at the Eller College of Management, University of Arizona, through their Small Research Grant Program.

Mechanical Turk platform for a 10-minute survey and paid $0.5) were told to imagine they had a chance to have their child's online conversations over the last 4 weeks assessed for free to ensure such conversations did not involve an online predator. In this Study, however, they were given two options – either an algorithm or a criminologist – and the accuracy of the two options was manipulated between subjects. In the equal condition, accuracy was identical (90%); in the algorithm condition, the algorithm (option 2) was more accurate (95% vs. 86%), and vice versa in the human condition (86% vs. 95%). Participants were then asked to indicate on a scale from 1 (Definitely Option 1, Human) to 7 (Definitely Option 2, Algorithm) which option they would trust the most to analyze and provide an expert opinion of their child's conversations. Low values of this measure, especially in the algorithm condition, would constitute evidence of algorithmic aversion. The remainder of the study was identical to Study 1.

B. Results

We recruited 292 participants. After screening out those who failed the attention check question (identical to Study 1), the usable sample consisted of 211 participants (58% male, M_{age} = 36.15, SD_{age} = 9.30). An analysis of the measure of choice between the algorithm and the human showed that participants valued accuracy and algorithmic aversion was not observed. A one-way ANOVA showed a significant difference across conditions ($F(2, 208) = 21.02$, $p < 0.01$), with all pairwise comparisons resulting significant after Bonferroni adjustment (M_{equal} = 4.37, SD_{equal} = 1.88, significantly lower than M_{algo} = 5.53, SD_{algo} = 1.40, and significantly higher than M_{human} = 3.51, SD_{algo} = 2.17; all p-values < 0.05).

The majority (68%) of participants expressed interest in having their child's conversations assessed, and there was no difference across conditions (43 out of 70 in the equal condition, 50 out of 68 in the algorithm condition, and 51 out of 73 in the human condition; a chi-square test reports a p-value of 0.29).

When participants imagined other parents were to decide whether they should trust the algorithm or a human more, the average answer was right in the middle of the scale (M = 4.31, indicating indifference between algorithm and criminologist). Differently from Study 1, and consistent with the lack of algorithm aversion, the accuracy manipulation significantly affected this variable: a one-way ANOVA shows significant differences across conditions ($F(2, 208) = 7.84$, $p < 0.001$), and Bonferroni-adjusted pairwise comparisons show a significant difference between the algorithm (M_{algo} = 4.91, SD_{algo} = 1.63) and the human condition (M_{human} = 3.70, SD_{algo} = 2.05, $p < 0.05$). There was also no evidence of algorithmic aversion in the healthcare scenario, where, just like in Study 1, the median response (and the mode) was 3 (Somewhat likely to seek out a doctor who uses AI for cancer diagnosis) and there was no difference across conditions. In particular, one should notice the null effect of the accuracy manipulation in this context: even participants exposed to an algorithm that is less accurate than a human showed no aversion to it assisting the human in the diagnosis.

IV. STUDY 3

A. Methods

Study 3 was an exploratory study which manipulated the location of the automated analysis of children's conversations within subjects. Specifically, participants were asked to imagine they were using an application that assesses their children's online conversations over the last 4 weeks for free, to ensure such conversations did not involve a cyberpredator. Similarly to the previous studies, participants were reassured that the conversations would be anonymized and that the children's identity would remain confidential. They were also informed that storage and analysis could be local on their personal device, with no transmission of data, or transferred to the cloud service. Participants were then asked to choose their preferred option to store and analyze the conversations (Local /Cloud service/ Indifferent).

To explore the reasons behind the choice, they were asked, on a scale from 1 (Strongly disagree) to 4 (Strongly agree), whether they agreed or disagreed with several statements. The first two measured concerns towards cloud-based technologies (Cloud-concern) and asked whether participants: believed data are secure when transferred to the cloud service over the internet (reverse-coded); believed data stored in the cloud are secure (reverse-coded). The third statement asked whether participants were concerned about data privacy when it comes to cloud storage (Privacy-concern). The fourth one asked whether they believed they had more control of information stored and analyzed locally on their personal devices (Control). The last three statements measured perceived efficacy of (or, reverse-coding all responses, concern towards) their own local technology (Local-concern) and asked whether participants believed their computer/device was reliable; it could adequately handle the tasks they need it to perform; it had sufficient storage. The rest of the study was identical to the previous two.

TABLE I. STUDY 3 REGRESSION RESULTS

	Coefficient	***SE***
Cloud_concern	0.94***	0.24
Privacy_concern	0.04	0.20
Control	0.28	0.24
Local-concern	-1.10***	0.32
N = 224, Pseudo R-squared = 0.11		

*** p < 0.01

B. Results

We recruited 292 participants in the same way as the previous studies and, after eliminating those who failed the attention check, the usable sample included 224 parents. The majority (58%) chose to use the algorithm locally, with 34% choosing the cloud service and the remaining being indifferent between the two options.

To investigate which reasons mattered most in the choice, we estimated a logit model using the choice of the cloud service as the binary dependent variable (equal to 1 for Cloud and 0 otherwise) and the four reasons described in the previous section as predictors. The first reason was Cloud-concern and it was built averaging data from first two statements (answers were strongly correlated, with a Chronbach's alpha of 0.87). The next two reasons were Privacy-concern and Control, and the last one was Local-concern, built by averaging data from the last three statements (answers were strongly correlated, with a Chronbach's alpha of 0.74). Perhaps unsurprisingly, results (reported in Table I) indicated that Cloud- and Local-concern were significant predictors of the choice of the cloud service over the local client. However, Privacy-concern and Control were not significantly correlated with the choice.

The majority (76%) of participants expressed interest in having their child's conversations assessed and were neutral regarding the choice by other parents to choose an automated or a human evaluation (M = 4.16, SD = 1.63). There was also no evidence of algorithmic aversion in the healthcare scenario, where, similar to the previous studies, the median response (and the mode) was 3 (Somewhat likely to seek out a doctor who uses AI for cancer diagnosis).

V. Discussion

Three experimental studies investigated the propensity of parents to use an automated system of detection of cyberpredators. Study 1 found a marginally significant level of algorithm aversion: by manipulating the availability of either a human criminologist or an automated system between subjects, parents were more trusting of the human even though the accuracy of the two methods was identical. This form of algorithm aversion may be due to parents not being willing to give up agency over decisions regarding their children. In fact, when asked about what other parents should choose, participants showed no sign of algorithm aversion. Furthermore, algorithm aversion was not detected in the context of a medical doctor using an AI to diagnose a disease, which provides further evidence for the mechanism of agency: when participants think they are leaving the decision to an algorithm, they tend to distrust it, but when the algorithm is only described as supporting the human in their decision, trust is restored. From a societal perspective, this is reassuring, as AI researchers and practitioners have been converging on the idea that the highest contribution AI can provide in the economy is to support rather than replace humans [21].

Study 2 manipulated relative accuracy of the algorithm versus the human between subjects, and asked participants to choose one or the other. When given the option to choose (rather than being assigned to one or the other) and when provided with salient information about accuracy, again, no algorithmic aversion is observed: parents rationally choose the more accurate option.

The empirical evidence of these first two studies suggests that the results found in existing literature, indicating widespread algorithm aversion, may be somewhat inflated by the specific experimental design used to investigate the phenomenon. Specifically, between- versus within-subject manipulations and salience of accuracy information can be important factors affecting the results.

Study 3 was an exploratory study aiming at understanding whether factors that existing literature indicates to be relevant in privacy decision making also affect the propensity to use a potentially privacy-invasive algorithm. Consistent with previous work (e.g., [22]), we find that perception of data security in the cloud is positively correlated with choosing to use an automated cyberpredator detector in the cloud, and efficacy of one's local machine is negatively correlated with it (intuitively: the better one's machine performs, the higher the propensity to choose it over a cloud-based service). However, contrary to our expectations (e.g., [23]), specific data privacy concerns and control concerns were not significantly correlated with the choice of the cloud-based service. The study only allows for speculation as to why that might be the case, but it is possible that those concerns matter at the extensive margin (whether to use the service or not, which was not an option in this study), but less so at the intensive margin (a choice over the location where the computation is performed, locally or in the cloud).

VI. Limitations and Future Directions

While several studies have shown that, with proper checks, the quality of data collected via Amazon Mechanical Turk and the representativeness of the sample as compared to national surveys increases (e.g., [24]), our studies may not replicate in different populations or different contexts.

Furthermore, we only explored a few of the mechanisms that may be responsible for algorithm aversion. For instance, future work may investigate whether attitudes toward algorithms change as a function of the person's mental model and understanding of what an algorithm actually is. Reference [25] provides a good overview of experimental evidence on trust and other attitudinal measures towards algorithms. The claim of the paper is that trust increases if some, but not too much information is provided regarding how the algorithm works. However, the dependent variables used do not actually measure trust towards the algorithm, but rather, they measure trust towards peers grading an essay in the context of an online class. In another set of experiments in the context of medical decision making [26], researchers have shown that algorithm aversion is, at least in part, caused by the fact that algorithms are perceived as inscrutable black boxes, more so than recommendations made by human doctors. This better understanding of human decisions is, however, an illusion, and once patients are exposed to more information about how algorithms reached a certain recommendation, aversion towards them decreases. In the context of predictions and recommendations, research has shown that people's understanding of recommendation systems is quite limited [27]. Furthermore, while algorithms are better than humans in various contexts, such as predicting which jokes people will consider funny, people are still averse to

trust the algorithm because they do not understand the automated process [28]. More generally, the evidence on the effect of transparency on trust and other related constructs is mixed, and in need of synthesis and critical review.

VII. Conclusion

Protecting children from dangerous online conversations with cyberpredators is fundamental to their safety but difficult for parents to do effectively, since children may use a wide variety of communication channels and it may be impossible for parents to constantly monitor them. The use of automated tools for detection of cyberpredators represents an effective alternative but it is important to evaluate whether parents would be willing to adopt them. The fact that such tools may be perceived by parents as taking agency away from them when it comes to protecting the safety of their children, and that they may be accessing potentially sensitive information could constitute an impediment to adoption. More generally, parents may not trust an algorithm to decide which situations may be dangerous for their children, and they may show algorithm aversion. In three experimental studies, we find that, while some degree of algorithm aversion may be detected, it is sometimes a consequence of the specific elicitation methods used, and when it is indeed present, it can be addressed by providing salient information regarding the accuracy of the algorithm and its role in support rather than in substitution of human judgment.

Acknowledgment

We thank former students at the University of Arizona's Artificial Intelligence Lab for help in the application for IRB approval and in the setup of the online studies. We also thank the Eller College of Management for funding support.

References

[1] S. Madigan, V. Villani, C. Azzopardi, D. Laut, T. Smith, J. R. Temple, D. Browne, and G. Dimitropoulos. The prevalence of unwanted online sexual exposure and solicitation among youth: A meta-analysis. *Journal of Adolescent Health*, 63(2):133 – 141, 2018.

[2] A. Glass, D. L. McGuinness, and M. Wolverton. Toward establishing trust in adaptive agents. In *Proceedings of the 13th International Conference on Intelligent User Interfaces*, 2008.

[3] P. Pu and L. Chen. Trust building with explanation interfaces. In *Proceedings of the 11th International Conference on Intelligent User Interfaces*, 2006.

[4] T. Rezvani, K. Driggs-Campbell, D. Sadigh, S. S. Sastry, S. A. Seshia, and R. Bajcsy. Towards trust- worthy automation: User interfaces that convey internal and external awareness. In *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 682–688, 2016.

[5] B. J. Dietvorst, J. P. Simmons, and C. Massey. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114, 2015.

[6] A. Prahl and L. Van Swol. Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, 36(6):691–702, 2017.

[7] J. Jung, H. Song, Y. Kim, H. Im, and S. Oh. Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. *Computers in Human Behavior*, 71:291–298, 2017.

[8] B. J. Dietvorst, J. P. Simmons, and C. Massey. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3):1155–1170, 2018.

[9] C. Longoni, A. Bonezzi, and C. K. Morewedge. Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4):629–650, 2019.

[10] Trusted third parties are security holes. Available at https://nakamotoinstitute.org/trusted-third-parties/. Last accessed on February 18, 2023.

[11] Y. D. Wang and H. H. Emurian. An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, 21(1):105–125, 2005

[12] N. Papernot. What does it mean for machine learning to be trustworthy? San Francisco, CA, Jan. 2020. USENIX Association.

[13] C. L. Corritore, B. Kracher, and S. Wiedenbeck. On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58(6):737–758, 2003.

[14] C. F. Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2011.

[15] J. Berg, J. Dickhaut, and K. McCabe. Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142, 1995.

[16] S. Y. Komiak and I. Benbasat. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, pp. 941–960, 2006.

[17] M. Promberger and J. Baron. Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5):455–468, 2006.

[18] P. M. Doney and J. P. Cannon. An examination of the nature of trust in buyer–seller relationships. *Journal of Marketing*, 61(2):35–51, 1997

[19] J. Wirtz, P. G. Patterson, W. H. Kunz, T. Gruber, V. N. Lu, S. Paluch, and A. Martins. Brave new world: Service robots in the frontline. *Journal of Service Management*, *29*(5): 907-931, 2018.

[20] G. Kaptchuk, E. Hargittai, and E. M. Redmiles. How good is good enough for covid19 apps? The influence of benefits, accuracy, and privacy on willingness to adopt. *arXiv preprint arXiv:2005.04343*, 2020.

[21] A. McAfee and E. Brynjolfsson. *Machine, platform, crowd: Harnessing our digital future*. WW Norton & Company, 2017.

[22] H. J. Smith, T. Dinev T, and H. Xu. Information privacy research: An interdisciplinary review. *MIS Quarterly*, 1:989-1015, 2011.

[23] L. Brandimarte, A. Acquisti, and G. Loewenstein. Misplaced confidences: Privacy and the control paradox. *Social psychological and personality science*, 4(3):340-7, 2013.

[24] N. Qureshi, M. Edelen, L. Hilton, A. Rodriguez, R. D. Hays, and P. M. Herman. Comparing data collected on Amazon's Mechanical Turk to national surveys. *American Journal of Health Behavior*, 46(5):497-502, 2022.

[25] R. F. Kizilcec. How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2390–2395, 2016.

[26] R. Cadario, C. Longoni, and C. K. Morewedge. Understanding, explaining, and utilizing medical artificial intelligence. *Nature human behaviour*, 5(12):1636-42, 2021.

[27] O. Alvarado, H. Heuer, V. V. Abeele, A. Breiter, and K. Verbert. Middle-aged video consumers' beliefs about algorithmic recommendations on youtube. *arXiv preprint arXiv:2008.03202*, 2020.

[28] M. Yeomans, A. Shah, S. Mullainathan, and J. Kleinberg. Making sense of recommendations. *Journal of Behavioral Decision Making*, 32(4):403–414, 2019.

Mobile Technologies and Live Streaming Commerce: A Systematic Review and Lexical Analysis

Curt A. Gilstrap, Ph.D. and Cristina M. Gilstrap, Ph.D.
University of Southern Indiana
cgilstrap1@usi.edu

***Abstract* - The proliferation of Live Streaming Commerce (LSC) across modern economies has lead organizations and individuals to adjust brands, marketing, and product/service offerings for exposure to live platforms. Afforded by strong mobile adoption trends, increasing shop-from-home behavior, and ever-increasing social media usage, LSC finds itself at the intersection of mobile and social within the expanding e-commerce field. Starting from 412 scholarly products and filtering for language, peer review, quality of indexing, and citation counts, this study provides a systematic review and lexical analysis of 93 academic peer-reviewed articles from the years 2018 through 2023 to explore the dimensions, contexts, and consumptive/collective behaviors of LSC users pertaining to mobile technologies. In particular, the nascent development of live streaming consumer behavior research and mobile technologies adoption are further compared to m-commerce and non-mobile streams of research. Findings reveal eight dialectics that define LSC research, diversity in theoretical approaches to LSC research, and ways to further explore the mobile technologies ubiquitous but underexamined within LSC literature.**

Keywords - Live Streaming Commerce, m-commerce, Leximancer, s-commerce, mobile technologies, systematic review

I. INTRODUCTION

The Live Streaming Commerce (LSC) market will be valued at nearly $200 billion globally by 2027 and the phenomenon already impacts millions of viewer-consumers per region in North America, Asia, and Europe [1]. LSC is electronic commerce that occurs when sellers synchronously broadcast on social platforms to sell products or services and interact with viewers while simultaneously providing benefits such as entertainment, education, brand identification, or additional social connection [2]–[5]. It is used widely on platforms such as Douyin and Taobao Live in Asia and YouTube and Twitch in North America. LSC also has been typified as live streaming embedded in sales platforms, live streaming platforms incorporating sales features, or social platforms offering live streaming features where users engage in selling live [6], [7]. Research examining LSC has focused on the phenomenon's growth, human interaction and purchase decision-making, marketing behavior, or brand advocacy; however, the current body of LSC research remains thin in comparison to the scale of global commercial activity [8]. Given the uniqueness of the format's user-synchronicity and accessibility as well as the capacity for myriad platform-apps to leverage mobile technologies in commercial-live pursuits, this study seeks to examine how mobile technologies are integrated into LSC research.

A. Research Background

Social commerce (s-commerce) is a subset of electronic commerce [9] where social platforms offer buying opportunities for consumers [10], [11]. Born from the advent of online commerce moving onto Web 2.0 social platforms [12], s-commerce affords user-generated content posting and relationship building between sellers and buyers within this new sales environment [10]. A subset of s-commerce, LSC leverages mobile apps and hardware, social platforms, and synchronous human interactivity through live video to extend commercial exchanges from websites and app-based sales into users' consumptive activities while on their devices [7], [13]. Further, the sudden growth in the LSC phenomenon matched with the early growth of LSC research means there is a need to examine what precisely LSC is relative to the field of s-commerce studies. Additionally, an observed paucity of named studies relative to mobile technology within the LSC literature should be investigated given the required technologies for LSC to expand across user bases globally and given LSC literature may assume mobile technologies to be homogenous in LSC experiences. As it stands, there is likewise a growing field of study in Mobile Commerce (m-commerce) [14], [15], yet almost no investigation has been conducted to examine the default assumptions in LSC literature regarding mobile technologies [16]. This study seeks to respond to these gaps.

B. Keyword and Automatic Content Analysis

A keyword search in Google Scholar reveals several hundred academic, peer-reviewed articles on LSC. However, very few incorporate mobile technology, smartphone platforms, or mobility generally in their article titles. Additionally, examining the keywords of the articles in this dataset further indicate mobile technology is not as central to the LSC conceptual space or research writ large as purchase behaviors, social identity concerns, marketing/branding strategy, or platform specificity. Keyword analysis is important for investigating research in a particular literature stream [10], although keywords

tend to be selected by authors or editors rather than generated by the abundance of terms naturally found in the body of research writing. For this study, the dataset was found using keywords and key phrases on Google Scholar. However, analysis also included a deeper lexical analysis [17] of the final dataset to fully assess and examine the LSC phenomenon in addition to author-selected keywords.

II. RESEARCH METHODOLOGY

A. Overview

In this study, a systematic review fully investigated current LSC literature relative to all emergent trends, methodologies, contexts, and implications. It examined the defined research questions by capturing and summarizing all empirical evidence that fit the articulated eligibility requirements. Additionally, a lexical map was produced of the data (e.g., abstracts, introductions) utilizing Leximancer [18] software for thematic counts and cluster analysis to further support the review findings.

A systematic review is conducted when a field of study produces enough research needed to stabilize a scholarly community's understanding of a phenomenon, and provide assurance as to what the phenomenon includes [19]. This knowledge stability alleviates uncertainty regarding what a field of study focuses on and regards as settled science. Modeling Esmaeili & Hashemi's (2019) systematic review of s-commerce, this review followed the previous study's process to analyze the most important and highest quality LSC research (see also, [20]). First, review questions were developed. Second, a review protocol was generated. Third, study selection criteria and processes were articulated. Fourth, the qualities of studies were assessed. Fifth, lexical and human analyses were applied, and summarized results of data extraction and synthesis were presented and interpreted.

B. Review Questions

The following research questions emerged from a deep reading of the literature, and a review of previous systematic reviews of social and mobile commerce [6], [21]:

RQ1. Over the LSC literature, what is the accepted definition of LSC?
RQ2. How is LSC unique from other S-Commerce types?`
RQ3. What are the main theories in LSC research literature?
RQ4. What are the research themes in LSC research?
RQ5. What are the methodologies & data collection approaches in LSC research?
RQ6. What are the limitations of LSC research?
RQ7. How is mobile technology located within the LSC research literature?

C. Review Protocol

Fig. 1 identifies the research steps conducted in this study. Using a modification of PRISMA approach [22], this study's review protocol followed a find/filter/check approach to capturing and cleaning the dataset of LSC articles for analysis. First, a Python based program crawled and captured all Google Scholar entries under the search term "live streaming commerce" in January of 2023 to build the data set of LSC research articles. The capture was saved into an Excel sheet including authors, titles, abstracts, and citation calculations (number of citations). Second, within the excel sheet, all candidate articles (by row) were organized by citation count, and article manuscripts were retrieved as PDF files from databases and saved to a cloud folder. Third, a first filter was applied with an "exclusion scope" [21] to remove all candidate rows not clearly peer-reviewed journal (PRJ) articles. A second filter's exclusion scope removed all non-English PRJs. A third filter's exclusion scope then removed all PRJs without a citation including those not currently cited by other journals, according to Google Scholar. A fourth filter's exclusion scope re-checked that titles included "live streaming" or "live" & "commerce" within them. A check determined the resulting PRJ data set was composed of those articles indexed by academic databases (e.g., Web of Science, Scopus, SCImago, EBSCO, Pubmed) and a final check removed any duplicates. The final dataset included 93 PRJs for review and analysis.

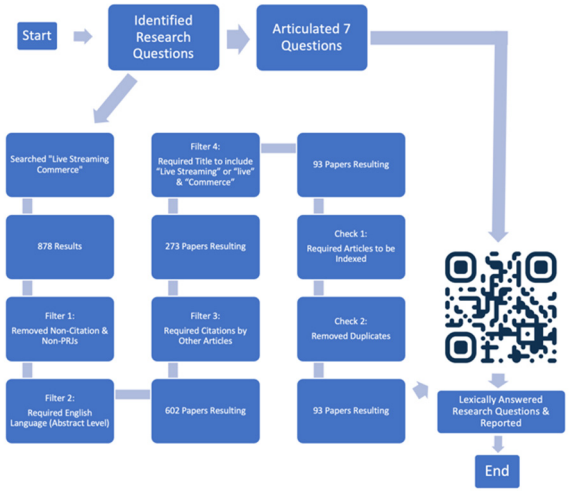


FIGURE 1. REVIEW PROTOCOL

D. Data Extraction and Analysis

This study also analyzed the entire dataset with the textual analysis software Leximancer to create outputs for research questions [23]. Leximancer uses a machine learning iterative process of seeding word definitions from frequencies and co-occurrences of words counted within blocks of text over a dataset to identify key concepts, which it groups into themes [18]. Having been used to assess scaled qualitative datasets in myriad disciplines, Leximancer was an appropriate tool to assess 93 articles for the clear concepts (direct counts) and themes (terms with strongest co-occurrences) that emerge in this LSC literature given 780124 words emerged in the dataset, roughly equivalent to the size of seven very large novels [24].

The temporal distribution of the final LSC PRJ dataset is as follows: 2018: 2; 2020: 4; 2021: 18; 2022: 67; 2023:

2. Clearly, the year 2022 saw significant growth in the number of LSC studies and inquiry. The major fields of study for these articles included education, engineering, e-commerce, marketing, and human-computer-interaction disciplines.

III. FINDINGS

A. LSC Definitions and Dialectics

The majority of dataset articles provided an explicit definition of LSC (see Fig. 1 QR code for Complete Dataset: https://bit.ly/3Y1vcRb), most often found in the introduction and literature review sections of articles. An aggregation of the explicit definitions across the dataset produced the following LSC definition:

LSC is the real-time interaction between streamers and followers across at least one live-video platform for the purpose of financial exchange—often the streamer selling a product or service and the follower consuming that product or service.

The totality of definitions also demonstrated dialectics for LSC research. Howard Sherman argued dialectics play an important role in scientific inquiry because they offer an "approach to problems that visualize the world as an interconnected totality undergoing minor and major changes due to…opposing forces" [25]. Over the dozens of LSC definitions within the dataset, researchers utilized eight specific dialectics to frame LSC including:

Communicator Dialectic—Streamer Versus Viewer: The majority of LSC studies focused on streamers, while others focused on the consumer, customer, viewer or follower. The two types of users, in this construct, make up a dialectic where one communicator is clearly focused on encouraging financial consumption in the other communicator. Streamers stream and sell. Viewers watch and buy. As LSC affords real-time dialogue for streamers/viewers, this dialectic was an inevitable focus of LSC research. Future LSC studies should extend this dialectical approach to examine communities of streamers compared to communities of viewers, as well as how collectives of viewing consumers impact the live nature of LSC relative to streamer/brand/product/service success.

Level of Profession Dialectic—Professional versus Amateur: Many LSC studies focused on professional popular streamers as influencers who contract with big brands or large retailers to sell products or push brand visibility. However, other studies revealed some live-video sellers engage in commercial exchange without professional relationships with big brands or larger retailers. Therefore, a clear dialectic emerged between professional streamers and amateurs who both use live video platforms to sell. Future LSC research should compare how these different types of streamers differently impact markets, build and sustain relationships, sell products/services, or endorse brands.

Channel Dialectic—Native Social Media Platforms Versus E-commerce Platforms: Within the dataset, some researchers focused on the channels created by e-commerce organizations where professional influencers are invited and/or contracted to sell on behalf of the organizations. However, other studies focused on individuals who use social media platforms to sell products or services. The dataset clearly demonstrated there is a difference between e-commerce platforms sellers and social media platform sellers. In some cases, successful social media platform sellers are recruited to e-commerce platforms (e.g., Instagram → Whatnot). In other cases, streamers may utilize both channels to sell and endorse, while building or sustaining relationships on one particular channel. Future LSC research should investigate how the exclusion of one or the other channel impacts streamers/viewers/brands/products/services and how various syntheses of both channel types impact streamers/viewers/brands/products/services.

Commercial Object Dialectic—Product Versus Service Versus Brand: Live streaming includes products, services, and overarching brands. Streamers sometimes only sell one of these though at other times they may sell two or more. LSC research in this dataset focused primarily on products, but there is room to investigate how selling products/services simultaneously or in relation to one another might impact streamers/consumers and associated brands. Additionally, future studies should examine how brands are impacted by product/services talk or when streamers combine these sales experiences across live-video channels.

Relational Dialectic—Commercial Versus Non-Commercial Relationships: Although several studies in the dataset focused on the relationships between streamers and consumers/customers, others defined LSC relative to its commercial aspect. Because streamers/sellers utilize LSC platforms for commercial reasons, studies also pointed up the need for streamers to persuade other users to buy or endorse brands/products/services. Overall, LSC researchers viewed the non-commercial and commercial relationships as distinct; but, moving forward researchers need to ask how slippage between the two types of relationships impacts both streamer success and product/service/brand success.

Viewing Intent Dialectic—Entertainment Versus Purchase: LSC research discussed viewers as both consumers of entertainment and customers of brands/products/services. Specifically, viewing intent, within the LSC dataset, differentiated viewers based upon their reasons for participating in the live-stream: intent to be entertained or intent to purchase. Future research must explore how switching intent impacts streamer/viewer relationship, brand/product/service success, and if intent switching impacts either of these research areas.

Streamer Disposition Dialectic—Professional Versus Authentic: LSC streamers can be a quality salesperson/endorsement vehicle or merely authentic in a way that gets more viewers to watch them live—thereby increasing sales. The LSC dataset viewed streamers as successful if they were a quality salesperson/endorsement vehicle or merely authentic enough to get more viewers to watch them live, and thus increase sales. However, both dispositions appear to be distinct in the literature. Additionally, some viewers perceived salespersons as

inauthentic. Future research may examine salesperson/authenticity ratios streamers should strive for to achieve success. Moreover, research should evaluate if such a ratio impacts purchase intentions and relational maintenance.

Frequency Dialectic—Single Purchase Versus Re-Purchases: The LSC literature highlighted when viewers/consumers/customers are impacted by streamers through live-video. For instance, some research focused on novel instances of purchasing intentions and behaviors. However, other studies concentrated on re-purchasing behavior over time. Future LSC research should investigate how streamers, live platforms, and other variables positively or negatively impact novel purchases versus repurchasing behaviors. Additionally, streamers or brands should prepare for the inevitable combination of novel and repurchasing behaviors relative to their business strategy.

B. Diversity of Theories in LSC Research

Tab. 1 demonstrates the expansion of LSC research in the last several years by identifying theories as well as example articles for each. Compared to a 2019 mobile applications literature review (column 3) and a 2022 LSC literature review (column 2), the present study identified a significant growth in LSC studies and more diversity in the theoretical frameworks applied across LSC research. In the current study, Purchase/Repurchasing Intention, Stimulus Organism Response (S-O-R), Customer/Consumer Engagement, Social Presence, and Affordance Theory were the top five most used theoretical frameworks based on occurrence. Although S-O-R and IT Affordance Theory were present in past studies, the remaining three are new theoretical frameworks in LSC literature.

TABLE I. LSC THEORIES – 2019 TO 2023 COMPARISONS

Theories in LSC Literature in Present Study	Theories in Prior LSC Literature Table Luo et al., 2022	Theories in Mobile App Literature Review Tang, 2019
Theories Shared Across Two or More Columns	**Theories Shared Across Two or More Columns**	**Theories Shared Across Two or More Columns**
Flow Theory (x4) [26]	Elaboration Likelihood Model	Elaboration Likelihood Model
IT Affordance Theory (x5) [27]	IT Affordance Theory	Flow Theory
Socio-technical System Theory [28]	Socio-technical System Theory (x2)	Technology Acceptance Model
Stimulus Organism Response (x10)[29]	Stimulus Organism Response (x2)	Theory of Planned Behavior
Technology Acceptance Model [30]	Trust Model/Theory	Uses & Gratifications Model
Theory of Planned Behavior	Uncertainty Reduction Theory	
Trust Model/Theory [31]		**Additional Theories**
Uncertainty Reduction Theory [32]	**Additional Theories**	
Uses & Gratifications Model [33]	Fit Theory	Adaptive Expectation Theory
	Source Traits	Ambivalence Theory
Additional Theories in the Current Study Literature		Big Five Trait Theory
		Cognitive Dissonance Theory
Abandonment [34]		Diffusion of Innovations Theory
Alternative Learning Theory [35]		Dimensional Theory of Emotion/Schema Theory
Attention-Interest-Desire-Action Model [36]		Environmental Psychology Perspective
Axiomatic Design Theory [37]		Expectation Confirmation Theory
Authenticity [38]		Habit Theory
Co-Creation Theory [39]		Identity Theory
Consistency-Driven Optimization Model [40]		Motivational Theory
Continuance Usage [41]		Perceived Value Theory
Customer/Consumer Engagement (x7) [42]		Personalization-Privacy Theory
Decision Support Systems [43]		Rational Addiction Theory
Dual-Systems Theory [44]		Resistance Theory
E-Commerce Success Model [45]		Self-Determination Theory
Evolutionary Game Theory/Stackelberg		Signaling Theory
Game Theory (x3) [46]		Telepresence Theory & Social Capital Theory
Gross Merchandise Value [47]		Theory of the Niche
Grounded Theory (x2) [48]		Unified Theory of Acceptance & Use of Tech.
Hypothesis/Parasocial Relationship Theory [49]		
Influencer Credibility [50]		
Information Dissemination Game Model [36]		
Key Opinion Leader [51]		
Latent Dirichlet Allocation [52]		
Matching Theory [53]		
Multi-task Learning [54]		
Pricing Strategy [55]		
Psychological Distance [56]		
Purchase & Repurchase Intentions (x13) [57]		
Semantic Network Analysis [40]		
Social Presence Theory (x6) [58]		
Source Credibility/Matchup [59]		
Switching Intention [60]		
Symbolic Interaction [61]		
Theory of Social Preference [46]		
User Experience Research [62]		
Value-Based Adoption Model [63]		
Value Perception [64]		

C. LSC Research Themes

The most frequent LSC thematic concepts that emerged from a Lexical Analysis of the 93 PRJ abstracts included influence of streamers, customer experiences during LSC consumption, the interactivity between streamers and customers, types of engagement most successful on live platforms, LSC's impact on sales, changes in shopping behavior, marketing modifications within the LSC ecosystem, purchase exchange roles, urging and nudging within LSC compared to other forms of commerce, and the impact of prices/pricing in LSC. Fig. 2 shows the lexical analysis that revealed these dataset trends based upon word count. Moreover, the cluster analysis visualized how these concepts co-occurred with additional important themes within the abstracts of this dataset.



FIGURE 2. LEXICAL ANALYSIS OF TITLES AND ABSTRACTS

D. LSC Methodologies and Data Collection

Fig. 3 identifies the data collection techniques found across the methodologies for the studies in the LSC

dataset. The most common methodologies used to examine LSC were survey analysis of viewers/consumers, survey analysis of streamers, analysis of specific platform features, and content analysis. Future research opportunities include comparative analysis of different kinds of live streaming content (e.g., text, video imagery, speed of interactions), different live-platforms relative to streamers/users communication styles, information exchanges, and consumptive intent. Researchers should also utilize biometric analysis to examine streamers from the vantage of viewers relative to commercial-based and non-commercial-based messaging. Brands and business, as well as streamers and consumers, need to know how the data available in live-streaming impacts trust and relationship building, purchase behaviors, streamer longevity, online community building, and clarity of live-exchanges. Additionally, more varied methodologies will provide important data about how various platform features operate together or separately to impact strategic outcomes.

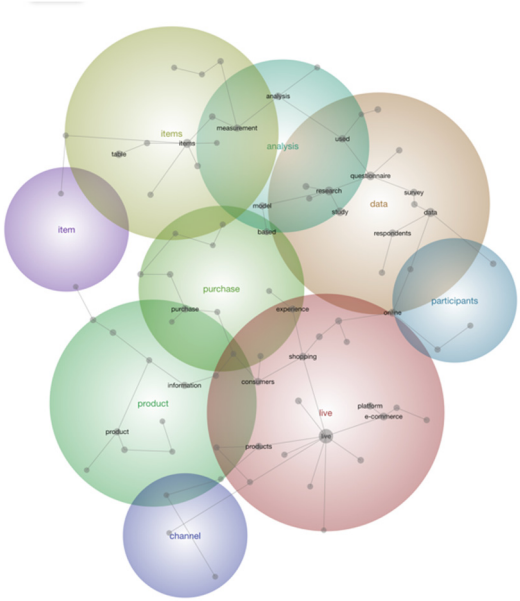


FIGURE 3. LEXICAL ANALYSIS OF METHODOLOGIES

E. LSC Research Limitations

Fig. 4 demonstrates the limitations and conclusions across the LSC dataset. After accounting for requisite terms (i.e., *live*, *study*, *research*, *results*), the top-ranked meaningful concepts of this analysis were *consumers/customers*, *purchase behaviors*, *products*, *trust*, *intentions*, *streamers*, *perception*, *information*/data, *relationships*, *engagement*, *platforms*, and *strategy* —in that order (Fig.4, list). The thematic analysis which incorporates co-occurrence calculations yielded the most important themes (Fig. 4, visual) to be *intention*, *products*, *trust*, *platform*, *sales*, and *strategy*. LSC research clearly has a diverse focus across myriad disciplines, and these limitations and conclusions themes bore out this reality. However, given that *platform* was still a less popular analytic term in the limitations and conclusions dataset than *products* and *trust* is important. First, the lack of mobile technology-specific research has not elevated these technologies to the forefront among scholars. This may indicate disciplines focused on mobile technologies (e.g., hardware, software applications), such as Engineering and Information Management, have yet to focus their analysis on LSC research. Additionally, disciplines currently leading the LSC research appeared to focus more on the human and business sides of LSC. Second, although *mobile technologies* did not appear in those sections of the dataset as ranked concepts, the more general term *platform* emerged as a concept and a theme. As platform analysis has yet to explore the parsing of technologies, processes, and human behavior to a greater degree, the mobile technologies research piece appeared to remain hidden in plain sight within LSC research. Future research should examine mobile technologies because these technologies play an important role in LSC human relationships and LSC business strategies.

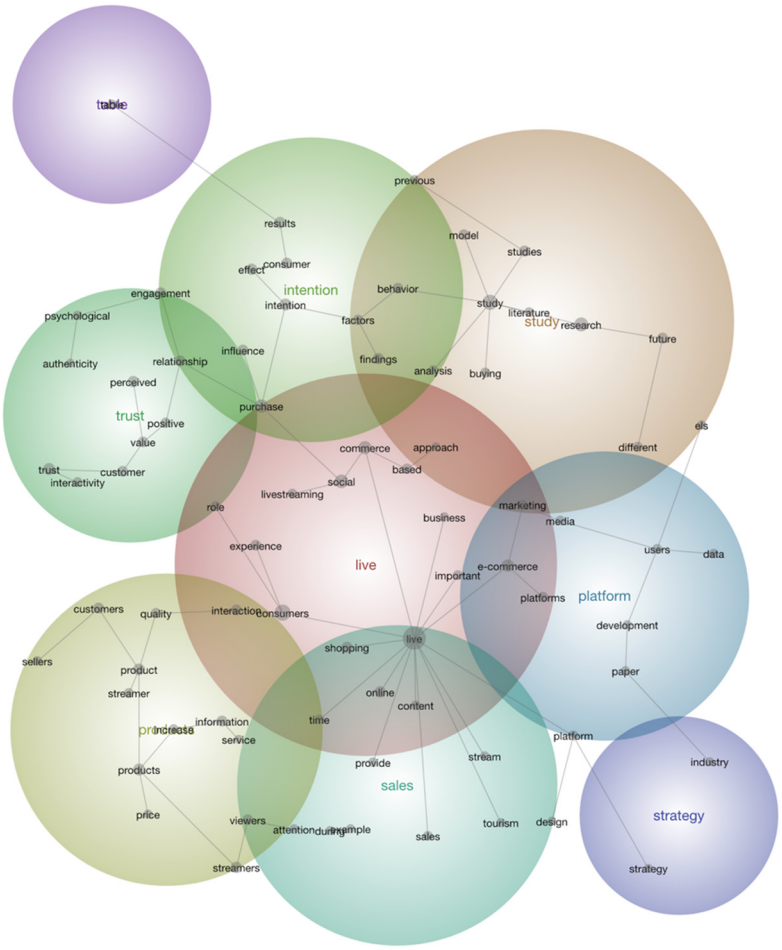


FIGURE 4. LEXICAL ANALYSIS OF LIMITATIONS AND CONCLUSIONS

F. Mobile Technologies and LSC Findings

Mobile technologies were not front and center in the titles of LSC articles, and mobile technologies were not the clear focus of most current LSC investigations. Rather, mobile platforms were assumed to be the primary modality of commerce or an imperative part of the LSC ecosystem. Fig. 5 shows a lexical analysis of the 780124 word dataset with abstracts, titles, references omitted, where *consumers*, *live*, and *research* were the largest themes. Within the *Mobile Technology Domains* notation, *mobile* was a smaller concept while *platform* was a fairly large theme. Of the 189 largest concepts within the dataset, the term *mobile* was ranked 123rd, with 345 instances counted at 5% importance. The theme *platform*, however, emerged as the 22nd largest concept at 16% importance, but co-occurred with other important terms 5427 times and was therefore calculated to be the fourth most important content theme. Moreover, the conceptual apparatus of platform served the dual purpose of keeping mobile technologies near the center of action while dismissing focus on mobile technology issues. Since these findings demonstrate the power and importance of mobile technologies within the LSC phenomenon, it should be expected that additional inquiry will be forthcoming regarding the technology that makes LSC possible. Whether for entertainment, consumption, or brand advocacy, LSC relies on mobile technologies in the form of cell phones, live video applications, and platforms for human interaction. Although it has not happened to date, additional inquiries will most likely pair previous concerns of streamers and consumers to the types, usages, and affordances of mobile technologies relative to the scope and scale of LSC platform use, the size and impact of products and/or services sold, growth of industries impacted by LSC activity, LSC platform adoption, and the evolution of how streamers and consumers use LSC platforms.

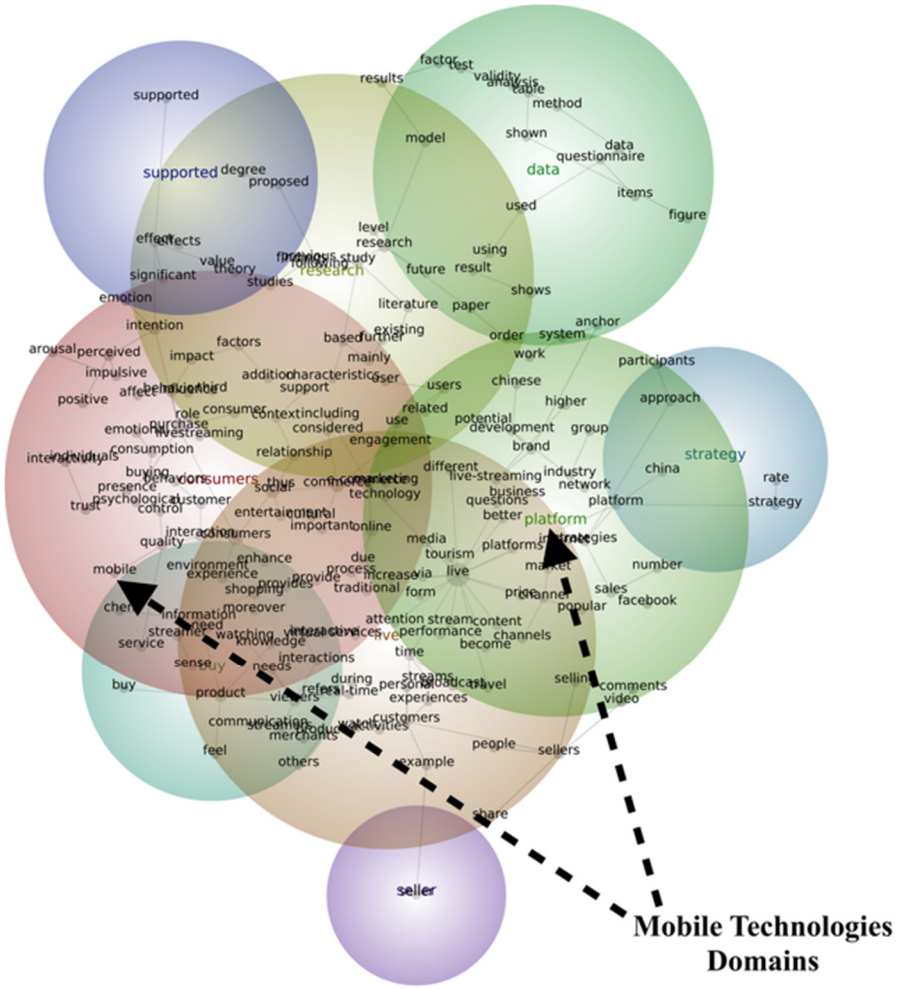


FIGURE 5. LEXICAL ANALYSIS OF COMPLETE 93 LSC ARTICLES DATASET – 189 CONCEPTS

G. *Conclusion - Future LSC Research*

The future of LSC research appears robust given the significant growth in the number of quality publications, the expansion and diversity of theoretical approaches, and the number of disciplinary perspectives in the present PRJ dataset. In dialectics that emerged within the dataset's definitions of LSC (levels of profession, channel, commercial object, relational, communicator, viewing intent, streamer disposition, and frequency) indicate myriad opportunities for further exploration within the LSC research community. However, the Mobility Dialectic should be examined more based upon the need to incorporate mobile technologies inquiry within LSC. This dialectic could show how the mobility of LSC technologies (i.e., cell phones versus web-wired studios, pad-based computers versus PC-based web cam outfits) impacts how streamers and viewers make sense of and interact with one another relative to outcomes (i.e., trust, purchase intentions), how products/services/brands are impacted by mobile technology choices in live streams, and if the greater world is visualized differently using mobile versus non-mobile technologies. This dialectic would also provide an exploration of streamer dynamism (i.e., sense of energy and movement) as well as perceived distance to/from streamers/viewers. Finally, the production value created for non-mobile hardware within LSC may impact consumption and quality perception, perception of streamer authenticity, and brand conversations. Currently, production value, environmental visualization, streamer dynamism, streamer location, and perceived distance are underexplored or not examined at all.

The limitations noted in the dataset did not indicate exploration of these dialectics nor did the conclusions indicate further study of these dialectics. Moreover, the PRJs in the dataset highlighted how researchers are overwhelmingly excluding investigation into or focus upon mobile technologies. They tacitly recognize the importance of mobile technologies as built into the platforms necessary for LSC to occur in the first place. Specifically, platform recognition was widespread in the dataset even though the mobile technologies permeating the platforms (e.g., phones, mobile devices, social media applications, live broadcasting software such as editing applications, wireless connectivity, Bluetooth) were not explored to any significant degree. Therefore, the LSC phenomenon is missing inquiry into one of its major features since mobile technologies offers a further space of inquiry for researchers. Addressing this gap will expand theoretical frameworks in future LSC literature and mobile

technology-based disciplines will add to the increasing richness of LSC research.

ACKNOWLEDGMENTS

We thank the USI Foundation and the USI TAPLab for their generous support of our work.

REFERENCES

[1] D. Ruby, "51+ Live Streaming Statistics For 2023 (Data & Infographics)," *DemandSage*, 28-Jan-2023. [Online]. Available: https://www.demandsage.com/live-streaming-statistics/#:~:text=Live Streaming Statistics 2023 (Top Picks)&text=There will be 163.4 million,to viewing gaming live streams. [Accessed: 28-Jan-2023].

[2] Z. Bao and Y. Zhu, "Understanding customers' stickiness of live streaming commerce platforms: an empirical study based on modified e-commerce system success model," *Asia Pacific J. Mark. Logist.*, 2022.

[3] A. F. Utami, I. A. Ekaputra, A. Japutra, and S. Van Doorn, "The role of interactivity on customer engagement in mobile e-commerce applications," *Int. J. Mark. Res.*, vol. 64, no. 2, pp. 269–291, Mar. 2022.

[4] C. Xie, J. Yu, S. (Sam) Huang, and J. Zhang, "Tourism e-commerce live streaming: Identifying and testing a value-based marketing framework from the live streamer perspective," *Tour. Manag.*, vol. 91, Aug. 2022.

[5] P. S. Lo, Y. K. Dwivedi, G. Wei-Han Tan, K. B. Ooi, E. Cheng-Xi Aw, and B. Metri, "Why do consumers buy impulsively during live streaming? A deep learning-based dual-stage SEM-ANN analysis," *J. Bus. Res.*, vol. 147, pp. 325–337, Aug. 2022.

[6] X. Luo, J.-H. Cheah, X.-J. Lim, and S. I. Ng, "A bibliometric review of user behaviors research in live streaming commerce," *J. Mark. Adv. Pract.*, vol. 4, no. 1, 2022.

[7] A. Wongkitrungrueng, N. Dehouche, and N. Assarut, "Live streaming commerce from the sellers' perspective: implications for online relationship marketing," *J. Mark. Manag.*, vol. 36, no. 5–6, pp. 488–518, Mar. 2020.

[8] A. Wongkitrungrueng and N. Assarut, "The role of live streaming in building consumer trust and engagement with social commerce sellers," *J. Bus. Res.*, vol. 117, pp. 543–556, Sep. 2020.

[9] T.-P. Liang and E. Turban, "Introduction to the Special Issue Social Commerce: A Research Framework for Social Commerce," *Int. J. Electron. Commer.*, vol. 16, no. 2, pp. 5–14, 2011.

[10] L. Esmaeili and S. A. Hashemi G., "A systematic review on social commerce," *J. Strateg. Mark.*, vol. 27, no. 4, pp. 317–355, May 2019.

[11] R. Curty and P. Zhang, "Website features that gave rise to social commerce: A historical analysis," *Electron. Commer. Res. Appl.*, vol. 12, pp. 260–279, 2013.

[12] S. H. Lee, D. DeWester, and S. R. Park, "Web 2.0 and opportunities for small businesses," *Serv. Bus.*, vol. 2, no. 4, pp. 335–345, 2008.

[13] J. Cai and D. Yvette Wohn, "Live streaming commerce: Uses and gratifications approach to understanding consumers' motivations," 2019.

[14] J. Sun and T. Chi, "Investigating the adoption of apparel m-commerce in the US market," *Int. J. Cloth. Sci. Technol.*, 2019.

[15] M. Ciupac-Ulici, D. G. Beju, V. P. Bresfelean, and G. Zanellato, "Which factors contribute to the global expansion of m-commerce?," *Electron.*, vol. 12, no. 1, Jan. 2023.

[16] C. Wen, N. Wang, J. Fang, and M. Huang, "An integrated model of continued m-commerce applications usage," *J. Comput. Inf. Syst.*, 2022.

[17] A. Rahmah, H. B. Santoso, and Z. A. Hasibuan, "Critical review of technology-enhanced learning using automatic content analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 1, pp. 385–394, 2022.

[18] Leximancer Pty Ltd, "Leximancer user guide: Relese 4.5," 2018. [Online]. Available: http://info.leximancer.com/. [Accessed: 03-Jul-2019].

[19] Center for Cognitive Ageing and Cognitive Epidemiology, "Systematic reviews and meta-analyses: a step-by-step guide," 2013. [Online]. Available: https://www.ccace.ed.ac.uk/research/software-resources/systematic-reviews-and-meta-analyses. [Accessed: 18-Jan-2023].

[20] B. Kitchenham and P. Brereton, "A Systematic Review of Systematic Review Process Research in Software Engineering," 2013.

[21] A. K. Y. Tang, "A systematic literature review and analysis on mobile apps in m-commerce: Implications for future research," *Electron. Commer. Res. Appl.*, vol. 37, Sep. 2019.

[22] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *The BMJ*, 29-Mar-2021. [Online]. Available: https://www.bmj.com/content/bmj/372/bmj.n71.full.pdf. [Accessed: 30-Nov-2021].

[23] K. Çakar and Ş. Aykol, "The past of tourist behaviour in hospitality and tourism in difficult times: A systematic review of literature (1978–2020)," *Int. J. Contemp. Hosp. Manag.*, vol. 35, no. 2, pp. 630–656, Jan. 2022.

[24] V. Wilk, G. N. Soutar, and P. Harrigan, "Tackling social media data analysis: Comparing and contrasting QSR NVivo and Leximancer," *Qual. Mark. Res.*, vol. 22, no. 2, pp. 94–113, Apr. 2019.

[25] H. Sherman, "Dialectics as a method," *Insurg. Sociol.*, vol. 6, no. 4, pp. 57–64, 1976.

[26] X. Liu, L. Zhang, and Q. Chen, "The effects of tourism e-commerce live streaming features on consumer purchase intention: The mediating roles of flow experience and trust," *Front Psychol*, vol. 13, Aug. 2022, doi: 10.3389/fpsyg.2022.995129.

[27] Z. Deng, P. Benckendorff, and J. Wang, "Travel live streaming: an affordance perspective," *Information Technology and Tourism*, vol. 23, no. 2, pp. 189–207, Jun. 2021, doi: 10.1007/s40558-021-00199-1.

[28] M. Zhang, Y. Liu, Y. Wang, and L. Zhao, "How to retain customers: Understanding the role of trust in live streaming commerce with a socio-technical perspective," *Comput Human Behav*, vol. 127, Feb. 2022, doi: 10.1016/j.chb.2021.107052.

[29] S. C. Lin, H. T. Tseng, F. Shirazi, N. Hajli, and P. T. Tsai, "Exploring factors influencing impulse buying in live streaming shopping: a stimulus-organism-response (SOR) perspective," *Asia Pacific Journal of Marketing and Logistics*, 2022, doi: 10.1108/APJML-12-2021-0903.

[30] H. Yang and H. Lee, "Exploring user acceptance of streaming media devices: an extended perspective of flow theory," *Information Systems and e-Business Management*, vol. 16, no. 1, pp. 1–27, Feb. 2018, doi: 10.1007/s10257-017-0339-x.

[31] Z. Wang, J. Li, and P. Chen, "Factors influencing Chinese flower and seedling family farms' intention to use live streaming as a sustainable marketing method: an application of extended theory of planned behavior," *Environ Dev Sustain*, vol. 24, no. 3, pp. 4299–4322, Mar. 2022, doi: 10.1007/s10668-021-01616-5.

[32] B. Lu and Z. Chen, "Live streaming commerce and consumers' purchase intention: An uncertainty reduction perspective," *Information and Management*, vol. 58, no. 7, Nov. 2021, doi: 10.1016/j.im.2021.103509.

[33] J. Cai and D. Yvette Wohn, "Live streaming commerce: Uses and gratifications approach to understanding consumers' motivations," 2019. Accessed: Jan. 27, 2023. [Online]. Available: https://scholarspace.manoa.hawaii.edu/items/58d44fca-57a8-4bc7-98dd-c6203ee6ecdc

[34] Z. Li, M. A. Kaafar, K. Salamatian, and G. Xie, "Characterizing and Modeling User Behavior in a Large-Scale Mobile Live Streaming System," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 12, pp. 2675–2686, Dec. 2017, doi: 10.1109/TCSVT.2016.2595325.

[35] X. Zhou and X. Tian, "Impact of Live Streamer Characteristics and Customer Response on Live-streaming Performance: Empirical Evidence from e-Commerce

Platform," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 1277–1284. doi: 10.1016/j.procs.2022.11.306.
[36] J. Lv, W. Yao, Y. Wang, Z. Wang, and J. Yu, "A game model for information dissemination in live streaming e-commerce environment," *International Journal of Communication Systems*, vol. 35, no. 1, Jan. 2022, doi: 10.1002/dac.5010.
[37] Q. Xiao, S. Wan, X. Zhang, M. Siponen, L. Qu, and X. Li, "How consumers' perceptions differ towards the design features of mobile live streaming shopping platform: A mixed-method investigation of respondents from Taobao Live," *Journal of Retailing and Consumer Services*, vol. 69, Nov. 2022, doi: 10.1016/j.jretconser.2022.103098.
[38] J. Sun, H. Dushime, and A. Zhu, "Beyond beauty: A qualitative exploration of authenticity and its impacts on Chinese consumers' purchase intention in live commerce," *Front Psychol*, pp. 1–17, 2022, doi: 10.3389/fpsyg.2022.944607.
[39] J. Yang, Y. Zeng, X. Liu, and Z. Li, "Nudging i nteractive cocreation behaviors in live-streaming travel commerce: The visualization of real-time danmaku," *Journal of Hospitality and Tourism Management*, vol. 52, pp. 184–197, Sep. 2022, doi: 10.1016/j.jhtm.2022.06.015.
[40] S. P. Wan, J. Yan, and J. Y. Dong, "Trust and personalized individual semantics based fusion method for heterogeneous multi-criteria group decision making and application to live streaming commerce," *Expert Syst Appl*, vol. 208, Dec. 2022, doi: 10.1016/j.eswa.2022.118151.
[41] H. X. Chong, A. H. Hashim, S. Osman, J. L. Lau, and E. C. X. Aw, "The future of e-commerce? Understanding livestreaming commerce continuance usage," *International Journal of Retail and Distribution Management*, Jan. 2022, doi: 10.1108/IJRDM-01-2022-0007.
[42] L. Guo, X. Hu, J. Lu, and L. Ma, "Effects of customer trust on engagement in live streaming commerce: mediating role of swift guanxi," *Internet Research*, vol. 31, no. 5, pp. 1718–1744, Nov. 2021, doi: 10.1108/INTR-02-2020-0078.
[43] Y. Wang, Z. Lu, P. Cao, J. Chu, H. Wang, and R. Wattenhofer, "How Live Streaming Changes Shopping Decisions in E-commerce: A Study of Live Streaming Commerce," *Comput Support Coop Work*, Dec. 2022, doi: 10.1007/s10606-022-09439-2.
[44] W. He and C. Jin, "A study on the influence of the characteristics of key opinion leaders on consumers' purchase intention in live streaming commerce: based on dual-systems theory," *Electronic Commerce Research*, 2022, doi: 10.1007/s10660-022-09651-8.
[45] Z. Bao and Y. Zhu, "Understanding customers' stickiness of live streaming commerce platforms: an empirical study based on modified e-commerce system success model," *Asia Pacific Journal of Marketing and Logistics*, 2022, doi: 10.1108/APJML-09-2021-0707.
[46] T. Chen, L. Peng, J. Yang, G. Cong, and G. Li, "Evolutionary game of multi-subjects in live streaming and governance strategies based on social preference theory during the covid-19 pandemic," *Mathematics*, vol. 9, no. 21, Nov. 2021, doi: 10.3390/math9212743.
[47] H. Chen, Y. Dou, and Y. Xiao, "Understanding the Role of Live Streamers in Live-Streaming E-Commerce," 2022. [Online]. Available: https://ssrn.com/abstract=4141864
[48] S. Zhang, C. Huang, X. Li, and A. Ren, "Characteristics and roles of streamers in e-commerce live streaming," *Service Industries Journal*, vol. 42, no. 13–14, pp. 1001–1029, 2022, doi: 10.1080/02642069.2022.2068530.
[49] Z. Deng, P. Benckendorff, and J. Wang, "From interaction to relationship: Rethinking parasocial phenomena in travel live streaming," *Tour Manag*, vol. 93, Dec. 2022, doi: 10.1016/j.tourman.2022.104583.
[50] M. Yan, A. P. K. Kwok, A. H. S. Chan, Y. S. Zhuang, K. Wen, and K. C. Zhang, "An empirical investigation of the impact of influencer live-streaming ads in e-commerce platforms on consumers' buying impulse," *Internet Research*, 2022, doi: 10.1108/INTR-11-2020-0625.
[51] W. Lyu, Y. Qi, and J. Liu, "Proliferation in live streaming commerce, and key opinion leader selection," *Electronic Commerce Research*, 2022, doi: 10.1007/s10660-022-09605-0.
[52] T. Chen, C. Tong, Y. Bai, J. Yang, G. Cong, and T. Cong, "Analysis of the Public Opinion Evolution on the Normative Policies for the Live Streaming E-Commerce Industry Based on Online Comment Mining under COVID-19 Epidemic in China," *Mathematics*, vol. 10, no. 18, Sep. 2022, doi: 10.3390/math10183387.
[53] W. Rungruangjit, "What drives Taobao live streaming commerce? The role of parasocial relationships, congruence and source credibility in Chinese consumers' purchase intentions," *Heliyon*, vol. 8, no. 6, Jun. 2022, doi: 10.1016/j.heliyon.2022.e09676.
[54] S. Yu *et al.*, "Leveraging Tripartite Interaction Information from Live Stream E-Commerce for Improving Product Recommendation," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2021, pp. 3886–3894. doi: 10.1145/3447548.3467151.
[55] C. Hao and L. Yang, "Resale or agency sale? Equilibrium analysis on the role of live streaming selling," *Eur J Oper Res*, 2022, doi: 10.1016/j.ejor.2022.09.021.
[56] L. T. Huang, "Exploring the Affective Way Leading to Impulse Buying in Social Media Live Streaming," in *2021 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1323–1327. doi: 10.1109/IEEM50564.2021.9672962.
[57] Y. Chen, F. Lu, and S. Zheng, "A Study on the Influence of E-Commerce Live Streaming on Consumer Repurchase Intentions," *Int J Mark Stud*, vol. 12, no. 4, p. 48, Nov. 2020, doi: 10.5539/ijms.v12n4p48.
[58] J. Chen and J. Liao, "Antecedents of Viewers' Live Streaming Watching: A Perspective of Social Presence Theory," *Front Psychol*, vol. 13, Mar. 2022, doi: 10.3389/fpsyg.2022.839629.
[59] Z. Jiang, H. Wang, J. Xue, and T. Zhai, "Propelling consumer engagement via entrepreneurs' live streaming?," *Front Psychol*, pp. 1–22, 2022, doi: 10.3389/fpsyg.2022.890707.
[60] D. Ye, F. Liu, D. Cho, and Z. Jia, "Investigating switching intention of e-commerce live streaming users," *Heliyon*, vol. 8, no. 10, Oct. 2022, doi: 10.1016/j.heliyon.2022.e11145.
[61] P. Liu, J. Zhao, and Y. Zhang, "Research on Symbolic Interaction in E-Commerce Live Streaming," *OAlib*, vol. 09, no. 11, pp. 1–12, 2022, doi: 10.4236/oalib.1109442.
[62] Q.-M. Ying and K.-R. Kim, "Effects of Customer Experience Factors on Beauty Product Satisfaction in Beauty Live Commerce: -For women in their 20s and 30s in Zhejiang Province, China-," *Journal of the Korean Society of Cosmetology*, vol. 27, no. 6, pp. 1437–1446, Dec. 2021, doi: 10.52660/jksc.2021.27.6.1437.
[63] C. Xie, J. Yu, S. (Sam) Huang, and J. Zhang, Tourism e-commerce live streaming: Identifying and testing a value-based marketing framework from the live streamer perspective," *Tour Manag*, vol. 91, Aug. 2022, doi: 10.1016/j.tourman.2022.104513.
[64] J. Cao, J. Li, Y. Wang, and M. Ai, "The Impact of Self-Efficacy and Perceived Value on Customer Engagement under Live Streaming Commerce Environment," *Security and Communication Networks*, vol. 2022, 2022, doi: 10.1155/2022/2904447.

Dehumanized Avatars: Unethical Behavior in the Metaverse

Carina Castagna, Sercan Demir, Markus Weinmann
University of Cologne, Cologne, Germany
castagna@wiso.uni-koeln.de, demir@wiso.uni-koeln.de, weinmann@wiso.uni-koeln.de

***Abstract*—Despite the rise of the metaverse, surprisingly, little research has explored the effects of immersion within the metaverse on individual user behavior. This study begins to fill this gap by examining the effect of dehumanized avatar identities—making avatars look less human-like—and its consequences for ethical behavior. We conducted an online experiment (N=130) to test if an avatar's level of dehumanization increases unethical behavior. Our study assessed three different levels of dehumanization (human-like vs. animal-like vs. abstract avatars). The results indicate that participants' willingness to cheat increases when engaging through an abstract and more dehumanized avatar. This study reveals the hidden risks of avatars in the metaverse and sheds light on the delicate relationship between unethical behavior and avatar dehumanization.**

Keywords—digital self, avatars, ethical behavior, dehumanization, metaverse

I. INTRODUCTION

The metaverse is expected to grow rapidly within the next years[1], so understanding users' behaviors within the metaverse is of utmost importance for companies and organizations. Users typically interact by choosing an avatar representing them in the digital sphere, also known as (digital) embodiment [1], [2]. In other words, they hide behind a digital self while interacting with others.

Hiding behind a digital self, however, may lead to unintended consequences; for example, the enhanced state of anonymity may increase unethical behavior such as cheating. Thus, this paper explores one unique form of embodiment in digital self-representations, anthropomorphism levels, and how this affects unethical behavior. We define a high anthropomorphized avatar as those who typically maintain a human-like form and low anthropomorphized avatars with dehumanized features (e.g., abstract, robotic, animal forms). Following the mechanistically dehumanized discussion [3], we raise the research question of whether ethical values found in the real world still apply to the virtual world for dehumanized avatars (i.e., low anthropomorphism). In particular, we study how ethical choices are affected by anthropomorphism levels of the digital embodiment of the self.

The present research starts to answer this important question by bridging the literature on ethical behavior and anthropomorphism to put forth the idea that when digital representations are less anthropomorphized, users may have more predisposition to cheat. Specifically, in this working paper, we begin to answer those questions by applying an initial online experiment (N = 130 participants) to show how constructions and definitions of the self (for example, as an avatar) may impact ethical behavior in digital environments. Specifically, we show that digital-self representations with lower levels of anthropomorphism—more abstract or animal-like avatars—lead users to cheat more.

In documenting these effects, this research makes significant contributions to theory and practice. To the best of our knowledge, this study is among the first to experimentally manipulate and test how avatars and their different anthropomorphism levels affect ethical behavior and decision-making. Next, we theorize about digital embodiment of the self, anthropomorphism, and cheating behavior. We then present the study that tests our prediction. The paper ends with a discussion of the theoretical implications of our findings.

II. THEORETICAL BACKGROUND AND HYPOTHESIS

We draw upon two distinct research streams: (1) Avatars and the digital embodiment of the self, and (2) Dehumanization and cheating behavior. We have chosen avatars as a study context because they are well-studied regarding their effect on users' decision-making [4], [5]. Additionally, users are increasingly interacting with avatars making decisions in the digital environment, and extending their identities to digital environments [1], [5]. This makes avatars a practically relevant and timely topic to study. Furthermore, as we are interested in the effect of dehumanization on online cheating, avatars are ideal for incorporating the concept of anthropomorphism, as avatars are malleable regarding their external appearance and individual styling.

A. Avatars and Digital Embodiment of the Self

Avatars are graphical representations of real-world users in a digital environment. They can be divided into generic and customized avatars [6]. While generic avatars are pre-defined, customized avatars are malleable and can be changed according to one's preferences. Our study builds on the extensive literature on avatar research [6].

A vast amount of literature found that users' decision-making in a digital environment is affected by an avatar's main appearance [7]. In this sense, users can extend themselves by using a digital representation (i.e., an avatar), which can also affect real-world decisions. Interestingly,

[1] https://www.statista.com/statistics/1288048/united-states-adults-reasons-for-joining-the-metaverse/

when people create an avatar, it can develop new constructions and definitions of themselves [1]. While users can alternate their avatar body, selecting and re-configuring its image, they feel free to become their ideal selves. For instance, people can have a new gender identity online or represent their aspirational self-image [8]. The idea of (digital) embodiment [1] explains changes in decision-making, as consumers are free to create new identities in the digital world, re-creating their selves and, consequently, changing their behavior.

B. Dehumanization and Cheating Behavior

Individuals' offline behavior can be affected by even the smallest changes between their real bodies and virtual selves. However, studies show that different levels of avatar anthropomorphism—making avatars look more human-like by attributing human-like qualities [9]—can diminish this impact, thus increasing the similarity between the virtual and offline behavior [10]. For instance, avatars wearing a user's real face increase their identification level with this avatar [11], which can lead to consequences in the real world. For example, people, who observed their own real faces on an avatar, had a significantly higher willingness to quit smoking, in comparison with the other-avatar condition [10]. While the relation between anthropomorphism and avatars is optimistic, dehumanized characteristics on avatars are increasing within the metaverse context. It is becoming more and more popular to create avatars with lower levels of anthropomorphism; we call this phenomenon dehumanization, which is the attribution of animal or object traits to oneself (i.e., avatars with geometric forms) [12].

Most previous literature focuses on making avatars more human-like, that is, increasing the levels of anthropomorphism; however, less research focuses on the opposite effect, that is, making avatars look less human-like, for example, the consequences of using animal-like or abstract avatars [12]. This effect is also referred to as "dehumanization," that is, removing human traits. Removing human traits might also have unintended consequences; for example, people acting via a dehumanized avatar might also act less morally or ethically. Given that users are typically free to choose any type of avatar in most virtual settings—human-like, animal-like or even something else—we investigate how dehumanizing avatars affects real behavior. Specifically, we will focus on ethical behavior for two reasons: (1) ethical behavior is usually associated with humans, and using dehumanized avatars may also lead to less ethical behavior; (2) understanding how avatar choice influences ethical behavior may help to design better avatars, which might lead to less cheating in virtual worlds.

Online cheating has been shown to be a major problem in digital environments [13]. Cheating comprises conscious unethical behavior, such as lying or manipulating performance standards to promote self-interests [14]. Cheating behavior can increase when users are behind their digitally reembodied selves. Previous research explains this effect as users feel more anonymous and lack accountability when being in a digital environment [15].

Thus, in the same way, that an anthropomorphized avatar endorses healthy offline habits [10], we posit that dehumanized avatars can possibly affect people's ethical behavior. Specifically, we predict that people can hide behind an avatar; and that the more dehumanized an avatar look, the more users tend to act less human-like. That is, moral standards go down, which leads users with dehumanized avatars cheating more. Therefore, we formally hypothesize:

Hypothesis 1 (H1): *Users exposed to more (less) dehumanized avatars cheat more (less).*

Our hypothesis highlights a significant contribution to avatars and ethical behavior discussion. Although previous studies applied avatars to verify ethical behavior, these investigations typically demonstrate the manipulation aspects of anthropomorphism and not dehumanization [16], [17]. Further, no manipulation of dehumanization was associated with unethical behavior, as most of the studies are also related to man vs. machine, that is, the comparison between AI and human interaction [18]. Another example is [19], who proposed to study realistic vs. unrealistic avatars. The authors manipulated both conditions using human features rather than geometric or animal features.

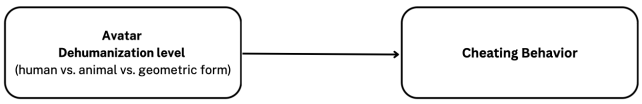


Fig. 1: Conceptual framework

III. METHOD

A. Material and Procedure

We conducted a between-group online experiment to test our hypothesis. We randomly assigned participants to one of three avatar conditions, only differing in the level of dehumanization: (a) the human-like condition, in which the avatar looked like a human and the level of dehumanization was relatively low, (b) the animal condition, in which the avatar was a dog, and the level of dehumanization was relatively moderate and (c) the abstract condition, in which the avatar was a white geometric form and the level of dehumanization was relatively high (see also Figure 2). All material was presented via Qualtrics.

We told participants that they were part of a metaverse study. After participants were randomly assigned to one of the three avatar conditions, we let them play the mind game [20], [21]. In this game, participants were asked to think of a number between one and six (the range of a die). Afterward, on a new page, a random numbers generator rolled a six-sided die. Participants were then asked to report "yes" if the number they thought of matched the

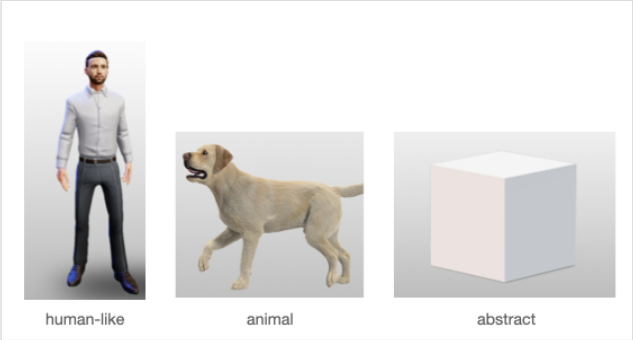


Fig. 2: Treatment conditions

die outcome, otherwise "no." To incentivize cheating, participants received £1 when they reported "yes," and £0 if they reported "no." At the end of the experiment, we gathered demographic data and debriefed the participants.

B. Participants

We used Prolific (https://prolific.co) to recruit participants who are older than 18 years and reside in the US. In total, we recruited 150 participants. Twenty participants did not pass an attention check, resulting in a final sample size of $N = 130$ (age = 40.3 years, 44% female). We paid all participants £1 (about US$1.21) for a 7-minute task, a payment equivalent to an hourly wage of £9.11 (US$11). In addition, they received a bonus depending on their decisions made in the experiment (described above).

C. Measures and Model Specification

The purpose of our analysis is to compare reported matches—i.e., the number people thought of and the results of the random numbers generator—across conditions. According to our theorizing, we expect more matches in more dehumanized conditions, that is, individuals cheat more when they are exposed to animal-like or abstract avatars. Our dependent variable was binary, that is if an individual indicated a match (=1) or not (=0). Hence, we specified a logistic regression model to estimate the effect of being in a particular condition—human-like vs. animal vs. abstract—on a participant's likelihood to report a match:

$$Pr(\text{match}_i = 1) = logit(p_i) = \alpha + \beta \times \text{condition}_\text{i}, \quad (1)$$

where α is the intercept representing reported matches in the human-like condition, which also served as baseline; β represents the differences of matches reported in the other conditions, either the animal or abstract condition. According to our hypothesis, we expected β to be positive and significant for the animal and abstract conditions. That is, participants report more matches in those conditions compared to the human-like condition.

IV. RESULTS

A. Descriptive Statistics

Figure 3 presents the descriptive statistics by condition. In the human condition, 47.91% of the participants reported a match. This indicates cheating, given an expected value of 16.67% (the probability that a participant's report matches the random numbers generator is 1 out of 6). In the other condition, the share was even higher: 63.26% in the animal condition, and 66.67% in the abstract condition reported a match.

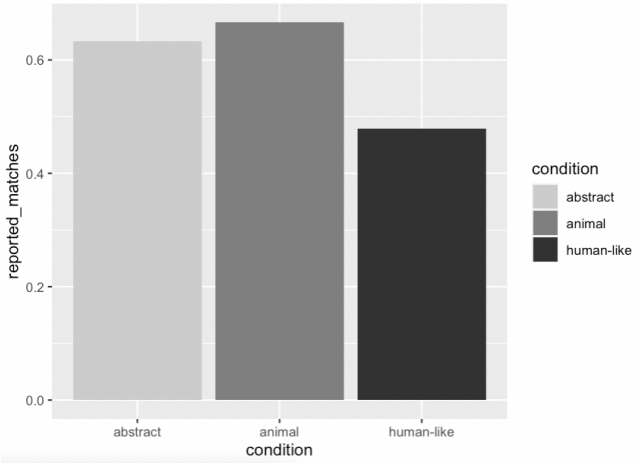


Fig. 3: Conditions and respective cheating share

B. Estimation Results

Table I presents our estimation results. Coefficients are presented on an odds scale. The results show that the odds of reporting a match in the abstract condition are 2.21 times higher than in the human condition (the reference group). Likewise, the odds of reporting a match in the animal condition are 2.13 times higher. Both coefficients are positive and significant, thus supporting our hypothesis.

TABLE I: Logistic Regression Results (Odds-Ratio)

	Dependent variable:
	Binary: reported match (yes = 1)
Abstract condition	2.21*
	(0.88–5.76)
Animal condition	2.13*
	(0.93–5.02)
Intercept	0.41
	(0.09–1.78)
Controls	yes
Observations	130
Log Likelihood	−84.316
Akaike Inf. Crit.	178.631
Notes:	*p<0.1; **p<0.05; ***p<0.01
	We used age and gender as controls.
	Confidence intervals in parentheses.

V. GENERAL DISCUSSION

Summary. The present research sheds light on the role of digital self-representations (i.e., avatars) on ethical

behavior. Specifically, we examined how different levels of dehumanized avatars—human-like, animal-like, abstract—can increase cheating behavior.

Theoretical contributions. The present research makes primary contributions to the disposal literature. First, we provide clear and causal evidence that users are more willing to cheat within a digital context when acting through a dehumanized avatar. Specifically, by manipulating three different levels of dehumanization (human-like vs. animal-like vs. abstract avatars), we found that users acting through a dehumanized avatar cheat more in 20% of the cases. Second, our results demonstrate that participants in a human-like (i.e., anthropomorphic avatar) condition act less dishonestly (approx. 16% less) than when assigned to an abstract or animal avatar. We contribute to the digital self-extension theory [1] by showing how the embodiment of a self on a dehumanized level can affect ethical behavior. The effect of anonymity on unethical behavior [22] can be even more substantial on the metaverse as users create and identify their selves with a new digital self, diminishing their ethical values.

We also contribute to avatar research [10], [11] by showing that anthropomorphism and dehumanized characteristics can influence ethical behavior. Dehumanized avatars (i.e., low anthropomorphism level) appear to boost unethical behavior. Thus, our results add to the novel research vain of dehumanization and digital selves [3], [7].

Limitations and future research. This research has some limitations that offer avenues for future research. In this study, we relied mainly on an online experiment, manipulating generic avatars, which participants could not actually customize. Despite manipulating a fictitious metaverse scenario, our online study was not tested on a real digital platform. Another limitation is the use of only one gender in avatar manipulation. Thus, future work may wish to change how avatars' pictures are displayed on the manipulation, emphasizing the avatar as the participant's identity.

Thus, an additional study could examine whether different genders impact our main effect. Further, the salience of the avatar image could also be increased by keeping it visible through each decision stage, thereby increasing the manipulation impact. This can be executed in two ways: First, by presenting avatar images during the mind game stage. Second, the avatar presence may also be highlighted by increasing the experiment's sense of reality. For instance, future research could examine the impact of digital self-representations on dishonest behavior by applying Virtual Reality scenarios. As users are able to customize avatars according to their own preferences, a personalized avatar could act as an extended self and thus could be strengthening the effect. Future research could further examine whether ethical behavior is affected by underlying processes such as socioeconomic vulnerability and personal control.

Managerial implications. Finally, our findings bring important implications for the design of digital environments for several stakeholder groups. As our results showcase the potential unethical behaviors of digital-self representations that are distant from the human form, we invite behavioral researchers and practitioners to consider how interventions might be used to improve ethical behavior in the metaverse. As the immersion in the metaverse may daze boundaries between the real world and the virtual world, we highlight the importance of ethical procedures during the primary phase of virtual worlds, the avatar creation and self-identification. In this sense, our results featured the need for ethical guidance associated with avatars' dehumanization level, increasing cheating probabilities. For instance, educative programs in order to increase users' identification with avatars, despite their dehumanized form, shall decrease this phenomenon. Further, companies may want to nudge metaverse users to consider the unethical cues that dehumanized avatars can portend for their virtual network. Cheating behavior could also be improved with the development of ownership nudges, making users more accountable for their actions even on dehumanized avatars. Our findings also suggest that consumers might want to consider shifting their avatars into a more anthropomorphized form to collaborate in an ethical digital environment. A universal ethical user certification could also be developed in order to create individual accountability and credibility online. Finally, we urge that ethical considerations and feasible solutions need to be discussed and built into the first step of the metaverse immersion, which has also shown to be one of the most important: the development and identification with an avatar.

VI. CONCLUSION

In this paper, we hypothesized and tested the effect of dehumanized avatars on online cheating behavior. In particular, results show that users are generally more willing to cheat when acting through an avatar in a digital context. A more dehumanized avatar (animal or geometric form) leads to more cheating behavior.

REFERENCES

[1] Belk, Russell W., "Extended self in a digital world," *Journal of Consumer Research*, vol. 40, no. 3, pp. 477–500, 2013.

[2] ——, "Possessions and the extended self," *Journal of Consumer Research*, vol. 15, no. 2, pp. 139–168, 1988.

[3] N. Castelo and D. R. Lehmann, "Be careful what you wish for: Unintended consequences of increasing reliance on technology," *Journal of Marketing Behavior*, vol. 4, no. 1, pp. 31–42, 2019.

[4] M. Seymour, K. Riemer, and J. Kay, "Actors, avatars and agents: Potentials and implications of natural face technology for the creation of realistic visual presence," *Journal of the Association for Information Systems*, pp. 953–981, 2018.

[5] A. Davis, J. Murphy, University of Nebraska at Omaha, D. Owens, D. Khazanchi, and I. Zigurs, "Avatars, people, and virtual worlds: Foundations for research in metaverses," *Journal of the Association for Information Systems*, vol. 10, no. 2, pp. 90–117, 2009.

[6] H. M. Aljaroodi, M. T. P. Adam, R. Chiong, and T. Teubner, "Avatars and embodied agents in experimental information systems research: A systematic review and conceptual framework," *Australasian Journal of Information Systems*, vol. 23, pp. 1–37, 2019.

[7] N. Yee, J. N. Bailenson, and N. Ducheneaut, "The proteus effect: Implications of transformed digital self-representation on online and offline behavior," *Communication Research*, vol. 36, no. 2, pp. 285–312, 2009.

[8] J. Martin, "Use-Value, Exchange-Value, and the Role of Virtual Goods in Second Life."

[9] N. Epley, A. Waytz, and J. T. Cacioppo, "On seeing human: A three-factor theory of anthropomorphism." *Psychological Review*, vol. 114, no. 4, pp. 864–886, 2007.

[10] H. Song, J. Kim, R. J. Kwon, and Y. Jung, "Anti-smoking educational game using avatars as visualized possible selves," *Computers in Human Behavior*, vol. 29, no. 5, pp. 2029–2036, 2013.

[11] Y. Seo, M. Kim, Y. Jung, and D. Lee, "Avatar face recognition and self-presence," *Computers in Human Behavior*, vol. 69, pp. 120–127, 2017.

[12] K. Karanika and M. K. Hogg, "Self–object relationships in consumers' spontaneous metaphors of anthropomorphism, zoomorphism, and dehumanization," *Journal of Business Research*, vol. 109, pp. 15–25, Mar. 2020.

[13] M. Weinmann, J. S. Valacich, C. Schneider, J. L. Jenkins, and M. Hibbeln, "The path of the righteous: Using trace data to understand fraud decisions in real time," *MIS Quarterly*, vol. 46, no. 4, pp. 2317–2336, 2022.

[14] M. S. Mitchell, M. L. Ambrose, R. Folger, M. D. Baer, and N. F. Palmer, "Cheating Under Pressure: A Self-Protection Model of Workplace Cheating Behavior," *Journal of Applied Psychology*, vol. 103, no. 1, pp. 54–73, 2018.

[15] J. Suler, "The online disinhibition effect," *CyberPsychology & Behavior*, vol. 7, no. 3, pp. 321–326, 2004.

[16] J. F. Nunamaker, D. C. Derrick, A. C. Elkins, J. K. Burgoon, and M. W. Patton, "Embodied Conversational Agent-Based Kiosk for Automated Interviewing," *Journal of Management Information Systems*, vol. 28, no. 1, pp. 17–48, Jul. 2011. [Online]. Available: https://www.tandfonline.com/doi/full/10.2753/MIS0742-1222280102

[17] R. M. Schuetzler, G. M. Grimes, and J. S. Giboney, "The effect of conversational agent skill on user behavior during deception," *Computers in Human Behavior*, vol. 97, pp. 250–259, Aug. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0747563219301311

[18] M. D. Pickard, C. A. Roster, and Y. Chen, "Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions?" *Computers in Human Behavior*, vol. 65, pp. 23–30, Dec. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0747563216305684

[19] M. Seymour, L. Yuan, A. R. Dennis, and K. Riemer, "Have We Crossed the Uncanny Valley? Understanding Affinity, Trustworthiness, and Preference for Realistic Digital Humans in Immersive Environments," *Journal of the Association for Information Systems*, vol. 22, no. 3, pp. 591–617, 2021. [Online]. Available: https://aisel.aisnet.org/jais/vol22/iss3/9/

[20] A. Kajackaite and U. Gneezy, "Incentives and cheating," *Games and Economic Behavior*, vol. 102, pp. 433–444, 2017.

[21] T. Jiang, "Cheating in mind games: The subtlety of rules matters," *Journal of Economic Behavior & Organization*, vol. 93, pp. 328–336, 2013.

[22] K. C. Yam and S. J. Reynolds, "The effects of victim anonymity on unethical behavior," *Journal of Business Ethics*, vol. 136, no. 1, pp. 13–22, 2016.

Tangible Data Exploration: Creating Card Games for Sensemaking

Annika Wolff and Antti Knutas
LUT University / School of Engineering Science (LENS), Lappeenranta, Finland
annika.wolff, antti.knutas@lut.fi

***Abstract*—Participatory sensing has largely focused on the collection and dissemination of data, paying little attention to the fact that people often struggle to use the data and relate it to their daily lives. However, data is strongly connected to the context, environment and community within which it is gathered, whether by human or machine. Thus, those who are local to the data - especially those who participate in its collection - can bring valuable insights and can themselves gain value from interpreting it. This paper offers an approach for designing interactive narrative games with civic datasets, which are designed to make data easier to interact with, discuss and share in informal settings. It describes the planning and execution of a public data exploration event based around these activities. This work will be of use to people who are designing civic interfaces to support participatory sensemaking from data, especially in informal settings.**

I. Introduction

Participatory Sensing (PS) - often also referred as citizen science, crowd-sensing or crowdsourcing - is a type of civic technology created to support civic participation and encourage exploration of the physical environment through the collection and interpretation of geospatial data about issues of shared concern [1]. Local communities bring important knowledge to both defining and finding solutions to the issues that concern them. They also have important local knowledge for making sense of the collected data. All data, especially geospatial data, are local and have complex attachments to the place from which they are derived [2], [3]. Thus, PS is powerful when used by communities to not only gather evidence, but to participate in the interpretation of data and to bring in their understanding of the local context. This in turn facilitates the general public to use data as evidence to advocate for change or to contribute to urban planning and decision-making.

Yet, even when they have a role in the collection of data, most people still struggle to participate in its interpretation and make sense of it in relation with their daily lives. Where once the majority of research around PS concentrated on the technologies to support data collection [4] and ways to engage volunteers in sensing initiatives [5], very recently the focus is beginning to shift to how this data can be made more usable by the general public, especially those who lack expertise in these types of activity [6].

This paper describes one approach to overcoming barriers to using data for urban change, by *using narrative principles and game mechanics to make geospatial data both easy and enjoyable to interact with and make sense of* [7]. We also consider how to support *participatory sense-making* from the data. In other words, how to foster the sharing of ideas amongst the community from which the data came (both those who contributed to its collection and those who did not) in order to draw on their different experiences, knowledge and goals. In this way, where PS harnesses the collective efforts of the community, so might the interpretation of the data also benefit from communal wisdom.

The setting for the work described here is an 'in the wild' study [8] of a civic data exhibition that was designed to engage the public with data that had been previously collected through a participatory sensing initiative, in which members of the local community had used an app to monitor their local area, by finding and marking abandoned or 'lost' items, invasive species and nice places to visit. In this event, the public would engage informally with the datasets, driven by their own interests and time constraints. Therefore, the data had to be presented in a way that would grab attention and be easy to engage with, even in a very limited time-frame. The approach we have taken is to frame a number of individual *data games* which are designed to familiarise people who are visiting the exhibition with the different datasets that have been collected, in a short time frame. These games are designed so that they adapt existing techniques for presenting data in narratively coherent ways [9] whilst introducing an element of interactivity to the data stories. We also explain how we designed the space to lead people on a journey of increasingly unconstrained exploration of data, increasing the agency that the visitors have in creating their own stories from data, while reducing narrative constraints.

The questions we aimed to answer during the design and running of this event, were:

1) Can narrative data games engage visitors with data in informal settings and within time constraints?
2) Does the design of the space and activities support participatory sense-making?

II. Related Work

Data on its own has little intrinsic value; the value is created by its use and by the meaning that people place on it in a given context. But data can be hard to find and use [10]. Thus, open data is often under-utilised, especially when considering its

use by citizens [11] and participatory sensing initiatives may find that people are not using the data as much as expected [6], [12], [13].

A range of research has looked at how to improve the data literacy of the general population [14], [15], [16], [17] to create a citizenship more able to understand and use data. Whilst definitions vary, data literacy is generally considered to be more than simply learning a set of technical skills, but also includes the ability to use data for civic empowerment [18], [19]. Data literacy skills support people to ask and answer questions from data, taking into account its possible limitations, the hidden assumptions behind the data and possible bias [20]. However, there are *limitations* to the data literacy approach. Many data literacy initiatives are designed for formal settings, such as in schools, where people have time to learn new concepts, tools and data literacy techniques. In a public setting where time is short a different strategy may be needed to help people who are not data professionals to make sense of large, or complex data sets. An alternative approach to building data literacy is to make the data itself easier to read and interact with [7]. Narrative principles are one approach to structuring new experiences in a way that helps people to make sense of them. The NAPA cards [21] are a set of design patterns that use narrative principles to structure data so that it can be presented in the form of *data comics*, made up of individual panels which they describe as 'a narrative element that captures a moment in the narration and focuses the readers attention'. Another approach to promote interaction with data is via games. Two fairly recent examples of this are the Datopolis board game designed by the Open Data Institute and the Datascape board game, designed to support learning how to match data and problems [22].

Based on the related work and given the goal of engaging people with data in a very short timeframe, our intentions were not to build data literacy but instead to reduce the need for data literacy skills by making the data easier to use and make sense of. We wanted to design data experiences that would a) entice the general public to explore the data collected during the monitoring initiative b) facilitate sensemaking of the data. We chose a *gamification* approach to engender the initial engagement and the use of *narrative* for sensemaking.

III. Designing Narrative Data Games

We utilised design patterns [9] to create individual narrative elements (or *panels*), presented as playing cards, that would frame the data in ways that highlighted the *key dimensions* (table 1). Instead of presenting cards in a fixed narrative, the goal was to design game mechanics in a way that would support people to find and tell stories from the data in an interactive manner. Narrative coherence was provided by fixing either the spatial or temporal aspect of the data within any given game and allowing the player to navigate only along one dimension at a time.

The following describes three different data games that were created based on the above principles. We also describe how they were presented in the context of the data exhibition.

Game	*Data Dimension*	*Design Template*
A	Type of data collected	Narrative (Expose)
B	Data collected across time	Temporal
C	Data collected across space	Faceting and Spatial

TABLE I
DATA DIMENSIONS



Fig. 1. Speed data-ing. The player must turn the last card and read quickly.

A: Speed data-ing is designed to help visitors to get to know the different collected datasets. Visitors have only 30 seconds to get to know each of the three main data types that were monitored, these being lost items, invasive species and nice places. A similar approach was used by [23] with personal data, revealing how this might change the relationship that people had towards data and the way they engaged with and questioned it. A short time period is used, as positive time based stress helps people to focus on most important aspects and as such helps productivity too [24]. Key information on the cards is a) the name and icon used to consistently identify the dataset in the platform and in the exhibition b) the locations where most instances of the data can be found c) the most likely time periods containing data. This game is very low in interactivity with the narrative presentation of data.

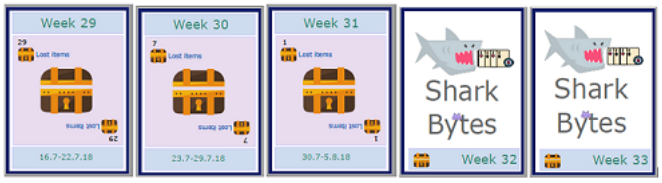


Fig. 2. Shark-bytes. The first 3 cards are shown face up, the player predicts the next two

B: Shark-bytes is based on the US television show Card Sharks (Play Your Cards Right in the UK). The game normally starts with a random playing card and the player has to guess if the subsequent card (facing downwards) would be higher or lower. In our version, each game focuses on one of the data types and each card shows one weeks worth of data. Players predict whether the value for that datatype went up, or down (in total) in each following week. A player 'wins' by getting to the end of the line of cards without error. We anticipated that players would discuss how they base their prediction, using their knowledge both of the city and also

knowledge of human behaviour. For example, by knowing popular holiday weeks, a player might predict lower values because the people participating in monitoring might be away. The aim of Shark-bytes was to support visitors in thinking about the importance of finding and analysing data trends - but also to cause reflection on how data is collected and how this may lead to errors or bias in data. To support the players memory, we also created a timeline of events in the city during the monitoring period captured by the playing cards and also an average of the weather conditions for each week. While Shark-Bytes can be played alone, we were also interested in the types of reasoning that people might bring to making a guess and whether people would collaborate and discuss how they made their predictions. This game maintains a fixed spatial aspect but supports the player to navigate the data across time. The player interaction slowly reveals the narrative, which is presented in a fixed order.

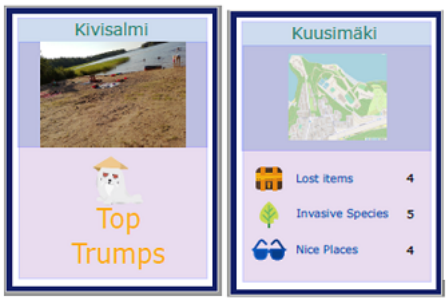


Fig. 3. Data Trumps, showing back of Kivisalmi card and front of Kuusimäki card, from which the player picks the category they think will 'win' the round

C: Data Trumps is based on the original Top Trumps card game. Data Trump cards relate to places in the city where the monitoring took place. Attributes are the three main data types that were monitored (lost items, invasive species and nice places) and the values are the total for the designated period (in this case, 5 weeks of monitoring). This game is designed to teach data comparison skills. Images taken by participants in the monitoring were used on the back of cards, to give a clue to the place (as even locals will not be familiar with all areas of the city). In addition, a map of the region was printed on the front of the card, along with the attributes and values. This is a competitive game, that is played in pairs. Each player takes a card and the player whose turn it is chooses an attribute that they think will have a higher value than their opponent. This game maintains a fixed temporal aspect but allows the player to navigate all three types of data across space. It has a much higher narrative interactivity than the previous two games, given the random nature of the cards and the player choice in choosing which attribute to find out about from the opponent.

A. Designing the visitor experience

As the card games were designed to be displayed and used in a public setting, it was important to consider the visitor's overall experience of visiting the exhibition space. Our approach was designed to lead the public through several distinct phases of interaction with data, starting with a high level of narrative constraint and low interactivity, leading to increasingly less constrained data exploration with lower narrative constraints imposed and higher agency for the visitor to find their own stories from the data. These stages are shown in Figure 2 and are now described in more detail, along with the activities within each.

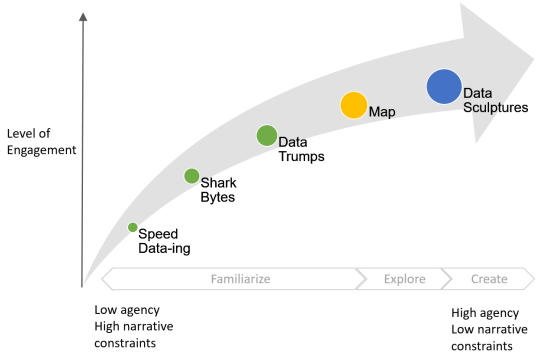


Fig. 4. Staged data exploration for increasing data engagement

Familiarise: The familiarisation stage consisted of the three interactive games; i) speed data-ing ii) shark-bytes and iii) top data-trumps, for visitors to the data exhibition to play. All three card games were presented in tangible form. There was an alternative digital version of the Data Trumps game that was made available on an Apple Mac Computer placed directly next to the physical version. The choice to present as cards that could be picked up and played was based on the premise that tangibility leads to deeper levels of engagement.

Explore: The exploration stage was designed to give visitors free access to the data, via a map-based interface presented on a large display over which they could easily collaborate. The map supported actions such as selecting:

1) One or more data sets to look at;
2) A geographical region (with panning and zooming);
3) A time period.

Create: The creation stage provided visitors with an artwork creation space to reflect a story they wanted to tell about what they had seen. Craft materials were provided, inspired by the data sculptures approach of D'Ignazio and Bhargava (https://databasic.io/en/culture/).

B. Guiding the visitor experience and capturing feedback

In the data exhibition, knowledge of museum curation strategies were used both to stage individual activities and to configure the layout to guide visitors along the intended path towards greater interactivity and storytelling from data. However, each activity was also designed to be self-contained and to make sense even if visited out of order.

Capturing feedback in a public place designed for fun can be problematic if it begins to detract from enjoyment of the overall experience of being in the place. As participation in the data

exhibition was based on semi-structured activities, evaluation could not happen in a controlled way. Also, attending the event and all engagement actions were entirely voluntary. Since intervening with questions or questionnaires could distract the attention from participation, the data capture was designed to be unobtrusive and directly fitted to event themes. There were two primary methods of collecting data. Firstly, directly from visitors and secondly through observation.

C. Visitor contributed

Each visitor was offered the opportunity to take a sheet which had a place for a stamp and a place to mark some feedback on a 5 point scale. The scale was presented as faces to make it appealing and easy to use, knowing in advance that there would be several schools visiting on that day. Visitors who collected stamps for all the activities were able to take a lollipop when they handed them in. Stamps with different colours and patterns were placed at each activity station. In this way, data could be collected about the order in which visitors did the activities.

D. Observation

The second method through which data was collected was by direct observation. This type of observation can be extremely challenging in a public space, especially when it is conducted in real time. The purpose of the observation data was to provide some additional support for the visitor contributed data. Observation data was captured directly onto maps of the exhibition layout. To measure engagement, we adapted a number of individual behaviour measures from Falk [25] to fit our context and then weighted them according to the levels of engagement within the Visitor Based Learning Framework (VBLF) developed by Barriault and Pearson [26]. Both scales were created to measure engagement within informal settings, such as museums.

Level	*Description*
1A	Doing activity in passing
1B	Doing activity somewhat completely
1C	Doing activity completely, but no further exploration
2	'Purposefully' watching others engaging in activity
3	Repeating the activity
4	Expressing positive emotional response
5	Referring to past experience while engaging in the activity

TABLE II
VBLF MEASURES

IV. RESULTS

A. Visitors

A total of 158 visitors (80 female, 78 male) indicated their presence at the exhibition by placing a LEGO piece onto a LEGO bar graph, used for capturing demographics. Breaking this down further, there were 28 visitors between 0-10 and 122 between 11-20, showing that children make up most of our sample. This is due to a number of schools bringing students to the event during the morning. Not everyone placed a piece of LEGO and this was particularly true of older visitors, so the true picture is not quite as out of balance as it appears here. Figure 5 shows pictures from the exhibition.

Level	*Description*	*Weight*
L1	Irrelevant behaviour (messing around, eating)	0
L2	Focusing attention on game cards, or map display	0
L4	Interacting randomly	0
L5	Visitors negotiate what to do or discuss the activity	1
L6	Visitors take turns in doing the activity	1
L7	Playing game as intended	1
L8	Two visitors do the activity at the same time	1
L9	Playing again, changing variables	2
L10	Smiling, pleased with exhibit	3
L11	Stronger signs of enjoyment	3
L12	Obvious signs of eagerness to participate	3
L13	Relating the activity to own life	4
L14	Can't tell	0

TABLE III
INDIVIDUAL BEHAVIOURS AND WEIGHTS



Fig. 5. Exhibition set up

The stamp sheets were analysed to evaluate the level of engagement visitors had with the different activities. 149 visitors completed a stamp sheet, meaning that they placed at least one stamp or one rating onto the sheet. While we do not have information about who completed these sheets, it was observed that the majority of these were from the under 20 demographics. 5 sheets were removed because the visitors had not understood where to get stamps and had visited other stalls within the carnival area and written feedback about those. 5 further sheets were removed as it was observed that the visitors had given feedback without engaging in the activities first. This left a total of 139 sheets for analysis. Analysis of these sheets revealed that 71 of the 139 visitors went to each activity in the intended order, 1-5, which is just over 51 percent of the sample. A further 5 visitors started in order and then did not complete remaining activities. In some of the remaining cases, there is only slight disruption in the ordering, for example visiting in the order 51234, or 15234. Others are completely disrupted (6 people visited in the order 53214) whilst 3 people

visited everything in reverse, going 54321. Some disruption can be explained by the busy times, when there were lots of visitors at once and people tended to visit the activity that had less people waiting. Only 16 people did not complete all 5 activities.

Each activity was rated between 1-5, where one was least enjoyment and 5 was maximum enjoyment. Not all stamped activities received a rating, leaving a number of missing values. Interestingly, it was usually the last rating that was left out, indicating a type of post-completion error where visitors were focused on the goal of collecting stamps and not on the rating that they gave. The activities were:

1) Speed Data-ing
2) Shark Bytes
3) Data Trumps
4) Map Interface, with questions
5) Data Sculptures

The mean ratings for the activities 1-5 (in order and with SD in brackets) were 4.2813 (1.03825), 4.3789 (0.67128), 4.4719 (0.61865), 4.4028 (.67989) and 4.5068 (.7092). The significance of the difference between ratings on the different activities was calculated using Wilcoxon Signed Ranks Test. A Bonferroni correction gives a significance level of 0.05. This showed a significant difference between the mean rating for activity 1 and activity 3 (Z = -3.050, p= .002), and also between activity 1 and activity 5 (z = 3.162, p = .002), but not between the other activities. To investigate this further, the same tests were conducted between the ratings given by people who encountered activities in intended order and those who did not. This showed a significant difference in ratings when people did activities 1 and 3 *in order* (z = 2.882, .004) but not when they did them in reverse, e.g. 3 and then 1. This result indicates that the ordering of at least these activities is important for fostering engagement.

Detailed observations of engagement were made by trained observers. In total, 39 people were observed, 20 female and 19 male. 8 were estimated to be over 20 and 31 were under 20. 16 participants visited the activities in the order 12345, a further 3 started in order but did not finish. Several of the observed visitors only completed one activity, which was the 4th activity (interactive map).

Each activity received an engagement rating between 0-4. This was calculated as a weighted average from the set of individual behaviours, adapting the equation used by Falk to obtain a weighted measure of engagement:

$$B_i = \frac{\sum_j b_j w_j}{n_i}$$

where:
b_j = the score for the behavioural category j
w_j = weighting for category b_j
n_i = total incidences of behaviour annotated for that activity

The mean ratings for the activities 1-5 (in order and with SD in brackets) were 0.6682 (0.65715), 0.8362 (0.72627), 1.1591 (0.67726), 0.9931 (.91204) and 1.1375 (.68704). The low mean scores are due to individual behaviours L1-L4, which in quantity can bring the overall engagement rating down. Given the age of the visitors and the fact that they were visiting in groups, there were many times when a visitor under observation was distracted from the activity. Another issue is that it was difficult for observers to identify L13 behaviour, where visitors related the experience to their own lives, without intruding on conversations.

However, the important finding is that the overall trend of the observation data is similar to that found from the self reported engagement and so for the purpose of being used to validate the accuracy of the self reported data, these findings do lend confidence. From both measures, activities 3 (Top Trumps) and 5 (Data Sculptures) were the most popular. The sample size of the observed data is quite small, therefore the more detailed comparison looking at order effects would not be reliable and so was omitted for this data.

V. Discussion

In answering the question *can narrative data games engage visitors with data in informal settings and within time constraints?* the self reported engagement measures, which were validated by formal and informal observation, showed a good level of engagement across all activities. The staged exploration did appear to support visitors to make sense of the data, this was shown through the effects of ordering on engagement. The explanation for these order effects could be that familiarisation with the datasets through the Speed Data-ing task prompted deeper types of engagement with the data via familiarisation when subsequently playing Data Trumps, rather than just focusing on surface elements of that task if coming to it without getting to know the data types first. However, Speed Data-ing itself was the least popular activity. This was the most narratively constrained activity, suggesting that the interactive elements were important in fostering engagement. In fact, there was a slight trend towards increasing engagement as agency increased, which was disrupted by the Data Trumps game. This was the last activity that had both game-like elements and structured narrative, lending credence to their importance.

Next we consider the question *How does the design of the space and activities support participatory sense-making?* While there was evidence of sense-making, it is harder to identify with any certainty whether this was individual or participatory, for reasons we now explain. We have taken the view that participatory sense-making is a form of collective intelligence that emerges from interaction between people, in this case around the collected data. While methods for formally measuring collective intelligence have been developed and used in laboratories and in online settings [27], there is no easy method available for capturing and measuring it at such a public event as the data exhibition. Another approach is to consider whether the conditions for collective intelligence were met. James Surowiecki [28] considers four conditions, which are: diversity, independence, utilization of decentralized knowledge, and effective aggregation of dispersed knowledge.

Within the data exhibition, there is evidence that people did try to relate the data to their daily lives, especially in the Shark-Bytes game. There may not have been diversity of age, but almost certainly a diversity of thinking across the school-age population of the city who visited with their schools. We can also assume that people would arrive at the exhibition with some level of independence of thought. From this, we might conjecture that we have conditions from which collective intelligence might emerge. But without mechanisms for formally measuring the extent to which this has caused dispersed knowledge to be aggregated, we have failed to recognise whether or not is was realized. This needs further and more formal study from which to draw concrete conclusions about what was occurring and why.

VI. Conclusions

This paper has described a data exploration event that has shown that gamified interactive data narratives can act as a first step towards supporting engagement and possibly participatory sense-making. Interfaces to data are important and should be appropriate to the setting. A data exhibition in a public place is better enjoyed through tangible and creative activities rather than through a screen. The novel contribution is in demonstrating how principles used to balance interactivity against narrative constraints for creating engagement can be applied to *data* storytelling as well as more conventional narrative forms. We also show how to use game elements to foster engagement and support narrative interaction, which can be particularly useful when there is only a short time in which people can engage with the data. Possible uses for this could be in empowering the general public to make sense of data and to bring their local and collective knowledge to its interpretation, for example in participatory sensing or citizen science activities.

VII. Acknowledgements

Aspects of this work have been supported by the Academy of Finland through the project 'Building the Civic Datascape'.

References

[1] S. Heggen, "Participatory sensing: repurposing a scientific tool for stem education," *interactions*, vol. 20, no. 1, pp. 18–21, 2013.

[2] Y. A. Loukissas, *All Data Are Local: Thinking Critically in a Data-Driven Society*. Mit Press, 2019.

[3] A. S. Taylor, S. Lindley, T. Regan, D. Sweeney, V. Vlachokyriakos, L. Grainger, and J. Lingel, "Data-in-place: Thinking through the relations between data and community," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 2863–2872.

[4] C. Golsteijn, S. Gallacher, L. Capra, and Y. Rogers, "Sens-Us: Designing Innovative Civic Technology for the Public Good," in *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 2016, pp. 39–49. [Online]. Available: http://dl.acm.org/citation.cfm?id=2901877

[5] I. Iacovides, C. Jennett, C. Cornish-Trestrail, and A. L. Cox, "Do games attract or sustain engagement in citizen science?: a study of volunteer motivations," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 1101–1106.

[6] S. Coulson, M. Woods, M. Scott, D. Hemment, and M. Balestrini, "Stop the noise! enhancing meaningfulness in participatory sensing with community level indicators," in *Proceedings of the 2018 on Designing Interactive Systems Conference 2018*. ACM, 2018, pp. 1183–1192.

[7] S. Kuznetsov and M. Tomitsch, "A study of urban heat: Understanding the challenges and opportunities for addressing wicked problems in hci," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 563.

[8] Y. Rogers and P. Marshall, "Research in the wild," *Synthesis Lectures on Human-Centered Informatics*, vol. 10, no. 3, pp. i–97, 2017.

[9] B. Bach, Z. Wang, M. Farinella, D. Murray-Rust, and N. Henry Riche, "Design patterns for data comics," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 38.

[10] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, "Benefits, adoption barriers and myths of open data and open government," *Information systems management*, vol. 29, no. 4, pp. 258–268, 2012.

[11] A. Wolff, A.-M. Valdez, M. Barker, S. Potter, D. Gooch, E. Giles, and J. Miles, "Engaging with the smart city through urban data games," in *Playable Cities*. Springer, 2017, pp. 47–66.

[12] K. Boehner and C. DiSalvo, "Data , Design and Civics : An Exploratory Study of Civic Tech," in *Proceedings of the 2016 ACM Conference on Human Factors in Computing Systems*. San Jose, CA: ACM, 2016.

[13] M. Balestrini, T. Diez, and F. Kresin, "From participatory sensing to making sense," in *Proceedings of the Workshop" Environmental Infrastructures and Platforms 2015–Infrastructures and Platforms for Environmental Crowd Sensing and Big*, 2015.

[14] S. Williams, E. Deahl, L. Rubel, and V. Lim, "City digits: Developing socially-grounded data literacy using digital tools," *Journal of Digital Media Literacy*, 2014.

[15] C. Anslow, J. Brosz, F. Maurer, and M. Boyes, "Datathons: an experience report of data hackathons for data science education," in *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 2016, pp. 615–620.

[16] C. D'Ignazio, "Creative data literacy," *Information Design Journal*, vol. 23, no. 1, pp. 6–18, 2017.

[17] R. Bhargava, R. Kadouaki, E. Bhargava, G. Castro, and C. D'Ignazio, "Data murals: Using the arts to build data literacy," *The Journal of Community Informatics*, vol. 12, no. 3, 2016.

[18] C. D'Ignazio and R. Bhargava, "Approaches to building big data literacy," in *Proceedings of the Bloomberg Data for Good Exchange Conference*, 2015.

[19] A. Wolff, D. Gooch, J. J. Cavero Montaner, U. Rashid, and G. Kortuem, "Creating an understanding of data literacy for a data-driven society," *Journal of Community Informatics*, vol. 12, no. 3, pp. In–press, 2017.

[20] S. Hautea, S. Dasgupta, and B. M. Hill, "Youth perspectives on critical data literacies," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 919–930.

[21] B. Bach, M. Stefaner, J. Boy, S. Drucker, L. Bartram, J. Wood, P. Ciuccarelli, Y. Engelhardt, U. Koeppen, and B. Tversky, " narrative design patterns for data-driven storytelling," in *Data-Driven Storytelling*. AK Peters/CRC Press, 2018, pp. 125–152.

[22] A. Wolff, M. Barker, and M. Petre, "Creating a datascape: a game to support communities in using open data," in *Proceedings of the 8th International Conference on Communities and Technologies*. ACM, 2017, pp. 135–138.

[23] C. Elsden, B. Nissen, A. Garbett, D. Chatting, D. Kirk, and J. Vines, "Metadating: exploring the romance and future of personal data," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 685–698.

[24] E. Salmela, A. Happonen, M. Hirvimäki, and I. Vimm, "Is time pressure an advantage or a disadvantage for front end innovation–case digital jewelry," *Journal of Innovation Management*, vol. 3, no. 4, pp. 42–69, 2016.

[25] J. H. Falk, "Time and behavior as predictors of learning," *Science Education*, vol. 67, no. 2, pp. 267–276, 1983.

[26] C. Barriault and D. Pearson, "Assessing exhibits for learning in science centers: A practical tool," *Visitor Studies*, vol. 13, no. 1, pp. 90–106, 2010.

[27] D. Engel, A. W. Woolley, I. Aggarwal, C. F. Chabris, M. Takahashi, K. Nemoto, C. Kaiser, Y. J. Kim, and T. W. Malone, "Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 2015, pp. 3769–3778.

[28] J. Surowiecki, "The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business," *Economies, Societies and Nations*, vol. 296, 2004.

Analytics Dashboards and User Behavior: Evidence from GitHub

J. Schilpp*, F. Pethig* and H. Hoehle*
* University of Mannheim, Germany
jan.schilpp@uni-mannheim.de

Abstract – **In open source software development it is essential to stimulate developers' contribution behavior in the absence of monetary incentives. But research on potential secondary effects of these stimuli, such as on contribution sentiment, remains scarce. Thus, we investigate how a promising external stimulus, an analytics dashboard, influences developers' number of contributions as well as associated message length and sentiment. Building on social comparison and goal setting theory we hypothesize that adopting the analytics dashboard increases developers' number of contributions but decreases contribution message length and sentiment. To test these hypotheses, we leverage granular data from a matched sample of 43,434 GitHub developers over two years who adopted an analytics dashboard on their personal developer page. Our difference-in-differences analyses reveal that adopting the analytics dashboard increases a developers' number of contributions and associated message length. However, after adopting the analytics dashboard users write more negative messages and this effect is primarily driven by users with low initial GitHub activity. These findings suggest that the performance increase through analytics dashboards comes at the cost of potentially harmful consequences for the developers or the platform. We discuss implications for theory and practice.**

Keywords – open; source; software; developer; analytics; dashboard; contribution; length; sentiment

I. INTRODUCTION

Nowadays, 90 percent of companies rely on open source software (OSS), making it indispensable for their system landscape [1]. However, OSS is predominantly developed by volunteers. Thus, it is a constant struggle for OSS platforms to motivate users' contributions.

Current research on OSS developers investigates the sources of developers' motivation and commitment when contributing to OSS [2]–[4] as well as potential ways how their contribution behavior can be altered [5]. Most studies investigate primary effects, i.e., how the number or quality of contributions is affected. But there is little research on potentially unintended secondary effects of external stimuli on OSS developers' contribution behavior, e.g., the contribution sentiment. This is a central research gap, in particular because insights from organizational software developers show that they might personally suffer from the idiosyncrasies of software development [6]–[8].

We aim to approach this gap by investigating how OSS developers are affected by a promising external stimulus, an analytics dashboard. The latter's popularity heavily increased in the past years given the high data availability [9] and similar stimuli were implemented on popular OSS platforms such as GitHub [10]. Furthermore, analytics dashboards have shown to have positive effects in other contexts, e.g., improved decision-making and performance [9]. Yet, the increased transparency associated with dashboards might have negative effects on the individual developer such as strategic effort allocation or stress. Therefore, we investigate the impact of analytics dashboards on OSS developers' number of contributions as well as the associated message length and sentiment. Overall, our goal is to answer the following research question: *How does an analytics dashboard influence developers' contribution behavior?*

We exploit a unique extension of the OSS development platform GitHub, which enables users to display a descriptive dashboard on their profile page[1]. This allows adopters to track their own activities and easily compare them with others. To understand how this impacts adopters' contribution behavior as well as what underlying mechanism might be we build on social comparison and goal setting theory. These theories provide rich insights into how the comparison with others, enabled by the analytics dashboard, impact developers' number of contributions, the contribution message length, and their sentiment. Thereby, this research is expected to make two important contributions to research on OSS developers. First, we investigate primary effects of analytics dashboards on OSS developers' contribution behavior. Second, we shed light on potentially unintended secondary effects, i.e., the contribution message length and the associated sentiment.

II. THEORETICAL BACKGROUND

In the following we first draw from research on OSS developers. Next, we introduce social comparison and goal setting theory to explain how analytics dashboards influence developer contribution behavior.

A. OSS Developers

The main difference between software developers in organizational and OSS settings is that software developers within organizations mainly contribute due to monetary incentives while OSS developers typically act on a voluntary basis. Therefore, OSS platforms must find alternative ways to influence their users' contribution behavior.

[1] https://github.com/anuraghazra/github-readme-stats

Correspondingly, studies on OSS developers focus on explaining why developers contribute to OSS. It is found that value and ideology congruence between a developer and the respective community is important [2], [11]. Moreover, OSS developers enjoy being an integral part of a community [3], [4]. Eventually, the contribution behavior of members in online communities can be altered by costly programs such as content contribution by the platform [5]. Overall, the identified determinants of OSS developer contribution behavior are mostly structural community features or costly efforts that cannot be easily influenced or implemented.

Nonetheless, it is important to incentivize OSS developers in a manner that sustainably alters their contribution behavior, i.e., considering primary effects such as the number of contributions as well as secondary effects such as the associated message length and sentiment [12]. One way this can be achieved is through an external stimulus on the respective OSS platform [10]. Given the high availability of data, descriptive analytics dashboards that display a developers' contributions and thereby allow to compare oneself with others are a possible path.

B. Social Comparison and Goal Setting Theory

We build on social comparison and goal setting theory to explain how analytics dashboards lead to a change in OSS developers' contribution behavior. Social comparison theory argues that people want to evaluate themselves, e.g., through comparison with others, especially people with greater abilities [13]. An analytics dashboard allows OSS developers to easily compare their number of contributions with peers. More specifically, analytics dashboard adopters transparently see that other users contribute more than oneself. Thus, we can expect that adopting the analytics dashboard incentivizes developers to contribute more to match more active users.

To explain the underlying mechanism, we borrow from goal setting theory. It argues that "people are motivated to strive towards goals" [14, p. 509] and regulate their effort to achieve their objectives accordingly. Reconsidering the adopters, it is their goal to match more active users' number of contributions. Consequently, the adopters direct effort toward improving the statistics displayed in the analytics dashboard. Based on this mechanism we propose our first hypothesis.

Hypothesis 1: *Adopting an analytics dashboard increases developers' number of contributions.*

The increased effort necessarily reduces the adopters' effort put into other things. More specifically, one can suspect that adopters direct less effort toward tasks that are not made visible by analytics dashboards. As the latter are of descriptive nature, they only display "what happened" [9, p. 1029] on a high level and properties of contributions such as the associated message length are not exhibited. Furthermore, previous research shows that more active developers might document less because they perceive changes to be clear just from the software code [15]. Thus, we hypothesize the following:

Hypothesis 2: *Adopting an analytics dashboard decreases developers' message length per contribution.*

Eventually, social comparison theory argues that the comparison with other people shapes one's self-perception [13]. That is, humans are found to have a more negative self-perception due to frequent comparisons with others because they consider other people to be better [16], [17]. Thus, we can assume that the social comparison enabled by the analytics dashboard negatively influences developer's self-perception. This should be visible in the sentiment of adopters' contributions. Our third hypothesis follows from this.

Hypothesis 3: *Adopting an analytics dashboard decreases developers' sentiment of contributions.*

III. Method

A. Research Setting

The three hypotheses are tested with data from OSS developers on GitHub. It is the most popular website for OSS development with more than 100 million users in February 2023 [18]. Leveraging the version control system Git, the platform allows developers to perform all necessary software development activities, i.e., the organization of single-person or collaborative projects, called repositories, through features like commits (i.e., changes to one or multiple files in a repository, typically software code accompanied by a short message that normally describes implemented changes), pull requests (i.e., proposed changes to a repository that are to be accepted by collaborators) as well as issues (i.e., a discussion thread related to a repository) [19]. Furthermore, GitHub offers social features such as user profiles and the possibility to follow other users.

The paper exploits the opportunity that some GitHub developers adopted a descriptive analytics dashboard on their profile page. The dashboard is depicted in Fig. 1. It displays the number of users' most important contributions on GitHub and ranks these users accordingly relative to other users.

B. Research Data

The analytics dashboard's initial launch was on July 9, 2020, and to implement it on a profile page users must add a string to their *Readme.md* in a commit to their own repository. We exploited this to determine if and when GitHub users adopted the analytics dashboard. Thereby, we identified 137,178 adopters of the analytics element in

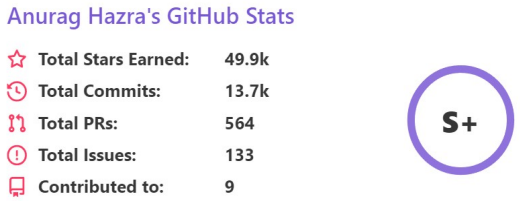


Figure 1. Analytics dashboard[2]

[2] Retrieved from https://github.com/anuraghazra/github-readme-stats, October 12, 2022

a two-year time window between July 9, 2020, and July 9, 2022. We then collected all public GitHub activities of these users between March 19, 2020, and October 29, 2022, from GhArchive[3].

We limit our analysis to commits for two reasons. First, commits are the most important type of contribution on GitHub as they are mostly software code, the main goal to be produced on an OSS platform. Second and given the enormous amount of data, this makes the analysis computationally more feasible and accessible for researchers and readers alike.

Furthermore, we filtered the users included in the analysis along multiple characteristics. First, we limited the sample to personal GitHub accounts of single developers by removing organizational accounts and bots. The latter is achieved by filtering users based on names and an unrealistic high number of contributions [10]. Second, we limited the sample to users that had at least ten commits in our timeframe to ensure that users were actively using GitHub. Third, we reduced the sample to developers who adopted the analytics dashboard and did not remove it within our study period. This allowed us to compare the effect of the dashboard across adopters. Fourth, we removed the commits users made to their own repository as this could bias our estimation because adopters must modify their repository to implement the dashboard.

We deployed a Bidirectional Encoder Representations from Transformers (BERT) machine learning model to determine the commit message sentiment [22]. A finetuned BERT model has shown to outperform other approaches to sentiment analysis in the software engineering domain [23]. Thus, we compile a ground truth dataset that is representative of the to be classified commit messages to finetune a BERT model for our data. This ground truth dataset is based on five publicly available datasets and encompasses 17,908 short messages from software development that are manually labelled as positive, negative, or neutral [24]–[28]. Our finetuned BERT model achieves an accuracy of 97%. To capture the nuances of a commit message we calculate the message's net sentiment by subtracting the predicted probability of having a negative message from the predicted probability of having a positive message.

Eventually, we aggregated the data on users' commits, i.e., the number of commits, the commit message length measured in number of characters, and net sentiment, on a monthly level to make the analysis computationally feasible and reduce noise in the data that stems from irregular contribution behavior. Thereby, we obtain a panel data set with one observation per user and month.

C. Research Method

We use a difference-in-differences (DiD) approach to compare the behavior of analytics dashboards adopters with the behavior of a suitable control group [29]. To account for the staggered adoption of the analytics dashboard we normalize the months relative to the treatment group's month of adoption.

As users deliberately choose to adopt the analytics dashboard, there is a high danger of self-selection into the treatment group of adopters. This might bias the results because users might have changed their behavior before adopting the analytics dashboard. For instance, users could become more active and then adopt the analytics dashboard to show their increased activity. To account for this issue, we deploy look ahead coarsened exact matching to identify a suitable control group [30]–[32]. This approach encompasses two elements that are best illustrated with an example of two GitHub users, A and B, with A adopting the analytics dashboard in month m and B in month m+10. We can expect these users to be similar along time-invariant unobservable characteristics such as their dedication to contribute to GitHub because both adopted the analytics dashboard eventually. Additionally, we match these users based on observable characteristics, i.e., the number of commits (log), mean commit message length, mean net sentiment per commit message, and tenure in months (log) with the help of coarsened exact matching. This matching is based on the activity of both users in month m-6 to account for the possibility that the users might have changed their contribution behavior before adopting the analytics dashboard [31]. Thus, our identification strategy aims to reduce differences between user A and B along unobservable as well as observable characteristics. In our setting, user B can be considered a suitable control group for user A between month m-5 and m+4. Consequently, the DiD analysis is based on data from month m-5 to m+4 (including m as part of the after-treatment period). We choose to analyze data from five months before and after the adoption to investigate the lasting impact of the analytics dashboard. Applying this matching procedure, we obtain a sample of 43,434 users (21,717 adopters and 21,717 future adopters) with 434,340 user-month observations[4].

IV. Results

A. Number of Contributions

To test whether the analytics dashboard influences developers' number of contributions (Hypothesis 1), we analyze how the number of commits changes before and after the adoption. Fig. 2. illustrates the mean number of commits for adopters and future adopters of the dashboard

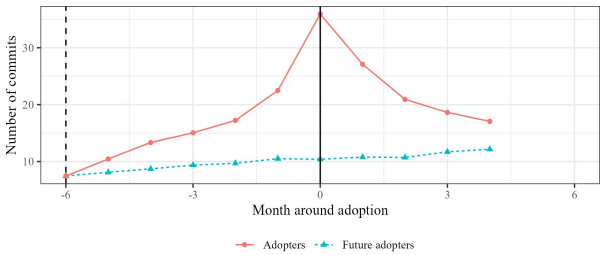


Figure 2. Number of commits around the adoption of the analytics dashboard

[3] The collection of this data was only feasible thanks to the Python Multiprocess module [20], [21].

[4] We repeat our main analysis with a matched sample of adopters and non-adopters as well with all adopters using an instrumental variable approach. We obtain similar results as for the look ahead matched sample.

around the adoption of the analytics dashboard. At the first glance, it is obvious that adopters commit most in the month of adoption. This intuitively makes sense as users must have been active in the adoption month to implement the dashboard, increasing the probability that they also contributed to other projects. Furthermore, the mean number of commits by adopters increases in the months leading up to the adoption. However, the graph suggests that the increased number of commits persists even after adoption as can be seen by the widening the gap between adopters and future adopters. That is, considering the same temporal distance from the adoption month, users commit more in the post- than in the respective pre-adoption period.

To further test this assumption, we estimate a Poisson DiD regression on the monthly number of commits with user and calendar month fixed effects. Using a Poisson model is the appropriate specification as the dependent variable follows a count data distribution. Moreover, we estimate separate regressions depending on the users' number of commits during the matching month to account for the idiosyncratic properties of user activity on GitHub. This contribution behavior follows a power law distribution where a minority of users is responsible for a majority of commits [33], [34]. The results are displayed in model 1 and 2 of Tab. I. The DiD coefficient (Adoption x After) is positive and significant for the low and high activity group which suggests that users commit more after adopting the analytics dashboard.

Consequently, we consider the first hypothesis to be supported and conclude that adopting the analytics dashboard increases developers' number of contributions.

B. *Message Length per Contribution*

As hypothesized, one might suspect that the higher number of commits comes at the cost that the adopters write shorter commit messages (Hypothesis 2). We investigate this proposition by considering the mean commit message length per commit. Fig. 3 displays the mean commit message length for adopters and future adopters around the adoption of the analytics dashboard. Surprisingly, we observe an increase in the mean commit message length per commit after the adoption of the analytics dashboard for the group of adopters.

Following the same structure as before, model 3 and 4 in Tab. I display the results of an ordinary least squared (OLS) regression on the mean commit message length. Again, we find both DiD estimators to be positive and significant suggesting that the commit message length increases after adopting the analytics dashboard.

From these findings, we conclude that adopting the analytics dashboard positively influences developers' message length per contribution and we reject the second hypothesis.

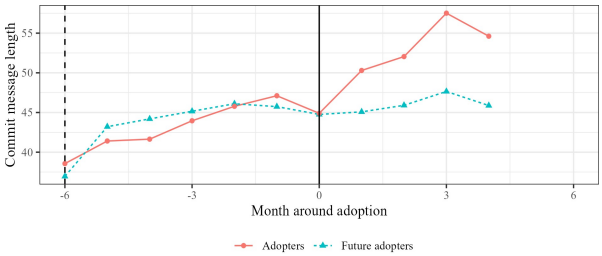


Figure 3. Mean commit message length around the adoption of the analytics dashboard

C. *Sentiment of Contributions*

Next, we examine the impact of the analytics dashboard on the commit message net sentiment as computed by our fine-tuned BERT model (Hypothesis 3). Model-free evidence in Fig. 4 suggests that the mean net sentiment of the commit messages among adopters is lowest during the month of adoption. This negativity partly diminishes in the following months but the commit message net sentiment for adopters stagnates on a visibly lower level than before the adoption and compared to the group of future adopters.

We complement this model-free evidence with an OLS regression on the mean net sentiment of the commit messages. The results are displayed in model 5 and 6 of Tab. I. Interestingly, the DiD estimator is negative and significant for low activity users but not for high activity users.

Overall, these findings provide partial support for our third hypothesis. We infer that adopting the analytics dashboard decreases the sentiment of contributions by users with low initial activity.

V. Discussion

Sustainably incentivizing OSS developer contribution behavior is essential but current research does not offer deeper insights into the impact of external stimuli such as analytics dashboards. To fill this important research gap, we build on social comparison and goal setting theory and argue that social comparison between developers impacts developers' number of contributions as well as associated message length and sentiment. Analyzing data from 43,434 developers, we find evidence for a positive effect of an analytics dashboard on developers' number of contributions and, surprisingly, associated message length. The latter might be driven by developers experiencing the need for longer documentation when committing more and collaborating with others [35]. Nonetheless, empirical evidence suggests a negative effect on contribution sentiment for initially less active developers. Two alternative mechanisms might explain the latter. First, less active adopters might experience higher stress or pressure mirrored in a more negative message sentiment. Second, these developers might adapt themselves to the tone of the platform.

We contribute to research on OSS developers by uncovering negative effects of altering contribution behavior as well as potential underlying mechanisms. For practitioners, our results imply that one should be careful

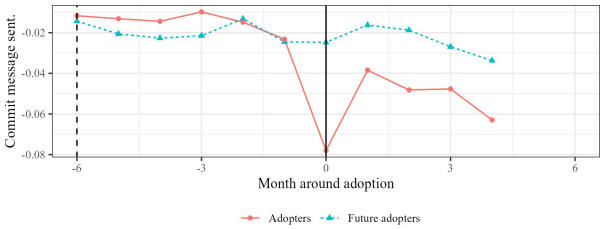


Figure 4. Mean commit message net sentiment around the adoption of the analytics dashboard

TABLE I. REGRESSIONS ON THE NUMBER OF COMMITS, MEAN COMMIT MESSAGE LENGTH, AND MEAN NET SENTIMENT

	Number of Commits (Poisson)		**Commit Message Length (OLS)**		**Net Sentiment (OLS)**	
	(1) Low	***(2) High***	***(3) Low***	***(4) High***	***(5) Low***	***(6) High***
After	0.4362*** (0.0370)	0.1101*** (0.0313)	0.3866 (0.5582)	1.4723 (0.9847)	0.0006 (0.0004)	0.0004 (0.0009)
Adoption x After	0.1043** (0.0444)	0.1875*** (0.0377)	4.0185*** (0.5428)	2.1188** (0.8865)	-0.0012** (0.0004)	-0.0000 (0.0008)
log(1+Tenure)	0.2644*** (0.0904)	-0.0384 (0.0938)	1.2579 (0.9629)	1.0388 (1.5167)	-0.0003 (0.0007)	-0.0032*** (0.0018)
Number of commits			0.0991*** (0.0263)	0.0584*** (0.0108)	-0.0000* (0.000)	-0.0000*** (0.0000)
User fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Calendar month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	284,960	104,030	328,700	105,640	328,700	105,640
Users	28,496	10,403	32,870	10,564	32,870	10,564
Adj. R^2			0.007	0.004	0.000	0.000

Note: The dependent variables are displayed in the first row. Users' activity level during the matching month according to a split along the 75th percentile are displayed in the second row. The number of users in column 1 and 2 differ because groups with zero observations are automatically dropped in a fixed effects poisson regression. Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

when incentivizing OSS developer contributions as this might come at a cost for the developers or the platform.

We are currently working on deepening these initial results to strengthen causal claims and to scrutinize underlying mechanisms. First, we are conducting additional robustness tests to test for parallel trends and the stable unit treatment value assumption. Second, we intend to conduct a randomized controlled trial with developers who are incentivized to adopt the analytics dashboard. Third, we are collecting additional qualitative evidence through interviews with adopters.

VI. CONCLUSION

Overall, our study suggests that the positive influence on contribution behavior triggered by analytics dashboards come at the cost of potentially harmful secondary consequences for the developers or the platform.

ACKNOWLEDGMENT

This work was supported by the University of Mannheim's Graduate School of Economic and Social Sciences. Furthermore, the authors acknowledge support by the state of Baden-Württemberg through bwHPC

REFERENCES

[1] GitHub Inc., "Octoverse 2022: The state of open source," *Octoverse 2022*, 2022. https://octoverse.github.com/ (accessed Feb. 02, 2023).

[2] L. M. Maruping, S. L. Daniel, and M. Cataldo, "Developer centrality and the impact of value congruence and incongruence on commitment and code contribution activity in open source software communities," *MIS Quarterly*, vol. 43, no. 3, pp. 951–976, Sep. 2019, doi: 10.25300/MISQ/2019/13928.

[3] G. von Krogh, S. Haefliger, S. Spaeth, and M. W. Wallin, "Carrots and rainbows: motivation and social practice in open source software development," *MIS Quarterly*, vol. 36, no. 2, pp. 649–676, Jun. 2012, doi: 10.2307/41703471.

[4] C. Zhang, J. Hahn, and P. De, "Research note—continued participation in online innovation communities: does community response matter equally for everyone?," *Information Systems Research*, vol. 24, no. 4, pp. 1112–1130, Dec. 2013, doi: 10.1287/isre.2013.0485.

[5] Y. Deng, J. Zheng, W. Khern-am-nuai, and K. Kannan, "More than the quantity: the value of editorial reviews for a user-generated content platform," *Management Science*, vol. 68, no. 9, pp. 6355–7064, Sep. 2022, doi: 10.1287/mnsc.2021.4238.

[6] S. Sarker, M. Ahuja, and S. Sarker, "Work–life conflict of globally distributed software development personnel: an empirical investigation using border theory," *Information Systems Research*, vol. 29, no. 1, pp. 103–126, Jan. 2018, doi: 10.1287/isre.2017.0734.

[7] V. Venkatesh, A. Rai, and L. M. Maruping, "Information Systems Projects and Individual Developer Outcomes: Role of Project Managers and Process Control," *Information Systems Research*, vol. 29, no. 1, pp. 127–148, Jan. 2018, doi: 10.1287/isre.2017.0723.

[8] V. Venkatesh, J. Y. L. Thong, F. K. Y. Chan, H. Hoehle, and K. Spohrer, "How agile software development methods reduce work exhaustion: Insights on role perceptions and organizational skills," *Information Systems Journal*, vol. 30, no. 4, pp. 733–761, Jul. 2020, doi: 10.1111/isj.12282.

[9] R. Berman and A. Israeli, "The value of descriptive analytics: evidence from online retailers," *Marketing Science*, vol. 41, no. 6, pp. 1074–1096, Nov. 2022, doi: 10.1287/mksc.2022.1352.

[10] L. Moldon, M. Strohmaier, and J. Wachs, "How gamification affects software developers: cautionary evidence from a natural experiment on GitHub," in *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, May 2021, pp. 549–561. doi: 10.1109/ICSE43902.2021.00058.

[11] L. M. Daniel, L. Maruping, M. Cataldo, and J. Herbsleb, “The impact of ideology misfit on open source software communities and companies,” *MIS Quarterly*, vol. 42, no. 4, pp. 1069–1096, Dec. 2018, doi: 10.25300/MISQ/2018/14242.
[12] S. Matook, G. Lee, and B. Fitzgerald, “Information systems development,” *MIS Quarterly Research Curations*, Dec. 2021. http://misq.org/research-curations (accessed Jan. 13, 2022).
[13] L. Festinger, “A theory of social comparison processes,” *Human Relations*, vol. 7, no. 2, pp. 117–140, May 1954, doi: 10.1177/001872675400700202.
[14] R. N. Landers, K. N. Bauer, and R. C. Callan, “Gamification of task performance with leaderboards: A goal setting experiment,” *Computers in Human Behavior*, vol. 71, pp. 508–515, Jun. 2017, doi: 10.1016/j.chb.2015.08.008.
[15] E. A. Alomar, A. Peruma, M. W. Mkaouer, C. D. Newman, and A. Ouni, “Behind the scenes: On the relationship between developer experience and refactoring,” *Journal of Software: Evolution and Process*; e2395, Oct. 2021, doi: 10.1002/smr.2395.
[16] S. Y. Lee, “How do people compare themselves with others on social network sites?: The case of Facebook,” *Computers in Human Behavior*, vol. 32, pp. 253–260, Mar. 2014, doi: 10.1016/j.chb.2013.12.009.
[17] D. A. de Vries and R. Kühne, “Facebook and self-perception: Individual susceptibility to negative social comparison on Facebook,” *Personality and Individual Differences*, vol. 86, pp. 217–221, Nov. 2015, doi: 10.1016/j.paid.2015.05.029.
[18] GitHub Inc., “GitHub About,” *GitHub*, Feb. 2023. https://github.com/about (accessed Feb. 02, 2023).
[19] GitHub Inc., “GitHub glossary,” *GitHub Docs*, Jan. 2022. https://docs.github.com (accessed Jan. 16, 2023).
[20] M. M. McKerns and M. Aivazis, “pathos: a framework for parallel graph management and execution in heterogeneous computing,” *pathos*, 2010. https://mmckerns.github.io/project/pathos/wiki.html (accessed Feb. 02, 2023).
[21] M. M. McKerns, L. Strand, T. Sullivan, A. Fang, and M. A. G. Aivazis, “Building a framework for predictive science,” in *Proceedings of the 10th Python in Science Conference (SciPy 2011)*, Austin, Texas, Jul. 2011. doi: 10.48550/arXiv.1202.1056.
[22] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding.” arXiv, May 2019. doi: 10.48550/arXiv.1810.04805.
[23] J. Wu, C. Ye, and H. Zhou, “BERT for sentiment classification in software engineering,” in *2021 International Conference on Service Science (ICSS)*, May 2021, pp. 115–121. doi: 10.1109/ICSS53362.2021.00026.
[24] F. Calefato, F. Lanubile, F. Maiorano, and N. Novielli, “Sentiment polarity detection for software development,” *Empirical Software Engineering*, vol. 23, no. 3, pp. 1352–1382, Jun. 2018, doi: 10.1007/s10664-017-9546-9.
[25] B. Lin, F. Zampetti, R. Oliveto, M. Di Penta, M. Lanza, and G. Bavota, “Two datasets for sentiment analysis in software engineering,” in *2018 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, Sep. 2018, pp. 712–712. doi: 10.1109/ICSME.2018.00084.
[26] B. Lin, F. Zampetti, G. Bavota, M. Di Penta, M. Lanza, and R. Oliveto, “Sentiment analysis for software engineering: how far can we go?,” in *Proceedings of the 40th International Conference on Software Engineering*, May 2018, pp. 94–104. doi: 10.1145/3180155.3180195.
[27] N. Novielli, F. Calefato, D. Dongiovanni, D. Girardi, and F. Lanubile, “Can we use se-specific sentiment analysis tools in a cross-platform setting?,” in *Proceedings of the 17th International Conference on Mining Software Repositories*, Sep. 2020, pp. 158–168. doi: 10.1145/3379597.3387446.
[28] G. Uddin and F. Khomh, “Automatic mining of opinions expressed about apis in stack overflow,” *IEEE Transactions on Software Engineering*, vol. 47, no. 3, pp. 522–559, Mar. 2021, doi: 10.1109/TSE.2019.2900245.
[29] J. D. Angrist and J.-S. Pischke, *Mostly harmless econometrics: An empiricist’s companion*. Princeton, NJ, USA: Princeton University Press, 2008. doi: 10.1515/9781400829828.
[30] R. Bapna, J. Ramaprasad, and A. Umyarov, “Monetizing freemium communities: Does paying for premium increase social engagement?,” *MIS Quarterly*, vol. 42, no. 3, pp. 719–735, Sep. 2018, doi: 10.25300/MISQ/2018/13592.
[31] J. Wang, G. Li, and K.-L. Hui, “Monetary incentives and knowledge spillover: Evidence from a natural experiment,” *Management Science*, vol. 68, no. 5, pp. 3549–3572, May 2022, doi: 10.1287/mnsc.2021.4048.
[32] M. Blackwell, S. Iacus, G. King, and G. Porro, “Cem: Coarsened exact matching in stata,” *The Stata Journal*, vol. 9, no. 4, pp. 524–546, Dec. 2009, doi: 10.1177/1536867X0900900402.
[33] S. L. Johnson, S. Faraj, and S. Kudaravalli, “Emergence of power laws in online communities: The role of social mechanisms and preferential attachment,” *MIS Quarterly*, vol. 38, no. 3, pp. 795–808, Sep. 2014.
[34] E. Kalliamvakou, G. Gousios, K. Blincoe, L. Singer, D. M. German, and D. Damian, “An in-depth study of the promises and perils of mining GitHub,” *Empirical Software Engineering*, vol. 21, no. 5, pp. 2035–2071, Oct. 2016, doi: 10.1007/s10664-015-9393-5.
[35] Y. Tian, Y. Zhang, K.-J. Stol, L. Jiang, and H. Liu, “What makes a good commit message?,” in *Proceedings of the 44th International Conference on Software Engineering*, Jul. 2022, pp. 2389–2401. doi: 10.1145/3510003.3510205

Using Hobbies as Proxy for Gamification Player Types

Semi Rantanen, Antti Knutas and Jussi Kasurinen
LUT University / School of Engineering Science (LENS), Lappeenranta, Finland
antti.knutas, jussi.kasurinen@lut.fi

***Abstract*—Many gamification approaches use a one size fits all approach. Different gamification player types and associated preferences have been proposed, such as the gamification player type Hexad. According to the Hexad, people prefer different gameful activities, such as exploration, social activities, or competition. Similarly in hobbies, different things motivate and interest different people, such as some finding board games interesting, while others like hiking and going on camping trips. We propose that a person's hobbies and interests provide an insight into what type of person they are, indicating what they might like, and could have a connection to gameful system preferences. In this paper, we explore the relationship between free-time activities and gamification player types, to enable earlier and faster detection of the player types and their preferences. The research was conducted by building a survey to measure the free-time activities of people and also measure the Hexad player type as defined by Marczewski and Tondello et al. Significant relationships were found between free-time activity variables "Active physical exerciser", "Audience of cultural events", and the player type Philanthropist. Overall, these results indicate that free-time activities can have correlations with player types and reported hobbies can be used as a proxy when selecting personalized gamification approaches.**

I. Introduction

Generally, the agreed definition of gamification is defined as the "intentional use of game elements for a gameful experience of non-game tasks and contexts" [1]. In practice, it's often about motivating users to interact with systems through the utilization of game design elements. Overall, often the goal of applying these elements is to improve human interaction with the system and improve intrinsic motivation to complete task at hand. Lately, a lot of research revolves around finding different contexts where to apply gamification and investigating possible benefits of applying game design elements in different applications. Gamification could play significant role in helping people find intrinsic motivators in contexts where it often might be difficult such as school or work environment [2], [3], [4], [5]. Utilization of game design elements in non-gaming contexts is about enhancing the experience that is being gamified. A hot topic within the research community has been finding proper use cases and even the right people that most benefit from these elements. The emergence of player types [6] that describe how and why people enjoy certain elements has led to questioning whether these player types connect to different aspects of our lives rather than just games. In other words, there could be hints in our everyday life into what sort of game design elements we enjoy and therefore would there be correlation between our free time non-gaming activities and the game design elements that we enjoy.

In this paper, we explore the identification process of player types for tailored gamification by using hobbies as the proxy. Overall, our research question is as follows: *Do the free-time activities of a person correlate with gamification player types and therefore have predictivity into the game design elements that they enjoy?*

To investigate the connection between hobbies and preferred enjoyment, this study is about connecting real-world activities with gamification player types and seeing if there are correlations where activities could predict the type of player a person is. For our study, we selected the gamification player type Hexad by Marczewski [7] and empirically validated by Tondello et al. [8], [9]. The study should help us understand how these player types connect with our behaviour in the real world and whether there are predictivity in free-time activities and enjoyed gamification elements. Overall, the goal is to extend our knowledge of gamification and how it is related to other activities that we enjoy and find outside factors that could explain our alignment with these player types.

II. Related research

In theory, the game design elements applied in non-entertainment applications should increase participation and intrinsic motivation to complete the task at hand, but research has shown that the application of gamification has varying results in different contexts [10], [2], [3], [4]. For example in education, the application of gamification elements seemed to have improved motivation and activity participation in a classroom setting but at the same time, these elements did not improve grades significantly [2]. Similarly, adding playfulness and robots to programming courses has failed [11], as has applying gamified learning platforms [3]. Its also important to understand that there are several ways to implement gamification, with the most game-element saturated systems practically being designed as serious games, whereas the minimal gamification can be something simple as showing a total number of likes for user's content, or percentage of users with less progress than the player, to motivate users to either catch up, or keep ahead [12]. In any case, the amount of gamified elements, and the type can vary in a wide spectrum, which means that quick identification of the most motivating elements that work,

and more importantly, those that don't work, for the target audience is important.

One important lesson to understand is that not all groups of people are motivated by gamification elements. Gamification for some students seemed to add another layer of complexity and did not seem to support their learning [3]. Conflicting results even within the same context seem to suggest that other factors could predict the successful outcome of gamification. The findings question whether context alone is enough to explain the successful outcome of gamification, or if other factors are more motivating, such as individual preferences. Often, applications of gamification take the one-fits-all approach, which depending on the application context may completely fail to account for differences in the way players want to experience their games [13]. Its also fairly common to blame misalignment with the user needs and gamification goals for the failures in product gamification [14].

The concept called adaptive gamification considers different playing motivations, and tries to explain different motivating factors by providing different gamer-profiles which try to explain why types of people enjoy certain gamification elements more [13]. Previous research indicates that a correlation between gamification feature categories and video game preferences exists. This means that people who enjoy certain types of games are more susceptible to enjoying similarly working gamification features [15], although empirical research validating and connecting player type to external attributes is quite rare. The connections between player types and personalities have been made and these comparisons try to answer how these traits play into gamification elements that people enjoy [16].

III. Research process

To study how users personal life and hobbies could correlate with the preferred gamification methods, a survey was designed to collect information. Overall, the research process consisted of three separate phases: survey construction, data collection, and data analysis, explained in detail in the following chapters.

A. Research methods

Surveys are used to gather generalized information about specific population. Different survey instruments are tools to conduct empirical research. Self-administered questionnaires allow people to answer survey by themselves and are ideal for quantitative research. Another tool is to interview people and often could be used to gather qualitative information [17], [18]. Purpose of these surveys is to gather data from smaller population and be able to extrapolate the results into larger population. To be able to do extrapolation, sufficient sampling is needed. Essentially, this means that there needs to be enough responses to allow data to be accurate and representative of a larger population. Data needs to be from heterogenous group of people that are part of so-called representatives. These are people who belong into target demographic that are being researched [19], [20]. In total, there are three objectives for surveys. Descriptive surveys are used to enable assertions about certain population. Explanatory survey tries to explain why certain population behaves the way they do. Finally, explorative aims to assure that important issues are not foreseen during pre-study phase [17]. The goal of this research is to find correlations between two distinct attributes within a population, namely free-time activities and gamification player types. This research is not trying to explain why people behave in certain way but rather assert that people who enjoy certain leisure time activities are also more likely to enjoy certain gamification elements. To collect data about this relationship, a survey was built.

The process of building the survey requires three distinct components. First, a basic demographic questionnaire that asks for basic information regarding respondent of the survey. This type of data is commonly collected to get an idea who is responding to the survey and in this case ensure that answers are from diverse group of people. Demographic data works as an additional analytical tool that can be used in analysis to see if different demographics deviate from overall results. Second, some sort of categorization of hobbies is needed. There are possibly thousands of different leisure time activities, and it is not feasible to allow people to choose too specifically. In nutshell, the more options you give to people the more data will be needed to confirm any hypotheses. Increased scope might not be possible due to budgeting and scheduling constraints. Third, a framework that can accurately predict person's player type accordingly to some player type typology is needed. This framework needs to be compatible with descriptive survey type and be simple enough that it can be self-administered. These three components provide data that is necessary to conduct this research.

B. Data collection

The survey was implemented as self-administered questionnaire hosted online. The survey was created using Webropol survey tool, which includes data management and has possibility to export data out of the service for analysis purposes. Goal of the research was to gather data from as wide group of people as possible and as such there is no specific target demographic specifically meant for this survey. While there was no specific target group, the questionnaire questions were in Finnish, limiting the respondents to those who know Finnish, albeit online translation tools do exist which means that submission is technically possible for non-speakers. In addition, there is not much need for detailed demographic data. For example, while it could be interesting to ask respondents for their occupation, getting that information might be redundant as its not directly related to the research question of this research. These are mainly used to get some basic understanding on who is filling out the survey and validate results against previous research. In total there are two points of data collected: age and gender. These were used to validate that not all answers came from same demographic group, but there was some diversity with survey respondents. Question about age had five categories: under 18, 18-30 years old, 31-50 years old,

51-70 years old, and those over 70. These age ranges are kept wide as more narrow ranging would likely not yield different results and this allows for more data entries per category. Question about respondent's gender included three options: male, female, and other.

Problem with categorizing different leisure time activities is that there are thousands of different hobbies and past time activities. Difficulty stems from finding general enough hobbies that apply to most respondents, while not being too general that all respondents share majority of the hobbies. To narrow down selection five distinct categories were chosen. These categories were collected from Finnish statistic agency Statistic Finland that collects data about different activities of Finnish people. Recreational data is one of the data groups collected by the agency. The five chosen categories were Active physical exerciser, Creative pursuit as a hobby, Audience of cultural event, Watching television programmes, and digital gaming. Even though these are very generalized activities there might be respondents who do not participate in any of these activities and therefore a sixth one was added titled "Something else." This section also included open-text field that could be filled out by the respondent. In case there would be enough responses with similar activities, these would be included in the analysis. Respondents were given option to choose any number of these activities, but at minimum one had to be selected. To help respondents find their hobbies from these categories, each option was given some example activities that are associated with that specific general category. For example, Creative pursuit as a hobby had examples that included cooking, dancing, or photographing. All given example activities are listed in Table I.

To determine respondent's player type, framework developed by Tondello et al. [8] in their research "The Gamification User Types Hexad Scale" was used. It is a 24-item survey that has four related questions for each six player types. This framework was built upon initial description of different player types that were first described in "Even Ninja Monkeys Like to Play: Unicorn Edition" by Marczewski [7]. Hexad scale framework was chosen mainly because there has been follow-up research that support validity of this framework [21], [9]. Another factor was the simplicity of the implementation. Questions regarding player types can be understood by people who might not be aware of gamification or phenomena related to it. Survey respondents will answer how well each question relates to them in a 7-point Likert scale. Total points from each category's four questions are summed together and the one with most points is respondent's primary player type. Interestingly, Tondello et al. [8] also noticed that there were some medium correlations between different player types. For example, if primary player type is Philanthropist, they are more likely to have more points in Socialiser category as well. Player type determining questions and translations are available in the online appendix [1]. The authors suggest that each item is presented without identifying the player type category they are related to. Another suggestion was to present these questions in randomized order. The approach described here is the recommended way to determine player type from this survey by the authors.

[1]Anonymized for review purposes

IV. Findings

In total, there were 59 responses to the survey. Demographic results are as follows: 70% of the respondents were in the 18-30 age group. This young adult age group had the most respondents for both male and female segments. Majority of the respondents were men, 40 out of 59 respondents. Rest were women and zero participants reported other as their gender.

First, factor analysis was used to reduce the data into clusters of variables and discover any underlying latent variables that cannot be measured directly. The first step of the process was to submit data into IBM's statistics software SPSS which is used to perform factor analysis. Then, linear regression was applied to confirm the strength of correlations between player types. Due to combining both exploratory and confirmatory methods, the entire statistical process should be considered exploratory and tentative.

A. Factor analysis

Kaiser-Meyer-Olkin (KMO) test indicates the suitability of data for factor analysis [22], [23]. Values less than 0.5 indicate that sampling is not adequate. With values between 0.5 and 0.6, researcher should evaluate the study and make the decision. For this dataset, a KMO test result outcome with a value of .563 is passable in certain situations. Bartlett's Test of Sphericity is used to validate the adequacy of the correlation matrix [23]. The test is significant at $p < .001$ which indicates that there are high correlations between some of the variables. Any significance value below $p < .05$ indicates that the dataset is suitable for factor analysis [23]. With these results, we decided to proceed, since the study is initial and exploratory.

Varimax with Kaiser Normalization was used to rotate resulting matrix and was used to ensure no correlation between factors. Factors found by the analysis are likely to have some level of correlation, but for the sake of analysis simplicity, orthogonal rotation method which eliminates correlations between factors was used. Factor analysis can provide as many factors as there are variables and therefore system is needed to only select the most meaningful factors out of the group. In this case, factors were selected based on their eigenvalues, and those with value over 1 were selected. Rotated factor matrix is presented in Table II.

After identifying the factors, the next step is to give names based on the underlying trait they try to explain. Factor 1 was named "Affection to sports and wellbeing" and factor 2 as "Appreciation of arts". Factor 1 describes a person that enjoys doing as well watching. At face value, it is difficult to find some underlying human characteristics that these variables explain. Therefore, it was chosen that these types of people are likely to enjoy sports by doing and watching as this is the common factor between the free-time activity two variables. Philanthropist is interested in well-being of others

TABLE I
LEISURE ACTIVITY CATEGORIZATION AND TRANSLATIONS

Name of English activity	Name of Finnish activity	Example activities
Active physical exerciser	Liikunnan harrastus	-
Creative pursuit as a hobby	Itsetekeminen	Writing, Dancing, Photographing, reading
Audience of cultural events	Kulttuuri	Visiting theater, cultural events, museums
Digital gaming	Pelaaminen	Video games, board games
Watching television programmes	Viihde	Watching sports, watching television, Netflix

TABLE II
ROTATED FACTOR MATRIX

Type	Factor			
	1	2	3	4
Philanthropist	0.685		0.380	
Active physical exerciser	0.553			
Watching television programmes	0.551			
Free Spirit		0.615		0.397
Distruptor		0.497		
Creative pursuit as a hobby		0.445		
Audience of cultural events		0.442		
Socialiser			0.739	
Player			0.465	
Achiever				0.775

and themselves and motivated by purpose and meaning [7]. As such, additional description of "and well-being" was given to factor 1 to highlight purpose and meaning aspect of the factor. Naming factor 2 was a little easier as these people are fond of cultural activities and like to partake in some themselves. Player types Free Spirit and Distruptor have already been found to positively correlate [8], [21], [24]. It is likely that this type of person enjoys artistic expression while being interested in finding their own path.

B. Linear regression analysis

The purpose of factor analysis was to cluster data into factors that at first glance display relationships between player types and free-time activities. This exploratory method was used to give advice on what free-time activity variables to test against player type variables. The next step is to validate the results by conducting multiple linear regression analysis where closer look will be taken into how these free-time activities impacted player type scores. Free-time activity variables that constituted the two relevant factors were used as independent variables to try and explain dependent player type variables.

Multiple linear regression analysis findings are presented in Table III. MLR1 variables are part of the first multiple linear regression equation and MLR2 variables are part of the second.

There were in total two relationships that had statistical significance between independent free-time activity variables and dependent player types. Both were present in the Philanthropist player type. First run, which was based on the factor "Affection of sports and well-being" contained both "Active physical exerciser" and "Watching television programmes" as independent variables. The results of this indicate that "Active physical exerciser" has statistically significant relationship and moderate strength relationship with Philanthropist.

What this analysis confirms is that there are some indications of free-time activities correlating with player types. Especially Philanthropists in this sample seemed to have two free-time activity variables that had a positive correlation with the player type. Simply put, a person who selects any of the "Active physical exerciser" or "Audience of cultural events" is likely to score higher in the Philanthropist type. As was mentioned in the previous chapter, these relationships are not strong, but they are significant. In this sample, other tests did not signify statistical significance.

V. DISCUSSION

The results of this study indicate that free-time activities can play minor role in person's player type alignment. This can have implications for the practical design of tailored gamification, because it is rarely possible for game designers or advertisers to be able to measure the player type of a person directly. Inference through correlation from other forms of data could help these people to tailor game mechanics to those who appreciate them the most, without having to resort to scientifically rigorous but difficult to answer surveys.

The main finding of our paper continues the line of research of gamification user analysis, which includes analysis of player behaviour Loria et al. [25] and evaluation of gamification preferences from player types [15]. This paper extends the state of the art by introducing a new potential correlation for inferring gamification preferences.

A second finding is that our study confirms earlier analysis on player types, where our analysis of Free Spirit and Distruptor correlation matches earlier studies by [8], [21], [24].

TABLE III
MULTIPLE LINEAR REGRESSION OUTCOMES BETWEEN FREE-TIME ACTIVITY VARIABLES AS INDEPENDENT VARIABLES AND PLAYER TYPE SCORES AS DEPENDENT

	Philantropist	Socialiser	Free Spirit	Achiever	Distruptor	Player
MLR1: Active physical exerciser	4.988 ($p < 0.01$; R^2=.177)	1.905		1.696		
MLR1: Watching television programmes	1.638					
MLR2: Audience of cultural event	1.835 ($p < 0.05$; R^2=.0.61)		1.007		0.899	

Due to limited dataset and not using an independent dataset for statistical inference (regression analysis), this paper can be considered an exploratory paper only and not hypothesis testing one. As future resesearch, we recommend further data collection and confirmatory research with an independent dataset, guided by these initial findings. Additionally, research contraints limited the size of the dataset and the survey research is based on a convenience sample. Howeever, despite these limitations, the findings are promising.

VI. CONCLUSIONS

In this paper, we explored the possibility of identifying player types by using hobbies as a proxy. We found that free-time activities can play a minor role, especially with the philanthropist player type. Our initial findings suggest that there is potential in designing tailored gamification by using personal preferences, such as hobbies, as a proxy for gamification player types.

We extend the state of the art similarly by pointing out the potential for this avenue of research. Our findings are exploratory, but these findings on player types guide the path for future hypothesis testing -based confirmatory research.

REFERENCES

[1] K. Seaborn and D. I. Fels, "Gamification in theory and action: A survey," *International Journal of human-computer studies*, vol. 74, pp. 14–31, 2015.

[2] G. Barata, S. Gama, J. Jorge, and D. Goncalves, "Engaging engineering students with gamification," in *2013 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES*, pp. 1–8,.

[3] K. Berkling and C. Thomas, "Gamification of a software engineering course and a detailed analysis of the factors that lead to it's failure," in *2013 International Conference on Interactive Collaborative Learning (ICL*, pp. 525–30,.

[4] J. Majuri, J. Koivisto, and J. Hamari, "Gamification of education and learning: A review of empirical literature," in *Proceedings of the 2nd International GamiFIN 2018 Conference*, p. 9.

[5] S. Reyssier, S. Hallifax, A. Serna, J.-C. Marty, S. Simonian, and E. Lavoué, "The impact of game elements on learner motivation: Influence of initial motivation and player profile," *IEEE Transactions on Learning Technologies*, vol. 15, no. 1, p. 42–54.

[6] A. C. T. Klock, I. Gasparini, M. S. Pimenta, and J. Hamari, "Tailored gamification: A review of literature," *International Journal of Human-Computer Studies*, vol. 144, p. 102495, 2020.

[7] A. Marczewski, *Even Ninja Monkeys Like to Play: Unicorn Edition*. Independently Published.

[8] G. F. Tondello, R. R. Wehbe, L. Diamond, M. Busch, A. Marczewski, and L. E. Nacke, "The gamification user types hexad scale," in *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, 229–43. CHI PLAY '16*. New York, NY, USA: Association for Computing Machinery.

[9] G. F. Tondello, A. Mora, A. Marczewski, and L. E. Nacke, "Empirical validation of the gamification user types hexad scale in english and spanish," *International Journal of Human-Computer Studies, Strengthening gamification studies: critical challenges and new opportunities*, vol. 127, p. 95–111.

[10] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work? – a literature review of empirical studies on gamification," in *2014 47th Hawaii International Conference on System Sciences*, pp. 3025–34,.

[11] B. S. Fagin and L. Merkle, "Quantitative analysis of the effects of robots on introductory computer science education," *Journal on Educational Resources in Computing (JERIC)*, vol. 2, no. 4, pp. 2–es, 2002.

[12] K. Werbach and D. Hunter, *For the win, revised and updated edition: The power of gamification and game thinking in business, education, government, and social impact*. University of Pennsylvania Press, 2020.

[13] I. Rodríguez, A. Puig, and A. Rodríguez, "Towards adaptive gamification: A method using dynamic player profile and a case study," *Applied Sciences*, vol. 12, no. 1, p. 486.

[14] K. Robson, K. Plangger, J. H. Kietzmann, I. McCarthy, and L. Pitt, "Game on: Engaging customers and employees through gamification," *Business horizons*, vol. 59, no. 1, pp. 29–36, 2016.

[15] L. Hassan, J. Rantalainen, N. Xi, H. Pirkkalainen, and J. Hamari, "The relationship between player types and gamification feature preferences," in *Proceedings of the 4th International GamiFIN Conference*, 2020.

[16] L. S. Ferro, S. P. Walz, and S. Greuter, "Towards personalised, gamified systems: An investigation into game design, personality and player typologies," in *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death, 1–6. IE '13*. New York, NY, USA: Association for Computing Machinery.

[17] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Berlin, Heidelberg: Springer.

[18] A. Fink, *The Survey Handbook*. SAGE.

[19] J. A. Krosnick, "Survey research," *Annual Review of Psychology*, vol. 50, no. 1, p. 537.

[20] K. KELLEY, B. CLARK, V. BROWN, and J. SITZIA, "Good practice in the conduct and reporting of survey research," *International Journal for Quality in Health Care*, vol. 15, no. 3, p. 261–66.

[21] J. Krath and H. F. Korflesch, "Player types and game element preferences: Investigating the relationship with the gamification user types hexad scale," in *HCI in Games: Experience Design and Game Mechanics*, ser. Lecture Notes in Computer Science, X. Fang, Ed. Cham: Springer International Publishing, p. 219–38.

[22] C. D. Dziuban and E. C. Shirkey, "When is a correlation matrix appropriate for factor analysis? some decision rules." *Psychological bulletin*, vol. 81, no. 6, p. 358, 1974.

[23] N. Shrestha, "Factor analysis as a tool for survey analysis," *American Journal of Applied Mathematics and Statistics*, vol. 9, no. 1, p. 4–11.

[24] A. Manzano-León, P. Camacho-Lazarraga, M. A. Guerrero-Puerta, L. Guerrero-Puerta, A. Alias, R. Trigueros, and J. M. Aguilar-Parra, "Adaptation and validation of the scale of types of users in gamification with the spanish adolescent population," *International Journal of Environmental Research and Public Health*, vol. 17, no. 11, p. 4157.

[25] E. Loria and A. Marconi, "Player types and player behaviors: analyzing correlations in an on-the-field gamified system," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 2018, pp. 531–538.

An Eye-Tracking Solution Using Consumer Grade Webcams for Potential Concussion Diagnosis and Evaluation

Nicholas Caporusso*, Gabriel Sanders*, Brett Thaman*, Elaina Hall*
* Northern Kentucky University, Highland Heights, Kentucky, USA
{caporusson1,sandersg1,thamanb1,halle14}@nku.edu

***Abstract*—In the past decade, eye-tracking technology has been increasingly utilized in healthcare settings to assess an individual's cognitive status under various circumstances to evaluate the presence of brain injuries and monitor the evolution of temporary or permanent cognitive illnesses. More specifically, emerging research has confirmed that disconjugate eye movements can be utilized as a predictor of a concussion and potentially return-to-play protocols.**
Currently, most healthcare applications use eye-tracking devices requiring more detailed signals at a higher sampling rate and with better camera resolution compared to the eye-tracking technology used in the consumer market. Unfortunately, this limits the availability of this type of diagnostic system to clinical settings and prevents its use in situations where early diagnosis is crucial (e.g., during a football game, where 300,000 concussions are reported yearly).
In this paper, we introduce a solution for potentially diagnosing and treating concussions based on images acquired with inexpensive and more available camera devices such as webcams, and we detail a performance evaluation study of a popular image segmentation and object detection machine learning model (i.e., MediaPipe Facemesh and Iris) applied to the acquisition and analysis of eye-related signals (e.g., eye movements and blinking) for healthcare applications.

Keywords—component, formatting, style, styling, insert (key words)

I. INTRODUCTION

A concussion is considered a mild traumatic brain injury that is caused by a bump or jolt to either the head or the body resulting in rapid brain movements inside the skull. Of the 1.4 million annual traumatic brain injuries (TBI) in the United States, 50,000 persons die and another 235,000 will require hospital admission while 50% of incidents go unreported or undetected [1]. More specifically, it is estimated that more than one-third of concussions go undiagnosed in athletes [2]. The lack of proper diagnosis of concussion may result in serious long-term consequences or risk of coma or death [3].

To avoid catastrophic injury and long-term negative health consequences due to undiagnosed concussions, accessible and affordable technology should be utilized and better developed for consumer and medical-grade consumption. Technology advances in the last decade have found eye-tracking can be used to assess brain injury and it can be utilized as a high-sensitivity biomarker compared to other types of subjective exams and assessments to determine concussion symptoms that are commonly administered by physicians [4] [5] [6].

Current eye-tracking solutions are expensive, are not accessible to many areas and people, and are only available via a healthcare provider. A frequent issue with non-computerized concussion assessments is the duration of the test and the likelihood of false-positive diagnoses or data to support a return to play or work [7]. For example, baseline testing batteries with reliable day-to-day change indices have merit, but invalid changes in a gait test or dexterity test can result in false-positive diagnoses due to faulty tests or interpretations [7] [8] [9]. Reliable and noticeable change indices can become even more challenging to interpret when implementing a multitest battery for concussion assessment [9]. There is emerging data and technology to support eye-tracking as a means for rapid diagnosis of concussions from youth to professional athletes [6] [10]. Eye-tracking can be a valuable tool to monitor subtle daily changes in eye movements that may indicate improvements in cognitive function resulting in well-calculated quantitative assessments for return to play, as opposed to subjective physician assessments. Neurocognitive testing is a method to assess brain function by implementing various noninvasive tests such as paper and pencil or computerized neurocognitive tests (CNT) [8] [11] [12]. More specifically, computerized neurocognitive testing can provide valuable concussion-related information only if a baseline test has been administered to serve as a reference for each individual [11]. In addition to a well-designed baseline test, these assessments can take a minimum of 15-30 minutes, need to be supervised by a clinical professional, and must be administered in an environment with no distractions with specific lighting that can also be administered on a sideline or in a locker room [13].

To avoid the complexity of cognitive assessments, rapid eye-tracking can be done more quickly, would not require a significant cognitive component, and could be used with minimal supervision as a computer program could guide the process. Rapid and reliable diagnosis of a concussion can occur on a field, in a gymnasium, or even at a residential home with portable and computerized technology designed to assess eye conjugation [5] [4] [14]. Accessibility to initial concussion diagnosis can further improve short and long-term health care by providing physicians with additional information regarding the extent of disconjugate eye movements [3] [10] [9] [15]. In addition to diagnosis,

accessible eye-tracking technology can improve recovery assessments for patients as data can be collected daily to observe and record improvements in eye movements [10]. The information can then be relayed to a physician for a more accurate determination of return to play or return to work.

II. RELATED WORK

Several companies have developed commercial eye-tracking solutions based on expensive dedicated hardware that can be utilized to aid the diagnosis and treatment of concussions. The Oculogica EyeBox is approved by the Food and Drug Administration (FDA) to detect concussions, does not require a baseline assessment, and can be completed in 4-5 minutes [4] [5]. Additionally, due to its cost and acquisition processes, clinical-grade eye-tracking technology is only available at certain clinics or medical facilities, which limits the possibility of reaching people in rural areas and results in underserved populations.

Nevertheless, eye tracking is an established technology that utilizes sensors such as infrared or RGB cameras to track the pupil and its movement, it can be accomplished with consumer-grade web cameras. Unfortunately, given the cost, eye-tracking is not widely available, especially in high schools and underserved locations, where most traumatic brain injuries and concussions are undiagnosed. Most eye-tracking applications involve the use of dedicated external devices that range in price from $250 to $10,000 USD and can only be accessed via a healthcare provider. As previously mentioned, eye-tracking is used as a tool to diagnose concussions in a non-invasive, accessible, and quantitative way. Eye-tracking can also collect and obtain novel information about the severity of concussions that is unavailable from traditional, noncomputerized examinations alone [4]. When compared to the Sports Concussion Assessment tool version 3 (SCAT3) [15], eye-tracking produced results with both specificity and sensitivity that could be useful as a baseline-free means of diagnosis, as well as having the potential to be the standard for the detection of brain injury [5]. In patients with head injuries, eye-tracking revealed a correlation between the severity of the oculomotor disruption and the severity of the symptoms displayed [5]. Results showed that both the horizontal and vertical conjugacy of both eyes can be compromised due to a traumatic brain injury that was not present in non-trauma patients [5]. Applying eye-tracking to young athletes also yielded different results than commonly used clinical assessments, suggesting that it may be an objective addition to concussion examinations [10] [15]. With the correlation between eye movement and concussive symptoms, eye-tracking technology can be used to detect concussions with reasonable accuracy [4] [5] [6]. Studies have shown that there is a need for objective concussion examinations outside of a hospital environment, as many concussions go undiagnosed due to a lack of testing accessibility as well as errors in objective concussion examinations [2] [3] [14]. The abnormalities in the eye movements of concussed patients can be very subtle, which can make them difficult to subjectively detect even when diagnostic examinations are performed by trained professionals on cooperating patients [4] [5]. Common causes of misdiagnoses in patients examined by medical professionals include lack of transparency from patients, miscommunication between patients and professionals, and bias on the part of medical professionals and their equipment or the type of test [2] [4] [8] [15] [16]. Undiagnosed concussions can lead to health complications, such as an increased likelihood of comorbid events [17]. Eye-tracking can be used to prevent the consequences of undiagnosed concussions by making accurate concussion technology readily available to a wider range of patients and reducing qualitative errors.

III. EYE TRACKING USING MEDIAPIPE

In the last years, several groups explored viable alternatives to dedicated eye-tracking devices using traditional RGB cameras and image processing algorithms to support the wider use of eye-tracking technology in healthcare. A potential solution resides with recent developments in camera technology, suggesting that standard webcams can be utilized to realize accurate eye tracking at sampling rates and with resolutions that would be appropriate for healthcare applications. However, standard consumer cameras such as the webcams that are commonly incorporated in notebooks or used for video conferencing, have a lower acquisition frame rate (i.e., 24-30 frames per second), provide only a monocular view of the subject's face, and process the acquired images via software. Moreover, as described in [18], eye-tracking based on RGB cameras involves a sophisticated workflow consisting of multiple steps, that is, detecting and tracking the user's face, locating their eyes, identifying the position of their pupils, and subsequently estimating the coordinates of the point of the screen where the user is looking, in real-time. Indeed, each step involves a different problem, specific image processing techniques, and computational concerns and requirements, which significantly affect performance (i.e., speed), accuracy (i.e., the distance between the predicted gaze location and actual target), and reliability (i.e., within-subject and cross-subject accuracy).

More recently, MediaPipe [19] [20] has been introduced as an efficient solution for real-time computer vision tasks. It comprises a set of models optimized for specific image segmentation and object recognition problems such as face, hand, posture, and iris detection. Moreover, MediaPipe Iris [21] focuses on identifying a subject's pupils, tracking their movement, and estimating their distance from the camera.

MediaPipe FaceMesh predicts the facial topology of the user based on a still image and represents it as an array of 468 key points, that is, vertexes each associated with a set of three-dimensional coordinates that enable describing the components of the face (e.g., silhouette, nose tip, cheeks, and iris). Each property provides useful information about the positioning and alignment of the subject with respect to the camera, including face posture and rotation.

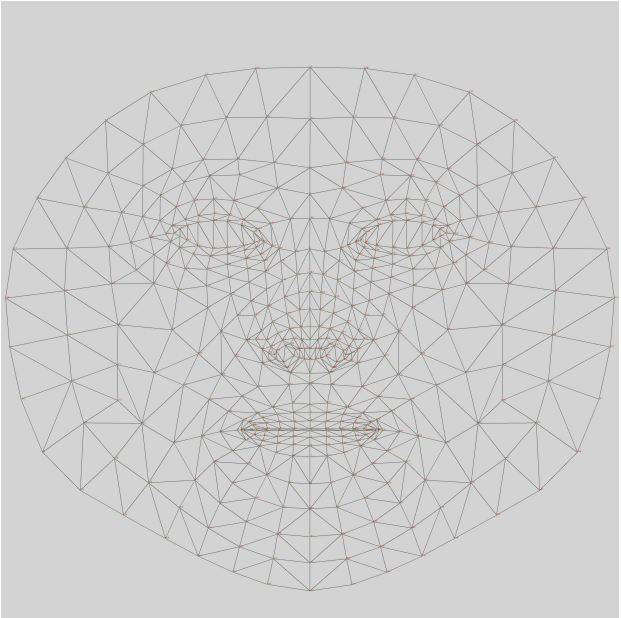


Fig. 1: The 478 landmarks utilized by MediaPipe FaceMesh and Iris to estimate the subject's facial geometry and the coordinates of their pupils.

In addition, MediaPipe Iris estimates the position of the user's pupils and represents them as an array of five landmarks for each eye (i.e., top, bottom, left, right, and center vertexes).
The opportunities offered by these models are remarkable, as MediaPipe FaceMesh and Iris are designed to be incorporated into cross-platform mobile applications and websites, which makes the library one of the promising candidates for exploring the feasibility of healthcare applications of eye-tracking based on standard RGB cameras. However, their clinical use has not been fully explored yet.

IV. Performance Evaluation

The purpose of our work is to enable the development of healthcare applications based on the use of ML models such as MediaPipe FaceMesh and Iris. The ultimate goal of the proposed project is to develop a concussion assessment solution that would increase the affordability and availability of exams. If computer-grade webcams and Machine Learning models can predict eye features with sufficient accuracy, they could be advantageous in reassessing the subjects' conditions and managing their rehabilitation without expensive equipment. To this end, it is crucial to assess the model's performance in effectively detecting and tracking eye movements with a signal-to-noise ratio (SNR) that is appropriate for supporting diagnosing and monitoring cognitive conditions and brain injuries.
Therefore, we realized a performance evaluation study aimed at understanding the key factors that affect signal quality in terms of accuracy and reliability when using MediaPipe FaceMesh and Iris to estimate the facial geometry and position of the pupils from the images acquired with a standard webcam.



Fig. 2: One frame collected by the experimental software.

A. Materials and methods

In our data acquisition, we utilized five different notebooks incorporating 720p (1280 x 720 pixels) webcams with standard lenses. Also, we recorded data using an external 720p webcam mounting a wide-angle lens to compare the performance and evaluate any differences. We developed a dedicated data collection software that estimated the distance of the subject based on their facial landmarks by using the focal length of the camera. Also, the software recorded the location of the center of each pupil using a coordinate system relative to the size of the frame. Figure 3 shows a frame captured by the experimental software: the distance and frames-per-second are shown at the top left corner, whereas the bottom right corner shows an enhanced view of the right (red) and left (green) pupils. At the beginning of each trial, the software asked the subject to position themselves at a distance of 10 centimeters from the camera. Subsequently, every five seconds, the software asked the subject to move one distance step (i.e., 1 cm) away from the camera while keeping their eyes aligned with the lens. The trial stopped when the subject was at a distance of 110 cm from the camera. This is because a range of 10-110 cm is the typical distance at which individuals either stay when they use their computers or can be asked to maintain for the duration of a quick assessment (i.e., shorter end of the range). We did not use any physical guidance or constraint to keep the subject aligned with the camera's lens as they moved away from it. However, data recording was supervised to prevent the subject from moving too fast or significantly deviating from the ideal alignment with the camera. As each camera recorded at a rate of 30 frames per second, we collected 90-150 frames for each distance step from each subject, which we considered sufficient to evaluate SNR and signal dispersion.
Five healthy subjects participated in our data collection. A smaller pool of subjects was sufficient because the purpose of our study was to evaluate the performance of the system in terms of signal accuracy and reliability of webcam-based eye tracking. Each subject realized four trials with each camera, for a total of 20 trials per subject. As a result,

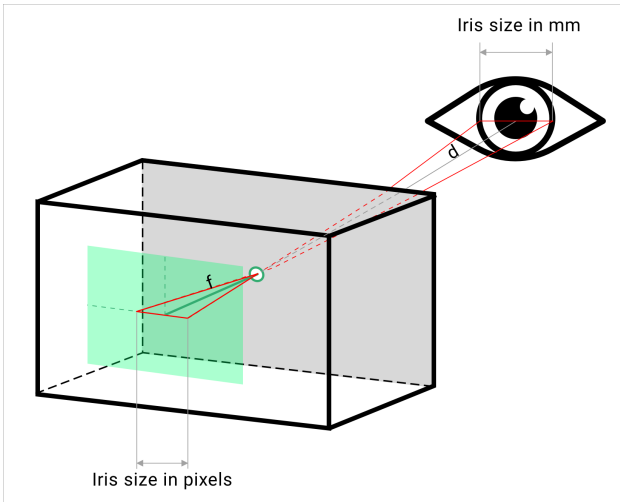


Fig. 3: Calculation of the distance of the subject based on the size of the iris and the focal length using similar triangles.

we acquired a total of 288,000 frames. After removing signal artifacts such as blinking, we had a total of 261,375 data points for each pupil (i.e., approximately 2600 frames for each distance step).

B. Results

Our data show that signal dispersion linearly increases with distance. The ideal SNR is achieved when the user is at a range of 10-50 centimeters from the webcam, where the dispersion is less than 4%, that is, less than 50 pixels on a 720p frame. Also, SNR degrades at a faster pace on the Y axis. This could be caused by vertical adjustments in the pose of subjects as they were moving away from the camera. Our findings are represented in Figures 4 and 5, which show the signal dispersion for the left and right eye, respectively.

Furthermore, by analyzing the variance of the signal dispersion, we can confirm that distances in the range of 10-50 centimeters are ideal for achieving more reliable SNR, with dispersion affecting the signal by less than 1 pixel, on average. Conversely, as shown by Figures 6 and 7, the signal shows wider dispersion when the subject is positioned more than 50 centimeters away from the camera. The dispersion variance, which is shown in pixels because of its small values, is contained within a range of 3.5 pixels on a 720pm webcam even at distances in the range of 100-110 centimeters, which shows the robustness of the facial geometry pipeline in accurately tracking the pupils.

Data from the wide-angle webcam show a statistically significant difference in the calculation of the distance of the subject from the camera, which is due to a different field of view (FOV) of the lens. However, our data did not show any statistically significant difference regarding SNR.

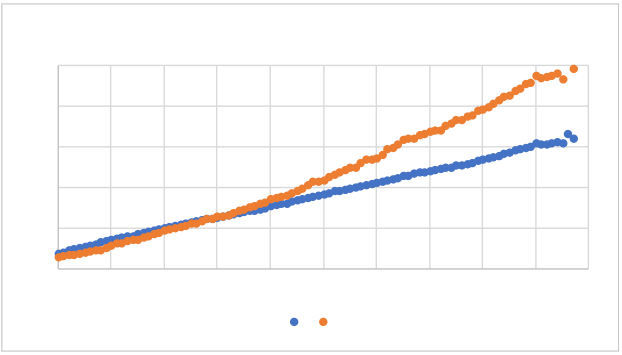


Fig. 4: Horizontal and vertical signal dispersion (percentage relative to the frame size) on the left eye.

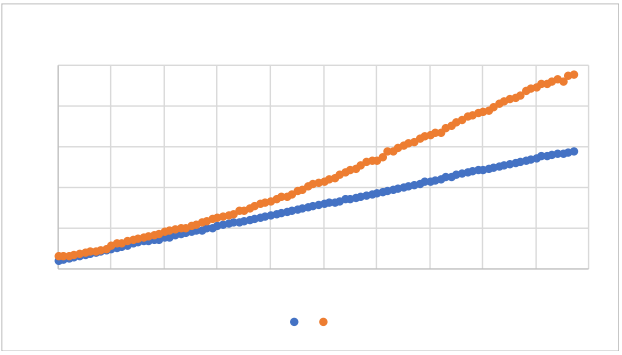


Fig. 5: Horizontal and vertical signal dispersion (percentage relative to the frame size) on the right eye.

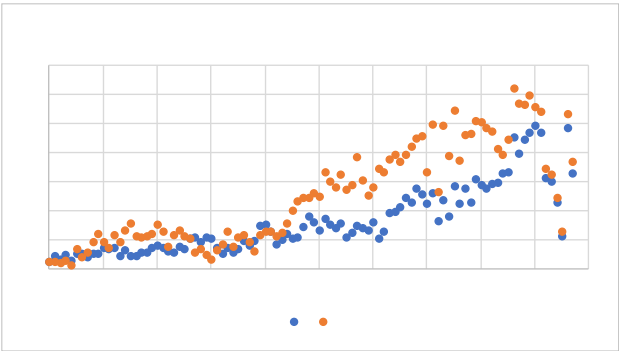


Fig. 6: Horizontal and vertical dispersion variance (in pixels) on the left eye.

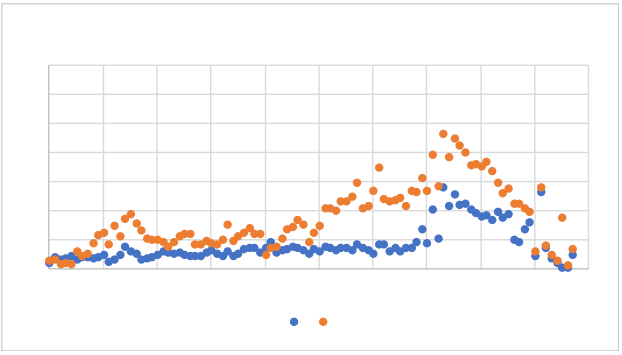


Fig. 7: Horizontal and vertical dispersion variance (in pixels) on the right eye.

C. Discussion

Our data show that MediaPipe FaceMesh and Iris are viable tools for potentially developing healthcare applications that leverage standard webcams for eye-tracking. Another contribution of our paper is that the subject's distance from the camera is a key factor. Based on our findings, the signal acquired by the system has sufficient accuracy and reliability when the subject is positioned at a distance ranging from 10 and 50 centimeters from the camera. Conversely, the SNR degrades significantly when the subject is more than 50 cm away from the camera, where noise values as high as 10% would not support the use of the system for the assessments of concussions. However, increased noise dispersion reported when the subject is more than 50 cm from the camera could also be caused by the subjects' movements as they were asked to distance themselves from the webcam.

Our experimental setup was designed to address the worst-case scenario, that is, circumstances in which subjects would not keep a fixed distance or hold their position in front of the camera. We were interested in evaluating the accuracy and reliability of MediaPipe FaceMesh and Iris at different distances from the camera and when the signal is affected by small movements of the subject to mimic real-world scenarios in which users would take their assessments at home, without supervision, or without adhering to recording protocols strictly. Even in this case, the system's performance supports tracking eye movements, including disjointed eye movements, fixations, and blinking, with sufficient accuracy. However, a higher precision assessment would require the subject to maintain their position at a specific distance (or within a defined range) from the camera, which can be obtained by implementing constraints based on the facial geometry estimated by MediaPipe FaceMesh. In addition to increasing the reliability of the signal acquired from the user, smoothing and denoising techniques could further enhance the signal. Moreover, we discovered that standard and wide-angle lenses can be used for data acquisition, provided that the focal length can be set as a configuration parameter to obtain more consistent results.

V. Conclusion

The goal of our work is to increase the availability of solutions for assessing and monitoring cognitive conditions and brain injuries such as concussions through the development of user-friendly technology that can be utilized more conveniently outside healthcare settings and without the direct supervision of clinical personnel.

In this paper, we presented a study in which we evaluated the performance of a Machine Learning model that could potentially be employed for estimating eye movements with a standard webcam. Specifically, we utilized MediaPipe FaceMesh and Iris, a cross-platform object segmentation and landmark detection library. Its advantage relies on leveraging standard webcams for estimating the facial geometry of the subject and tracking their eyes instead of requiring dedicated eye-tracking devices. However, as clinical assessments require high-quality signals, in our work we primarily focused on analyzing the accuracy and reliability of the data acquired by MediaPipe FaceMesh and Iris.

To this end, we realized a preliminary experiment in which we identified the ideal conditions for obtaining high-quality signals from MediaPipe FaceMesh and Iris. Our findings show that when the user is positioned at a range of 10-50 cm from the camera, the signal obtained by the webcam and processed by MediaPipe FaceMesh is sufficiently accurate and reliable for achieving a detailed assessment of eye movements, including disjointed eye movements, as discussed in previous studies. However, our data show that the SNR obtained by the system becomes very poor when the subject is located more than 50 cm away from the camera. Nevertheless, the information produced by the system is enough for adding constraints that provide the subject with guidance on the ideal positioning for signal collection.

In our future work, we will compare the performance of the proposed system with clinical-grade devices and realize control studies involving healthy subjects and individuals affected by brain injuries or cognitive conditions to assess the validity of our system and its potential uses in healthcare settings.

Acknowledgment

This project was realized with the financial support of NKU's Institute for Health Innovation and thanks to Kentucky Commercialization Ventures' IMPACT grant.

References

[1] L. M. Gessel, S. K. Fields, C. L. Collins, R. W. Dick, and R. D. Comstock, "Concussions among united states high school and collegiate athletes," *Journal of athletic training*, vol. 42, no. 4, p. 495, 2007.

[2] W. P. Meehan III, R. C. Mannix, M. J. O'Brien, and M. W. Collins, "The prevalence of undiagnosed concussions in athletes," *Clinical journal of sport medicine: official journal of the Canadian Academy of Sport Medicine*, vol. 23, no. 5, p. 339, 2013.

[3] M. Putukian, "The acute symptoms of sport-related concussion: diagnosis and on-field management," *Clinics in sports medicine*, vol. 30, no. 1, pp. 49–61, 2011.

[4] U. Samadani, M. Li, M. Qian, E. Laska, R. Ritlop, R. Kolecki, M. Reyes, L. Altomare, J. Y. Sone, A. Adem *et al.*, "Sensitivity and specificity of an eye movement tracking-based biomarker for concussion," *Concussion*, vol. 1, no. 1, 2016.

[5] U. Samadani, R. Ritlop, M. Reyes, E. Nehrbass, M. Li, E. Lamm, J. Schneider, D. Shimunov, M. Sava, R. Kolecki *et al.*, "Eye tracking detects disconjugate eye movements associated with structural traumatic brain injury and concussion," *Journal of neurotrauma*, vol. 32, no. 8, pp. 548–556, 2015.

[6] A. B. Zahid, M. E. Hubbard, J. Lockyer, O. Podolak, V. M. Dammavalam, M. Grady, M. Nance, M. Scheiman, U. Samadani, and C. L. Master, "Eye tracking as a biomarker for concussion in children," *Clinical journal of sport medicine*, vol. 30, no. 5, pp. 433–443, 2020.

[7] L. D. Nelson, "False-positive rates of reliable change indices for concussion test batteries: a monte carlo simulation," *Journal of Athletic Training*, vol. 50, no. 12, pp. 1319–1322, 2015.

[8] J. Bogner, L. Brenner, B. Kurowski, J. Malec, D. R. Howell, D. L. Hunt, J. R. Oldham, S. E. Aaron, W. P. Meehan, and C. O. Tan, "Postconcussion exercise volume associations with depression, anxiety, and dizziness symptoms, and postural stability: preliminary

findings," *Journal of head trauma rehabilitation*, vol. 37, no. 4, pp. 249–257, 2022.

[9] J. R. Oldham, D. R. Howell, C. A. Knight, J. R. Crenshaw, and T. A. Buckley, "Single-task and dual-task tandem gait performance across clinical concussion milestones in collegiate student-athletes," *Clinical journal of sport medicine*, vol. 31, no. 6, pp. e392–e397, 2021.

[10] J. R. Oldham, C. L. Master, G. A. Walker, W. P. Meehan III, and D. R. Howell, "The association between baseline eye tracking performance and common concussion assessments in high school football players," *Optometry and vision science: official publication of the American Academy of Optometry*, vol. 98, no. 7, p. 826, 2021.

[11] W. P. Meehan III, P. d'Hemecourt, C. L. Collins, A. M. Taylor, and R. D. Comstock, "Computerized neurocognitive testing for the management of sport-related concussions," *Pediatrics*, vol. 129, no. 1, pp. 38–44, 2012.

[12] J. E. Resch, M. W. Schneider, and C. M. Cullum, "The test-retest reliability of three computerized neurocognitive tests used in the assessment of sport concussion," *International journal of psychophysiology*, vol. 132, pp. 31–38, 2018.

[13] R. C. Lynall, K. G. Laudner, J. P. Mihalik, and J. M. Stanek, "Concussion-assessment and-management techniques used by athletic trainers," *Journal of athletic training*, vol. 48, no. 6, pp. 844–850, 2013.

[14] M. Putukian, "Clinical evaluation of the concussed athlete: a view from the sideline," *Journal of athletic training*, vol. 52, no. 3, pp. 236–244, 2017.

[15] J. R. Oldham, W. P. Meehan III, and D. R. Howell, "Impaired eye tracking is associated with symptom severity but not dynamic postural control in adolescents following concussion," *Journal of sport and health science*, vol. 10, no. 2, pp. 138–144, 2021.

[16] S. V. Ettari, E. Roden, V. Ahuja, and U. Samadani, "Oculogica: An eye-catching innovation in health care and the privacy implications of artificial intelligence and machine learning in diagnostics for the human brain," *SMU Sci. & Tech. L. Rev.*, vol. 25, p. 23, 2022.

[17] T. McAllister and M. McCrea, "Long-term cognitive and neuropsychiatric consequences of repetitive concussion and head-impact exposure," *Journal of athletic training*, vol. 52, no. 3, pp. 309–317, 2017.

[18] M. S. Mounica, M. Manvita, C. Jyotsna, and J. Amudha, "Low cost eye gaze tracker using web camera," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2019, pp. 79–85.

[19] I. Grishchenko and V. Bazarevsky, "Mediapipe holistic—simultaneous face, hand and pose prediction, on device," *Retrieved June*, vol. 15, p. 2021, 2020.

[20] Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann, "Real-time facial surface geometry from monocular video on mobile gpus," *arXiv preprint arXiv:1907.06724*, 2019.

[21] A. Ablavatski, A. Vakunov, I. Grishchenko, K. Raveendran, and M. Zhdanovich, "Real-time pupil tracking from monocular video for digital puppetry," *arXiv preprint arXiv:2006.11341*, 2020.

Do I Trust My Medical Professional on a Virtual Consultation? Biosensors, Human Computer Interaction and Trust-Building

Bohan Ye*, Chad Milewicz*, Sudesh Mujumdar*, Mohammed Khayum*
* University of Southern Indiana, Evansville, United States
mkhayum@usi.edu, cmmilewicz@usi.edu, smujumda@usi.edu, bye@usi.edu

***Abstract*— Telehealth usage in 2021 was 38-times higher compared to the average pre-pandemic level [1]. Patient interaction with a medical professional mediated by a screen – whether it's a computer screen, tablet screen or a mobile phone screen - has implications for trust-building [2], which, in turn, impinges on health outcomes [3]. In this paper, given that over 55% of communication is non-verbal [4], we focus on understanding how screen-mediated non-verbal cues – eye-movements, hand gestures and facial expressions - influence feelings of trust. This understanding is sought through the design of an experiment where we vary non-verbal cues in standardized images/videos and examine the effects on feeling of trust as reflected not only in self-reports by subjects but also in more objective data captured by biosensors.**

Keywords*— *biosensors, human computer interaction, trust-building

I. Introduction

The pandemic has resulted in an explosive growth in screen mediated (virtual) meetings or interactions. As the linguistic anthropologist, Elizabeth L. Keating [5] notes, in moving from face-to-face interactions to screen-mediated interactions, we lose key pieces of non-verbal information. In drawing on Goffman [6, 7], Keating further notes that "we lose the ability to observe others observing us, because we often don't know where their gaze is directed. (They're looking at their screens, not us)." This ability is of particular import since about two-thirds of the social meaning of a conversation emanates from non-verbal communication. The dissipation of important pieces of non-verbal cues impinges on feelings of trust – especially in the patient-healthcare-provider relationship. As noted by Kimball and Morgan [8], "…on a very basic subconscious level, we can see a person in front of us but cannot locate them in space. This strain on our proprioception increases our stress levels — and, for many patients, visiting a doctor is already a stressful proposition." This 'low quality' communication is not conducive to trust-formation and adversely impacts receptivity to healthcare provider's advice and treatment, as pointed out by [3]; "Trust is typically associated with high quality communication and interaction, which facilitates disclosure by the patient, enables the practitioner to encourage necessary behaviour changes and may permit the patient greater autonomy in decision-making about treatment."

This paper focuses on how the most elemental non-verbal cue, viz., just a portrait-icon of the person one is interacting with or will prospectively interact with, impacts trust-levels, when this interaction is mediated through a computer screen. A closely related paper [9] examines the differential impacts of nonverbal cues in a web-based chat (with only "semantic information" being the source of non-verbal cues and no visual information about the person one is interacting with is available) versus face-to-interactions, where the face-to-face interactions occur between human-human and human-robot. The human-human interactions help uncover non-verbal cues that are linked to trust, where trust is measured in similar fashion to the standard trust-game in experimental economics [10]. The robot is then programmed with these non-verbal cues for the human-robot interactions. Results show higher trust-levels for relevant non-verbal cues in human-human and human-robot interactions compared to web-based chat interactions. Our paper is different not only in that it examines the effect of a visual non-verbal cue in screen-mediated interactions on trust-levels, but it also seeks to understand the underlying biological mechanisms with the help of biosensor technologies.

II. Hypothesis

This paper tests how an elemental non-verbal cue – such as a visual of the person one is interacting with – impacts trust-levels. So, our hypothesis is that a scenario where one can see the picture of the person one is interacting with will produce higher trust than the scenario where there is no picture as a result of a positive emotion.

III. Testing Methodology and Apparatus

To test our hypothesis the methodology of experimental economics is employed with the following key features:

a) Controlled environment in a laboratory setting

b) Randomized control trial (RCT) approach to test hypotheses

c) Monetary incentives as the underlying anchor

Interaction between individuals (subjects) is mediated by a computer screen and enabled with the commonly-used software; z-tree [11]. Eye-tracking devices, cameras, and the face-reading software [12] for detecting emotion complete our testing apparatus.

IV. Design of Experiment

We leverage the standard trust game a la Berg et.al. [10] to measure trust. At the beginning of the experiment,

participants are randomly assigned to the role of either trustor or trustee. The experiment will consist of multiple sessions. Each session is comprised of two periods. At the beginning of period one, the trustor receives an endowment of US$10 and makes a decision on how much of the endowment to transfer to the trustee. The experimenter triples the amount that is transferred to the trustee. In period two, the trustee makes a decision on how much of the tripled amount is to be transferred back to the trustor. Trustor and trustee are made aware of their respective net amounts from the session. The actual transfers of trustors constitute a measure of trust, and the reaction of the trustees, a measure of trustworthiness. This two-period game (with payoffs) is diagrammatically presented in figure 1.

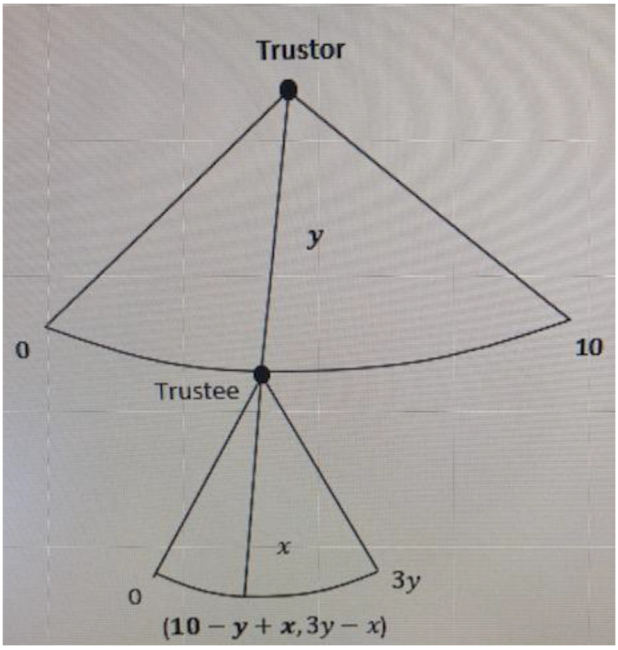


Figure 1. Two-period game (with payoffs)

Participants who have been assigned to the role of trustor complete the task in the BiometriX Discovery Lab of the Romain College of Business on a computer monitor screen. The trustor's console is equipped with eye-tracking, camera, and face-reader technologies. Participants who have been assigned to the role of trustee complete the task in the Griffin Experimental Economics Lab on a computer monitor screen. All participants' screen-mediated interactions are enabled using the z-Tree program through intranet.

The Control condition follows the standard trust game where the trustor and trustee remain anonymous are not identified by any portrait-icon

In the Treatment condition, in period 1, the trustee selects one out of four facial (gender-neutral) emojis/portrait-icons expressing, respectively, emotions of happiness, anger, neutrality and sadness. This portrait-icon is revealed to the trustor prior to the trustor making the decision to send money to the trustee. The trustor then makes the decision on how much money to send back to the trustee. In period 2, the trustor decides how much of the tripled amount to return to the trustee.

After the completion of the two periods, the trustee is given a survey asking them whether they know the trustor they were matched with prior to both of them participating in the experiment. A demographic survey is also be given to all participants.

During the entire experiment, eye-trackers are activated on the trustor's side in the BiometriX lab to determine the trustor's attention allocation. The expressed emotion during the decision-making process of the trustor will be analyzed as well using Noldus Facereader software. The trustee's portrait-icon will be analyzed in Noldus Facereader to determine the expressed emotion in the Treatment condition.

V. Data Collection and Results

We ran a pilot study in April, 2023 at the University of Southern Indiana. Eight subjects were recruited to complete a trust game for 4 rounds. Among the subjects, 62.5% were females and 37.5% were males. The pilot session lasted for 20 minutes. One of the rounds was randomly selected for the actual payment. The average earning of subjects was 12.75 USD. During the experiment, webcams were used to collect facial expression data. Subjects were randomly assigned into the role of either trustor or trustee. Their roles were fixed throught the experiment. Based on the data from the pilot session, we have the following findings:

a) A portrait-icon, on average (across types), generated greater trust compared to no-icon. This is reflected in the finding that trustors seeing portrait-icons, on average, sent 7.25 USD to the trustees compared to sending only 5.25USD, on average, when no portrait-icon was observed.

b) Among the portrait-icons sent to the trustors, the portait-icon showing a sad emotion engendered the most trust. On average, 8 dollars out of an endowment of 10 dollars was sent by trustors to the trustees.

c) Facial-expression analyses reveal that the portrait-icon for sadness generated the highest level of emotion anong the trustors compared to any other portrait-icon and compared to the non-icon condition. Importantly, in this case, the trustor's emotion mirrored that of the trustee – one of sadness – indicating that the sadness portrait-icon had the power to activate the highest relative level of empathy.

VI. Conclusions

This paper demonstrates how the most elemental non-verbal cue of a portrait-icon in a screen mediated interaction can generate a higher level of trust compared to the exclusion of a non-verbal cue. Further, it it revealed that a portrait-icon reflecting some degree of sadness, by activating greater levels of empathy than portrait-icons associated with other emotions (including happiness), generates the highest level of trust. To our knowledge, this is the first paper to shed light, scientifically (without relying on self-reported reasons), on the underlying mechanism through which a certain non-verbal cue produces a certain level of trust. This has important implications for strategically managing non-verbal cues in telehealth and other kinds of screen-mediated interactions.

Future studies can expand this pilot research by simulating health-related contexts in additional controlled laboratory environments and utilizing field data from telehealth services. First, recruiting a health care provider to record a simulated telehealth consultation would allow for more extensive analysis of responses to computer-screen-mediated communications. In this study, there is the opportunity to manipulate the healthcare provider's eye gaze and facial expressions while controlling for the information communicated, environmental cues, and vocal tone. The purpose is to determine if variations in a healthcare provider's eye gaze and facial expressions influence participants' emotional responses and trust in the healthcare provider. In this study, participants' facial expressions are analyzed using AFFDEX software to measure emotional responses while eye-tracking technology captures which video elements attract participants' attention. A survey utilizing validated psychological scales to measure affective responses and trust capture self-report measures from participants. This study design follows the research of Helou et al. [13] which solely focused on a doctor's camera gaze and utilized Global Consultation Rating Scales to capture participant responses. Second, a field study using doctor-patient interactions via telehealth would provide greater external validity. The goal of such research would be to analyze patient affective responses to the telehealth experience, via facial expression analysis, combined with survey responses following the experience. This form of research can connect subconscious responses to telehealth video experiences with cognitive responses. As such, it would add to research which indicates positive affective responses to human-to-human interaction are an important determinant of telehealth success [14]. The challenge of such research is to develop a scientific study that meets all patient confidentiality required by the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the United States. One such study using recorded patient videos of emotional-focused-therapy videos has been identified [15] , but is unclear what approval would be needed to evaluate doctor-patient telehealth video consultations. Most studies assessing field data in telehealth use employee follow-up questionnaires, which do not allow for biometric data from the interaction to be analyzed.

REFERENCES

[1] Bestsennyy, O., et. al. (2021). "Telehealth: A Quarter-Trillion-Dollar Post-Covid-19 Reality? McKinsey: https://www.mckinsey.com/industries/healthcare/our-insights/telehealth-a-quarter-trillion-dollar-post-covid-19-reality.

[2] Moss, J. (2022). "The Pandemic Changed Us. Now Companies Have to Change Too," Harvard Business Review.

[3] Rasiah, S. et.al. (2020). "A Study of the Nature and Level of Trust Between Patients and Healthcare Providers, its Dimensions and Determinants: A Scoping Review Protocol." British Medical Journal (BMJ) Open 2020;10:e028061. doi:10.1136/ bmjopen-2018-028061

[4] Ivens, M. "Over 55% of Communication is Nonverbal – This is Why Video Conferencing is Vital." https://insights.ricoh.co.uk/empowering-people/video-conferencing-vital

[5] Keating, E. (2021). "Why Do Virtual Meetings Feel So Weird?" American Scientist: https://www.americanscientist.org/article/why-do-virtual-meetings-feel-so-weird

[6] Goffman, E. (1961). "Encounters; Two Studies in the Sociology of Interaction. Indianapolis:" Bobbs-Merrill.

[7] Goffman, E. (1963). "Behavior in Public Places; Notes on the Social Organization of Gatherings." New York: Free Press of Glencoe.

[8] Kimball A.B. and Morgan N. (2021). Building Trust into Telehealth, https://hbr.org/2021/03/building-trust-into-telehealth

[9] DeSteno, et.al. (2012). "Detecting the Trustworthiness of Novel Partners in Economic Exchange," Psychological Science: 23(12), pp. 1549-1556.

[10] Berg, J., et.al. (1995). "Trust, Reciprocity and Social History," Games and Economic Behavior: 10(1), pp. 122-142.

[11] Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," Experimental Economics: 10, pp. 171–178.

[12] Noldus Facereader software for emotion-detection; https://www.noldus.com/facereader

[13] Helou, S., El Helou, E., Evans, N., Shigematsu, T., El Helou, J., Kaneko, M., & Kiyono, K. (2022). Physician eye contact in telemedicine video consultations: A cross-cultural experiment. International Journal of Medical Informatics, 165. https://doi.org/10.1016/j.ijmedinf.2022.104825

[14] Seitz, L., Bekmeier-Feuerhahn, S., & Gohil, K. (2022). Can we trust a chatbot like a physician? A qualitative study on understanding the emergence of trust toward diagnostic chatbots. International Journal of Human - Computer Studies, 165. https://doi.org/10.1016/j.ijhcs.2022.102848

[15] Bailey, G., Halamová, J., & Vráblová, V. (2023). Clients' Facial Expressions of Self-Compassion, Self-Criticism, and Self-Protection in Emotion-Focused Therapy Videos. International Journal of Environmental Research and Public Health, 20(2), 1129. MDPI AG. Retrieved from http://dx.doi.org/10.3390/ijerph20021129

The Influence of the Internet and Social Networks on Behavioral Problems of Primary and Secondary School Students in Croatia

Natali Flego, Lucija Josipović Deranja and Tihomir Orehovački
Juraj Dobrila University of Pula, Faculty of Informatics, Pula, Croatia
{natali.flego, lucija.josipovic.deranja}@student.unipu.hr
tihomir.orehovacki@unipu.hr

Abstract - **Social networks are a form of Internet service designed to connect users but nowadays they are used for many other purposes. Considering the amount of time spent interacting with social networks, they have a major impact on the development of children and adolescents. An empirical study presented in this paper aims to determine the extent to which social networks contribute to problems in the behavior of primary and secondary school students. Data were collected using the questionnaire whereas hypotheses were explored with t-tests. The sample of study participants was composed of 105 primary school students enrolled in seventh and eighth grades and 106 secondary school students enrolled in third and fourth grades from Istria County and Primorsko-Goranska County. The analysis of collected data revealed that neither primary nor secondary school students perceive themselves as overly addicted to the Internet and social networks. As a consequence of the impact of social networks on their lives, secondary school students are more prone to aggressive behavior and are more likely to be depressed. However, no statistically significant differences between primary and secondary school students were found in terms of the degree to which they are feeling lonely due to the frequent use of social networks.**

Keywords - Behavioral Problems; Internet and Social Network Addiction; Aggressive Behavior; Cyberbullying; Loneliness; Depression; Primary School; Secondary School; Croatia; Empirical Study

I. INTRODUCTION

Social networks are web services that enable people to communicate synchronously or asynchronously, collaborate, and share a variety of multimedia content [15]. Facebook, Instagram, Snapchat, Twitter, TikTok, WhatsApp, Viber, and YouTube are nowadays the most popular social networks [5]. Young people are growing up with social media and often use them as an escape from what may seem like an ugly reality. Although social networks have many benefits, their use can also harm youth and adolescents. The content presented on social networks often creates unrealistic expectations while their excessive use results in reduced physical activity and lack of in-person social interactions which in turn leads to changes in behavior and insufficiently developed motor skills of young people. Various breeds of social interaction enable young people to develop a better understanding of themselves and their relationships with others. However, interaction on social networks requires young people a plethora of skills that are quite different from those necessary for casual conversation with friends [4]. Eight- to eighteen-year-olds spend more time using social networks than on any other activity except (perhaps) sleeping. On average, they are spending seven and a half hours a day interacting with social networks [17]. Considering the amount and questionable integrity of content delivered through social networks, young people can easily become confused, deceived, and deluded which prevents them from developing their own opinion, character, and rational thinking.

Social network algorithms are becoming more sophisticated to keep users' attention on content provided by them for as long as possible. The insistent sounds and vibrations of notifications attract the attention of users and create an addiction to constantly checking the status and newly posted content on social networks. Due to the fast-paced lifestyle in which the Internet and social networks play an important role, young people are becoming antisocial, depressed, and unable to communicate well in real life. In addition, due to the inappropriate and unnatural position of the body in which they search the Internet and use social networks, young people are increasingly experiencing neck and spine health issues. Therefore, it is of great importance to educate young people to become more careful and responsible in interactions with the Internet and social networks. The focus in that respect is primarily on prevention through actions that promote positive youth development rather than intervention among delinquents [4].

II. METHODOLOGY

A. Research Questions

Current studies indicate that behavioral disorders such as the Internet and social network addiction, aggressiveness, loneliness, and depression are interrelated. Kırcaburun [13] discovered that daily Internet use and social media addiction affect depression in adolescents. Findings of a study conducted by Giordano et al. [6] revealed that male adolescents who spend much time online and have a high level of social media addiction are very likely to commit acts of cyberbullying. Giumetti and

Kowalski [7] emphasized that cyberbullying via social media can have negative effects on well-being, including psychological distress, decreased life satisfaction, and suicidal ideation. Jiang et al. [10] found that loneliness significantly affects depression. The objective of this paper is to determine the extent to which the use of the Internet and social networks contributes to behavioral disorders in primary and secondary school adolescents in Croatia. In that respect, the following research questions were raised: (1) Does the extent of an addiction to the Internet and social networks differ significantly between primary and secondary school students? (2) Does using social networks contributes to aggressive behavior in primary and secondary school students? (3) Does frequent use of social networks cause loneliness in primary school students? (4) Does using social networks leads to depression in secondary school students?

B. Hypotheses

Adolescents' addiction to the use of social media and the Internet has become a pressing societal problem [16]. Aimless surfing on the Internet can over time result in the neglect of socialization, primarily with parents and then with friends. According to the findings of a study on Internet addiction of secondary school students in Međimurje County in Croatia [3], there is a correlation between the time spent on the Internet and the neglect of school duties. Considering that adolescents easily abandon their daily activities and replace them with time spent online [11], we are proposing the following hypotheses:

H1. Primary school students are more addicted to the Internet and social networks than secondary school students.

H2. Primary school students are more likely to neglect their daily activities due to the use of social networks than secondary school students.

Aggressive behavior in the form of peer violence and harassment is an increasing social problem in the lives of young people. As they grow up, adolescents more often use language to express aggression. One of the reasons for this is their awareness that physical aggressiveness is more often condemned by society [2]. When applied in an online environment, aggressive behavior is most commonly referred to as cyberbullying and mainly consists of sending abusive or humiliating messages. Girls are usually cyber victims, whereas boys are commonly cyber perpetrators [9]. Although Aizenkot et al. [1] found that cyberbullying occurs more often in primary school compared to secondary school, we are proposing the following hypothesis:

H3. Secondary school students who use social networks are more likely to engage in online aggressive behavior than primary school students.

Since bullying can easily expand from virtual to real social spaces such as classrooms, violate classroom social climate, and disrupt students' sense of belonging in class [12], we are proposing the following hypothesis:

H4. Secondary school students who use social networks are more likely to engage in aggressive behavior in real life than primary school students.

Depression in adolescents is characterized by feelings of loneliness, helplessness, anger, and disappointment with family and friends who do not understand them [19]. Loneliness caused by an external factor such as the use of social networks can lead a person to the point where (s)he can no longer reason and needs to seek professional help to get better. Nowadays there are more and more young people who feel lonely and rejected by the people around them. According to Yu et al. [22], loneliness is a direct risk factor for social networking addiction among youth. In that respect, the following hypothesis is proposed:

H5. Primary school students who use social networks feel more lonely than secondary school students.

Being a common mental disorder, depression has been recognized as a leading cause of disability worldwide [21] which affects more and more people every year [14]. Depression often occurs at the beginning of secondary school, or even earlier [18]. Findings of a recent study [20] indicate that gender, social media usage, and mattering are strongly associated with adolescent depressive symptom reporting. Therefore, we are proposing the following hypothesis:

H6. Secondary school students who use social networks are more prone to depression than primary school students.

C. Sample and Apparatus

Participants in the empirical study were seventh and eighth-grade primary school students and third and fourth-grade secondary school students. The study was conducted in six schools located in Istria County and six schools from Primorsko-Goranska County in Croatia. Such a distribution of students was chosen because seventh and eighth graders of primary school are entering puberty thus representing early adolescents while third and fourth graders of secondary school belong to the group of middle to late adolescents. Data was collected using an online questionnaire that was created and administered with Google Forms. The study participants needed 10 minutes to complete the questionnaire which consisted of two parts. The first part of the questionnaire was created for gathering respondents' demographics (8 items) while the second part was designed for examining students' perception of (1) their addiction to the Internet and social networks (P1; 6 items), (2) their neglect of daily activities for the sake of using social networks (P2; 5 items), (3) their aggressive behavior on social networks (P3; 5 items), (4) their aggressive behavior in real life (P4; 5 items), (5) their feeling of loneliness due to the use of social networks (P5; 5 items), and (6) their tendency to depression due to the interaction with social networks (P6; 6 items). Responses to the items in the second part of the questionnaire were modulated on a five-point Likert scale (1 – strongly agree, 5 – strongly disagree). Data that were collected with questionnaire items designed for measuring facets of the aforementioned constructs were combined into six (P1 to P6) single-score discrete numerical values, each of them representing the sum of all responses to all items allocated to a particular construct. The internal consistency of the aforementioned

constructs was explored with Cronbach's alpha coefficient [8]. The proposed hypotheses were examined with independent samples t-tests in SAS 9.4 statistical software where the first sample (GRUPA 1) refers to primary school students while the second sample (GRUPA 2) pertains to secondary school students.

III. RESULTS

A. Participants

A total of 211 subjects took part in the study where 105 of them were primary school students while 106 were secondary school students. They ranged in age from 12 to 19 years (M = 14.87, SD = 2.1852). The majority of them (62.6%) were girls. As shown in Figure 1, Viber (90.2%) and Instagram (86.9%) are the most commonly consumed whereas Facebook (31.8%) and Twitter (20.1%) appeared to be two of the least used social networks by study participants.

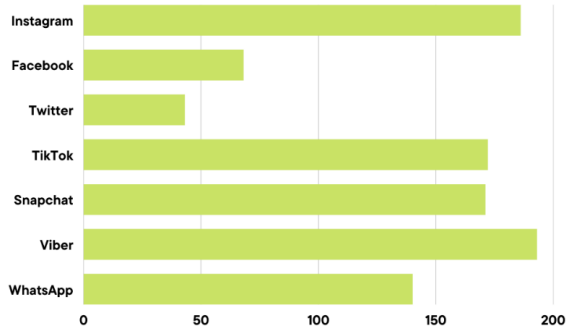


Figure 1. The popularity of social networks among study participants

Study results presented in Figure 2 indicate that 79.4% of students are using social networks at least one to three hours a day while findings reported in Figure 3 point out that 68.5% of respondents is interacting with social networks at least four to seven times a day.

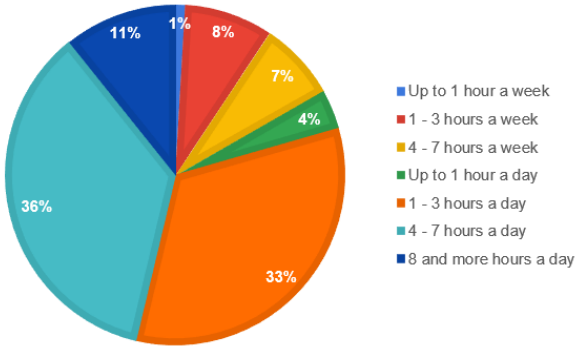


Figure 2. Frequency of using social networks expressed in hours

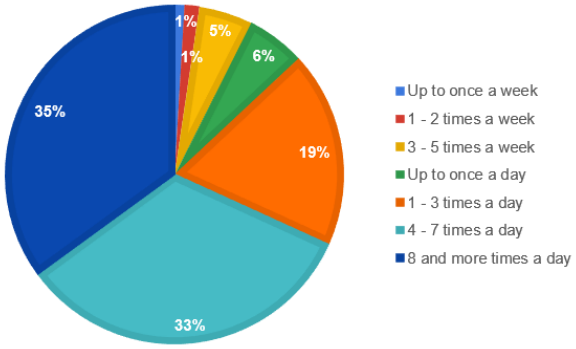


Figure 3. Frequency of using social networks expressed in the number of interactions

B. Findings

As shown in Table 1 in the Appendix, Cronbach's alpha values were in the range between 0.684 and 0.962 which is acceptable for exploratory studies [8] and implies that the reliability of scales is deemed adequate.

The analysis of collected data presented in Figure 4 uncovered that there is no significant difference (t(209)=1.11, p=0.2698) between primary (M = 19.0667) and secondary (M = 18.2830) school students in the extent to which they perceive their addiction to the Internet and social networks. Nevertheless, compared to primary school students, secondary school students are spending slightly more time on the Internet and checking posts on social networks which is why they are feeling somewhat more irritated when they are prevented from doing so. In that respect, the H1 hypothesis was rejected.

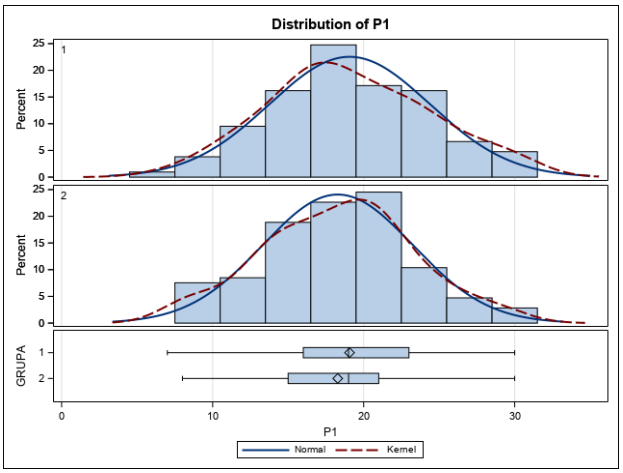


Figure 4. Students' perception of their addiction to the Internet and social networks

As shown in Figure 5, no significant difference (t(209)=1.43, p=0.1545) exists between primary (M = 18.8571) and secondary (M = 18.0094) school students in the degree to which they are neglecting their daily activities for the sake of using social networks. Although secondary school students are using the Internet and social networks a little less than primary school students when they are around their peers, they are staying awake a bit longer to check posts on social networks and are disregarding their study obligations a little more than

primary school students. Therefore, hypothesis H2 was rejected.

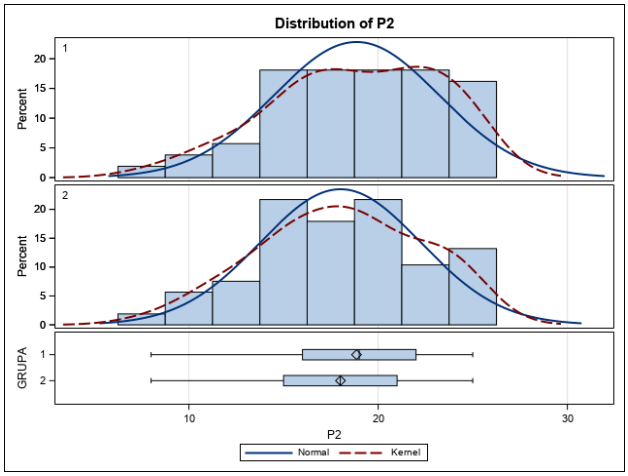


Figure 5. Students' perception of their neglect of daily activities for the sake of using social networks

Results of data analysis presented in Figure 6 indicate that secondary school students (M = 21.3774) who use social networks are significantly (t(209)=2.90, p=0.0045) more likely to engage in aggressive behavior than primary school students (M = 23.5429) thus providing support for H3 hypothesis. Although most of the study participants are not inclined to aggressive behavior on social networks, significantly more secondary school students tend to create fake profiles to insult and harass others or put them in embarrassing situations.

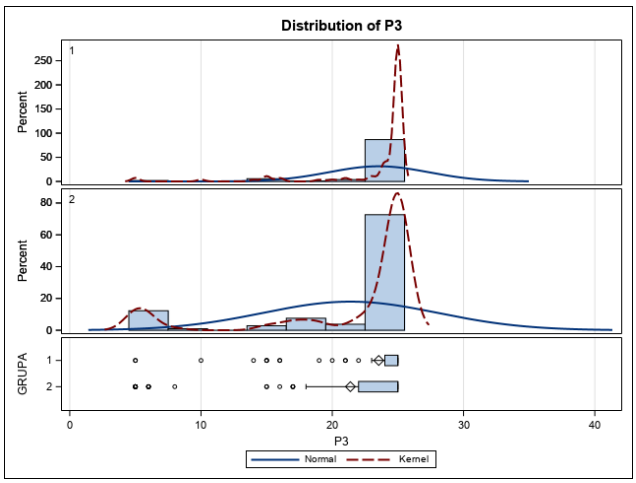


Figure 6. Students' perception of their aggressive behavior on social networks

As illustrated in Figure 7, we also discovered that secondary school students (M = 21.1981) are significantly more prone (t(209)=2.55, p=0.0122) to aggressive behavior in real life than primary school students (M = 22.8667) which provides support for H4 hypothesis. While neither primary school nor secondary school students consider their behavior or actions to be particularly aggressive, secondary school students are more likely to copy the aggressive behavior they have been exposed to and have experienced on social networks into their in-person relationships with others.

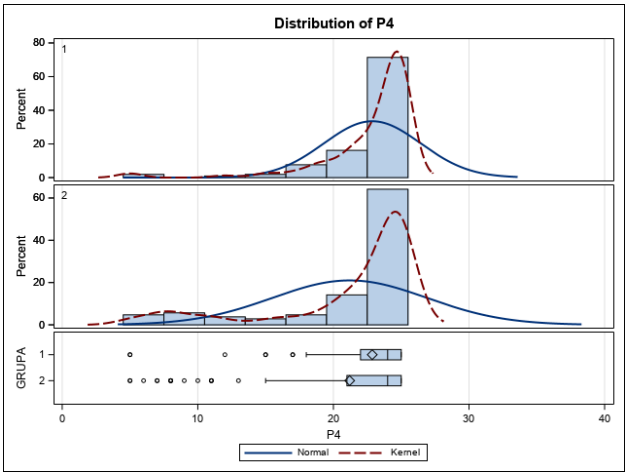


Figure 7. Students' perception of their aggressive behavior in real life

Study findings reported in Figure 8 point out that no significant difference (t(209)=0.94, p=0.3499) was found between primary (M = 19.0667) and secondary (M = 18.4057) school students concerning the extent to which they are feeling lonely due to the use of social networks thus rejecting the H5 hypothesis. While primary school students are more likely to use social networks when feeling lonely, secondary school students are going to feel lonelier and more rejected because they think they are not up to their peers who present themselves on social networks.

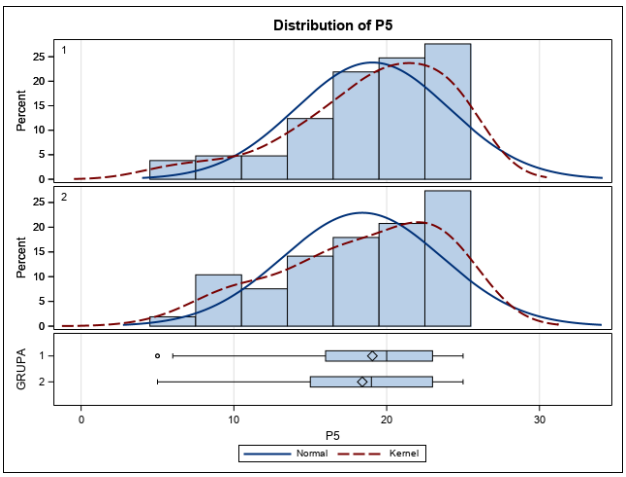


Figure 8. Students' perception of their feeling of loneliness due to the use of social networks

Study findings reported in Figure 9 point out that secondary school students (M = 21.2264) who use social networks are significantly more inclined (t(209)=2.43, p=0.0158) to depression than primary school students (M = 23.2857) thus providing support for H6 hypothesis. This is because secondary school students are more susceptible to changes in mood and interests and run away from reality into the virtual world much faster when they feel misunderstood and hurt.

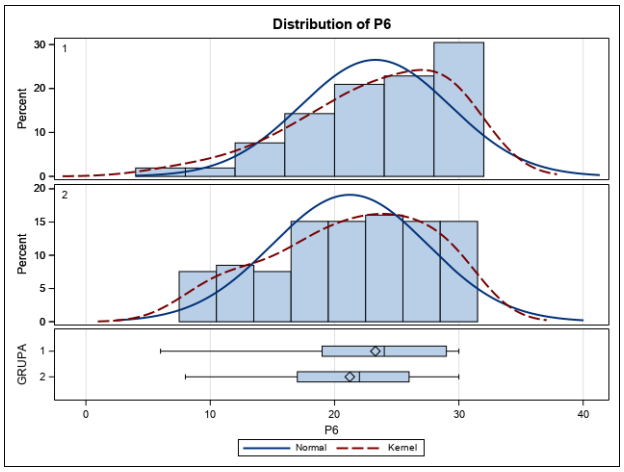


Figure 9. Students' perception of their tendency to depression due to the use of social networks

Responses of primary and secondary school students to questionnaire items designed for examining the perceived effect of social networks on their behavioral issues are summarized in Table 1 (see Appendix).

IV. CONCLUSION

Social networks have nowadays infiltrated all aspects of human activity, and the content posted on them is massively consumed by people of all age groups. The use of social networks offers many benefits to young people, from the faster acquisition of knowledge and development of interpersonal skills to networking, collaboration, and exchange of artifacts with their peers. However, interaction with social networks also has its dark side because it significantly affects an individual's feelings, self-confidence, and behavior. This is particularly true for primary and secondary school students who grow up with social networks.

The objective of this paper was to discover to what extent the use of the Internet and social networks contributes to behavioral problems in primary and secondary school students. The analysis of collected data revealed that secondary school students, when compared to primary school students, spend more time on the Internet by posting and consuming content on social networks, which is why they often stay up late and neglect their duties such as studying. However, no significant difference was found between these two groups of study participants related to the extent to which they perceive themselves as addicted to the Internet and social networks as well as to the degree to which they are disregarding their life habits due to using social networks. We also found that secondary school students are more likely to be violent towards others through social networks, especially when no one knows their identity, and transfer such behavior to their in-person interactions with others. Finally, while no significant difference was discovered between primary and secondary school students concerning the degree to which they are feeling lonely as a result of using social networks, it appeared that secondary school students who use social networks are significantly more prone to depression than primary school students.

The reported findings can be used by researchers as a backbone for future studies in the field while practitioners can consider them when developing and optimizing algorithms for displaying content that is adapted to the age of social network users. They can also serve as a foundation for developing preventive programs and intervention programs for students, teachers, parents, and all other interested parties. The proposed questionnaire can be employed for measuring the effect of social networks on the behavioral problems of their users at all levels of education and adapted for the same purpose outside the educational ecosystem.

Like all other empirical studies, this one also has limitations that need to be acknowledged. Considering that study participants were students from only two Croatian counties and that a more heterogeneous sample structure in terms of students' origin and gender representation could provide significantly different responses, reported findings should be interpreted cautiously.

To draw generalizable sound conclusions and to examine the robustness of findings, in our future work, we are planning to further examine the psychometric features of the measuring instrument and explore the interplay of proposed constructs.

REFERENCES

[1] D. Aizenkot and G. Kashy-Rosenbaum, "Cyberbullying Victimization in WhatsApp Classmate Groups among Israeli Elementary, Middle, and High School Students," Journal of Interpersonal Violence, vol. 36, no.15–16, pp. 8498–8519, 2021.

[2] T. Antičević, D. Protić, and V. Pranjić, "Agresivnost u adolescenciji: istraživanje nekih oblika agresivnog ponašanja srednjoškolaca," Nastavnička revija, vol. 2, no. 2, pp. 3–23, 2021.

[3] N. Carević, M. Mihalić, and M. Sklepić, "Ovisnost o Internetu među srednjoškolcima," Socijalna politika i socijalni rad, vol. 2, no. 1, pp. 64–81, 2014.

[4] J. Cotterell, Social Networks in Youth & Adolescence. New York, NY: Routledge, 2007.

[5] S. R. Ghodke and S. M. Kulkarni, "Social Media and Youth: Study on its Impact, Benefits and Challenges", https://www.researchgate.net/profile/Shubhada-Kulkarni/publication/333973984_Social_Media_and_Youth/links/5d10adeaa6fdcc2462a040f4/Social-Media-and-Youth.pdf

[6] A. L. Giordano, E. A. Prosek, and J. C. Watson, "Understanding Adolescent Cyberbullies: Exploring Social Media Addiction and Psychological Factors," Journal of Child and Adolescent Counseling, vol. 7, no. 1, pp. 42–55, 2021.

[7] G. W. Giumetti and R. M. Kowalski, "Cyberbullying via social media and well-being," Current Opinion in Psychology, vol. 45, 101314, 2022.

[8] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M Ringle, "When to use and how to report the results of PLS-SEM," European Business Review, vol. 31, no. 1, pp. 2–24, 2019.

[9] T. Heiman and D. Olenik-Shemesh, "Cyberbullying Experience and Gender Differences Among Adolescents in Different Educational Settings," Journal of Learning Disabilities, vol. 48, no. 2, pp. 146–155, 2015.

[10] Q. Jiang, F. Zhao, X. Xie, X. Wang, J. Nie, L. Lei, and P. Wang, "Difficulties in Emotion Regulation and Cyberbullying Among Chinese Adolescents: A Mediation Model of Loneliness and Depression," Journal of Interpersonal Violence, vol. 37, no. 1–2, pp. 1105–1124, 2022.

[11] S. Karacic and S. Oreskovic, "Internet Addiction Through the Phase of Adolescence: A Questionnaire Study," JMIR Ment Health, vol. 4, no. 2, e11, 2017.

[12] G. Kashy-Rosenbaum and D. Aizenkot, "Exposure to cyberbullying in WhatsApp classmates' groups and classroom climate as predictors of students' sense of belonging: A multi-level analysis of elementary, middle and high schools," Children and Youth Services Review, vol. 108, 104614, 2020.

[13] K. Kırcaburun, "Self-Esteem, Daily Internet Use and Social Media Addiction as Predictors of Depression among Turkish Adolescents," Journal of Education and Practice, vol. 7, no. 24, pp. 64–72, 2016.

[14] Đ. Ljubičić, I. Ljubičić Bistrović, R. Ljubičić, and J. Ivelić, Depresija i duhovnost. Rijeka: Hrvatski institut za duhovnu psihijatriju, 2012.

[15] T. Orehovački, G. Bubaš, and A. Kovačić, "Taxonomy of Web 2.0 Applications with Educational Potential," in Transformation in Teaching: Social Media Strategies in Higher Education, C. Cheal, J. Coughlin, and S. Moore, Eds. Santa Rosa: Informing Science Press, 2012, pp. 43–72.

[16] M. Peris, U. de la Barrera, K. Schoeps, and I. Montoya-Castilla, "Psychological Risk Factors that Predict Social Networking and Internet Addiction in Adolescents," International Journal of Environmental Research and Public Health, vol. 17, 4598, 2020.

[17] V. J. Rideout, U. G. Foehr, and D. F. Roberts, Generation M^2: Media in the Lives of 8- to 18-Year-Olds. Menlo Park, CA: The Henry Kaiser Family Foundation, 2010.

[18] M. E. P. Seligman, Optimistično dijete. Zagreb: IEP d.o.o., 2005.

[19] A. Vulić-Prtorić, Depresivnost u djece i adolescenata. Jastrebarsko: Naklada Slap, 2003.

[20] J. C. Watson, E. A. Prosek, A. L. Giordano, "Distress among Adolescents: An Exploration of Mattering, Social Media Addiction, and School Connectedness," Journal of Psychoeducational Assessment, vol. 40, no. 1, pp. 95–107, 2022.

[21] World Health Organization, "Depression", 2021, https://www.who.int/news-room/fact-sheets/detail/depression

[22] S. Yu, A. M. S. Wu, I. J. A. Pesigan, "Cognitive and Psychosocial Health Risk Factors of Social Networking Addiction," International Journal of Mental Health and Addiction, vol. 14, pp. 550–564, 2016.

APPENDIX

TABLE I. RESPONSES OF STUDY PARTICIPANTS TO QUESTIONNAIRE ITEMS RELATED TO PERCEIVED INFLUENCE OF SOCIAL NETWORKS ON THEIR BEHAVIORAL PROBLEMS

	Primary school students		**Secondary school students**	
	M	**SD**	**M**	**SD**
Perceived Addiction to the Internet and Social Networks (Cronbach's alpha = 0.684)	**19.07**	**5.312**	**18.28**	**4.972**
ADD1. I spend my free time on the Internet.	2.76	1.287	2.68	1.263
ADD2. I am trying to reduce the time spent on the Internet, but I am not succeeding.	3.17	1.444	3.10	1.317
ADD3. Connecting to the Internet and checking social networks has become my first activity of the day.	3.16	1.493	2.93	1.452
ADD4. Every post on social networks is important to me.	4.00	1.310	3.79	1.445
ADD5. When I have nothing to do, I use social networks to avoid being bored.	2.41	1.394	2.32	1.329
ADD6. I feel irritable when I do not have access to the Internet.	3.51	1.376	3.44	1.333
Perceived Neglect of Daily Activities (Cronbach's alpha = 0.696)	**18.86**	**4.371**	**18.01**	**4.246**
NEG1. When I hang out with friends, I constantly use the Internet and social networks.	3.79	1.133	3.87	1.133
NEG2. I scroll through new posts on social networks until late at night.	3.47	1.443	3.28	1.280
NEG3. Due to the use of social networks, I neglect my life habits.	3.72	1.257	3.48	1.334
NEG4. I do not study enough because I am constantly checking new posts on social networks.	3.61	1.351	3.36	1.341
NEG5. Using the Internet is so important to me that I avoid in-person communication with friends.	4.27	1.154	4.01	1.314
Perceived Aggressive Behavior on Social Networks (Cronbach's alpha = 0.962)	**23.54**	**3.811**	**21.38**	**6.654**
AGS1. I have a fake profile on social networks that I use to harass others.	4.65	1.020	4.20	1.450
AGS2. I like to insult others through social networks because they do not know who I am.	4.70	0.860	4.26	1.403
AGS3. I am violent towards others through social networks.	4.72	0.877	4.35	1.367
AGS4. I send abusive or humiliating messages through social networks.	4.71	0.847	4.26	1.403
AGS5. I intentionally forward messages, photos, or recordings that may embarrass others.	4.66	0.857	4.24	1.400
Perceived Aggressive Behavior in Real Life (Cronbach's alpha = 0.893)	**22.87**	**3.574**	**21.20**	**5.700**
AGR1. I put the violence that is shown on social networks into action.	4.62	0.958	4.36	1.269
AGR2. I manipulate other people easily.	4.36	1.159	3.91	1.357
AGR3. I deliberately treat others badly because I have seen it on the Internet and social networks.	4.73	0.808	4.30	1.361
AGR4. My behavior has become aggressive due to exposure to social networks.	4.52	0.965	4.31	1.269
AGR5. Because of using the Internet and social networks, I easily get into conflicts.	4.53	0.872	4.28	1.265
Perceived Feeling of Loneliness (Cronbach's alpha = 0.792)	**19.07**	**5.024**	**18.41**	**5.222**
LON1. The feeling of loneliness disappears when I use social networks.	3.72	1.365	3.60	1.302
LON2. When I feel lonely, I reach for social networks.	3.03	1.545	3.15	1.344
LON3. I feel lonely and rejected because I do not think I am like my peers portrayed on social networks.	4.21	1.221	4.07	1.399
LON4. I like to be at the center of attention as promoted by social networks.	4.07	1.344	4.07	1.312
LON5. After excessive use of the Internet and social networks, I feel quite lonely.	4.03	1.299	3.52	1.604
Perceived Tendency to Depression (Cronbach's alpha = 0.832)	**23.29**	**6.017**	**21.23**	**6.268**
DEP1. I feel uneasy among friends when I cannot use social networks.	4.34	1.173	3.78	1.436
DEP2. I find it easier to type with friends than to communicate with them in person.	3.83	1.356	3.83	1.307
DEP3. I spend a lot of time switching between social networks.	3.34	1.400	2.94	1.420
DEP4. I constantly check social networks because reality seems unbearable to me.	3.98	1.317	3.70	1.382
DEP5. I easily lose interest in activities that used to make me happy and prefer to spend time on the Internet instead.	3.99	1.307	3.48	1.469
DEP6. I feel sad when I cannot achieve an appearance similar to the one shown on social networks.	3.82	1.535	3.50	1.544

Cloud Accounting as a Factor in Protection from Cyber Attacks and Theft of Accounting Data

PhD Valentina Vinšalek Stipić & Mile Vičić
Business department
University of Applied Sciences "Nikola Tesla" in Gospić
Gospić, Croatia
vvs@velegs-nikolatesla.hr
mvicic@velegs-nikolatesla.hr

Abstract* - Incidents such as cyber-attacks are becoming more frequent in the business world. Cyber-attacks cause significant damage to the operations and accounting and financial data of companies. The question arises whether cloud accounting is the solution to reduce cyber-attacks and theft of accounting data. Many companies or individuals still do not trust cloud accounting. The problem of this research is the acceptance of accounting in the cloud by accounting staff and insufficient understanding of the benefits of working in the cloud to prevent cyber-attacks and theft of accounting data. Research is conducted among accounting professionals in Croatia, randomly distributed, via anonymous survey, where 183 responses were gathered. Statistical analysis of the collected data confirmed the set hypotheses, i.e. the frequency of cyberattacks is significantly statistically related to the method of data storage ($r = 0.807$; $p < 0,001$), and partially ($r = 0.199$; $p < 0,001$) to the method of authentication in accounting software and the method of creating a backup copy of accounting data.

***Keywords* - cloud accounting; cyber attacks; protection of accounting data; perception about cloud accounting**

I. INTRODUCTION

Today's competitive business environment forces fast-growing companies to constantly monitor progress, change business strategy and adapt to the risks of the business environment. Over the past decades, the role of information technology has changed significantly [1]. The advancement of cloud technology represents one of the biggest technological trends at the moment. The cloud is a platform for data availability anytime, anywhere, from almost any device that has an Internet connection [2]. The most famous term for cloud technology is the term cloud computing [3]. Cloud computing is computing resources that companies and users can access from remote locations via the Internet [4]. Such solutions occupy a very important place in the optimization and improvement of all business processes of the company. Like other business sectors, accounting has also embraced cloud technology solutions. Cloud accounting partly changes the nature of the accounting business, so the slow adoption of cloud accounting is expected. A significant resistance to working with accounting in the cloud is due to the loss of self-control over the company's accounting and financial data and the fear of hacker attacks and complete data loss [5]. The question arises as to the cause of frequent hacker attacks on accounting information systems and is business data sufficiently protected in cloud accounting? The answers to these questions were obtained through empirical research on a sample of 183 respondents, accounting employees in the Republic of Croatia. The aim of this research is to prove that the frequency of hacker attacks in accounting depends to a significant extent on the method of data storage and the method of authentication in accounting software.

II. CLOUD ACCOUNTING

The use of new information technology, such as cloud accounting, is a challenge for users (accountants) and companies, this is due to the constant evolution of technology and the frequent appearance of new technologies that appear every day. A big challenge is the technical language and its features. Some of the computer programs come in languages that users such as accountants cannot easily understand. Therefore, accountants must raise the level of training before use, to enjoy the benefits of cloud accounting. Another challenge is the difficulty of adapting to such frequent changes [5]. Some workers may lose interest and this may reduce work morale. Cloud computing has been compared and equated to the industrial revolution. However, its transformational nature is associated with significant security and privacy risks [6]. Cloud accounting gives accountants instant and mobile access to financial information and completely changes the way accountants work. Because SMEs can quickly respond to changes in technology demand, they can be the fastest adopters of cloud accounting services [7]. Companies whose employees work remotely may also prefer cloud accounting for convenience and availability, companies that cannot ensure data security and companies that want to avoid all potential physical accidents with technology, destruction of hard drives, and thus accounting data [8].

A. Advantages and disadvantages of cloud accounting

Traditional accounting programs and cloud accounting solutions have their advantages and disadvantages. The use of cloud accounting in business operations brings many advantages, the most prominent of which are [9]:

Lower costs – the price paid for using cloud accounting services is generally lower than traditional accounting

programs. The user does not pay installation fees or have to purchase new updates in case accounting rules or tax regulations change because the monthly or annual subscription costs include the cost of updates. Also, the user doesn't have the costs of purchasing hardware, licenses for antivirus programs and operating systems, and server maintenance costs, as is the case with traditional accounting programs. All you need is a stable internet connection.

Security – by using cloud accounting, there is no risk of destruction or theft of hardware and software. All data is stored away from the company on secure servers that are under the control of the company that manages it. Such companies usually have stronger antivirus programs, which makes financial data relatively safer in the cloud than stored locally, on the user's computer.

Availability – in cloud accounting, data is always available to all authorized users, wherever they are. All that is required is a stable internet connection and a device from which to access the software. This is especially important in companies that prefer remote work. In addition, it is very easy to add new users - by setting up an authorized profile and password, which facilitates collaboration and because everyone can see the relevant financial information at any time.

Up to date – the latest and updated versions of the software that the cloud accounting service provider upgrades are immediately available to the user, which is especially important when it comes to changes in accounting rules and tax regulations. In this way, the user saves time and focuses on his business, knowing that these tasks will be done for him by the service provider.

Automatic Data Backup and Restore – With cloud accounting, there is no need to manually back up accounting data because the software does it automatically, reducing the possibility of human error and forgetfulness.

The advantages of cloud accounting overshadow the disadvantages, however, there are potential risks that users of this software may encounter, some of which are:

Internet connection - cloud accounting requires a constant and strong internet connection; cloud accounting software does not work well at low speed. Internet speed depends on the line capabilities of the user's location and can vary from place to place, which is a big disadvantage.

Security – as cloud services becomes more popular, data becomes more and more the target of attacks, which is why some businesses such as banks benefit more from keeping their sensitive financial data "in house" than with a cloud computing service provider.

Reduced control – loss of control over the accounting software, which is completely managed by the service provider, is one of the main problems of cloud accounting application. Maintaining the software in the event of failure or downtime and performing application updates is the sole responsibility of the service provider, not the user, which makes companies afraid of losing control over accounting data.

B. Cloud accounting features

The greatest impact of IT on accounting is the ability of companies to develop and use computerized systems for up-to-date recording of accounting data. Most popular accounting systems can also be tailored to specific industries or businesses [10]. This allows businesses to create individual reports quickly and easily for management and/or decision-making. Second, the advantages of computerized accounting systems can be summarized as follows: increased functionality, improved accuracy, faster processing, better external reporting. However, today's accountants need to be familiar with software tools that can help them perform accounting functions more efficiently.

III. CYBERSECURITY IN CLOUD ACCOUNTING

Cyber-attacks are deliberate attempts to disrupt, steal, alter or destroy data stored in IT systems. Attacks are often motivated by profit and many intruders are technically sophisticated and have a nuanced understanding of system functioning [11]. Cybersecurity has become more urgent as malicious actors develop sophisticated techniques. But quantifying the risk or resilience of institutions to cybersecurity incidents is difficult [12]. The lack of standardized data on such incidents and company control is a challenge for the protection of accounting systems, but above all, companies are responsible for their own security against cyber-attacks [13].

A. Forms and effects of cybersecurity incidents in accounting

Like any new technology, cloud computing is subject to security threats and vulnerabilities, including network threats, information threats, and underlying infrastructure threats [14] in the form of cyber-attacks. Attack tactics include finding weaknesses in software to get into IT systems, redirecting to fake email servers to steal passwords, redirecting to fake websites that infect users with malware, and software that wipes users from their own systems. Detailed data on the frequency, tactics and results of cybersecurity incidents are scarce. Data is scarce in part because financial firms avoid reporting incidents due to reputational concerns [15]. Concerns about cyber-attacks are growing [16]. However, an effective review creates a basis for advancing knowledge [17]. It is necessary to synthesize studies related to cyber security in the field of accounting and auditing. The number and severity of cyber threats have been unprecedented in recent years, and successful cyber-attacks are regularly reported [18]. The costs of the consequences of cyberattacks are enormous; therefore, cyber security risk management is considered extremely important for organizations [18]. In this regard, Hausken (2006) states that the intensity of cyber warfare has increased through the Internet revolution. [19]. Also, Gordon et al. (2003) suggest that the Internet revolution has

dramatically changed the way individuals, businesses, and government communicate and do business. They concluded that the widespread interconnection of industrial branches increased the vulnerability of computer systems [20]. Gansler and Lucyshyn (2005) state that the growing dependence of the public and private sectors on Internet-based technologies and networks for their financial management systems comes at a price, and that price is increased vulnerability [21]. Therefore, in 2016, the financial services sector was attacked more than any other industry [22]. In this global information society, where information travels through cyberspace, it is crucial to manage it effectively [23]. Effective management, on the other hand, is linked to awareness of increasing vulnerabilities, such as cyber threats and information warfare. Effective cyber security management is essential. Investments in cyber security are profitable if the benefits of additional activities for information security exceed its costs [24]. Although cyber security does not always benefit the organization, cyber-attacks are one of the main risks that organizations must control [25]. Earlier review studies related to this topic discussed research opportunities in IT and internal auditing [26] and the impact of information security events on the stock market [27]. Based on the above arguments, it is crucial to synthesize the previous literature related to cybersecurity and identify the research streams of the articles under review.

B. Effects of cyber security deficiency in accounting

Cyber security is often used as an analogous term for information security. However, cyber security is not necessarily only the protection of cyberspace itself, but also the protection of those who function in cyberspace and any of their assets that can be accessed through cyberspace [28]. Cyber security encompasses technologies, processes and controls designed to protect systems, networks and data from cyber-attacks. Effective cyber security reduces the risk of cyber-attacks and protects societies, organizations and individuals from unauthorized exploitation of systems, networks and technologies. Cyber security is an umbrella concept that encompasses information security and information assurance [29]. Thus, cyber security includes the protection of information that is evaluated and transmitted through any computer network [24].

While cybersecurity threats impose direct costs on businesses, incidents in cyberspace can also pose a broader risk to financial stability. Financial companies work in complex networks and rely on electronic transactions, often representing rapid management in time. They are digitally connected to other and non-financial entities, including third-party service providers. However, defending decentralized networks against cyber-attacks with many entry points can be extremely difficult [30]. Cloud has already conquered almost every business area, but it seems that the accounting profession is sceptical about this new model. According to some accountants, cloud-based software is an obvious threat [31]. Of course, it can be seen as a threat to those unwilling to adapt and clearly understand the benefits involved. Therefore, it is necessary to pay attention to cybersecurity in accounting. In the absence of cybersecurity in accounting, three effects are possible [32]:

Lack of interchangeability – the financial services industry relies on a robust IT infrastructure to complete transactions and move payments. In many financial networks, several companies or central e-services servers serve as hubs. Their services will be difficult to replace if lost or interrupted. Problems at key hubs can increase stability concerns. Policies that encourage the expansion of the financial system can reduce those risks. Regulators should consider such policies.

Loss of trust – cyber-attacks often target customer account information and financial assets. Most of these attacks were one-off events, injuring only the victim and their clients. However, the excessive occurrence of cyber theft can cause a wider loss of trust. For example, after the cyber theft of customer names, credit card information and phone numbers of a retail chain, many clients call or visit their banks to ask for information about the presence of their money in their accounts. Then many customers cancel their credit cards, which can cause a banking crisis.

Lack of data integrity – financial data integrity is critical. Many financial markets operate in real time. Financial companies need a robust data backup that can be recovered shortly after a cybersecurity incident. However, there has been a trade-off between rapid recovery and ensuring that recovered data is secure, accurate and does not spread the cyber risks, especially for markets that process orders quickly. Data corruption can disrupt market activity and can be difficult to initiate or recover from.

IV. GOALS AND HYPOTHESES OF RESEARCH

The previously presented theoretical approach shows the importance of cybersecurity in accounting. This research emphasizes the development of cybersecurity in accounting, ways of storing data and protecting them from possible cyber-attacks. Therefore, this paper tried to investigate the correlation between the frequency of cyber-attacks in accounting with the method of data storage, then with the method of authentication in accounting software, and the intensity of cloud accounting protection against cyber-attacks. Accordingly, the following research hypotheses were set:

H_1 – The frequency of cyber-attacks is statistically significantly related to the method of data storage

H_2 – The frequency of cyber-attacks is statistically significantly related to the method of authentication in the accounting software

H_3 – The frequency of cyber-attacks on accounting is statistically significantly related to the way accounting data is backed up

V. METHOD OF THE RESEARCH

Data for the implementation of this empirical research were collected through a survey on a sample of 183 accounting employees in the Republic of Croatia. The survey was conducted anonymously, through invitation for participation within accountant's professional association groups on social and business networks in March 2022 and by invitation in January 2023. The survey consisted of 20 questions, divided into three groups, including general

information about examinee, familiarity with cloud accounting and cyber security application, it used a Likert measuring scale with five intensities, from which the following variables were defined:

a) independent variables, which were obtained by the weighted average rating of the responses from the survey questionnaire:

X_1 – data storage mode (DataA)

X_2 – authentication mode (Authent)

X_3 – data backup mode (ProtectA)

b) dependent variable frequency of cyber-attacks on accounting data (CyberA) => more than three times a year; two to three times a year; once a year; once every few years; never.

The conducted research showed a correlation of the frequency of attacks on accounting information systems with the methods of storing accounting data, the methods of authentication in accounting software and the method of data backup. To prove the set hypotheses, the statistical methods of correlation and regression analysis were applied using the SPSS 26.0 statistical program. The obtained results of the statistical data processing are shown below.

VI. RESULTS OF THE RESEARCH

A total of 183 accountants from Croatia took part in the research, of which 86.9% were female, while the majority were 31 to 45 years old (48.6%), followed by 18 to 30 years old, which made up 26.2% of respondents. and the rest (25.1%) to respondents aged 46 to 65. According to the level of education, 40.4% of the respondents have a higher vocational education, followed by a bachelor's degree in 30.1%, while 26.2% of them have a secondary vocational education and the rest are university specialists or Doctor of Science. The structure of respondents according to the level of work experience in accounting is shown in table I.

TABLE I: YEARS OF WORKING EXPERIENCE IN ACCOUNTING

Work experience in years	Number of respondents	%
until 2	48	26,2
3 to 5	27	14,8
6 to 10	25	13,7
11 to 15	29	15,8
16 and over	54	29,5

Performed simple regression analysis between the observed variables (DataA and CyberA) shows a strong significant linear correlation between the frequency of cyberattacks and the way data is stored in accounting. From Table II., a correlation coefficient of 0.807 is visible, while the F ratio is significantly higher than the theoretical value (65.2% of variations in the dependent variable are the result of variations in the independent variable), with a significance level of 0.05 and the number of degrees of freedom 1.182, it can be concluded that the first hypothesis is confirmed. Also, Durbin-Watson is 2.070 which means no existence of autocorrelation of relationship errors.

TABLE II. STATISTICAL RELATIONSHIP OF THE FREQUENCY OF CYBER ATTACKS AND THE METHOD OF STORAGE OF ACCOUNTING DATA

Model Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
1	,807[a]	,652	,650	,663
a. Predictors: (Constant), DataA b. Dependent Variable: CyberA				

By analysing the statistical connection between the frequency of cyber-attacks and the method of authentication in the accounting software shown in Table III. a positive but weak significant correlation is visible ($r = 0.199$; $p < 0,001$). The coefficient of determination R2 is closer to zero than to one, therefore we cannot talk about good representativeness of the model, with a significance level of 0.05 the second hypothesis is partially confirmed, there is a weak positive correlation between the frequency of cyber-attacks and authentication in accounting software.

TABLE III. STATISTICAL RELATIONSHIP OF THE FREQUENCY OF CYBER ATTACKS AND THE METHOD OF AUTHENTICATION IN ACCOUNTING SOFTWARE

Model Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
1	,199[a]	,040	,034	1,101
a. Predictors: (Constant), Authent b. Dependent Variable: CyberA				

Below, from table III. is visible a positive but weak statistical correlation between the frequency of cyber-attack and the method of backup of accounting data. Durbin-Watson has a value close to 2, which indicates the absence of autocorrelation of relational errors. With a significance level of 0.05 and from the coefficient of determination, it is visible that 40% of the variations of the dependent variable are the result of variations of the independent variable. The obtained results partially confirm the third hypothesis.

TABLE IV: STATISTICAL RELATIONSHIP OF THE FREQUENCY OF CYBER ATTACKS AND THE BACKUP METHOD OF ACCOUNTING DATA

Model Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
1	,199[a]	,040	,034	1,101
a. Predictors: (Constant), ProtectA b. Dependent Variable: CyberA				

The collected responses to the conducted survey revealed that 60.9% of respondents use a username and password for authentication to work in accounting software, and 20.7% of respondents use only a password. The fact that 8.7% of respondents do not have any security authentication for accessing accounting data is worrying, which can be considered inadequate protection of accounting data. It is also worrying that 19.6% of respondents save a backup copy of accounting data locally, that is, on the same computer on which they use the accounting program.

VII. CONCLUSION

By conducting empirical research on a sample of 183 accounting employees in Croatia, this research wanted to prove that the frequency of cyber-attacks in accounting depends to a significant extent on the method of data storage and the method of authentication in accounting software. Statistical analysis of the collected data confirmed the set hypotheses, and it can be concluded that the frequency of cyberattacks is significantly positively statistically related to the method of data storage ($p = 0.807$; $p < 0{,}001$) and there is a weak positive correlation between cyberattacks and authentication methods in accounting software and the method of creating a backup copy of accounting data ($r = 0.199$; $p < 0{,}001$). Earlier research showed a significant correlation between cyber-attacks and the age structure of accounting employees and their level of education, that is, older accounting employees with a lower level of education are more significantly exposed to cyber-attacks [33, 34]. This research led to the realization that 20.7% of respondents use only a password for authentication to work in accounting software, but a more worrying fact is that 8.7% of respondents do not have any security authentication for accessing accounting data, which can be considered inadequate protection of accounting data. It is also alarming that 19.6% of respondents save a backup copy of accounting data locally, that is, on the same computer on which they use the accounting program. This research confirmed that more complex authentication and data storage outside the location of the accounting program itself reduces the possibility of cyber-attacks in accounting. Also, storing accounting data on the cloud significantly reduces the possibility of cyber-attacks.

The shortcomings of this research are reflected in the impossibility of identifying the total number of accountants in the Republic of Croatia in relation to the number of respondents included in this research. Also, due to the unavailability of data, it was not possible to compare the obtained data on the frequency of cyber attacks in the Republic of Croatia (data obtained through this research) with the frequency of cyber attacks in other countries. It is recommended to repeat the research on a more representative sample with a more detailed investigation of the type of cyberattacks and their frequency.

REFERENCE

[1] F. Altınay, G. Dagli, Z. Altınay, „The Role of Information Technology in Becoming Learning Organization," Procedia Computer Science, 102, 2016, pp. 663–667.

[2] P. Kumar Paul and M. K. Ghose, „Cloud Computing: Possibilities, Challenges and Opportunities with Special Reference to its Emerging Need in the Academic and Working Area of Information Science," Procedia Engineering, 38, 2012, pp. 2222–2227.

[3] N. Bačanin Džakula and I. Štrumberger, "Klaud računarstvo," Beograd: Univerzitet Singidunum, 2018.

[4] N. Ž. Budimir, "Primjena oblak računarstva u računovodstvo," Putokazi: Sveučilište Hercegovina, 7(1), 2019, 137–148.

[5] V. Vinšalek Stipić and M. Vičić, "An analysis of accountants' resistance to cloud accounting," Journal of Economic and Business Issues, 2(2), 2022, pp. 15–23.

[6] N. Kshetri, „Privacy and security issues in cloud computing: The role of institutions and institutional evolution," Telecommunications Policy, 37(4–5), 2013, pp. 372–386.

[7] C. Christauskas, and R. Miseviciene, "Cloud–Computing Based Accounting for Small to Medium Sized Business," Engineering Economics, 23(1), 2012, pp. 14–21.

[8] T. Khanom, "Cloud Accounting: A Theoretical Overview," IOSR Journal of Business and Management, 19(6), 2017, pp. 31–38.

[9] P. A. Abdalla and A. Varol, "Advantages to Disadvantages of Cloud Computing for Small-Sized Business," 2019 7th International Symposium on Digital Forensics and Security (ISDFS), Barcelos, Portugal, 2019, pp. 1–6.

[10] O. Dimitriu and M. Matei, „Cloud Accounting: A New Business Model in a Challenging Context," Procedia Economics and Finance, 32, 2015, pp. 665–671.

[11] J. Jang-Jaccard and S. Nepal, „A survey of emerging threats in cybersecurity," Journal of Computer and System Sciences, 80(5), 2014, pp. 973–993.

[12] D. P. David, L. Maréchal, W. Lacube, S. Gillard, M. Tsesmelis, T. Maillart and A. Mermoud, „Measuring security development in information technologies: A scientometric framework using arXiv e-prints," Technological Forecasting and Social Change, 188, 2023, 122316.

[13] White House. Critical Infrastructure Security and Resilience. Presidential Policy Directive (21). Washington: White House, Feb. 12, 2013. https://www.whitehouse.gov/ the-press-office/2013/02/12/ presidential-policy-directive-critical-infrastructure-security-and-resil

[14] R. M. Jabir, S. I. R. Khanji, L. A. Ahmad, O. Alfandi and H. Said, "Analysis of cloud computing attacks and countermeasures," 2016 18th International Conference on Advanced Communication Technology, PyeongChang, Korea (South), 2016, pp. 1–1.

[15] Office of Financial Research. 2015 Financial Stability Report. Washington: OFR, Dec. 15, 2015.

[16] Symantec Corp. Internet Security Threat Report. Herndon, Va.: Symantec, April 2016.

[17] J. Webster, and R. Watson, "Analysing the past to prepare for the future: writing a literature review", MIS Quarterly, 26(2), 2002, pp. 13–23.

[18] M. S. Islam, N. Farah, and T. S. Stafford, "Factors associated with security/cybersecurity audit by internal audit function: an international study", Managerial Auditing Journal, 33(4), 2018, pp. 377–409.

[19] K. Hausken, "Income, interdependence, and substitution effects affecting incentives for security investment", Journal of Accounting and Public Policy, 25(6), 2006, pp. 629–665.

[20] L. A. Gordon, M. P. Loeb and W. Lucyshyn, "Sharing information on computer systems security: an economic analysis", Journal of Accounting and Public Policy, 22(6), 2003, pp. 461–485.

[21] J. Gansler, and W. Lucyshyn, "Improving the security of financial management systems: what are we to do?", Journal of Accounting and Public Policy, 24(1), 2005, pp. 1–9.

[22] The World Bank (2018), "Financial sector's cybersecurity: regulations and supervision", available at: http://documents.worldbank.org/curated/en/686891519282121021/pdf/123655-REVISED-PUBLIC-Financial-Sectors-Cybersecurity-Final-LowRes.pdf

[23] J. W. Lainhart, "COBIT™: a methodology for managing and controlling information and information technology risks and vulnerabilities", Journal of Information Systems, 14(1), 2000, pp. 21–25.

[24] A. L. Gordon, and P. M. Loeb, "Managing Cybersecurity Resources: A Cost–Benefit Analysis," New York: McGraw Hill, 2006.

[25] E. Amir, S. Levi, and T. Livne, "Do firms underreport information on cyber-attacks? Evidence from capital markets", Review of Accounting Studies, 23(3), 2018, pp. 1177–1206.

[26] M. Weidenmier and S. Ramamoorti, "Research opportunities in information technology and internal auditing", Journal of Information Systems, 20(1), 2006, pp. 205–219.

[27] G. Spanos and L. Angelis, "The impact of information security events to the stock market: a systematic literature review", Computers and Security, 58, 2016, pp. 216–229.

[28] Von Solms, R. and van Niekerk, J. (2013), "From information security to cyber security", Computers and Security, Vol. 38, pp. 97-102.

[29] Gyun No, W. and Vasarhelyi, M.A. (2017), "Cybersecurity and continuous assurance", Journal of Emerging Technologies in Accounting, Vol. 14 No. 1, pp. 1-12.

[30] E. Rosengren, "Cyber Security and Financial Stability." Speech to the Basel Committee on Banking Supervision, Cape Town, Jan. 30, 2015, available at: www.bostonfed.org/news/ speeches/rosengren/2015/013015/ 013015text.pdf

[31] CCH. Cloud Computing - A matter of survival for the accounting industry. CCH Research Report. 2013, available at: http://www.cchifirm.com/why_ifirm/cloud-computing-and-the-accounting-industry

[32] Office of Financial Research. 2022 Financial Stability Report. Washington: OFR, available at:
https://www.financialresearch.gov/reports/

[33] Boban, M., Vinšalek Stipić, V. & Grabić, J. (2018). Infuence of IT on Accounting Practice and Exposure to Cyber Attacks, 2nd ISIP 2018, Symposium on Information Security and Intellectual Property, September 13 – 15, 2018. in Split – Supetar in the frame of the 26th International Conference on Software – SoftCom 2018, No. 1570485850, available at: https://www.bib.irb.hr/996659

[34] Boban, M. & Vinšalek Stipić, V. (2020), „Cloud accounting – security, reliability and propensity of accounting staff to work in cloud accounting.", 28th International Conference on Software, Telecommunications and Computer Networks - SoftCOM 2020, September 17 – 19, 2020. Hvar, No. 1570674979, available at: https://www.bib.irb.hr/1091437

Principles for Designing for Data Exploration for a Non-Expert Audience

Natasha Tylosky*, Antti Knutas*, Annika Wolff*
* LUT University, Department of Software Engineering, Lappeenranta, Finland
Natasha.Tylosky@lut.fi, Antti.Knutas@lut.fi, Annika.Wolff@lut.fi

***Abstract*—A gap currently exists between how data is presented by researchers and how the average person understands data. Thus, although an abundance of good research data is available for free online, it can be difficult to impossible for laypeople to access and make sense of it on their own. The issue of data accessibility and interpretability is consequently preventing open science goals from being reached. Data exploration systems can provide people with limited data expertise an entrance point into complex data and help remedy this problem. This paper presents design principles that can be applied when creating a data exploration software system.**

Keywords*—*Data Exploration, Interactive Data Visualization, Sensemaking, Systematic Mapping Study, Design Principles

I. INTRODUCTION

In today's world, access to open data is readily available to anyone with an internet connection. There is, however, a gap in how the scientific community presents data and how the average person can interpret that data. Consequently, open scientific data sets and open data portals are not always findable or usable for those with less data expertise [1], [2]. Furthermore, there is research potential in inviting those with less data expertise to give their own input and interpretations of data [3], [4]. While there have been recent literature reviews on human data interaction systems and user interfaces in the field of HCI [5]–[7], the focus of these reviews has been on usability as a whole without considering the intended users of such systems and interfaces. Given that traditional data systems often assume that users possess good knowledge of the meaning and contents of a database, and that users are certain of the information they are looking for [8], this lack of consideration given towards non-expert audiences when designing data exploration systems can be considered a potential hurdle for users who are new to such systems and could discourage non-expert users from engaging with open data.

In this work, we conduct a systematic mapping review that gives a perspective on the types of data exploration systems that have been created with a non-expert audience in mind. From this systematic mapping review, we seek to synthesize a set of design principles for designing data visualization systems that encourage data exploration in a non-expert audience. In this paper, we define a "non-expert audience" as anyone who can be considered as not being a domain expert of the information found in a visualization. For example, in the context of medical data visualization, medical doctors would be considered an expert audience, whereas patients would be considered non-expert. If a data visualization system is designed for patients, it would be designed for a non-expert audience. In addition to the above, we also examine the domains in which data visualization research is being applied and the types of visualizations that are being used, as it is possible that there may be interesting links between the domain and visualization model or audience and visualization model.

In the following Section, we review related research on the topic. In Section 3, we present the systematic mapping process and continue with thematic synthesis of mapping outcomes in Section 4. We conclude with discussion of outcomes in Section 4 and conclusions in Section 5.

II. RELATED RESEARCH ON DATA VISUALIZATION AND EXPLORATION

When hearing the term "data visualization", the first thing that springs to mind may be 2D visualizations such as line graphs and bar charts, and while these are indeed common methods of visualization [9], data visualization does not have to be limited to such methods. For example, the Peruvian knot tying technique known as quipu is an indigenous method of data visualization that predates computers [10]. In this paper, we are interested in uncommon methods of data visualization, specifically interactive data visualizations that enable sensemaking of data. Sensemaking can be described as the process of constructing meaning from information and is an iterative process that involves linking different pieces of information into a single conceptual representation [11], [12]. Here, we will be using the inclusion of sensemaking in a data visualization system as a marker for whether the system is exploratory [8], i.e., it is a data exploration system or tool.

III. SYSTEMATIC MAPPING STUDY

As a starting point for our research, we conducted a systematic mapping study of trends in interactive data visualizations in the Human Computer Interaction (HCI) field. The aim of a systematic mapping study is to gain a comprehensive overview of a particular research topic, present unbiased assessment of current literature, identify research gaps and collect evidence for future research directions [13].

This study been funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 872500.

A. Keyword and database selection

We initially did deep search testing of various keywords and search strings in Scopus and Science Direct. The goal of this deep search was to find the keywords that produced the best results; a variety of keywords were needed in this review as there are science communication related papers that may contain pertinent research which may not appear under computer science focused terms. After the deep search, another search was done across 4 databases, namely, Scopus, Science Direct, IEEE, and ACM, using the keywords we had chosen. The final search terms were as follows.

1) ("Data Storytelling" OR "Participatory Storytelling" OR "Data Visualization" AND "Storytelling")
2) ("Data Exploration" AND "HCI" OR "Data Exploration" AND "Human Computer Interaction" OR "Science Education" AND "Data Exploration")
3) ("Sensemaking in Interactive Space")

B. Inclusion criteria and selection process

Following the literature search, a set of inclusion criteria was defined that selected papers must meet.

1) Inclusion criteria:

- The paper must research data dissemination aimed at a non-expert audience. This review excludes data dissemination directed at domain experts. [14]
- The methods used in the paper must include a data visualization aimed at a non-expert audience. This is necessary as the goal of this paper is to look at data exploration systems for non-expert audiences.
- The audience must interact with the data in some way. The intended audience must be able to interact with the data in some way. In the case of static visualizations, this could be an activity that does not change the visualization itself.
- The paper needs to be peer reviewed.
- The paper is published in English.
- The paper must be published between 2016 and 2021. This criterion is included to focus on recent research, we are only interested in recent research as data exploration is a new concept.

2) Paper selection process:

1) The title of the paper was read to determine if the paper potentially matched the search.
2) Duplicate studies were removed.
3) During the first pass the abstract of each paper was read to determine if the study fits the inclusion criteria.
4) During the second pass the remaining papers were reviewed more deeply.

C. Search outcomes

A total of 24 papers were ultimately chosen for inclusion in the literature review. The stages of the paper selection process are shown in Fig. 1.

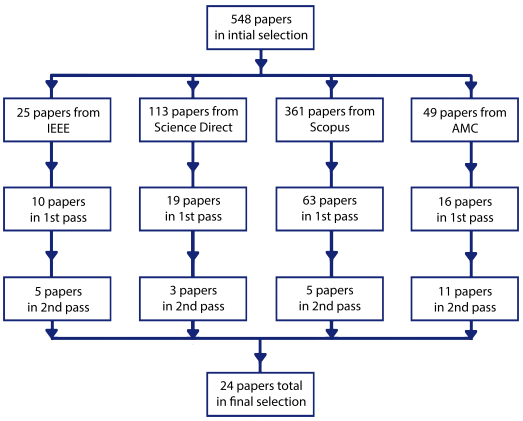


Fig. 1: Flowchart showing the paper selection process for the literature review

Full list of articles included in systematic mapping study are available at the Zenodo open data repository [1].

IV. THEMATIC SYNTHESIS

Data was extracted and summarized using the thematic synthesis method defined by Cruzes and Dyba [15]. The approach comprises a process of steps for extracting, coding, and translating codes into themes. The themes are then distilled into higher order themes and subsequently synthesized to shed light on the research question. The categories selected based on thematic synthesis were: 1. Visualization Model, 2. [9] Intended Audience, and 3. Application Domain [16]. The category of visualization model referred to words and phrases related directly to the visualization itself, such as 3D, interactive, or time series; intended audience included any words or phrases that mentioned participants, users, and such synonyms; and application domain included words and phrases related to the purpose of a visualization. The initial coding process produced 16 distinct codes for visualization models, 14 distinct codes for intended audience, and 14 distinct codes for application domains. These codes were then grouped into further codes based on similarity. 3 final codes related to the visualization model category, 4 codes related to the intended audience, and 4 codes related to the application domain. Diagrams of each thematic synthesis can be seen in Fig. 2.

From the thematic synthesis, we compared the application domains and intended audience in relation to the visualization models to examine if there were any trends in how these concepts were used. The results can be seen in the Table 1.

V. PRINCIPLES BASED ON THE SYSTEMATIC MAPPING STUDY

Next, we searched for common themes in the findings of the 24 selected papers, and from these common themes, we then synthesized three higher-order themes related to

[1] https://doi.org/10.5281/zenodo.7575802.

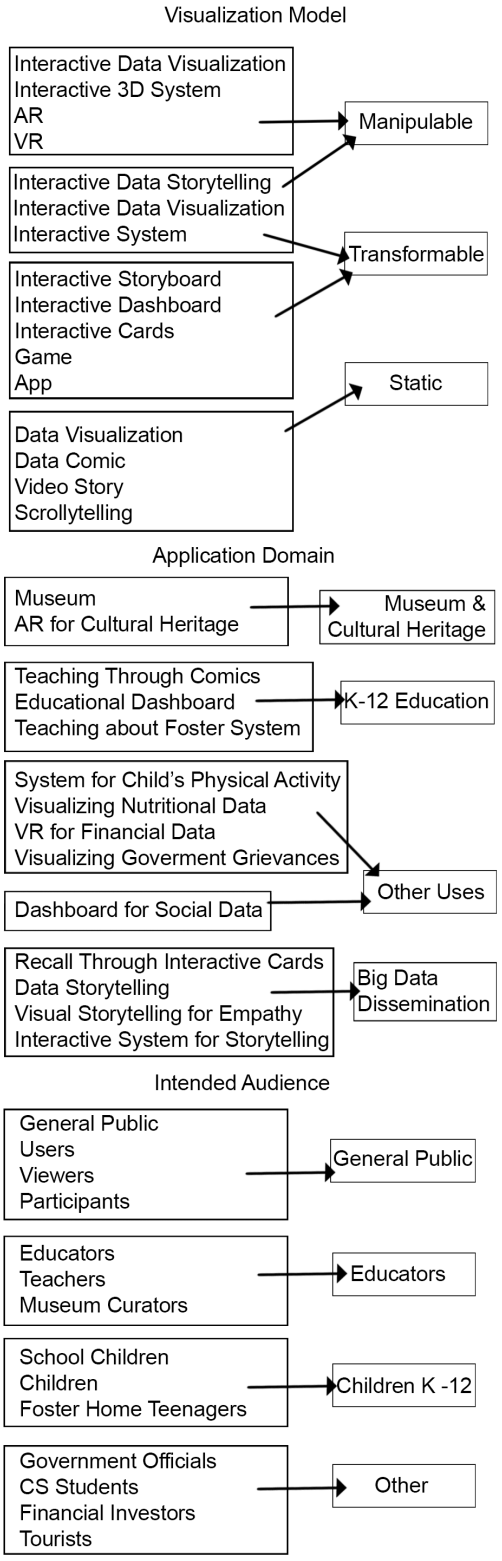


Fig. 2: Thematic synthesis process

TABLE I: Relationships between thematic synthesis categories

	Visualization Models		
	Manipulable	**Transformable**	**Static**
Intended Audience			
General Public	3	3	4
Children	2	1	2
Educators	0	2	1
Other	3	2	1
Application Domain			
Museum & Cultural Heritage	3	2	0
K-12 Education	0	3	1
Big Data Dissemination	3	2	5
Other	2	2	1

sensemaking of data with a non-expert audience. From this we ascertained three themes which provided the principles for designing data exploration software systems.

1) Principle 1: Allow users to find their own patterns in the data. Data exploration requires that a user not only be able to interact with the data but come to their own conclusions about what may be going on in the said data by allowing users to find their own patterns within the data [8], [17]–[20]. The same data often allows for different interpretations depending on the user [21], which can be seen as an advantage in the context of data exploration.

2) Principle 2: Visualize relationships between data sets. It is important to provide a way to visualize the potential connections between different data sets, as data exploration is about finding connections in data [17]. A data exploration system should allow users to sort information via a set of concepts or ideas [22]–[24] rather than having to investigate each data set in isolation.

3) Principle 3: Use artistic expression to engage non-expert users. One of the three principles is that a data visualization should use artistic expression [25], [26] and drive the audience's attention by using visual and narrative elements [27], [28]. An example of this could be data comics which make use of storytelling applied to sequential art or using info-graphics. [29]–[31]

VI. DISCUSSION

In this study, we synthesized literature on data exploration into a set of design principles for creating new data exploration systems for non-expert users. Data exploration is a fairly new sub-field and currently no prescriptive guidelines exist for the process of creating data exploration systems. The design principles introduced in this paper may fill that gap when designing such systems for a non-expert audience and extend the state of the art in prescriptive design knowledge.

Our recommendations, grounded in literature from the field, underline the importance of: 1) empowering the user and allowing the freedom to explore, 2) providing visual ways of interacting with the field, and 3) using artistic expression to engage non-expert users.

Related literature studies exist for example on information systems usability and prescriptive knowledge in the form of design patterns [5]–[7]. In comparison to earlier

studies, we have three novel contributions to the start of the art: 1) We consider the specific intended audience of a visualization or a system, 2) provide prescriptive knowledge specifically to designers of data exploration systems for general audiences, and 3) provide an updated descriptive review to academic audiences.

The main limitation of the study is that while the principles are strongly grounded in literature, they need further testing in practice together with designers and system users. The second limitation of this study is that the systematic mapping review only incorporated papers published between the years of 2016 and 2021, meaning that some seminal papers may not have been included in the review.

VII. CONCLUSION

In this paper, we sought to define a set of design principles by conducting a systematic mapping review, followed by a thematic synthesis. This gave us a perspective on the types of data exploration tools that have been made, and we then derived a set of principles for designing a data exploration system for a non-expert audience, based on the findings and perspectives in the 24 papers selected in the review. These design principles were: 1. Enable users to find their own patterns in the data 2. Visualize the relationships between different data 3. Use artistic expression to attract the attention of a user. In future study we will further evaluate these principles via user testing of a prototype data exploration tool that makes use of these principles in its design.

REFERENCES

[1] E. Osagie, M. Waqar, S. Adebayo, A. Stasiewicz, L. Porwol, and A. Ojo, "Usability evaluation of an open data platform," 06 2017, pp. 495–504.

[2] V. Weerakkody, Z. Irani, K. Kapoor, U. Sivarajah, and Y. Dwivedi, "Open data and its usability: an empirical view from the citizen's perspective," *Information Systems Frontiers*, vol. 19, 04 2017.

[3] L. Garcia, R. Riedy, J. Polman, R. Swanson, T. Nuessle, and N. Garneau, "Expanding public participation in science practices beyond data collection," *Citizen Science: Theory and Practice*, vol. 6, p. 11, 04 2021.

[4] J. Shirk, H. Ballard, C. Wilderman, T. Phillips, A. Wiggins, R. Jordan, E. McCallie, M. Minarchek, B. Lewenstein, M. Krasny, and R. Bonney, "Public participation in scientific research: A framework for deliberate design," *Ecology and Society*, vol. 17, pp. 29–48, 06 2012.

[5] E. Victorelli, J. dos Reis, H. Hornung, and A. Prado, "Understanding human-data interaction: Literature review and recommendations for design," *International Journal of Human-Computer Studies*, vol. 134, 09 2019.

[6] T. Zaki and M. N. Islam, "Neurological and physiological measures to evaluate the usability and user-experience (ux) of information systems: A systematic literature review," *Computer Science Review*, vol. 40, p. 100375, 05 2021.

[7] L. Punchoojit and N. Hongwarittorrn, "Usability studies on mobile user interface design patterns: A systematic literature review," *Advances in Human-Computer Interaction*, vol. 2017, pp. 1–22, 11 2017.

[8] S. Idreos, O. Papaemmanouil, and S. Chaudhuri, "Overview of data exploration techniques," ser. SIGMOD '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 277–281. [Online]. Available: https://doi.org/10.1145/2723372.2731084

[9] R. Mazza, *Introduction to Information Visualization*, 1st ed. Springer Publishing Company, Incorporated, 01 2009.

[10] A. Hosne, "Assessing indigenous forms of writing," *Journal of Jesuit Studies*, vol. 1, pp. 177–191, 03 2014.

[11] A. Blandford and S. Attfield, *Interacting with Information*, 01 2010, vol. 3.

[12] L. Koesten, K. Gregory, P. Groth, and E. Simperl, "Talking datasets – understanding data sensemaking behaviours," *International Journal of Human Computer Studies*, vol. 146, 11 2020.

[13] M. Salama, R. Bahsoon, and N. Bencomo, *Managing Trade-offs in Self-Adaptive Architectures: A Systematic Mapping Study*, 08 2016, pp. 249–297.

[14] D. Cvetkovic, *Introductory Chapter: Multimedia and Interaction*, 04 2019.

[15] D. S. Cruzes and T. Dyba, "Recommended steps for thematic synthesis in software engineering," in *2011 International Symposium on Empirical Software Engineering and Measurement*, 2011, pp. 275–284.

[16] R. Glass and I. Vessey, "Focusing on the application domain: everyone agrees it's vital, but who's doing anything about it?" vol. 3, 02 1998, pp. 187 – 196 vol.3.

[17] Y. Loukissas, *All Data Are Local: Thinking Critically in a Data-Driven Society*, 01 2019.

[18] W. Zeng, A. Dong, X. Chen, and Z.-l. Cheng, "Vistory: Interactive storyboard for exploring visual information in scientific publications," *J. Vis.*, vol. 24, no. 1, p. 69–84, feb 2021. [Online]. Available: https://doi.org/10.1007/s12650-020-00688-1

[19] F. Bosello, C. Colli, P. Clini, I. Dražić, D. Dražina, A. Pugliese, R. Quattrini, and G. Vettorel, "Remember: an adriatic network of virtual museums to highlight ports culture, traditions and history," *IOP Conference Series: Materials Science and Engineering*, vol. 949, p. 012082, 11 2020.

[20] D. Blezinger and E. van den Hoven, "Storytelling with objects to explore digital archives," ser. ECCE '16. New York, NY, USA: Association for Computing Machinery, 2016. [Online]. Available: https://doi.org/10.1145/2970930.2970944

[21] J. Liem, C. Perin, and J. Wood, "Structure and empathy in visual data storytelling: Evaluating their influence on attitude," *Computer Graphics Forum*, vol. 39, 06 2020.

[22] S. Podworny, Y. Fleischer, S. Hüsing, R. Biehler, D. Frischemeier, L. Höper, and C. Schulte, "Using data cards for teaching data based decision trees in middle school," 10 2021.

[23] Z. Koukopoulos and D. Koukopoulos, "A participatory digital platform for cultural heritage within smart city environments," 12th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), 2016, pp. 412–419.

[24] A. Beheshti, A. Tabebordbar, and B. Benatallah, "istory: Intelligent storytelling with social data," 04 2020, pp. 253–256.

[25] S. Ketchell, W. Chinthammit, and U. Engelke, "Situated storytelling with slam enabled augmented reality," in *Proceedings of the 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, ser. VRCAI '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3359997.3365681

[26] Z. Wang, H. Dingwall, and B. Bach, "Teaching data visualization and storytelling with data comic workshops," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–9. [Online]. Available: https://doi.org/10.1145/3290607.3299043

[27] V. Echeverria, R. Martinez-Maldonado, R. Granda, K. Chiluiza, C. Conati, and S. Buckingham Shum, "Driving data storytelling from learning design," 03 2018, pp. 131–140.

[28] V. Echeverria, R. Martinez-Maldonado, and S. Buckingham Shum, "Towards data storytelling to support teaching and learning," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, ser. OZCHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 347–351. [Online]. Available: https://doi.org/10.1145/3152771.3156134

[29] B. Bach, N. Henry Riche, S. Carpendale, and H. Pfister, "The emerging genre of data comics," *IEEE Computer Graphics and Applications*, vol. 38, pp. 6–13, 05 2017.

[30] D. Seyser and M. Zeiller, "Scrollytelling – an analysis of visual storytelling in online journalism," 07 2018, pp. 401–406.

[31] Z. Wang, S. Wang, M. Farinella, D. Murray-Rust, N. Henry Riche, and B. Bach, "Comparing effectiveness and engagement of data comics and infographics," 02 2019.

Stranded Away: Implementation and User Experience Evaluation of an Indie Platformer Game Developed Using Unity Engine

Luka Blašković, Alesandro Žužić and Tihomir Orehovački
Juraj Dobrila University of Pula, Faculty of Informatics, Pula, Croatia
{luka.blaskovic, alesandro.zuzic}@student.unipu.hr
tihomir.orehovacki@unipu.hr

Abstract - **Platformer games have been a staple in the video game industry for decades, dating back to the early days of arcade games such as Donkey Kong and Super Mario Bros. These games continue to be popular today and have evolved to include a wide range of subgenres and gameplay styles. This paper introduces Stranded Away - an indie platformer game developed in the Unity engine. The implemented game retains most of the main aspects of platformers, such as collectibles, moving platforms, different types of enemies, and obstacles, along with puzzle video game elements such as solving riddles to progress. However, the emphasis is placed on core game mechanics, namely player movement and puzzles, as well as tools we have developed to design game levels more efficiently. An empirical study was conducted to examine the user experience of the introduced game. The sample of study respondents consisted mainly of students who first played the game and then completed the post-use questionnaire. The reported findings uncovered to which extent Stranded Away has met the requirements of relevant user experience dimensions.**

Keywords – Videogames; Unity Engine; Indie Game; User Experience; Human-Computer Interaction; 2D Platformer; Puzzle Games; Pixel-Art; Post-Use Questionnaire; Empirical Study

I. INTRODUCTION

Platformer games are a video game genre that involves navigating a character through a series of obstacles and enemies. Originating in the early 1980s, they played a key part in the evolution of video games and even helped game consoles to thrive. Typically, they are defined by their side-scrolling gameplay, where the player controls a character who must navigate through levels filled with obstacles, enemies, and puzzles. The usual objectives include reaching the end of each level, defeating bosses, or collecting certain items or power-ups along the way. These games often rely on precise controls, timing, and strategy to successfully navigate through each level.

In recent years, indie game developers have been able to create and distribute video games on their own, through various channels, including digital distribution platforms such as “Steam” and “App Store”. These indie platformer games often offer unique and creative takes on the classic platformer formula while introducing new mechanics, art styles, and storylines [1].

This paper presents "Stranded Away," an indie 2D platformer developed in Unity. The versatile, user-friendly engine supports cross-platform game creation with built-in components like physics and rendering. Stranded Away blends classic platformer and puzzle elements for a unique, challenging gameplay experience. To evaluate the user experience of Stranded Away, an empirical study was conducted with a sample of mainly student participants who played the game and completed the post-use questionnaire. The results of this study will provide valuable insights for future indie game developers and contribute to the understating of how platformer games can effectively engage and challenge players.

The remainder of this paper is structured as follows. The second section briefly describes the Unity game engine. In the third section, the methodology and theoretical basis of the proposed UX dimensions are presented, while details on the main aspects of the game and its implementation are provided in the fourth section. The findings of an empirical study are reported in the fifth section. Discussion and implications are covered in the sixth section, and the conclusion and future work directions are provided in the final section.

II. UNITY GAME ENGINE

Unity is a cross-platform development environment for video games, developed by Unity Technologies. The environment provides fundamental functionalities, including a rendering engine, sound importation and utilization, physics simulation, animation capabilities, and networking support. Unity provides tools for designing 2D and 3D games for a variety of devices and operating systems. The extensive availability of learning resources and pre-written libraries for frequently performed tasks is one of the reasons for Unity's popularity among beginner programmers. Unity utilizes the object-oriented programming language C#. All languages supported by Unity are object-oriented scripting languages. The object referred to as “GameObject” in Unity serves as the primary building block within the Unity environment. It allows game developers to encompass various elements such as a *character*, *object*, *camera*, *light source*, *terrain*, *image*, *text*, etc. To endow an object with attributes like objects in the real world, such as light, window, or camera, it is necessary to add components to the object.

The highly flexible and user-friendly user interface of Unity plays a crucial role in the utilization and functioning of the software. For instance, it allows developers to further modify and even create custom new inspectors. That can be very useful for adding frequently needed functionalities directly to the editor interface [2].

III. METHODOLOGY

A. Research Framework

To gain a comprehensive understanding of the user experience in Stranded Away, it is crucial to examine various dimensions that contribute to the overall enjoyment and success of the game. By drawing from established models such as the Game Experience Questionnaire (GEQ) by Ijsselsteijn et al. [3] and the Player Experience of Need Satisfaction (PENS) model by Ryan et al. [4], we can formulate research questions that encompass a wide range of user experience dimensions. These dimensions include gameplay mechanics, learnability, user interface sensibility, visual elements, audio elements, engagement, enjoyment, and behavioral intentions. By investigating these aspects, we aim to answer the following research questions:

1. How do players perceive Stranded Away?
2. How does the overall user experience of Stranded Away differ between more experienced and less experienced platformer players?

To evaluate the user experience of Stranded Away, several quality dimensions must be considered, as they contribute to the overall enjoyment and success of the game. Gameplay mechanics (GMH) refer to the rules and systems governing player interaction and actions within the game and are crucial for player engagement and enjoyment, as highlighted by Fabricatore et al. [5]. Learnability (LRN) measures the degree to which the game is easy to learn and master, playing a significant role in player satisfaction and motivation, as discussed by Kiili [6]. User interface sensibility (UIS) refers to the design, organization, and intuitiveness of the game's user interface, which enhances the player experience, as suggested by Desurvire et al. [7] in their heuristics for evaluating playability. Visual elements (VIS) encompass the visual design of the game and its impact on the player's overall experience, as emphasized by Miszal and Schild [8] in the context of video game art and visual design. Audio elements (AUD) pertain to the overall feel, mood, and ambiance created by in-game sound, contributing to the player's immersion and emotional connection to the game, as explored by Jørgensen [9]. Engagement (ENG) is the state of focus, concentration, and immersion experienced by players and is a key factor in maintaining player satisfaction, as reported by Ijsselsteijn et al. [3]. Enjoyment (ENY) reflects the degree to which the player experiences pleasure from playing the game and is crucial in determining the overall success of a video game, according to Vorderer et al. [10]. Finally, behavioral intentions (BEH) represent the likelihood of the player recommending the game, playing again, or purchasing similar games in the future, reflecting the player's overall satisfaction and experience with the game, as supported by Hamari et al. [11] in their study on the behavioral intentions of players in the context of online games. By examining these dimensions, we can better understand the user experience of Stranded Away and identify potential areas for improvement.

B. Apparatus

The user experience in Stranded Away was examined through an empirical study involving primarily students from the Juraj Dobrila University of Pula, along with other individuals with gaming experience across Istria County in Croatia. Data collection was carried out online using a questionnaire created and administered with Google Forms. The first part of the questionnaire consisted of demographic questions, as well as inquiries regarding participants' past gaming experience and habits. The second part of the questionnaire aimed to measure user experience, featuring 40 items divided into eight dimensions. Responses to these items were collected using a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The user experience evaluation procedure was adopted from previous studies [12,13] and tailored to the context of Stranded Away. After participants played the first two levels of the game, which are slow-paced and allowed for thorough exploration and unlocking of the game's features, they were asked to complete the online questionnaire. Finally, the internal consistency of the questionnaire was examined using Cronbach's alpha coefficient, ensuring the reliability of the findings.

IV. GAME INTRODUCTION

Stranded Away is a 2D platformer/side-scrolling game with puzzle elements. The player controls a mysterious space traveler who lands on the planet „Athion" with his spaceship. He is in search of the human species after the galactical apocalypse created by mad scientist dr. Hone - the main antagonist in the game. The player is introduced to the game in the first out of three game levels. Figure 1 represents the opening scene in the first level of the game. All game scenes and visuals are produced in the pixel-art style of graphics.

The story builds up as the player finds data files located on each level. Accordingly, the user unlocks different weapons and power-ups to successfully pass the upcoming levels. Core game mechanics, including player movement, shooting, object interaction, inventory system, and puzzle solving, will be described in the following subsections.



Figure 1. The opening scene in-game

A. Player character movement

Player character movement is an important aspect of every game. Platformers are based on the player traversing through levels and overcoming obstacles thus making character movement one of the core gameplay mechanics.

Player movement in Stranded Away consists of two main parts:

a. Finding the target velocity and then gradually changing towards it over time. The vector is smoothed out by a spring-damper-like function. A spring damper refers to a mathematical model used to simulate physics behavior. It is commonly used in game physics to create realistic motion, such as character jumping, as well as for smoothing out a follow camera [2].

b. Attaching several *colliders* to the player, as shown in Figure 2. In unity, colliders are components that define the shape of an object for physical collisions. For instance, the rounded circle-cast collider will indicate that the player is grounded if it hits anything designated as ground.

B. *Shooting system*

The shooting system in platformers allows players to attack enemies or obstacles in the game world. This is implemented by creating an *arm pivot* object, a common mechanism that allows players to shoot projectiles more realistically and intuitively. It works by using an animation of an arm pivoting at the shoulder joint, along with a *ray cast*, to determine the direction of the shot.

C. *Object interaction*

Object interaction refers to how the player character interacts with various objects within the game world. In Stranded Away, this includes opening doors, pushing, and picking up boxes, and stepping on the pressure plates or jump pads. Object interaction plays a key element in the game since it consists of many riddles that must be solved to progress throughout the game. An example of interacting with an object is presented in Figure 2.



Figure 2. Highlighting doors with white outline

D. *Inventory system*

Inventory systems in video games enable players to manage and organize collected items. With crafting elements, players gather raw materials and finished products for equipping and selling. Crafting mechanics add depth and strategy to gameplay, as players must manage resources and decide which items to craft for easier progression. For example, players collect raw materials for crafting ammunition. Figure 3 illustrates the game's inventory and crafting grid.



Figure 3. Inventory (right) and crafting grid (left)

E. *Gameplay elements*

Gameplay in Stranded Away includes typical platforming elements such as jumping, platforming, coin pickup, puzzle-solving, power-ups, and enemies. The main user interface, shown in Figure 4, consists of ten parts: *1 - ammunition, 2 - selected weapon, 3 - health bar, 4 - shield bar, 5 - oxygen bar, 6 - energy bar, 7 - total coins, 8 - settings, 9 - stats, and 10 - inventory*. Initially, players have a 4-point health bar and a basic blaster gun which has unlimited ammunition, but it's the weakest of them. As they progress, they unlock weapons and can purchase power-ups with collectible coins. The interface also includes settings, stats, and inventory buttons, which have keyboard shortcuts.



Figure 4. User interface in-game

The core game mechanic involves *energy box* manipulation. Players can move them around to activate pressure plates, which trigger objects like *laser doors* and *moving platforms* by emitting visible signals. Four scenarios exist: (1) One pressure plate → Single object, (2) One pressure plate → Multiple objects, (3) Multiple pressure plates → Single object, (4) Multiple pressure plates → Multiple objects. This allows for the design of more complex and challenging riddles for the player.

F. *Enemy AI*

Enemy AI is vital for creating engaging virtual opponents in video games. In Unity, ray casting is a common technique for implementing enemy AI, where a line or "ray" is emitted from a starting point to detect the player's character. This enables dynamic adjustments of AI behavior based on the player's actions and position, as illustrated in Figure 5.



Figure 5. Enemy raycasting

The *Enemy AI* script utilizes a state machine to dynamically control the behavior of enemy characters in the game. The state machine has multiple states including *Idle*, *Roaming*, *Chasing*, *Attacking*, *Spitting*, and *Flying*. Each state represents a distinct behavior or action that the enemy can perform.

G. *Tilemap painter*

Designing levels in a 2D video game can be repetitive and time-consuming, as it often involves manual placement of individual tiles. To address this, a custom system for automatic terrain and tilemap generation was developed, called "Tilemap Advanced Ruleset." This tool streamlines the process within the Unity game engine by allowing level designers to use a prototype tile palette to create a tilemap and apply predefined rulesets. Rulesets are based on the relationship between tiles and their

surroundings, creating a system of logic similar to cellular automata. The tool automatically generates tilemaps using appropriate tiles from pre-defined palettes based on the applied rulesets (Figure 6). Additional tools, like "Tilemap Advanced Generation," were developed to generate decorative elements, such as pillar structures in caves, to enhance the visual appearance of the levels.

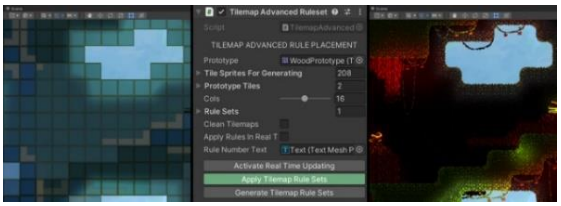


Figure 6. Before and after applying the tilemap generation tool

H. Level generation

The game includes a "challenge" level, "The floor is lava," where players must climb infinitely, dodging rising lava and obstacles while collecting power-ups. A "Level generation" tool with a procedural generation algorithm was implemented for fair terrain generation, using predefined "platforms" to create random yet well-designed sections (Figure 7). The tool also tracks the player's maximum height, enabling additional elements like "money barriers" and height-based effects.

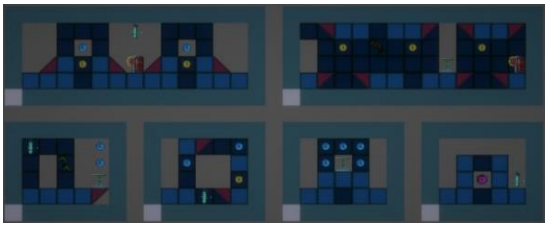


Figure 7. Predefined sets of tiles

This approach in combination with “Tilemap Advanced Generation” not only keeps the level interesting and fresh for the player but also ensures that the player's path is not blocked.

V. USER EXPERIENCE EVALUATION

A. Participants

To gather an understanding of to which extent Stranded Away has met the requirements of relevant user experience dimensions, an empirical study was carried out. A total of 75 participants, mostly students from the Juraj Dobrila University of Pula, Faculty of Informatics took part in the study. The sample was composed of 72% male and 25.3% female respondents while the remaining 2.7% declined to provide an answer in that respect. The majority (44%) of participants were enrolled in the first year of undergraduate study, 24% were enrolled in the third year of undergraduate study, 20% were enrolled in the second year of graduate study while the remaining 8% percents were enrolled in the second year of undergraduate and first year of graduate study. Three participants (4%) were not students. The average age of respondents was 22.15 years (SD = 3.86).

Regarding previous gaming experience and gaming habits, the participants were rather diverse, ranging from casual gamers to avid gamers, with different preferred genres. Participants’ responses related to the most important game feature are summarized in Figure 8. The majority of participants (40%) agreed that enjoyment is the most important aspect of a video game.



Figure 8. The most important game features

Platformers might pose a challenge for players with little to no experience for several reasons, such as precision jumping, timing-based obstacles, challenging enemies, and increasing difficulty. As shown in Figure 9, a majority (59%) of players had no to limited previous experience with platformers, while others were moderate to very experienced in such games.

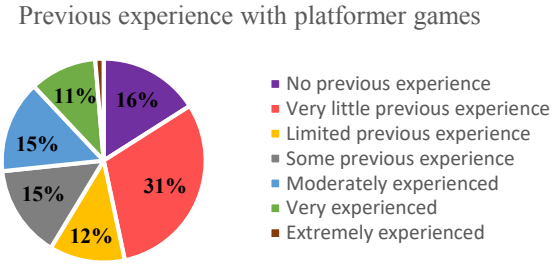


Figure 9. Previous experience with platformer games

Finally, a bar chart presented in Figure 10's bar chart illustrates the most popular video game genres based on a multiple-choice survey. Action and shooter games are the top choices, preferred by 50% of respondents. About a third enjoy survival, strategy, and RPG games, while 25% prefer horror, racing, sports, fighting, and other genres. Platformer games attract only 12% of participants.

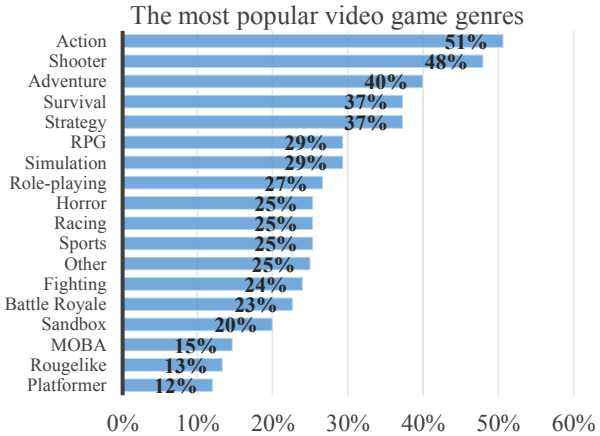


Figure 10. The most popular video game genres

B. Findings

As demonstrated in Table 1, Cronbach's alpha values ranged from 0.639 to 0.880, which is acceptable for exploratory studies and indicates that the reliability of the scales is considered adequate [14].

The independent t-test was conducted to compare the perceived user experience among players with varying degrees of platformer game experience. Players were divided into two groups: the first group consisted of those with "No previous experience," "Very little previous experience," and "Limited previous experience," while the second group included players with "Some previous experience," "Moderately experienced," "Very experienced," and "Extremely experienced." First, the scores were calculated by adding up the values of each item in the construct, and the overall user experience score was determined by summing the scores across all dimensions, as indicated in bold in Table 1. The results of the t-test revealed no significant difference ($p = 0.137$) in the perceived user experience between the two groups, with p-values above the threshold level of 0.05 [15]. This suggests that Stranded Away provides a consistent user experience across players with different levels of platformer game experience, highlighting its accessibility and appeal to a diverse audience.

TABLE I. CRONBACH'S ALPHA AND T-TEST P VALUES ACROSS UX DIMENSIONS

Constructs	Cronbach's alpha	Independent Samples Test p-value
Gameplay mechanics (GMH)	0.785	0.423
Learnability (LRN)	0.639	0.704
User interface sensibility (UIS)	0.768	0.410
Visual elemenets (VIS)	0.878	0.109
Audio elemenets (AUD)	0.843	0.924
Engagement (ENG)	0.848	0.700
Enjoyment (ENY)	0.880	0.156
Behavioral intention (BEH)	0.800	0.067
User Experience (UX)		**0.137**

Utilizing descriptive statistics, the analysis of the collected data revealed that 77% of participants agreed on the well-designed game mechanics, with 21% remaining natural. A significant majority, 90.7%, found Stranded Away easy to learn. Furthermore, 96% of participants perceived the user interface as well organized. In terms of visual elements, 80% found them appealing, 13.3% stayed neutral, and 7% expressed a negative opinion. For the audio experience, 76% of players felt that sound and music enhanced their gaming experience. 68 % found the game to be engaging, and 78.7% considered the game enjoyable. As for future play intentions, 38.7% of participants remained neutral, but over half agreed they would recommend the game to their gaming community and try other games from the same development team. These findings are summarized in Figure 11 (refer to the Appendix).

VI. DISCUSSION AND IMPLICATIONS

Understanding the impact of user experience dimensions on player satisfaction enables developers to cater to diverse preferences and expectations. Study constructs were validated using background work and Cronbach's alpha test, enhancing the research findings. In the context of this study, one strength of using students as a sample is their familiarity with digital technologies and gaming, making them an appropriate and relevant population for evaluating user experiences in video games. Moreover, students tend to have diverse gaming habits and preferences, which can provide valuable insights into the factors that contribute to a positive gaming experience across a wide range of player profiles However, one limitation of using students as a sample is the potential lack of generalizability to other populations, such as older adults or individuals with different cultural backgrounds [16]. The study's findings may not apply to all player demographics, and further research is needed to confirm these insights across a wider participant range. In summary, the insights offered by this study provide valuable guidance for both practitioners and researchers in the field of video game design and development. Incorporating these learnings and addressing study limitations can help the industry create engaging experiences for players across various genres and platforms, informing new game design principles that consider genre and context nuances. Additionally, further research could explore the relationships between user experience dimensions and player demographics or gaming habits to better understand how different player profiles might influence the user experience. Ultimately, the study's findings provide a valuable starting point for both practitioners and researchers to enhance the design, development, and evaluation of video games, ensuring that they continue to provide engaging and satisfying experiences for players.

VII. CONCLUSION

Platformer games, like Stranded Away developed in Unity, have regained popularity due to indie developers. Unity's versatility and user-friendliness make it a popular choice. Stranded Away combines classic platformer and puzzle elements, offering a unique and challenging experience. Various tools were built for tasks like terrain generation and puzzle creation. An empirical study evaluated user experience, demonstrating Unity's potential and the game's engaging gameplay. This study serves as a starting point for developing a UX research framework for video games and reveals positive user perceptions of Stranded Away, primarily due to its well-designed game mechanics, and visual, and audio elements. However, the study's limitations include a small, homogenous sample of students mostly from one university.

REFERENCES

[1] P. E. Dickson, "Using Unity to Teach Game Development", Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education, pp. 75–80. June 2015 https://doi.org/10.1145/2729094.2742591

[2] Unity Technologies. (n.d.). Unity Documentation. Unity. Retrieved January 31, 2023, from https://docs.unity.com/

[3] W. A. Ijsselsteijn, Y. A. W. de Kort, and K. Poels, "The Game Experience Questionnaire," Human Technology Interaction, 2013.

[4] R.M. Ryan, C.S. Rigby, and A. Przybylski, "The Motivational Pull of Video Games: A Self-Determination Theory Approach," Motiv Emot 2006, vol. 30, pp. 344–360. https://doi.org/10.1007/s11031-006-9051-8

[5] C. Fabricatore, "Gameplay and game mechanics design: a key to quality in videogames," In: ENLACES (MINEDUC Chile) -OECD

Expert Meeting on Videogames and Education, pp. 29-31 October, 2007, http://dx.doi.org/10.13140/RG.2.1.1125.4167

[6] K. Kiili, "Digital game-based learning: Towards an experiential gaming model," The Internet and Higher Education 2005, vol. 8, pp. 13-24. https://doi.org/10.1016/j.iheduc.2004.12.001

[7] H. Desurvire, M. Caplan, and J. A. Toth, "Using heuristics to evaluate the playability of games," In Proceedings of CHI EA '04: CHI '04 Extended Abstracts on Human Factors in Computing Systems, Vienna, Austria, April 2004, ACM: New York, NY, USA, pp. 1509-1512. https://doi.org/10.1145/985921.986102

[8] S. Miszal, and J. Schild, "Visual Delegate Generalization Frame – Evaluating Impact of Visual Effects and Elements on Player and User Experiences in Video Games and Interactive Virtual Environments," In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April 2022. https://doi.org/10.1145/3491102.3501885

[9] K. Jørgensen, "How Audio Affects Player Action. Mellen Press," In A Comprehensive Study of Sound in Computer Games, 1st ed.; Jørgensen, K., Jensen, K. B., Eds.; Edwin Mellen Pr: Lewiston, NY, USA, 2009; vol. 3, pp. 120-157.

[10] P. Vorderer, C. Klimmt, and U. Ritterfeld, "Enjoyment: At the heart of media entertainment," Communication Theory, vol. 14, pp. 388-408, January 2004. https://doi.org/10.1111/j.1468-2885.2004.tb00321.x

[11] J. Hamari, K. Alha, S. Jrvel, J. M. Kivikangas, J. Koivisto, J. Paavilainen, "Why do players buy in-game content? An empirical study on concrete purchase motivations, " Computers in Human Behavior, vol. 68, pp. 538-546. https://doi.org/10.1016/j.chb.2016.11.045

[12] M. Bošnjak, and T. Orehovački, "Measuring Quality of an Indie Game Developed Using Unity Framework," In Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 21-25 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1820-1825. http://dx.doi.org/10.23919/MIPRO.2018.8400283

[13] R. Šajina, and T. Orehovački, "User Experience Evaluation of 2D Side-Scrolling Game Developed Using Overlap2D Game Editor and LibGDX Game Engine," In Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 21-25 May 2018, IEEE: Piscataway, NJ, USA, pp. 1826-1831. https://doi.org/10.23919/MIPRO.2018.8400284

[14] D. Russo, K.-J. Stol, „PLS-SEM for Software Engineering Research: An Introduction and Survey," ACM Comput. Surv. 2022, 54, 1–38. https://doi.org/10.1145/3447580

[15] H. Levene, "In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling," I. Olkin et al. eds., Stanford University Press, 1960. pp. 278-292.

[16] D. Compeau, B. Marcolin, H. Kelley, and C. Higgins, "Research commentary—Generalizability of information systems research using student subjects - A reflection on our practices and recommendations for future research," Information systems research vol. 23 2012. pp. 1093-1109.

APPENDIX

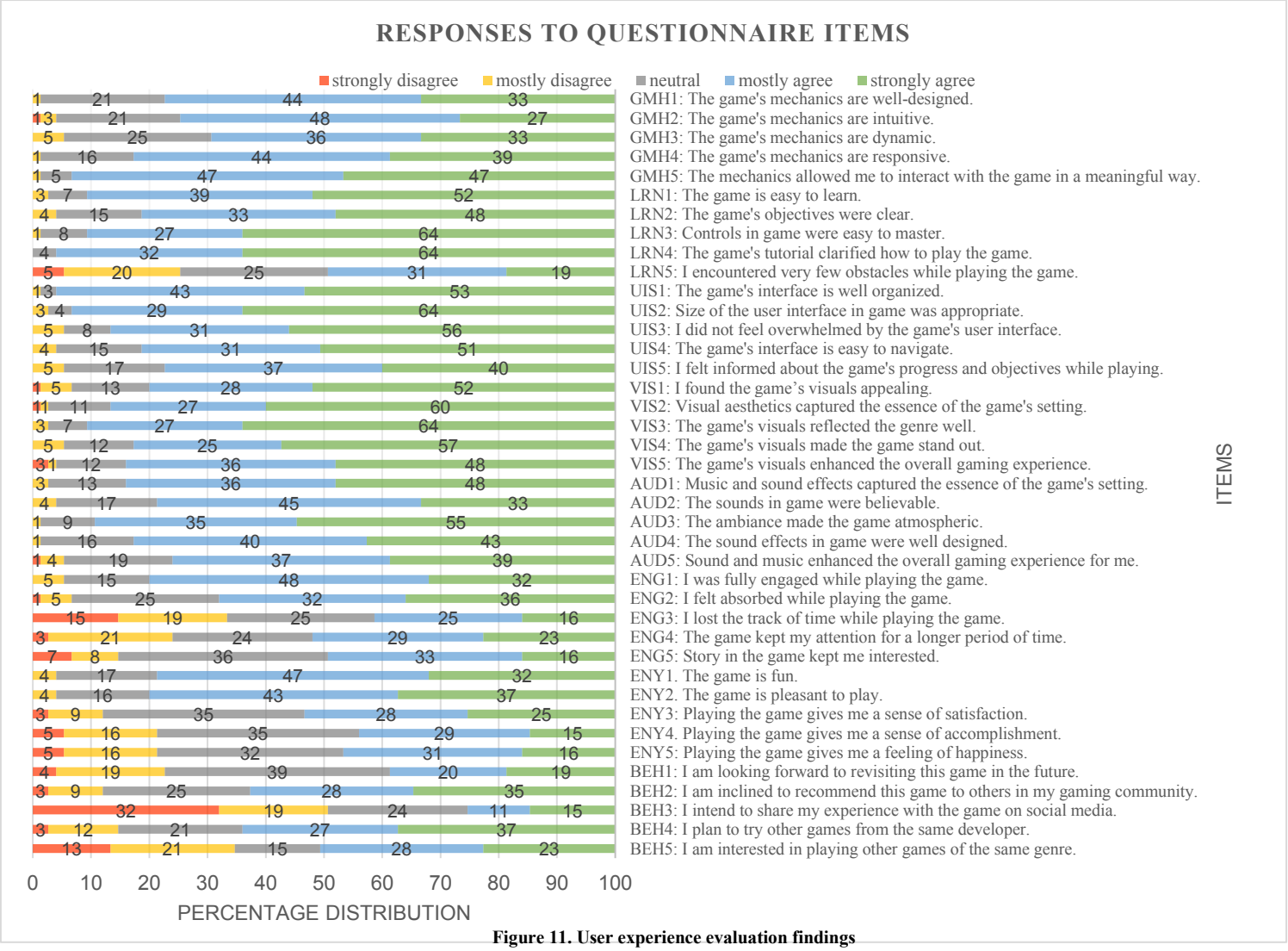


Figure 11. User experience evaluation findings

Analyzing the Impact of Digitality and AI on HCI: Four Dimensions of Intelligent Digitality That Afford Innovative Designs

D. Te'eni
`Coller School of Management, Tel Aviv University, Israel
teeni@tau.ac.il

Abstract - **Interactive games on the Internet often reflect aspects of 'serious' behavior and attitude found in the real and virtual worlds. They can be used to analyze how digitality may be designed to attain positive individual and social impact. Consider a game such as Wordle and think about how different it is from a printed word puzzle, and moreover, think about how even more different you can make it by adding more digitality in the game's form and the way it is played. This paper analyzes the distinct, albeit interrelated, dimensions in which intelligent digitality can innovate the product, the process, the agency and the environment of human-computer interaction. We build on previous HCI research and metaphors to provide a framework for incorporating digitality, especially AI, to produce innovative designs, demonstrating these ideas with the (imaginative) design of Wordle.**

Keywords – **digitality; HCI design; AI**

I. INTRODUCTION

Online behavior has become pervasive in our lives to the extent that we cannot ignore it in understanding human behavior and attitude online and offline. The distinct core of online behavior is the digitality than enables new behaviors it and impacts the human-computer interaction (HCI). While there are numerous studies on how elements of digitality, such as media richness, affect behavior, designers need a systematic approach to identify the potential innovations to HCI that digitality affords. Furthermore, in light of the growing pervasiveness of AI in online behavior, the potency of 'smart' (intelligent) digitality should be emphasized in innovating human-machine interaction.

This paper proposes a four-dimensional framework that identifies innovative functionality and usability afforded by digitality. Additional functionality, however, may bring higher complexity to the HCI, which underscores the roles of usability and flow in design but also plays a role in the allocation of tasks and responsibilities between human and AI-based agents. The framework, therefore, is also meant to assess complexity on each dimension and on the combination of dimensions.

Games are good for analyzing discretionary online behaviors, their antecedents and their effects [1]. This paper uses the online game Wordle to conceptualize the dimensions of the technology's impact on behavior, with special attention to the potential contribution of AI on each of the dimension. The analysis highlights several aspects that determine the potential of digitality to innovate designs and determine their resulting complexity. Furthermore, games underscore the hedonistic expectations of users that determine intentions to play the game and the user's experience with the game [2].

Wordle (https://www.nytimes.com/games/wordle) is a popular game, albeit a very specific type of online games, very different from video games. It is a puzzle type of game that has grown very popular in the past year to become a daily practice for millions of users round the world. The game has been translated to several languages, and has been studied in several domains, e.g., [3], [4].

Players can play the game only once a day. The player has six attempts to guess the correct five-letter word of the day. In each attempt, after considering all previous attempts, the player offers a legal candidate word. The candidate word is checked and the player receives the following feedback: each of the five letters is painted in black (to designate a letter that does not appear in the correct word), yellow (to designate a letter that appears in the correct word but in a different position), or green (to designate a letter that appears in the right position of the correct word). After receiving the feedback, the player proceeds to the next attempt. When the correct word is found, the player is congratulated according the number of attempts made (e.g., on the sixth attempt the feedback is 'phew'). Indeed, special attention will be given to the role of feedback. If the six attempts are exhausted before the correct word is identified, the player is notified of the correct word. In both cases, the new score is recorded and added to the player's history, which can easily be shared with others.

II. FOUR DIMENSIONS

We use this game to demonstrate four dimensions in which digitality can affect the way the game is played: product, process, agency and environment. These dimensions correspond to four questions about the game, namely, 1) what is the form (structure) of the game, 2) what is the process of playing the game, 3) who is

playing the game, and 4) who and what surrounds the player playing the game.

A. *Product*

For the first dimension, product, we ask what is the form and what is the functionality of the digital product. Potentially, the digital product can be very different from the non-digital product due to additional forms and functions afforded by digitality. Asking what in Wordle (the digital product) would not be feasible in a similar (non-digital) printed puzzle, demonstrates some of the digital characteristics in play. In Fig. 1, the screen has two parts: the upper part is a 5*6 matrix (columns for five letters and rows for six attempts), and the lower part is a keyboard presentation of the letters (QWERTY style) to show the accumulative feedback across attempts per letter.

Several aspects of digitality are relevant to the design of the form and functionality of the product such as interactivity, connectivity, channel capacity and adaptivity. We use the first two attributes to demonstrate how to consider digitality in the product.

Interactivity is an obvious digital characteristic that affords feedback [5]. The feedback is at the level of a single character, at the level of a word after each attempt, and at the level of a game after the correct word is found or after six failed attempts. The first type of feedback is instant feedback when typing a single character in a cell (e.g., a character that is not an English letter). A second feedback generated after each attempt consists of word level feedback shown in the upper part of the screen (Fig. 1), and additionally, the feedback is aggregated across attempts and shown at the level of a letter in the lower part. A third form of feedback is performance (outcome) feedback provided at the end of the game, as shown in Fig. 2. The performance statistics calculated by aggregating past games from previous days substantially change the non-digital game.

Feedback changes the process of playing the game (as explained in the next section) but it also changes the intention to use the game. In hedonic systems, perceived enjoyment and anticipated psychological wellbeing are as influential as perceived usefulness and ease of use in determining intentions to use [2,6,7]. Players become accustomed (if not addicted) to seeing their performance scores (Fig 2) to the extent they will not play if they cannot see their scores, much like joggers who will not jog unless they have their smart watches. (Wordle shows past scores only on your regular device.)

Connectivity is another digital characteristic critical for innovating HCI [8]. It affords links and relationships with others that are infeasible in the non-digital games. In particular, the easy ability to share results invites interactions with friends through social media. The additional functionality enabled by connectivity changes substantially the product, as does the feedback, affecting the user's perceived enjoyment or value, and, as a result, motivation to play.

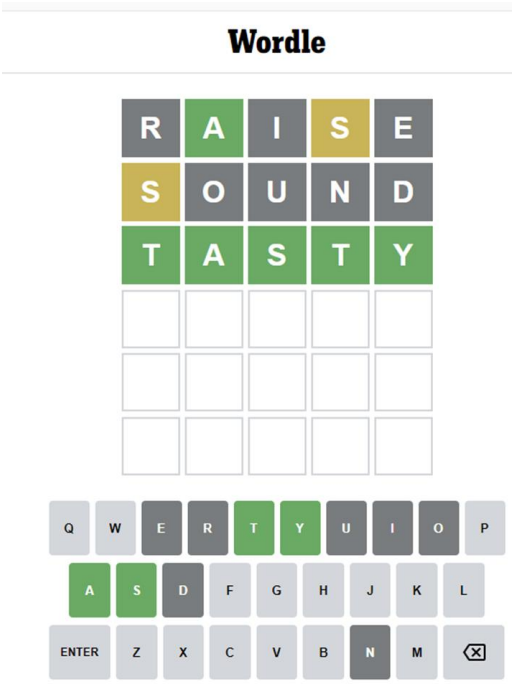


Figure 1. On the left, the word of the day 'Tasty' is guessed on the third attempt. On the right, six attempts failed to guess. Green is a letter in the word in its correct position, a Yellow letter is included in the word but is not in its correct position and Black letters are not in the word

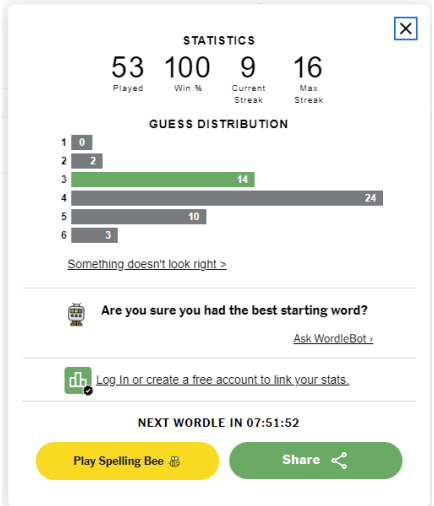


Figure 2. Additional functions in the digital product are the performance statistics and ability to share

B. *Process*

For the second dimension, process, we ask what is the user's process of using the product to achieve some goal. In the case of Wordle, what is the process of playing the game? One way of describing the process is to break it down into several steps such as deciding on a strategy, executing the strategy, studying the feedback, and revisiting the selected strategy and either continuing to execute it or revising the strategy and then executing it. Thus playing Wordle can be seen as following up to six iterations in a cycle of strategy-execution-feedback. For example, a strategy may consist of using the first two

attempts to find all or most of the vowels that appear or do not appear in the correct word and accordingly suggesting words that maximize the reduction of uncertainty (possible words given prior attempts). In executing the strategy, the first attempt could begin with a word that includes vowels, and according to the feedback continue with a second word of different vowels (see first two attempts in Fig. 1).

In the game represented in Fig. 3, the player revisited the strategy, assuming that most of the vowels had been identified so that there was no need to use the second attempt to identify further vowels. Accordingly, the second attempt began guessing the correct word. The third attempt failed to choose a word that minimized uncertainty, and resulted in a failed game (the correct word 'TASTY' is shown after six attempts).

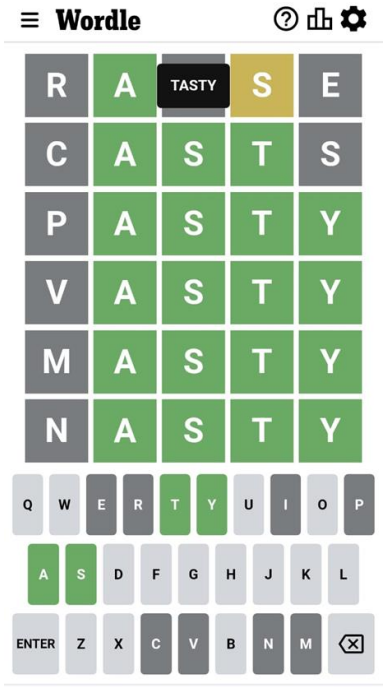


Figure 4. Inappropriate execution of a strategy resulting in a failed game

A comprehensive analysis of how digitality affects the process when it involves HCI must address the allocation of tasks between the human and the machine as explained in the next section on agency. Nevertheless, we can begin the analysis by ignoring for now which agent performs the tasks. The tasks in the strategy described above clearly rely on feedback afforded by digitality, primarily on the feedback at the level of a word after each attempt. More on feedback in the next section.

Actual behavior does not always follow the planned strategy-execute-feedback cycle flawlessly. Intelligent agents are expected to deal effectively with unplanned events or erroneous implementation of strategies. Some players request tips how to progress when strategy execution becomes too difficult. The three daily tips generated automatically for the correct word 'TASTY' were the following: 1) This word is an adjective; 2) It is used to describe food that has a pleasant, distinct flavor, and 3) It features a 'T'.

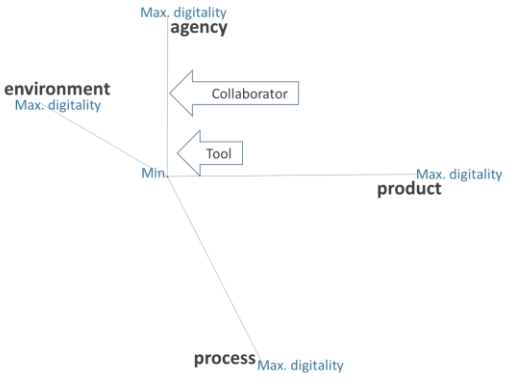


Figure 3. Four dimensions to incorporate digitality

C. *Agency*

A third dimension on which digitality affects HCI is agency [9]. For agency, we ask who governs, initiates, controls and bears the responsibility for the process and its sub-processes. Agency can be characterized on a continuum according to the latitude of decision making granted to the agent, e.g., [10].

In the case of Wordle, complete human autonomy is when the player has complete control over the process and sub-processes so that the player alone initiates and takes charge of all actions, even though execution of the player's orders may be executed by the machine. An autonomous machine is one that acts as its own agent. An autonomous game would be one that plays and displays the results when the game is over. In between these extreme cases are many combinations of human-machine configurations [11], often referred to as 'Keep the Human in the Loop' [12]. These configurations call for a shift away from the metaphor of computer as tool to a metaphor of collaboration between human and machine.

Under the collaboration metaphor too, human autonomy is seen as an important value in the design of HCI [13]. Indeed, in games, enjoyment and human autonomy should be design goals [14]. Nevertheless, the human-machine collaboration configuration introduces several complications that need to be addressed in their design. Collaboration assumes task allocation and intelligible communication between human and machine [11]. Smart digitality can enable or support both allocation of tasks requiring intelligence to the machine and intelligible communication between human and machine. For instance, a possible division of labor could be the player setting a strategy of first identifying which vowels appear in the correct word and communicating the message to the machine, and consequently, the machine suggesting the best words to implement the strategy. In such a division of labor, the player's sense of autonomy would not necessarily decrease because the higher level decision remains with the human [15].

A more complex collaboration would be to adapt the machine's algorithm to the human priority on which of the word's five positions to concentrate. For example, say the player wants to resolve positions 4 and 5, and asks the

machine to do so. Most machine algorithms for solving Wordle operate at the level of reducing uncertainty for the entire word, e.g., [16]. The incompatibility between the methods would require challenging algorithms to integrate the two and enable effective communication.

D. Environment

For the fourth dimension, environment or context, we ask who and what surrounds the user using the system. In the case of Wordle, in what virtual environment is the game played and what agents are part of the environment that interact or perhaps compete with the player, e.g., avatars that offer tips when asked.

Say the game is played in a virtual world, in which the player, represented by a similar looking avatar, can buy tips from other avatars dressed various experts. Thus an element of bargaining can add to the game's challenges.

Fig 4. sums up the four dimensions, product, process, agency and environment, that represent the opportunities to incorporate digitality in the design of HCI.

III. SMART FEEDBACK AND ADAPTIVITY

Feedback that utilizes AI ('smart feedback') can adapt its content and timing according to the player's state, skills and style, as well as to the game situation, e.g., the number of attempts already made. For instance, if the player is risk aversive, tips or suggested words can be tailored to be more conservative relative to suggestions compatible with an 'I'm feeling lucky' mood and ready to make a wild bet.

Adapting (tailoring) tips to the particular player and game situation is another example of smart responses. The three daily tips mentioned above, could be tailored to the first few attempts. For example, after the first two according to the history of attempts. The third tip "it features a 'T'." would be relevant after the first two attempts in Fig. 1 but not after the first two attempts in Fig. 3. The tips could additionally be tailored to the player. For example, the first tip ('… is an adjective') would fit a person who generally does not use adjectives. Moreover, if the machine is given privileges to decide when to provide a tip, it could detect signs of a user facing complexity, e.g., an unexpected pause, errors in writing a candidate word, or obviously redundant attempts that bring no new information. Thus based on a smart analysis of both past behavior and current situation game, the feedback or help can be adapted to the user and the game situation in terms of content and timing.

Adaptation that relies on intelligence is relevant to each and all four dimensions. For example, when an AI-based agent selects a strategy, it may (unlike the human strategy described above) select a sequence of candidate words that consistently and sequentially reduce uncertainty (as noted above in [16]). The feedback to the agent will be adapted accordingly. Furthermore, agency itself may change dynamically according to the results.

IV. COMPLEXITY

Fig. 4 specifies the opportunities for digitalization in HCI, but the four dimensions can also serve to assess the complexity of the human-computer interaction introduced into the new design. Firstly, digitality affords new functionality that almost always brings with it new complexity on the relevant dimension, and secondly, the interrelationship between the dimensions results in a combined effect that may further increase the complexity of the HCI.

Recently, a new version of Wordle has quadrupled it into a game called Quordle, greatly increasing the complexity of playing the game. In Quordle, the player has nine attempts to guess correctly *four* daily five-letter words simultaneously. The screen shown in Fig. 4 is divided into four matrices, each showing the history of attempts for one of the four correct words ('WRYLY' 'APPLE' 'SCOUR' 'CLUNG'). Accordingly, a matrix has five columns and nine rows. As in Wordle, the player offers a candidate word (In Fig. 4 the first candidate word is 'RAISE'), which is applied to all four matrices. This expansion would be practically infeasible in a non-digital version. The increased complexity on the product dimension is clearly visible.

Part of Quordle's additional complexity compared with Wordle (with a single matrix for a single correct word) is the need to manage attention between four correct words and the respective matrices. Consider the challenge in thinking of a candidate word so as to progress according to the feedback on previous attempts towards one correct word versus four correct words simultaneously. The players must manage their attention shifts from one matrix to another, recalling their thoughts about the previous matrices. Moreover, consider the complexity of reading the integrated feedback on a common keyboard (not shown in Fig. 5). The increased complexity compared with the original Wordle is the result of higher digitality on the dimension of the product (the four matrices) and on the dimension of the process (the attention and comprehension of the integrated feedback).

The complexity of the human-computer interaction created by the additional functionality afforded by digitality is due not only to complexity on each dimension but also to that generated by the combination of interrelated dimensions. For instance, the higher decision-making latitude of AI-based agents may require better coordination between processes. In Wordle, the differences in human and artificial strategies may require extra coordination and mutual understanding if the machine works at the level of a whole word and the human is interested in a candidate word to resolve a question about vowels at the end of the word.

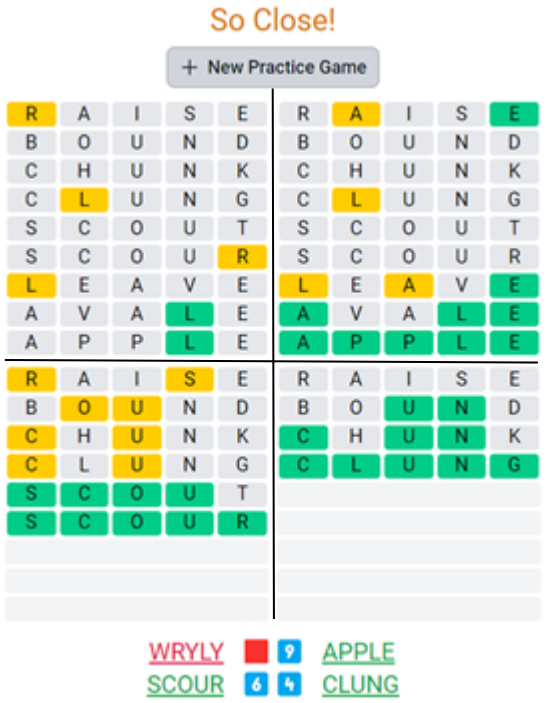


Figure 5. . Quordle with four words to guess simultaneously is highly complex

CONCLUSION

The four dimensions shown in Fig. 4 direct the designer's attention to opportunities that leverage digitality to innovate the design of HCI. Furthermore, taking feedback as one of the most impactful attributes of digitality, we show how AI in particular can leverage the capacity of digitality to innovate by adapting the content and timing of feedback to the user and the game situation. With added functionality afforded by smart digitality comes, however, added complexity, which can be reduced and managed, partly by adapting to the user's needs and preferences and by adjusting agency. Hence, the need to use the framework to consider at once both the potential impact of digitality and it consequent complexity.

This work can be seen as a first step towards a more general framework for leveraging digitality in the design of HCI. It took an individual perspective of a user and a machine that should be expanded to collective use of multiple agents. The framework should also be extended to incorporate specific procedures to identify possible contributions of digitality to each dimension. For example, specifying in advance likely attributes of digitality (e.g., interactivity, connectivity, capacity and adaptivity) that affect each dimension. More generally, the framework can be expanded by taking two perspectives: HCI and AI. For example, the HCI perspective considers both complexity and perceived complexity as two separate factors. In the analysis of Wordle, perceived complexity would be expected to affect the user's perceived enjoyment but also decisions of agency. The AI perspective for example, should consider the algorithmic implications of considering responsibility alongside authority when tasks are delegated to AI-based agents. Such extensions and expansions of the framework will hopefully emerge as the framework is put to use.

REFERENCES

[1] P. Penix-Tadsen, "Why we should take video games seriously (and when we shouldn't)", Latin American Research Review, 48(1), pp. 174 - 190 (2013).

[2] D. Peters, R.A. Calvo, R.M. Ryan, "Designing for motivation, engagement and wellbeing in digital experience." Frontiers in psychology, pp. 1-6, article 797, 2018.

[3] C.L. Liu, "Using wordle for learning to design and compare strategies." 2022 IEEE Conference on Games (CoG). IEEE, 2022.

[4] D. Bertsimas and A. Paskov, "An Exact and Interpretable Solution to Wordle." Available at URL.(Accessed: 12 February 2023) (2022).

[5] D. Te'eni, "A cognitive-affective model of organizational communication for designing IT." MIS Quarterly (2001), pp. 251-312.

[6] H. Van der Heijden, "User acceptance of hedonic information systems," MIS Quarterly, 28 (4), pp. 695-704, 2004.

[7] H. Van der Heijden, H. Ogertschnig, and L. van der Gaast, "Effects of Context Relevance and Perceived Risk on User Acceptance of Mobile Information Services". In: *Proceedings of ECIS 2005*. Regensburg, Germany, 2005.

[8] P. Wang, "Connecting the parts with the whole: Toward an information ecology theory of digital innovation ecosystems." MIS Quarterly 45.1 pp. 397-422. (2021).

[9] J. Tholander, M. Normark, and C. Rossitto. "Understanding agency in interaction design materials." Proceedings of the SIGCHI conference on human factors in computing systems. 2012.

[10] A. Baird and L.M. Maruping, "The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts." MIS quarterly 45.1 (2021).

[11] L. A. Suchman, "Human-machine reconfigurations: Plans and situated actions", Cambridge University Press, 2007.

[12] A.Holzinger, E. Weippl, A. M. Tjoa, P. Kieseberg, "Digital transformation for sustainable development goals (sdgs)-a security, safety and privacy perspective on AI". In Machine Learning and Knowledge Extraction: 5th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2021, Virtual Event, August 17–20, 2021, Proceedings 5 2021 (pp. 1-20). Springer International Publishing.

[13] B. Friedman, "Value-sensitive design", Interactions 1996, 3, pp. 16–23.

[14] F. da Rocha, T. Filho, P. Mirza-Babaei, B. Kapralos, and G. M. Mendonça Junior, "Let's Play Together: Adaptation Guidelines of Board Games for Players with Visual Impairment", Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–15.

[15] D. Kumar and N. Srinivasan, "Naturalizing sense of agency with a hierarchical event-control approach", PLoS One. 2014 Mar 18;9(3):e92431.

[16] S. Bhambri, A. Bhattacharjee, and D. Bertsekas. "Reinforcement Learning Methods for Wordle: A POMDP/Adaptive Control Approach." arXiv preprint arXiv:2211.10298 (2022).

.

Evaluating the Impact of Different User Engagement Elements on Software Applications

Mohammad Hajarian, Paloma Diaz, Ignacio Aedo
Universidad Carlos III de Madrid, Department of Computer Science, Leganes, Spain
Mohammad.Hajarian@uc3m.es

***Abstract*— User engagement elements are important aspects of any software application. The presence or absence of these elements can have different levels of impact on applications' success or failure. However, to the best of our knowledge, the presence or absence of these elements in successful and unsuccessful software applications has not yet been analyzed to help software engineers understand the impact each of these elements can have on applications. To fill this gap, in this paper, based on the reviewed literature, we have categorized the user engagement elements into eight main elements and evaluated them against 50 successful and unsuccessful software applications. Evaluation results indicate that utility, interactivity, and design quality have been highly used in successful software applications, while gamification, personalization, and social networking have been almost absent in unsuccessful applications. We have discussed the implications of these findings and given some recommendations for improving user engagement in software applications.**

Keywords*— *User engagement, software application, utility, interactivity, design quality, social networking, personalization, gamification, content quality, privacy

I. INTRODUCTION

One of the essential factors of any software application is user engagement. Collaborative software applications such as social media or instant messaging can quickly fail without appropriate user engagement, and applications with less user engagement are less widely used [1]. As a result, user engagement can tremendously affect applications' success or failure [2]. User engagement is crucial for enabling some applications to make billions [3]. Having said that, any application can succeed by increasing user engagement, which can be done by adequately implementing user engagement elements. Increasing user engagement not only makes the application's current users keep coming back and using it, but it also gains new users for the application by increasing word of mouth (WOM) about the application [4].

Considering the importance of user engagement in software applications, knowing the elements that can affect it is necessary. In some literature, user engagement was defined as the quality of the user interaction and experience with the application [5]. However, beyond this definition, many criteria can impact the quality of user interaction. Some of these criteria and elements were mentioned in several studies [6, 7]. The design quality is among those elements that can impact user engagement. Ramírez-Correa et al. (2019) [8] found that the perceived ease of use, enjoyment, and usefulness increases user engagement in social networks for older people. Other criteria that can affect the quality of design and interaction are interface aesthetics, personalization, reinforcement, communication, navigation, credibility, and message presentation [9]. The interactivity of the application is another element that impacts user engagement. Criteria such as the presence of clickable items, interaction prompts, reminders, and different types of notifications can determine the quality of interactivity, which can impact increasing user engagement [9, 10].

Torous et al. (2018) [1] suggest that to have an engaging application, it should solve the user's problems and give the user the service their desire. In other words, it should have a good utility for the users. Similarly, Tian et al.'s (2021) study on mobile travel apps revealed that compatibility with users' expectations and needs is an important factor that impacts user engagement along with an attractive user interface and ease of use [11]. The importance of application utility and other criteria such as customization, privacy, user interface, and the credibility of the information was also mentioned in [12]. The study results in [13] suggest that privacy concerns lead to decrees in user engagement in the long run. The findings of [13] imply that the users who enjoy the application are more concerned with privacy issues, and if these concerns do not address properly, user engagement will be decreased. Some criteria that impact user engagement because of privacy concerns in the software application are clear opt-in and opt-out policies [14] and transparent data protection notices [13].

In addition to elements such as utility, privacy, trust, enjoyment, and type of language used in the application, personalization was another element that impacted applications' user engagement [15, 6]. Personalization has been mentioned in several recent studies as another element that can increase user engagement in software applications [16]. The finding in [17] reveals that personalizing application content based on the users' personalities and emotions can increase user engagement with the application. Criteria such as giving personalized feedback to the users and allowing them to customize when and how to receive reminders can be considered personalization [9]. Personalization can also be made using different recommender systems, such as collaborative or content-based ones [18]. Moreover, personalization can be made on the user interface, known as customization, by enabling users to change their background music or other interface features [10].

Nevertheless, other elements, such as gamification, are also playing a role in software applications' user engagement [19]. Today gamification is used in various software applications with different utilities, such as

educational applications, to change user behavior and increase user engagement [20]. Amo et al. (2020) [21] found that gamification criteria such as points and leaderboards can affect user competence and increase user engagement. In another study, the gamification level effectively increased user engagement [22]. Another element used in software applications to increase user engagement is the quality of information or content. In [23], the authors found that the quality of the information in the application increases user satisfaction in health application sites by increasing site reuse and perceived benefits. This aligns with [24], which stated that the quality of information, system, and service positively impacts user engagement. Sharing content and spreading information is among other criteria that increase user engagement. These criteria can be translated to social networking features [7, 25]. Social networking allows users to connect and like each other posts and profiles [26, 27]. These types of interactions and communications among the application's users can increase user engagement and improve the users' trust in the application [28].

As can be seen, different user engagement elements have been mentioned and studied in different literature. However, some of these features are present in some software applications, while others are not present in other applications. Hence, it raises the question of which of these user engagement elements is more important and can highly affect the success and failure of an application. The answer to this question is important because software engineers and developers can use this information to prioritize their tasks and objectives in developing software applications. However, to the best of our knowledge, this question remained unanswered in the literature. Hence, this research aims to fill this gap by investigating the presence of criteria mentioned in the literature that affect user engagement in software applications to find the most important and less important user engagement elements. Understanding the impact of implementing each user engagement element is the first step to developing a successful application with high user engagement. As a result, by evaluating the presence of user engagement elements in both successful and unsuccessful software applications with respect to their criteria in previous works, we can identify the impact of each of these elements and then give appropriate recommendations for implementing user engagement in software applications based on these findings. Therefore, we define our research questions as follows:

- RQ1: which user engagement elements are mostly used in successful software applications?
- RQ2: which user engagement elements were missed in unsuccessful software applications?
- RQ3: What mistakes did unsuccessful software applications make, and how may they address these mistakes to improve user engagement?

To answer these questions, in the next section, we have described how we have used our theoretical framework to evaluate several successful and unsuccessful software applications.

II. Research Method

To evaluate the impact of different user engagement elements on the success or failure of software applications, first, using the reviewed literature in section 1, we have categorized the user engagement elements into eight categories, including interactivity, social networking, gamification, personalization, content quality, design quality, utility, and privacy [1, 4, 6, 9, 12, 7, 25, 10, 29]. Then based on the criteria of each user engagement element that was mentioned in Section 1, we evaluated several successful and unsuccessful mobile applications by analyzing the degree of adaptation of these criteria and the effectiveness of adaptation. To this end, during the evaluation, we inspected which user engagement elements have been used inside each application. To this end, we have calculated the usage level and visibility level of each engagement element inside the application, and we wrote down a score for each user engagement element for each application on a scale from 0 to 10.

In our study, we downloaded 50 apps from the google play store for evaluation. Selecting this number of applications for evaluation is inspired by a study in [10] about user interaction with iOS apps. We have considered download counts and user rankings to determine whether an application is successful or unsuccessful. The reason for using a combination of both download counts and ranking was that some apps are new to the store or might have a low number of installations. Hence, they are not well-established to understand their success or failure. Moreover, applications with very low download counts might not have a correct ranking since few users might rank them, and they could be biased. Figure 1 shows the theoretical framework of this study.

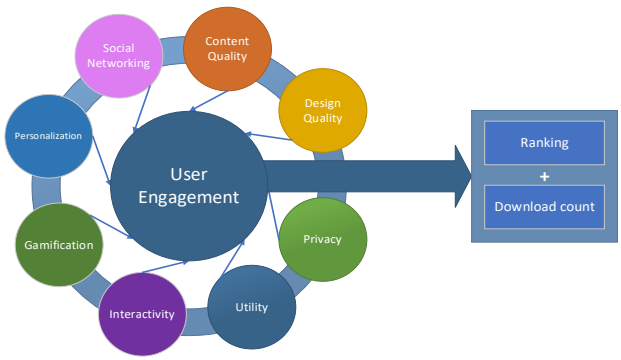


Figure 1. The theoretical framework for evaluating the impact of different user engagement elements on software applications.

Of the 50 evaluated applications, 30 were selected from successful applications with over 100 million download counts and ranked more than 3.7 out of 5. These applications were selected using Google's top application list, Wikipedia's most downloaded apps list, and the businessofapps.com list. We have also excluded games and preinstalled apps. Moreover, 20 apps were selected from unsuccessful applications with 1K to 50K download counts and user ranking below 2.5. However, finding unsuccessful apps was difficult because app stores usually promote top apps, not worse ones. To find such applications, we conduct a custom search in the google play store by entering the subject of the apps based on the title of the apps reviewed in the first stage. Type of applications used for

the evaluation was health and fitness, communications, photography, social media, travel, tools, media streaming, cloud storage, auto and vehicles, news, magazines, music and video, entertainment, dating, lifestyle, weather, shopping, productivity, social, simulation, and education.

III. Results

After evaluating all 50 successful and unsuccessful applications, the average score was 5.37 (S.D = 3.43), with a median of 6.12. The highest score was 9.62, and the lowest score was 0. Figure 2 gives us an overview of the usage level of each user engagement element in all evaluated applications.

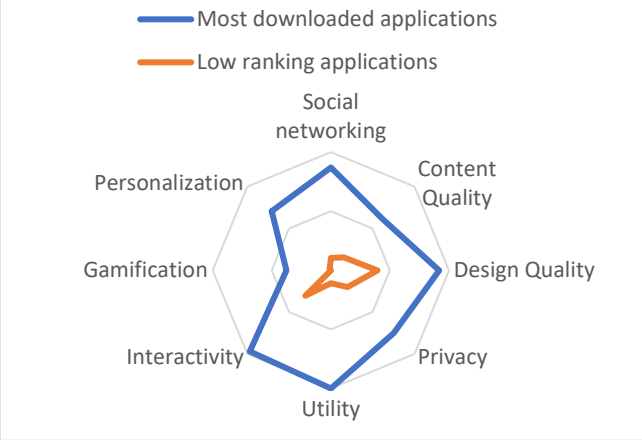


Figure 2. Visual presentation of using each user engagement element in successful apps versus unsuccessful apps.

To inspect this in detail, figure 3 compares the overall percentage of using each engagement element in 30 successful and 20 unsuccessful applications.

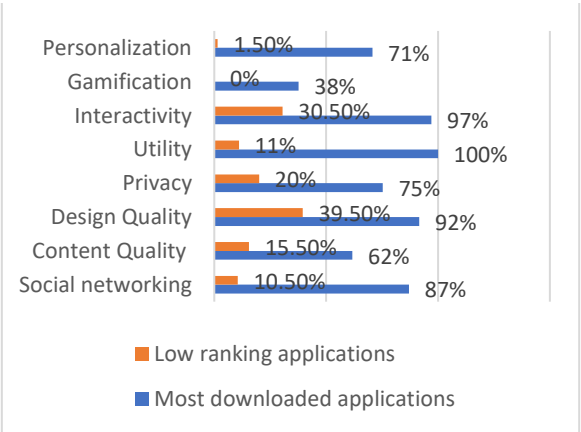


Figure 3. Comparing the user engagement element usage between successful applications and unsuccessful applications.

Moreover, figure 4 illustrates the boxplot of received scores for each user engagement element in successful and unsuccessful applications (on a scale of 0 to 10).

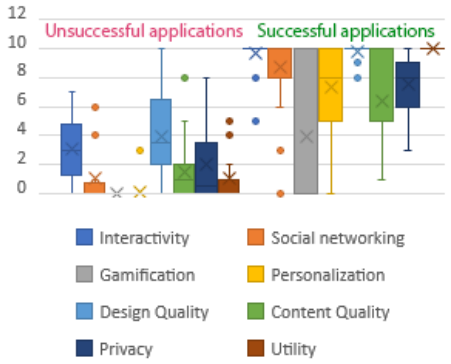


Figure 4. Boxplot of engagement elements usage in successful and unsuccessful applications.

Moreover, table 1 shows the obtained score differences between the successful and unsuccessful applications.

Table 1. Difference between the highest and lowest scores.

User engagement element	Difference
Utility	89%
Social networking	76.5%
Personalization	69.5%
Interactivity	66.5%
Privacy	55%
Design quality	52.5%
Content quality	46.5%
Gamification	38%

Figures 2, 3, and 4, in addition to Table 1, gave us insights into the usage of user engagement elements in successful and unsuccessful software applications. During the evaluation, we observed that all the user engagement elements of the theoretical framework have been present in successful applications. However, some were missing in unsuccessful applications. To understand what this data means in action and how successful applications benefit from the user engagement elements and unsuccessful applications lack them, in the next section, we will discuss these results more deeply and talk about their implications.

IV. Discussion and Implications

Since each user engagement element has been implemented differently in different applications, this section will discuss the results and lessons learned from evaluating the most successful and unsuccessful mobile applications.

While evaluating the user engagement of applications, we noticed that unsuccessful software applications result from incomplete development and sometimes the owners' greed to make money with unthoughtful elements. Most unsuccessful applications did not consider the users' needs and requirements. These applications struggle with engagement elements such as utility and interactivity, and engagement elements such as social networking, gamification, and personalization are almost absent in them. However, some of them have the potential to become popular, yet they have to increase their utility and

interactivity in addition to other user engagement elements. Then a wide range of online marketing techniques like methods mentioned in [30] can help them reach new customers. The following subsections will discuss our findings concerning each user engagement element studied.

A. Utility and user engagement

As Figure 3 shows, based on evaluating successful applications, the utility has been fully present in them (100%), meaning they fully address their users' needs and expectations. While in unsuccessful applications, the utility was observed only by 11%. Figure 4 shows that only a few unsuccessful applications had utility. According to Table 1, there is a significant difference (89%) between utility usage in successful applications and unsuccessful applications. This implies the crucial impact of utility on user engagement of software applications. Concerning Figure 2, we can see some apps with a low level of personalization, gamification, and even social networking features are among the best apps in the world because, interestingly, all of them have a high utility by perfectly solving human problems. In fact, they are so good at solving the problem that they did not implement other engaging features. SHAREit and Android Auto are among these apps that solve users' problems uniquely and efficiently without implementing several other engagement elements. However, because of having a high level of innovation and functionality, and for some of them, no competitor, they are successful and among the most downloaded apps on google play. Several studies, such as [11] and [1], emphasize utility as the only important factor of user engagement without mentioning other factors. However, we observed that other user engagement factors are also crucial in highly compatible applications, such as movie streaming or communication.

Figure 3 shows that unsuccessful apps generally had lower utility scores. Functionality problems, not loading the content, and crashing aside, some of these applications provide functionality only after buying a VIP membership, and without that, the application is completely useless. This implies that in order to have an engaging application, limiting the application's functionality is not a good solution for making revenue because having high utility is essential. Thus, software engineers should seek more innovative approaches to make revenue from their applications. Using user-centric design is also important for increasing the utility of applications [1]. Moreover, the application should regularly be tested to ensure they function correctly [31].

B. 4.2. Social networking and user engagement

Figure 3 shows that social networking features were highly used in successful applications (87%), while it was rarely used in unsuccessful applications (10.5%). At the same time, its usage score was only 10.5% in unsuccessful applications. Table 1 shows a significant difference (76.5%) between the usage of social networking in successful and unsuccessful applications, which implies that this feature can significantly increase user engagement in software applications. Some unsuccessful applications not only had a very basic social networking feature but also limited these features and required payment to allow users to access them. The presence of social networking features in unsuccessful applications implies the fundamental change in the nature of software applications since the early 2000s and the emergence of Web 2.0. However, evaluated applications and statistical reports about social networking usage show that applications that want to succeed in today's world must implement social networking features [32].

C. Gamification and user engagement

Figure 3 shows gamification was not used much in successful applications (38%) and was utterly absent in unsuccessful applications. Moreover, gamification was used differently in each successful application. In most cases, gamification was designed based on the desired action application wants the users to take. The fact that gamification has lost 60% scores in evaluating successful applications implies that some of these apps provide a unique utility to their customers, and they do not have any competitors.

Although successful applications may still use classic gamification elements like those mentioned in Amo et al. [21], the findings of this study imply that software applications do not necessarily need to implement these gamification elements directly [21]. It is more meaningful for applications to innovate new game elements based on the functionality they want their users to perform. This can engage users similarly or better than using well-known game elements. For example, Snapchat, among the most successful applications, previously used trophies to reward users for their achievements. However, this feature was recently replaced by snap streaks, which encourages users to keep engaging with each other in an innovative gamification way. Moreover, instead of using leaderboards, Snapchat uses Snap Score, a general score for posting new content in snap chat. It allows users to compare themselves by going to other users' profiles and viewing other users' scores without using a leaderboard. Combining gamification with other user engagement elements, such as personalization, similar to the study [22], can improve its long-term effects on user engagement. Nevertheless, classic gamification elements have also been presented in some applications. Google street view benefits from gamification to reward users by badge, level, and points. It also shows the user progress and contributions to motivate users to contribute more to the application. As a result, using classic gamification elements in crowdsourcing applications and directly rewarding users like the described method in [33] can benefit applications' user engagement.

D. Privacy and user engagement

Figure 3 shows that privacy was highly present in successful applications (75%), while in unsuccessful applications, it had a low presence (20%). Figure 3 also shows that privacy notes were present in all the successful apps (mainly in their settings section). There was a devoted section for privacy settings in some successful apps with many options for customizing the user's privacy, while there were very few in others. Nevertheless, in unsuccessful applications (figure 4), many apps did not mention user privacy or did it poorly. This finding implies the important role of privacy in user engagement.

E. Content quality and user engagement

Content quality in unsuccessful applications (Figures 3 and 4) was very low (15.5%), and most of these applications had very low-quality content. On the other hand, successful applications have a higher content quality (65%) based on Figure 3. That said, it is important to note that some successful applications are not content based. For example, the Google Auto application is not dependent on content. Moreover, social networking applications like Instagram heavily depend on their users for their content, and they have little content by themselves. However, Figure 4 shows that many successful applications have tremendous high-quality content. Applications such as Netflix and Spotify are among those which align with what Shim and Jo (2020) stated in [23], the quality of content increases user engagement. Hence, these findings imply that applications can have different content quality depending on their utility to have a high user engagement. However, if applications' users generate their content, personalization can play an important role in increasing the quality of content for users.

F. Design quality and user engagement

Figure 3 shows that the design quality in successful applications was very high (92%). In most successful apps, the design quality criteria were simple, and thanks to the user interface frameworks, most successful applications followed similar design principles and elements mentioned in [34] that increased their user engagement (Figure 4). For example, fonts and icons used in most of the apps were similar, and the same metaphors were used, which is suitable for users to remember functionalities. These criteria were also observed in the unsuccessful software applications (Figure 4), and they had almost a high-quality design. These findings imply that the design quality alone is not enough to achieve high user engagement in software applications, although it is an essential element.

G. Interactivity and user engagement

Figure 3 shows that in successful applications, the interactivity score is very high (97%). While in unsuccessful applications, the interactivity score is low (30.5%). Nevertheless, even the unsuccessful applications have some interactivity (figure 4), and users can interact with the application, similar to what [10] mentioned. However, our findings imply that this interactivity is not enough to keep the user engaged with the application. We noticed that some evaluated unsuccessful applications limit users' interactivity intentionally. In some of these applications, users cannot freely move around the application or interact with it without making a payment. Also, in some cases, the interactions are interrupted by advertisements, and sometimes they have delays in responding, or it seems the application is stopped. Hence, these interactivity problems made unsuccessful applications frustrating and useless.

H. Personalization and user engagement

As shown in Figure 3, personalization was high in successful applications (71%), while in unsuccessful applications, it was very low (1.5%). According to Table 1, there is a significant difference in the presence of personalization in successful applications compared with unsuccessful applications (69.5%). The high level of personalization in the successful application (figure 4) implies that personalization is essential for the application's success. Several different types of personalization were used in evaluated applications, including customization of the app's utility, privacy, and user interface. These customizations were also mentioned as important user engagement criteria in [10]. However, one of the important personalization types used in these applications was personalizing the content presented to the user based on their interest. For example, on Netflix, users can get personalized recommendations, and this personalization is very effective in increasing user engagement. Recently it was reported that the Netflix personalization system alone generated a billion-dollar income for the company [3].

V. Conclusions

In conclusion, we have evaluated 50 successful and unsuccessful software applications against different user engagement elements to understand the level of impact of each of these elements on the success of applications. The study results showed that utility, interactivity, and design quality are respectively highly present in successful software applications (figure 3). Moreover, utility, social networking, and personalization elements are significant indicators of software applications' success since they are less present in unsuccessful applications (table 1).

To sum up, we answered the mentioned research questions in section one by conducting the evaluation and discussing the implication of our findings to help the software engineers understand the important aspects of user engagement. We also discussed how user engagement elements could be combined and implemented in software applications and why some software applications fail. We shed some light on the importance of each user engagement element with this article to help software engineers select more appropriate user engagement elements and assist them in developing successful software applications with high user engagement.

One interesting outcome of this study was that gamification has the most negligible impact on the success of software applications, and it was completely absent in unsuccessful applications. Thus, investigating how gamification can make more impact on the success of software applications would be an interesting topic to study in the future. Evaluating a similar study on applications of the same type (e.g., only communication apps) might be another interesting study.

ACKNOWLEDGMENTS

Researcher 1 (Mohammad Hajarian) acknowledges support from the CONEX-Plus programme funded by Universidad Carlos III de Madrid and the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 801538. And Spanish State Research Agency (AEI) under grant Sense2MakeSense (PID2019-109388GB-I00).



REFERENCES

[1] Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evidence-based mental health, 21(3), 116-119.

[2] McGwin, M. (2020). HiveGuide: An Education and Crowdsourcing App for Bee Conservation.

[3] Golbeck, J. (2020). Optimizing for Engagement Can Be Harmful. There Are Alternatives. IEEE Annals of the History of Computing, 35(04), 117-118.

[4] Bitrián, P., Buil, I., & Catalán, S. (2021). Enhancing user engagement: The role of gamification in mobile apps. Journal of Business Research, 132, 170-185.

[5] Attfield, S., Kazai, G., Lalmas, M., & Piwowarski, B. (2011, February). Towards a science of user engagement (position paper). In WSDM workshop on user modelling for Web applications (pp. 9-12).

[6] Patel, V., Das, K., Chatterjee, R., & Shukla, Y. (2020). Does the interface quality of mobile shopping apps affect purchase intention? An empirical study. Australasian Marketing Journal (AMJ), 28(4), 300-309.

[7] Oh, J., Bellur, S., & Sundar, S. S. (2018). Clicking, assessing, immersing, and sharing: An empirical model of user engagement with interactive media. Communication Research, 45(5), 737-763.

[8] Ramírez-Correa, P., Grandón, E. E., Ramírez-Santana, M., & Belmar Órdenes, L. (2019). Explaining the use of social network sites as seen by older adults: The enjoyment component of a hedonic information system. International journal of environmental research and public health, 16(10), 1673.

[9] Wei, Y., Zheng, P., Deng, H., Wang, X., Li, X., & Fu, H. (2020). Design Features for Improving Mobile Health Intervention User Engagement: Systematic Review and Thematic Analysis. Journal of medical Internet research, 22(12), e21687.

[10] Soni, N., Aloba, A., Morga, K. S., Wisniewski, P. J., & Anthony, L. (2019, June). A framework of touchscreen interaction design recommendations for children (tidrc) characterizing the gap between research evidence and design practice. In Proceedings of the 18th ACM International Conference on Interaction Design and Children (pp. 419-431).

[11] Tian, Z., Shi, Z., & Cheng, Q. (2021). Examining the antecedents and consequences of mobile travel app engagement. Plos one, 16(3), e0248460.

[12] Melcher, J., Camacho, E., Lagan, S., & Torous, J. (2020). College student engagement with mental health apps: analysis of barriers to sustained use. Journal of American College Health, 1-7.

[13] Jozani, M., Ayaburi, E., Ko, M., & Choo, K. K. R. (2020). Privacy concerns and benefits of engagement with social media-enabled apps: A privacy calculus perspective. Computers in Human Behavior, 107, 106260.

[14] Hollo, Z., & Martin, D. E. (2023). An equitable approach to enhancing the privacy of consumer information on my health record in Australia. Health Information Management Journal, 52(1), 37-40.

[15] Woldaregay, A. Z., Issom, D. Z., Henriksen, A., Marttila, H., Mikalsen, M., Pfuhl, G., ... & Hartvigsen, G. (2018, January). Motivational Factors for User Engagement with mHealth Apps. In pHealth (pp. 151-157).

[16] Desai, D. (2019). An empirical study of website personalization effect on users intention to revisit E-commerce website through cognitive and hedonic experience. In Data Management, Analytics and Innovation: Proceedings of ICDMAI 2018, Volume 2 (pp. 3-19). Springer Singapore.

[17] Amit-Danhi, E. R., & Shifman, L. (2020). Off the charts: user engagement enhancers in election infographics. Information, Communication & Society, 1-19.

[18] Hajarian, M., & Hemmati, S. (2020, September). A gamified word of mouth recommendation system for increasing customer purchase. In 2020 4th International Conference on Smart City, Internet of Things and Applications (SCIOT) (pp. 7-11). IEEE.

[19] Hajarian, M., & Diaz, P. (2021, November). Effective Gamification: A Guideline for Gamification Workshop of WEEF-GEDC 2021 Madrid Conference. In 2021 World Engineering Education Forum/Global Engineering Deans Council (WEEF/GEDC) (pp. 506-510). IEEE.

[20] Hajarian, M., Díaz, P., & Aedo, I. (2023, March). Work-in-Progress: The Social Networks Users' Behavior Change Using Gamified Education and Their Personality. In Learning in the Age of Digital and Green Transition: Proceedings of the 25th International Conference on Interactive Collaborative Learning (ICL2022), Volume 1 (pp. 982-989). Cham: Springer International Publishing.

[21] Amo, L., Liao, R., Kishore, R., & Rao, H. R. (2020). Effects of structural and trait competitiveness stimulated by points and leaderboards on user engagement and performance growth: A natural experiment with gamification in an informal learning environment. European Journal of Information Systems, 29(6), 704-730.

[22] Hajarian, M., Bastanfard, A., Mohammadzadeh, J., & Khalilian, M. (2019). A personalized gamification method for increasing user engagement in social networks. Social Network Analysis and Mining, 9(1), 1-1

[23] Shim, M., & Jo, H. S. (2020). What quality factors matter in enhancing the perceived benefits of online health information sites? Application of the updated DeLone and McLean Information Systems Success Model. International Journal of Medical Informatics, 137, 104093.

[24] Ali, F., Terrah, A., Wu, C., Ali, L., & Wu, H. (2021). Antecedents and consequences of user engagement in smartphone travel apps. Journal of Hospitality and Tourism Technology.

[25] Wei, J., Seedorf, S., Lowry, P. B., Thum, C., & Schulze, T. (2017). How increased social presence through co-browsing influences user engagement in collaborative online shopping. Electronic Commerce Research and Applications, 24, 84-99.

[26] Hajarian, M., Bastanfard, A., Mohammadzadeh, J., & Khalilian, M. (2017). Introducing fuzzy like in social networks and its effects on advertising profits and human behavior. Computers in Human Behavior, 77, 282-293.

[27] Hajarian, M., Bastanfard, A., Mohammadzadeh, J., & Khalilian, M. (2019). SNEFL: Social network explicit fuzzy like dataset and its application for Incel detection. Multimedia tools and applications, 78, 33457-33486.

[28] Santos-Vijande, M. L., Gómez-Rico, M., Molina-Collado, A., & Davison, R. M. (2022). Building user engagement to mhealth apps from a learning perspective: Relationships among functional, emotional and social drivers of user value. Journal of Retailing and Consumer Services, 66, 102956.

[29] Elsholz, E., Chamberlain, J., & Kruschwitz, U. (2019, March). Exploring language style in chatbots to increase perceived product value and user engagement. In Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (pp. 301-305).

[30] Hajarian, M., Camilleri, M. A., Díaz, P., & Aedo, I. (2021). A taxonomy of online marketing methods. In Strategic corporate communication in the digital age (pp. 235-250). Emerald Publishing Limited.

[31] Walkinshaw, N. (2018). Testing functional black-box programs without a specification. In Machine Learning for Dynamic Software Analysis: Potentials and Limits (pp. 101-120). Springer, Cham.

[32] "28 Mobile App Usage & Revenue Statistics To Know In 2021." MindSea Development, https://www.facebook.com/MindSea, 1 Oct. 2020, https://mindsea.com/app-stats/.

[33] Hajarian, M., & Hemmati, S. (2021, May). A Crowdsourcing and Gamification based Product Ranking Method for E-Commerce. In 2021 7th International Conference on Web Research (ICWR) (pp. 197-201). IEEE.

[34] Garett, R., Chiu, J., Zhang, L., & Young, S. D. (2016). A literature review: website design and user engagement. Online journal of communication and media technologies, 6(3), 1.

Interaktivna potpuno imerzivna virtualna stvarnost računalne simulacije brodske strojarnice

Dario Ogrizović
Sveučilište u Rijeci / Pomorski fakultet, Rijeka, Hrvatska
dario@pfri.uniri.hr

***Sažetak* - S obzirom na mogućnosti visokoškolskih pomorskih institucija i potreba studenata, te sadašnjih i budućih pomoraca kao i sve veće konkurencije na svjetskom pomorskom tržištu više je nego očita moguća uloga uvođenja i primjene novih informacijsko komunikacijskih tehnologija u modernizaciji procesa obrazovanja i usavršavanja studenata te pomoraca. Računalno generirano digitalno okruženje, dostupno kroz potpuno imerzivnu virtualnu stvarnost, nudi korisniku realističan osjećaj odnosno percepciju fizičkog prisustva u kojem korisnici mogu vršiti interakciju s virtualnim objektima i svojim radnjama odrediti što se događa u njemu. Kako je stjecanje iskustva sastavni dio učenja što se odvija u nastavnom procesu kroz uporabu stvarnih sustava na brodu i specijaliziranih simulatora razvijena je računalna simulacija brodske strojarnice korištenjem tehnologije virtualne stvarnosti. Provedeno je korisničko testiranje kako bi se istražio utjecaj potpuno imerzivne virtualne stvarnosti na spoznaju i rad korisnika, osjećaj prisutnosti, psihološki utjecaj, učinkovitost izvedbe u odnosu na tradicionalne sustave te interes za upotrebu nove tehnologije.**

Ključne riječi - virtualna stvarnost; simulacija brodske strojarnice; potpuna imerzija; korisničko testiranje; VR

I. UVOD

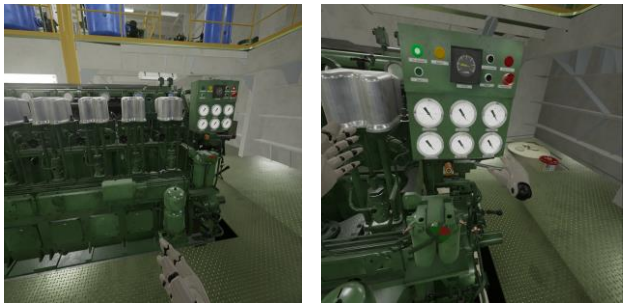
Uporaba stvarnih sustava na brodu i specijaliziranih simulatora, kao oblik iskustvenog učenja, je sastavni dio obrazovanja studenata pomorskih učilišta i pomoraca. Stvarni brodski sustavi nude istinsko iskustvo, ali i sigurnosne rizike i posljedice te su iznimno skupi za rad i imaju prostornu te vremensku ograničenost korištenja. Specijalizirani simulatori imaju mogućnost replikacije specifičnih zadataka uz neograničen broj ponavljanja, no također su relativno skupi, zahtijevaju instruktora da upravlja sustavom i imaju ograničenu dostupnost studentima pomorcima kao polaznicima specijaliziranih tečajeva. Specijalizirani simulatori za brodsko strojarstvo zasnovani su na upotrebi tradicionalnih MERS (engl. marine engine-room simulator) simulatora, a Međunarodna pomorska organizacija (engl. International Maritime Organization – IMO) zahtjeva da budući časnici stroja moraju dobiti profesionalnu obuku na simulatoru brodske strojarnice (MERS) prije rada na brodu [1]. Tradicionalni MERS (engl. marine engine-room simulator) simulatori imaju malu razinu imerzije što onemogućuje korisnicima prepoznavanje stanja i korištenje pripadajuće opreme kao da se nalaze u stvarnoj, realnoj brodskoj strojarnici. Kako bi se poboljšala percepcija realne okoline brodske strojarnice, razvijen je koncept interaktivne potpuno imerzivne virtualne stvarnosti računalne simulacije brodske strojarnice. U tom računalnom generiranom digitalnom okruženju provedeno je korisničko testiranje potpuno imerzivne virtualne stvarnosti među studentima prediplomskog i diplomskog studija brodostrojarstva Sveučilišta u Rijeci.

II. VIRTUALNA STVARNOST

Virtualna stvarnost nudi korisniku realističan osjećaj ili percepciju fizičkog prisustva unutar umjetnog, računalno generiranog digitalnog okruženja u kojem korisnici mogu vršiti interakciju s virtualnim objektima [2]. Virtualna stvarnost stvara se uz pomoć 3D računalne grafike te VR ulaznih i izlaznih uređaja sa kojima se ostvaruje interakcija i dobivaju podražaji iz virtualnog okruženja u realnom vremenu. Postoje tri različite vrste virtualne stvarnosti: neimerzivna, poluimerzivna i potpuno imerzivna virtualna stvarnost. Neimerzivni virtualna stvarnost temelji se na standardnom monitoru računala gdje korisnik može neizravno manipulirati 3D okruženjem pomoću tipkovnice, miša i kontrolera. MERS simulatori bazirani na ovaj način nude jednostavnu vizualizaciju i pomažu razumjeti logiku stvarnog brodskog sustava. Poluimerzivni VR je nešto između neimerzivnog i potpuno imerzivnog VR-a, a primjeri su simulatori leta, vožnje ili pomorski simulatori s konkavnim monitorima i zidnim projektorima. Cave (engl. Cave Automatic Virtual Environment) je primjer takvog sustava koji je posebno konstruirana prostorija sastavljena od zidnih, stropnih i podnih projektora i projekcijskih platna [3]. Na ovaj način se koriste polufizički MERS simulatori koji se sastoje od niza projektora i dijela stvarne brodske opreme. Potpuno imerzivna virtualna stvarnost u potpunosti ispunjava korisnikovo vidno polje sa zaslonom u obliku naočala i kontrolerima za stvaranje interakcije sa virtualnim objektima unutar 3D računalno generiranog digitalnog okruženja. Iako ogromna količina sadržaja virtualne stvarnosti dolazi iz videoigara i industrije zabave, virtualna stvarnost nudi priliku u mnogim područjima obrazovanja [4-6] od STEM-a, medicine do humanističkih znanosti. Virtualna stvarnost kao vrsta učenja je podskup e-učenja gdje interaktivna i dinamična virtualna okruženja potiču aktivno sudjelovanje što dovodi do većeg angažmana, motivacije i interesa.

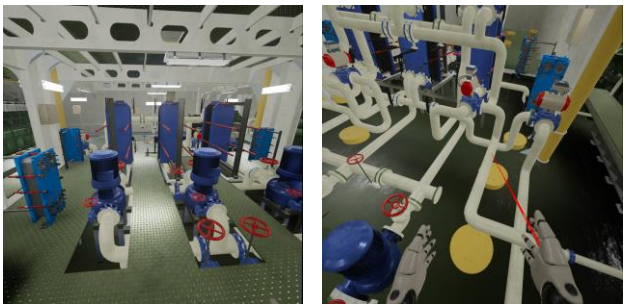
III. RAČUNALNA SIMULACIJA BRODSKE STROJARNICE

Svi 3D modeli brodske strojarnice izrađeni su prema stvarnim nacrtima, specifikacijama opreme i fotografijama stvarne brodske strojarnice i opreme, a sve pod nadzorom upravitelja stroja sa dugogodišnjim iskustvom. 3D model brodske strojarnice sastoji se od dva glavna i dva pomoćna pogonska stroja, dva generatora i različite pomoćne opreme. Za izradu 3D modela korišteno je razvojno okruženje Blender 3.5 [7] koje je bazirano na otvorenom izvornom kodu (engl. Open Source), a za izradu interaktivne potpuno imerzivne virtualne stvarnosti računalne simulacije brodske strojarnice korišteno je Unreal Engine 5 [8] programsko okruženje. Programsko okruženje prilagođeno je VR HMD-u (engl. Virtual Reality Head-Mounted Display). Izgled glavnog pogonskog stroja i cjevovoda prikazan je na slici 1.



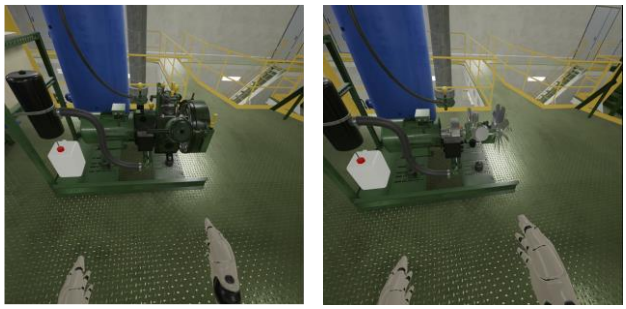
Slika 1. Prikaz glavnog pogonskog stroja

Izgled brodskih pumpi i cjevovoda sa mogućnošću pomicanja poda prikazan je na slici 2.



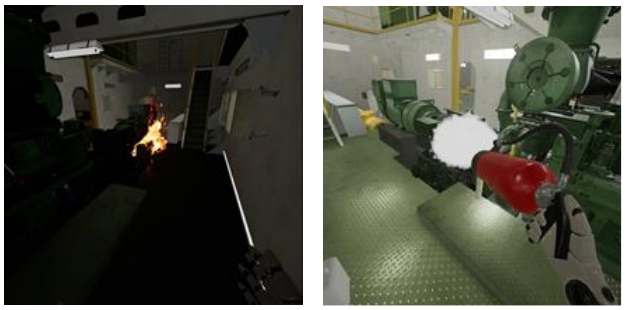
Slika 2. Prikaz brodskih pumpi i cjevovoda

Na pojedinim uređajima postoji mogućnost pregledavanja unutrašnjosti uređaja pomicanjem kućišta kao što je prikazano u slučaju kompresora na slici 3.



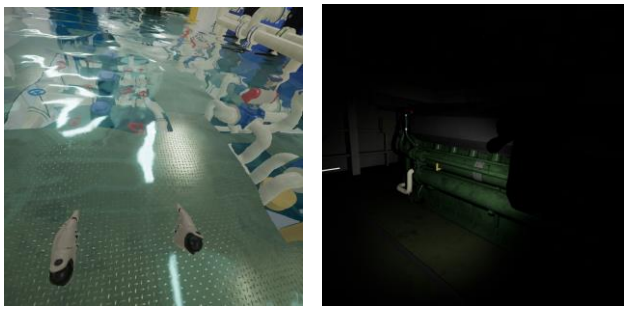
Slika 3. Prikaz kompresora

Implementirano je nekoliko scenarija, a jedan od njih je i nastanak te gašenje požara, slika 4.



Slika 4. Nastanak i gašenje požara

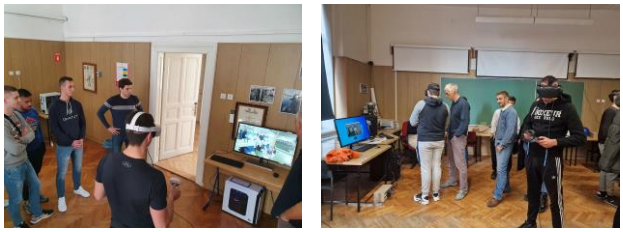
Slijedeći scenarij je poplava brodske strojarnice i nestanak električne energije gdje se koristi ručna svjetiljka, slika 5.



Slika 5. Poplava brodske strojarnice i nestanak električne energije

IV. KORISNIČKA TESTIRANJA

Korisnička testiranja, slika 6, napravljena su sa ukupno 31 studentom druge godine preddiplomskog studija Brodostrojarstvo te diplomskog studija Brodostrojarstvo i tehnologija pomorskog prometa Pomorskog fakulteta Sveučilišta u Rijeci. Testiranje se odvijalo u računalnom laboratoriju, studenti su koristili 5 VR HMD-ova (Meta Quest 2 i HTC Vive) i ručne kontrolere. Također su koristili 5 računala s Intel i7 procesorima, 32 GB RAM-a, 500 GB SSD-om, 6 TB SATA diskovima i NVIDIA RTX 3080 GPGPU karticama.

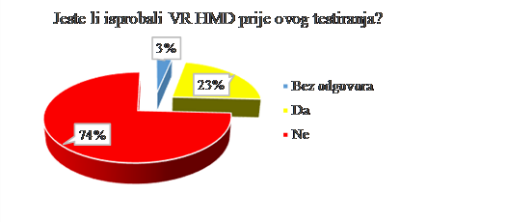


Slika 6. Korisnička testiranja

Povratne informacije studenata prikupljene su i procijenjene na strukturiran i sustavan način pomoću upitnika [9]. Upitnik se sastojao od ukupno 12 pitanja. Google Forms [10], web aplikacija korištena je za izradu upitnika i prikupljanje podataka. Nakon što su studenti koristili VR HMD-ove, ispunili su upitnik koji se odnosi na cjelokupno korisničko testiranje. Nakon svakog

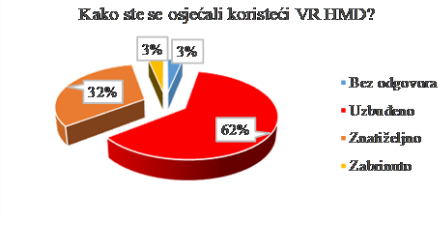
pojedinog korisničkog testiranja analizirani su prikupljeni podaci i napravljen je završni izvještaj.

Prvo pitanje iz upitnika bilo je korištenje VR HMD-a, slika 7. 23 učenika nikada prije ovog testiranja nije isprobalo VR HMD. 7 studenata isprobalo je VR HMD prije ovog testiranja te su komentirali su da su isprobali VR HMD na konferencijama videoigara ili da imaju vlastiti VR HMD.



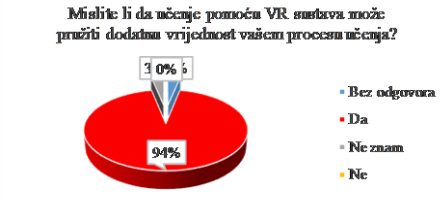
Slika 7. Korištenje VR HMD-a

Jedno od pitanja unutar upitnika odnosilo se na osjećaje prilikom korištenja VR HMD-a, slika 8. Iz popisa ponuđenih pojmova (osjećaja) u odgovorima je značajno prisustvo uzbuđenja (19), znatiželje (10) i zabrinutosti (1). Bilo je 7 dodatnih komentara o prvom potpuno imerzivnom VR osjećaju (odgovori su bili: super, uživancija, oduševio sam se, sjajno, impresivno, top, zakon).



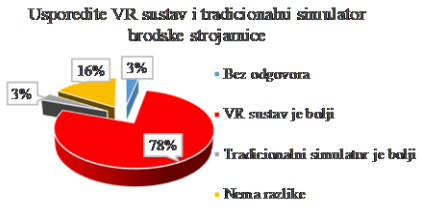
Slika 8. Osjećaji prilikom korištenja VR HMD-a

Na pitanje ako misle da učenje pomoću VR sustava može pružiti dodatnu vrijednost u njihovom procesu učenja 29 učenika izjavilo je da VR sustav može pružiti dodanu vrijednost njihovom procesu učenja, 1 student je odgovorio sa "Ne znam", slika 9. Među studentima koji nikada prije ovog eksperimenta nisu isprobali VR HMD, svi su odgovorili da im VR HMD može pružiti dodatnu vrijednost u njihovom procesu učenja.



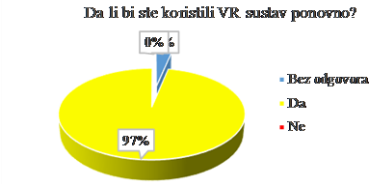
Slika 9. Dodatna vrijednost u procesu učenja pri korištenju VR HMD-a

Na slici 10, kod usporedbe VR sustava i tradicionalnog simulatora brodske strojarnice 24 studenta je odgovorilo da je VR sustav bolji, 1 student da je klasični simulator bolji i 5 studenata da nema razlike između jednog i drugog.



Slika 10. Usporedba VR HMD-a i klasičnog simulatora

Gotovo svi studenti odgovorili su da bi ponovno koristili VR HMD, slika 11.



Slika 11. Ponovno korištenje VR HMD-a

V. ZAKLJUČAK

Nakon testiranja VR HMD-a i korištenja interaktivne potpuno imerzivne virtualne stvarnosti računalne simulacije brodske strojarnice, studenti su iznesli svoj pozitivan stav prema ideji. Svidjelo im se jer im je sadržaj bio poznat jer su imali iskustva s klasičnim simulatorom te su pokušali istražiti i isprobati interaktivne virtualne objekte koliko god su mogli. Studenti su bili uzbuđeni, zainteresirani i znatiželjni kada su koristili VR HMD i gotovo svi su se složili da bi ponovno isprobali VR brodsku strojarnicu. Također su smatrali da bi to bila izvrsna dodana vrijednost za proces učenja. Kako bi obrazovanje i usavršavanje studenata i pomoraca bilo što učinkovitije važno je razviti i eksperimentirati s odgovarajućim razvojnim alatima, istražiti njihove koristi i mogućnosti te ih integrirati u postojeće strategije učenja na preddiplomskoj i diplomskoj razini studija. Dinamična priroda virtualnih okruženja zahtjeva aktivno sudjelovanje pojedinih korisnika što uzrokuje veći angažman, motivaciju i interes što je uzrokovano direktnom interakcijom i izazovima unutar računalno generiranog digitalnog okruženja.

ZAHVALE

Rad je podržan kroz program Obzor Europa (Horizon Europe), ID projekta: 101087348 (INNO2MARE).

LITERATURA

[1] International Maritime Organization, STCW: Including 2010 Manila Amendments: STCW Convention and STCW Code: International Convention on Standards of Training, Certification and Watchkeeping for Seafarers. London: International Maritime Organization, 2011.

[2] LaValle S. M., Virtual Reality. Cambridge: Cambridge University Press, 2023.

[3] Cruz-Neira C., Sandin D.J., DeFanti T.A., Kenyon R.V., Hart J.C. The CAVE: Audio Visual Experience Automatic Virtual Environment. Communications of the ACM.1992. pp 65-72.

[4] Radianti, J., et al. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 2020.

[5] Hamilton, D., et al. Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. Journal of Computers in Education, 2021, 8.1: 1-32.

[6] Wu, Bian; YU, Xiaoxue; GU, Xiaoqing. Effectiveness of immersive virtual reality using head - mounted displays on learning performance: A meta - analysis. British Journal of Educational Technology, 2020, 51.6: 1991-2005.

[7] Blender 3.5 https://www.blender.org (12.02.2023.)

[8] Unreal Engine 5, https://www.unrealengine.com (12.02.2023.)

[9] Fink A. Survey Research Methods. *International Encyclopaedia of Education (Third Edition)*. 2010. pp 152-160. DOI: 10.1016/B978-0-08-044894-7.00296-7

[10] Google Forms, https://workspace.google.com/products/forms (03.02.2023.)

SIDE

International Conference on

SMART INDUSTRIES AND DIGITAL ECOSYSTEMS

Steering Committee

Chairs: Neven Vrček, University of Zagreb, Varaždin, Croatia
Karolj Skala, Ruđer Bošković Institute, Zagreb, Croatia

Members: Darinko Bago, Croatian Exporters Association, Zagreb, Croatia
Hrvoje Balen, Algebra University College, Zagreb, Croatia
Vladimir Olujić, Storm Computers, Zagreb, Croatia
Nedjeljko Perić, University of Zagreb, Croatia
Tadej Slapnik, Hashnet, Slovenske Konjice, Slovenia
Marija Šćulac Domac, Croatian Chamber of Economy, Zagreb, Croatia
Ivan Vidaković, Infodom, Zagreb, Croatia
Slavko Vidović, Infodom, Zagreb, Croatia
Ernest Vlačić, Algebra University College, Zagreb, Croatia

Program Committee

Members: Zoran Aralica, Institute of Economics, Zagreb, Croatia
Boris Blumenschein, Ilba, Zagreb, Croatia
Darko Bosnar, Infodom, Zagreb, Croatia
Bojan Jerbić, University of Zagreb, Croatia
Goran Marković, Infodom, Zagreb, Croatia
Mladen Mrvelj, Infodom, Zagreb, Croatia
Matija Srbić, University of Zagreb, Croatia
Marija Šutina, Croatian Chamber of Economy, Zagreb, Croatia
Domen Verdnik, Infodom, Zagreb, Croatia
Marko Vidović, Infodom, Zagreb, Croatia

PAPERS

Digitalization of Innovation Networks: Theoretical and Empirical Issues

V. Omelyanenko[*, **, ***], I. Pidorycheva[***], H. Shevtsova[***], O. Omelianenko[**], N. Rudenko[****], O. Yurchenko[*****]

[*] Estonian Entrepreneurship University of Applied Sciences, Tallinn, Estonia
[**] Sumy State Pedagogical University, Sumy, Ukraine
[***] Institute of Industrial Economics of National Academy of Sciences of Ukraine, Kyiv, Ukraine
[****] Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
[*****] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine
omvitaliy@gmail.com

***Abstract* – The article is devoted to the analysis of digital aspects of innovation policy, in particular, based on the use of international ratings, the determination of the features of the construction of the digital component of innovation networks, and the deepening of scientific and methodological approaches to the digital support of innovation processes. Data from the Global Innovation Index and the ICT Development Index were used to assess the impact of digitalization on the development of innovation processes. Based on the results of the calculations, significant positive correlation estimates were obtained with the key components of the innovation system and the level of cluster development. Marketing, communication, infrastructural, international dimensions of digitalization of innovative processes are defined. It has been proven that the growing availability of ICT makes it necessary to rethink many processes in the innovation field in a timely manner. On the basis of world experience, the role of universities is argued, which should assume a significant share of responsibility for the formation of a new technological order by improving the quality of research and implementing their results in the real economy within the framework of S2B-B2S (Science-to-Business – Business-to-Science).**

Keywords - innovation network, project, tech transfer, ICT

I. INTRODUCTION

The development of digital technologies, the formation of a digital economy and society is one of the main megatrends of today. According to Pricewaterhouse Coopers experts [1] by 2030 digitalization and smart automation will provide 14% of global GDP growth, which is equivalent to approximately US$15 trillion. at current prices. In the conditions of Industry 4.0, issues of digital transformation concern not only service services, but also a wide range of processes, including those based on the cooperation of various economic agents. Digitization significantly helps the participants of innovation networks to develop new products based on the results of scientific research, conduct the necessary testing or measurements, model and develop prototypes, improve existing technology, ensure communication on the implementation of scientific and technical solutions, technology transfer and innovation marketing.

Digital transformation as a result of the application of information and communication technologies (ICT) is expressed in the creation of unique business models, a qualitative change in the perception of a product or service by customers, a fundamental acceleration and simplification of the operational activities of the participants of the innovation network. It should be taken into account that in modern conditions, the company's success is determined not only by the use of unique technologies, but also by the digitalization strategy. Companies that apply ICT without a previously formed strategy solve only tactical tasks. It is important to remember that ICT is only a means, not an end. In the case of an innovation system, the goal is always a unique product (service) created using such technologies.

The report of Industrie 4.0 Working Group [2] notes that innovative development strategies in the context of Industry 4.0 should be based on a new approach to production as a conglomerate of large industrialists, experts, economists and scientists. With the inclusion of customers and contractors in a single open organizational structure, the boundaries between interacting enterprises become quite blurred and dynamic, and the success of some is a necessary condition for the success of others. At the same time, partners are increasingly taking an active part in the development of new products, services, and even business strategies.

Therefore, the question arises as to the design of appropriate innovative networks that will ensure the synergy of interaction, and the creation of their effective infrastructure support. Given the comprehensive impact and large-scale prospects of digitalization processes, the task of working out the conceptual issues of managing digital transformations in the context of innovative trends becomes especially relevant, in particular for developing countries. Both the macro level of digitization of innovative processes and the specifics of using ICT innovations for the transformation of business models based on a network mechanism require attention.

The article is devoted to the analysis of digital aspects of innovation policy, in particular, based on the use of international ratings, the determination of the features of the construction of the digital component of innovation networks, and the deepening of scientific and methodological approaches to the digital support of

innovation processes.

II. Review of Recent Studies

Transformations of business models become a critical factor of innovation, but are poorly studied from a theoretical and practical point of view. Scientists note the lack of tools that allow studying business models in combination with the complex relationship between their change and market results. The study introduces a comprehensive framework for these issues, recognizing the relevance of tools for the practical implementation of value integrated management of critical components of the business model. The changes caused by Industry 4.0, in particular its digital capabilities, lead to the transformation of production, economic and social systems [3; 4; 5]. One of the main trends in the development of companies is information and network interaction, through which they learn to manage information, exchange knowledge, form and develop the necessary competencies.

The report of UNCTAD [6] notes that digitalization has given rise to a new wave of innovation that will have profound consequences for humanity, changing the relationship between society, the state and business, and will also lead to the transformation of the structure of society and the economy. The pace of economic growth, labor productivity and the development of human potential will be increasingly determined by the level of integration into the digital economy. Indeed, digitalization not only creates new business opportunities, but also create a number of problems and risks.

The study [7] examines the case of the Chinese software and services industry. It is shown that it has become a driving factor of economic growth and leading innovations. The study of the structure of an intercity network based on patent cooperation between cities helped to understand the positioning of cities from the point of view of innovations related to service software (software service-related innovation).

The issue of the development of innovation platforms is very relevant in the framework of the ever-increasing cooperation in the creation of innovation value (co-creation), which makes it possible to introduce new products (services), increasing the efficiency of business processes and creating new value for customers [8]. Article [9] illustrates the fact that increasing economic attention to knowledge concerns the development of knowledge-intensive business services (KIBS).

The work [10] emphasizes the importance of the level of self-sufficiency and recognition of the status of research structures. In addition, the study demonstrates that the mass of necessary knowledge is constantly increasing due to constant technological progress, due to which opening through cooperation (open up through cooperation) is necessary for the company. In these conditions, digital transformation is considered as a factor of access to information and new opportunities that can improve the technological potential of the firm. The articles [11] investigated the essence of the ICT category, its relationship with such concepts as the Internet of Things, the Industrial Internet of Things, cyber-physical systems, identified and analyzed the factors that determine the implementation of information and digital technologies within the framework of the concept smart manufacturing. Digitization enables the creation of administrative solutions, data protection systems, storage of research data and educational resources, as well as opportunities for better collaboration [12; 13].

The study [14] analyzed the impact of digitalization on the formation of advantages of new technologies and access to them, as well as determined their impact on organizational transformations in terms of cooperation and competitive advantages. In this context, for a better understanding of "agent-agent" interactions (actor-to-actor interactions), elements of systems thinking, determining the degree of openness of systems and the concept of managing communication flows are used.

Research results [15] empirically support the hypothesis that ICT acts as a catalyst for innovation. However, the impact of digitalization on innovation activity depends on the type of ICT innovations, as well as the scale of their application. In particular, a strong positive relationship was found between ICT integrating administrative and industry software and innovative productivity. The identified pattern does not apply to market-oriented ICT, such as e-commerce or customer relationship management software.

One of the options for ICT innovations from the point of view of machine-to-machine and human-to-machine interaction problems in the innovation space is presented in the research [16]. Among the main issues covered are problems of agents, their interaction with each other.

Another direction is considered in the works [17], which refers to Big Data as a new paradigm that provides a lot of data and opportunities for improving or expanding application networks research and decision support for digital applications including business, science and technology. For the purposes of our research, the following conclusions are quite useful:

- cloud computing and big data make it possible to make scientific discoveries and applications;
- cloud computing provides solutions for Big Data;
- big data, spatio-temporal thinking and various fields of applications contribute to the advancement of cloud computing and related technologies with new requirements;
- the internal spatio-temporal principles of Big Data and geospatial sciences serve as a source of finding technical and theoretical solutions for optimizing big data processing;
- the open availability of Big Data and processing capabilities create social problems of geospatial significance;
- there is a transformation of innovation waves.

Therefore, it can be concluded that digitalization promotes the exchange of accumulated knowledge and experience, which provides opportunities for various stakeholders to make more informed decisions in the framework of innovation management and marketing. ICT support for the vide range participation tasks of agents in innovation networks [18; 19; 20; 21; 22], is an important factor in the formation of modern innovation systems on

at different levels.

III. Empirical Estimations

Data from the Global Innovation Index and the ICT Development Index were used to assess the impact of digitalization on the development of innovation processes. In the table 1 analyzed the impact of digitalization on the relevant components of the innovation system and provided a quantitative assessment of such impact based on the data of 2019. Based on the results of the calculations, significant positive correlation estimates were obtained with the key components of the innovation system and the level of cluster development (column 4 of Table 1). Also the dependence of the Global Competitiveness Index and the ICT Development Index (the correlation coefficient between them is 0.92) confirms the significant impact of digitization on the global competitive positions of countries. The dependence of the innovation capability indicator of the economy (Innovation capability) and the ICT development index (the correlation coefficient between them is 0.772) confirms the positive impact of digitalization on the innovation potential of countries.

IV. Trends Analytics and Policy Recomendations

In the conditions of a change in the paradigm of innovative development and increased competition for ideas, it is difficult for developers and owners of new technologies to organize the marketing of their developments and find partners to create joint productions. From the point of view of developing countries, this aspect is a challenge, since the problems of technology transfer, in particular international, are largely due to the lack of a modern technology transfer infrastructure, which would include specialized centers and information systems for the exchange of technological requests and proposals.

Another aspect of the need for accelerated digitization is related to the creation of conditions for overcoming technological dependence on other countries and the development of one's own innovation ecosystem. In other words, we are talking about the formation and satisfaction of the domestic technology transfer market.

It should be noted that the growing availability of ICT makes it necessary to rethink many processes in the innovation field in a timely manner. A wave of disruptive innovation is not only sweeping through technology markets, but also disrupting a number of established and forecasted industries. In particular, in this situation, the role of universities is important, as they must take on a significant share of responsibility for the formation of a new technological order. It is advisable to simultaneously pay attention to both improving the quality of research and the issue of implementing their results in real economy.

Analysis of the experience of leading foreign universities shows that each university has its own innovation ecosystem. We can mention Stanford, LU Innovation System, Cambridge Enterprise, KU Leuven Research & Development, Polytechnic School of Lausanne (EPFL) IV-VP TTO. Each of these university innovation ecosystems has an important ICT component, which has an individual character depending on the types of projects, industry specifics, type of partnerships, etc.

TABLE I. Correlations between ICN and Innovation Indicators

Indicators of the Global Innovation Index	Components of the indicator	Impact of ICT	Correlation with the ICT Development Index
Knowledge creation	Patenting (national and international), publications, citations	Access to international patent and scientometric databases, scientific networks, research and forecasts, reports and overviews of trends in industries and technological areas. Informational support for the marketing of ideas, covering all activities aimed at forming a network that will further implement the relevant innovative project(s).	0,667
Knowledge impact	GDP growth, new business creation, software acquisition, high-tech business, ISO 9001 certification	Informational support for the marketing of innovations as an activity of enterprises for the creation of innovative products that allow ensuring the high-quality satisfaction of needs or the formation of a new market.	0,564
Knowledge diffusion	Fees for intellectual property rights, high-tech export, export of ICT services, foreign direct investment	Technology services and databases, information systems for the exchange of technological requests and proposals, venture project databases, news archives and company databases. Informational support for market activities of organizations aimed at achieving the key goals of technology transfer.	0,603
State of cluster development	Geographical concentration of firms, suppliers, manufacturers related products and services, as well as specialized institutions	The use of digital infrastructure for the formation of interregional and international clusters, in particular virtual clusters, as well as the management of innovative interaction processes.	0,630

Noting this successful experience, it is necessary to especially emphasize the educational function of science within the framework of S2B-B2S (Science-to-Business – Business-to-Science) networks, which is important not only for education, but also for science itself, since in the process of teaching knowledge their development and approval often takes place. Thus, a comprehensive approach to the creation of digital innovation ecosystems requires the study of both institutions and participants, networks of their interaction, the specifics of environment, resources, technologies, etc.

Innovation networks in modern conditions are considered as an effective mechanism of mutually beneficial interaction between research organizations, startups and innovative businesses interested in the practical implementation of scientific and technical developments and projects both on the territory of their state and at the international level. One of the leading

international tools for assessing the spread of the network approach with the use of ICT is the Network Readiness Index (NRI). The inset presents the positioning of Ukraine according to the results of a recently updated global study of network readiness [23].

We propose to consider the role of ICT in the development of innovation networks based on a combination of the following modern approaches to the development of innovations:

1) multidisciplinary engineering analysis (PLM Ural, 2019), which involves parametric optimization and statistical analysis, sensitivity analysis, topological optimization through high-performance computing.

2) system design methodology, which consists in the formation of an interdependent balanced system "goals – projects – measures – resources – result", which is determined on the basis of an iterative procedure. It combines the results of analysis and modeling of the targeted application of complex technical systems and their components with the use of simulation models - design models of means and functioning of the system as a whole and its components (their external and internal relationships), innovation models based on targeted complex programs with taking into account limitations on resources (levels and rates of financing, production and experimental capacities, etc.).

3) the methodology of parallel engineering (concurrent engineering or C-engineering), which provides information interaction of remote participants of the innovation network during the design of complex technical systems. At the same time, the exchange of information takes place at all stages of design, which makes it possible to organize reduction of deadlines due to the parallel solution of project tasks. The methodology of parallel engineering involves the implementation of development and design processes simultaneously with the modeling of manufacturing and operation processes.

It is assumed that the C-engineering project team must solve various design tasks in parallel to ensure the achievement of project goals. Each such task is associated with the implementation of certain important requirements, which in turn affect a number of other requirements. It is these relationships that are often not taken into account in the traditional (sequential) approach to the development of new products.

The basis of the application of C-engineering for the development of innovative networks is based on conceptual provisions about the possibility of a hierarchical representation of object systems, as well as the possibility of a hierarchical representation of the goals and tasks of the designed complex system.

The hierarchical basis of these methods is based on the concept of distributed access to product information. It includes a set of data that is created and used during its life cycle and includes information about its configuration, characteristics and properties, organizational information (a description of the processes associated with changing product data, necessary resources, etc.), information about testing and all product data from the moment of design to sales and subsequent service.

An important element of the digitalization of innovation processes is the infrastructural provision of information exchange between various network participants (customers, researchers, developers, manufacturers, operators, etc.). When transferring project data from one ICT component to another, universal formats and appropriate organizational efforts are required. Under the conditions of a low information culture, these processes are accompanied by a large expenditure of labor and time, in particular, for repeated coding, which leads to numerous errors and loss of time.

The tasks of creating a new product, especially in high-tech areas, are characterized by high requirements for reliability and are complex in nature. Therefore, the cost of making changes to the project increases exponentially until the final stages of the project. In the case of using parallel engineering, many problems that may arise at later stages of the life cycle are detected and solved already at the appropriate design stage.

Studies of the development of high-tech areas and technology support networks have shown that the main trend in the development of the modern innovation system is a new type of infrastructure – virtual, which functions in the virtual world of the Internet, complements the existing ecosystem, overcoming its limitations and providing radically new opportunities. A number of projects were developed for informational support of such entrepreneurship. As part of the National Industrial Information Infrastructure Protocol (NIIIP) project in the USA, the development of open software protocols for industry was carried out, which allowed manufacturers and suppliers to interact effectively. In EU, a number of ESPRET (ESTimation and PRediction of Execution Time) projects have been developed to create an appropriate infrastructure for virtual entrepreneurship, in particular for small and medium-sized firms.

In the work [24], the system of joint design and production for virtual organizations CDMS (Collaborative design and manufacturing system) was investigated, the architecture of the agent for the implementation of such virtual organizations in CDMS was revealed, and an example of joint planning of several enterprises was presented. A vivid example of a virtual scientific organization is the program Partnerships for Advanced Computational Infrastructure (RASI), founded by the US National Science Foundation, which includes more than 50 institutes and thousands of scientists. This association is aimed at the development of technologies in the field of informatics for the general use of resources in the process of scientific research. On the basis of the above, a virtual innovation structure can be defined as an association of organizations participating in the innovation process, the key competencies of which form the innovation cycle, and the interaction is carried out in the virtual space on the basis of technical and semantic compatibility, and whose activities go beyond geographical boundaries and cover stages of development and promotion of innovative ideas. Virtual integration is aimed at openness and active exchange with the external environment within the paradigm of "open" innovations.

Virtual integration assumes a relationship of complete

information transparency between partners, maximum flexible interaction, which allows to optimize material flows and reduce costs. The functional environment is formed as necessary – the same entity can simultaneously be a participant in international innovation projects or their parts, through participation in active networks.

V. CONCLUSIONS

The practical value of the presented proposals should be considered in the context of the practical implementation of the recommendations of the analytical report of OECD [25], dedicated, in particular, to the development of DSIP (Digital science and innovation policy) initiatives in many countries. The report notes that digital platforms and solutions in the field of science and innovation policy increasingly connect different sources of information and use new technologies and applications for innovation analytics. For example, government structures can use DSIP systems for policy design, implementation, monitoring and evaluation, and financial and credit institutions for planning, coordination, monitoring and evaluation of their activities. As the volume of research and innovation data increases, DSIP systems will help determine ways to transform science and technology policy and provide public services.

The results of the study show that DSIP-systems can be presented in different versions in the context of national policies. Thus, the creation of DSIP-systems taking into account the specific nature of digital management in the field of scientific and innovation policy is an urgent task for the development of competitive national innovation systems.

ACKNOWLEDGMENT

Research was funded by Ministry of Education and Science of Ukraine within the projects «Organizational and economic support of the post-war sustainable development of territories based on the infrastructure and service methodology of innovation communities' development», LET EDU, "Territory of innovations: best practices for sustainable development at the local level" and Estonian National scholarship program for foreign students, researchers and lecturers in Estonian Entrepreneurship University of Applied Sciences.

REFERENCES

[1] R. Geissbauer, E. Lubben, S. Schrauf, and S. Pillsbury, "Digital Champions. How industry leaders build integrated operations ecosystems to deliver end-to-end customer solutions," PricewaterhouseCoopers, 2018.

[2] Industrie 4.0 Working Group, "Recommendations for implementing the strategic initiative Industrie 4.0," 2013.

[3] R. Morrar, H. Arman, and S. Mousa, "The Fourth Industrial Revolution (Industry 4.0): a social innovation perspective," Technology Innovation Management Review, № 7 (11), pp. 12-20, 2017.

[4] L. Melnyk, O. Kubatko, I. Dehtyarova, O. Matsenko, and O. Rozhko, "The effect of industrial revolutions on the transformation of social and economic systems," Problems and Perspectives in Management, Vol. 17 (4), pp. 381-391, 2019.

[5] L. Melnyk, O. Kubatko, V. Piven, K. Klymenko, and L. Rybina, "Digital and economic transformations for sustainable development promotion: A case of OECD countries," Environmental Economics, 12(1), pp. 140–148, 2022.

[6] UNCTAD, "Information Economy Report 2017: Digitalization, Trade and Development," United Nations publication. Sales No. E.17.II.D.8. New York, Geneva, 2017.

[7] L. Zhang, Zh. Cao, G. Chen, and Zh. Wang, "A study of China's inter-city networks for innovation cooperation within software and service firms," Eurasian Geography and Economics, Vol. 60, Iss. 5, pp. 582-615, 2019.

[8] A. Aine, T. Björkroth, and A. Koponen, "Horizontal information exchange and innovation in the platform economy – a need to rethink?" European Competition Journal, Vol. 15. Iss. 2-3, pp. 347-371, 2019.

[9] R. Huggins, "The growth of knowledge-intensive business services: innovation, markets and networks," European Planning Studies, Vol. 19. Iss. 8, pp. 1459-1480, 2011.

[10] J.P. François, F. Favre, and S. Negassi, "Competence and organization: two drivers of innovation," Economics of Innovation and New Technology, Vol. 11. Iss. 3, pp. 249-270, 2002.

[11] M. Ghobakhloo, "Determinants of information and digital technology implementation for smart manufacturing," International Journal of Production Research, 2019.

[12] J. Khalid, B.R. Ram, M. Soliman, A.J. Ali, M. Khaleel, and M.S. Islam, "Promising digital university: a pivotal need for higher education transformation," International Journal of Management in Education, Vol. 12, № 3, pp. 264-275, 2018.

[13] V. Omelyanenko, O. Prokopenko, O. Kudrina, M. Biloshkurskyi, and O. Omelyanenko, "Digital Component of Innovation Landscapes: Con of Sustainable Development at the Local Level," 2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 – Proceedings, pp. 1324–1328, 2021.

[14] B. Quattrociocchi, M. Calabrese, X. Hysa, and E. Wankowicz, "Technology and innovation for networks," Journal of Organisational Transformation & Social Change, Vol. 14, Iss. 1. pp. 4-20, 2017.

[15] P. Santoleri, "Diversity and intensity of information and communication technologies use and product innovation: evidence from Chilean micro-data," Economics of Innovation and New Technology, Vol. 24. Iss. 6, pp. 550-568, 2015.

[16] R. Kattel, V. Lember, and P. Tõnurist, "Collaborative innovation and human-machine networks," Public Management Review, 2019.

[17] Ch. Yang, Q. Huang, Zh. Li, K. Liu, and F. Hu, "Big Data and cloud computing: innovation opportunities and challenges," International Journal of Digital Earth, Vol 10. Iss. 1, pp. 13-53, 2017.

[18] T. Tirto, Y. Ossik, and V. Omelyanenko, "ICT support for Industry 4.0 innovation networks: education and technology transfer issues," Lecture Notes in Mechanical Engineering. Springer, Cham, 2020.

[19] V. Omelyanenko, and D. Volodin, "Nanoinformatics application framework for R&D and industrial analisys," 2017 IEEE 7th International Conference Nanomaterials: Application & Properties (NAP), Odessa, UKraine, pp. 01NNPT03-1-01NNPT03-4, 2017.

[20] Y. Robul, I. Lytovchenko, L. Tchon, O. Khanova, and O. Omelianenko, "Digital marketing tools in the value chain of an innovative product," International Journal of Scientific and Technology Research, 9(4), pp. 158–165, 2020.

[21] V. Omelyanenko, "Analysis of strategycal aspects of technology transfer in metallurgy," Metallurgical and Mining Industry, No.12, pp. 394–397, 2015.

[22] V. Omelyanenko, "Preconditions analysis of using of technological package concept for development strategy of space metallurgy," Metallurgical and Mining Industry, No. 8, pp. 508–511, 2015.

[23] Portulans Institute, "The Network Readiness Index 2019: Towards a Future-Ready Society," S. Dutta & B. Lanvin (Eds). Washington D.C., USA, 2019

[24] W. Shen, "Virtual organization in collaborative design and manufacturing system," Proceedings of the 4th International Workshop on CSCW in Design. Compiegne. pp. 177-184, 1999.

[25] OECD, "OECD Science, Technology and Innovation Outlook 2018: Adapting to Technological and Societal Disruption," OECD Publishing, Paris. 2018.

Sensor System for Real-time Water Quality Monitoring

Elena Simonoska*, Dijana Capeska Bogatinoska*, Ile Dimitrievski*, and Reza Malekian**,***
* University of Information Science and Technology "St. Paul the Apostle", Ohrid, N. Macedonia
** Department of Computer Science and Media Technology, Malmö University, Malmö, 205 06, Sweden
*** Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, 0002, South Africa
elena.simonoska@cse.uist.edu.mk, dijana.c.bogatinoska@uist.edu.mk, ile.dimitrievski@uist.edu.mk, reza.malekian@ieee.org

Abstract **- Water pollution is a global issue that has an impact on the entire ecosystems' life cycles. Traditional sampling and laboratory testing techniques are labor-intensive and error-prone, making them ineffective for quickly detecting changes in water quality. This paper presents the development of a low-cost, portable and efficient prototype sensor-based system for monitoring water quality in real-time. The system consists of a microcontroller, temperature, turbidity, pH, and distance sensors, and an application for a visual representation of the data. Extensive testing was carried out to ensure uninterrupted operation. The prototype is a user-friendly sensor system that can be positioned close to the target area in order to assist in preventing environmental and biological harm. This can ensure safe, healthy, and sustainable water supplies for the communities, environment, and the economy. Continuous monitoring of water parameters can also help avoid critical situations. The experimental results demonstrate a successful development of a smart water quality monitoring system with potential applications in various scenarios.**

Keywords - Internet of Things; sensor system; water quality; water parameters; community services.

I. INTRODUCTION

Water is an essential resource for the survival of all living organisms and is vital for human. However, ensuring safe and clean water is a challenging task with the ever-increasing urbanization and the introduction of new solutions. The rapid pace of society has led to a deterioration in the quality of water, which has been affected by various factors, including sewage, industrial discharges, agricultural activities, and natural disasters such as floods and droughts. Regardless of the purpose of water usage, whether for drinking, domestic use, recreation or food production, it must be safe and readily available. Water is essential for public health, and it is crucial to establish and maintain a water quality balance. Otherwise, not only human health, but also the entire ecosystem can be severely damaged. Providing pure drinking water has become a major challenge worldwide, and many international governing bodies, such as the United Nations (UN) and the World Health Organization (WHO), are involved in efforts to ensure safe water everywhere in the world. In 2010, the UN General Assembly adopted a resolution stating that access to safe and clean water is a human right and is essential for full enjoyment of life [1]. The UN also highlighted five reasons for clean water to help communities to thrive, which include sustainable development, socio-economic development, energy and food production, health and survival, and healthy ecosystems [2]. The consequences of water pollution are numerous, and include reducing biodiversity, scarcity of drinkable water, diseases, mortality, destruction of animal and plant habitats, and acid rain [3].

Methods for monitoring quality of water in real-time [4] can ensure safe, healthy, and sustainable water supplies for the communities, environment and the economy. Additionally, monitoring water parameters in a watershed, rivers or lakes demonstrates how numerous natural and anthropogenic stressors may have an effect on quality of water [4].

Traditionally, water quality has been assessed by collecting water samples and testing in a laboratory. However, there is a need for a low-cost and quick solution to monitor water quality because real-time results are required. Our research work focuses on the development of an application for monitoring water quality by leveraging sensors that measure a range of water parameters and display them on the application. IoT solution increase the automation and reduce the impact of human errors in industrial applications. The obtained results are with increased reliability because testing is done at the spot and results are simultaneously delivered to user. The proposed system is versatile, as it can be adapted to various scenarios and use cases related to water quality. The application is available on both web and mobile platforms, and it provides users with the ability to track changes in water parameters in real-time. Furthermore, users can receive alerts when values are out of the range, view data in the form of charts, and constantly update values as they are received from the sensors. The proposed low-resources and low-time-consuming application will help people, especially those who require a cost-effective solution, to make informed decisions about water quality.

The rest of this paper is organized as follows. We summarize the literature review in section II. Section III

discusses the proposed system, including the software and hardware requirements. The development of the system is detailed in Section IV, while testing and results are demonstrated in Section V. Next, we conclude the research work.

II. Literature Review – an Overview

The extensive literature review conducted in order to initially formulate a solution to the problem, as well as to decide on the most suitable methodology for designing the individual components of the proposed system, led to the utilization of the works developed in [3] and [5]. These works were used to gain an understanding of the impact of water on both the environment and living beings, with the WHO [5] providing numerous studies and figures to illustrate the effects of water pollution on human health. Several recent research works [6], [7], [8], and [9] consulted to gain an in-depth understanding of what water quality is and its parameters. Hassan Omer et al. [9] provided a summary of water classifications based on the values of its parameters. Furthermore, various research works were consulted to explain the hardware used in the proposed system. Research works in [10], [11], and [12] provided detailed information regarding the origin, operating capabilities, and development of models of the Raspberry Pi. The authors in [13], [14], and [15] provided details on the sensors, which form the core of the system, and their classification. Azzuni et al. [16] analysed various alternatives for measuring water quality, ultimately leading to the conclusion that the proposed system outperforms other solutions in several aspects. The first alternative is laboratory testing, which is inconvenient due to the need for sample transportation and is thus open to external factors which may affect the results. Additionally, it is a slow process as the procedure is not done on-site. The second option is remote sensing, with remote sensors attached to satellites or aircrafts to take images of water bodies at different wavelengths, allowing the detection of the presence of certain substances. This process is contactless and can only detect visible changes. The third option is less common and involves observing the behaviour of organisms in normal conditions, followed by controlled changes to observe how the organisms respond. However, the results may be unreliable as the organisms may be affected by other factors.

The manual nature of the aforementioned alternatives, which require human involvement, increases the likelihood of human error. This necessitates the development of a smart, automated system which can monitor water in real-time, thus increasing reliability as external factors are not present.

III. Proposed System

Our proposed system for smart and real-time water quality monitoring is an Internet of Things (IoT) solution which utilizes a Raspberry Pi microcontroller (MCU), four sensors, and additional hardware for the successful interconnection between them. This is further supported by a web application with unlimited access from a variety of devices.

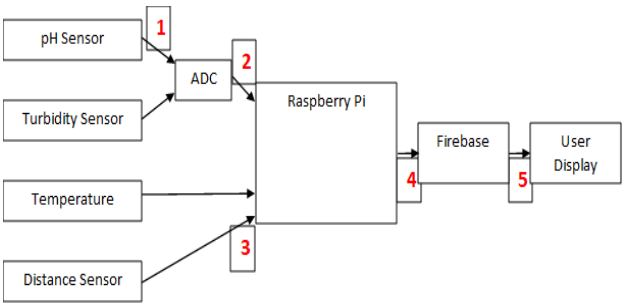


Figure 1. System's interfaces

The system has five types of connections, each with a distinct purpose and every one of which is necessary for the proper functioning of the system. The system's interfaces are given in Fig. 1.

The user can observe the measured data on their device, and the data is organized by the system, but must be taken from a database. The data in the database is retrieved from the MCU, which interacts with the sensors.

To ensure the successful functioning of the project, there are four categories of requirements that must be considered:

A. *Functional Requirements*

Functional requirement defines the actions that must be performed by the system to allow users to accomplish their tasks and utilize its capabilities. There are two ways to access the application; as an administrator and as a user.

B. *Non-Functional Requirements*

Non-functional requirements define the quality attributes of the system and describe how well the system should function. The attributes that are evaluated are: usability, accessibility, interoperability, learnability, portability, scalability, and security.

C. *Performance Requirements*

The performance requirement refers to the operating capabilities of the system in terms of accuracy, response time, and memory. The performance of the hardware and software components is essential for the proper functioning of the system and managing data correctly. This includes:

- Programming the sensors to read every 5 minutes.
- Updating the database with new data every 5 minutes.
- Refreshing the dashboard page every 1 minute.
- Comparing the measured values with reference values to classify the sample.
- Organizing the data into a table for better visual representation.
- Displaying the maximum, minimum value, their difference, and current value on the parameter page.

D. *Hardware and Software Requirements*

Hardware and software requirements include components that are necessary for the system to operate. These requirements are prerequisites for the system.

These components are described below:

A. *Hardware Description*

The main components of the project are the Raspberry Pi 4 Model B microcontroller and the sensors. The sensors are used to measure four characteristics of the water quality: pH, temperature, turbidity, and depth.

a) pH Sensor

This sensor is used to measure how acidic or basic the water is, ranging from zero to 14. All solutions with pH ranging from zero to 7 are classified as acidic, while samples between 8 to 14 are categorized as basic. Solution of pH 7 is considered as a neutral. Most water plants and animals require an environment of a specific pH, and water with a pH of 11 or higher causes irritation. The pH sensor used in this system has a measuring range from 0 to 14, a temperature range from 0 to 60°C, and a response time of ⩽ 1 minute.

b) Temperature Sensor

This sensor is used to measure the temperature of the water. The reference value of drinking water is different from the temperature of a fishpond, for example, or the temperature of industrial water cannot be compared with the temperature of the lake. Temperature is important because it affects the status of the water, such as conductivity, chemical and biological reactions, maximum dissolved oxygen concentration, and pH. Temperature is also crucial for organisms in the ecosystem, as they have preferred regimes and cannot adjust to changes. The temperature sensor used in this system has a measuring range from -55 to 125° C and a response time of ⩽ 750 ms.

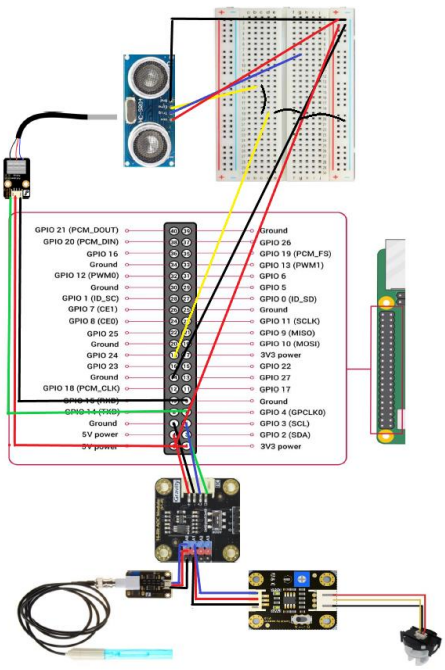


Figure 2. System schematic diagram

c) Turbidity Sensor

This sensor is used to measure the clarity of the water. Excess turbidity can cause illnesses, especially if the water is used for drinking. High turbidity water also affects the overall wellbeing of organisms living in the water. The turbidity sensor used in this system has an operating temperature of 5 to 90° C and a response time of ⩽ 500 ms.

d) Distance Sensor

This sensor is used to measure the level of water in a closed system or in water surfaces such as rivers, lakes, seas, or oceans. Monitoring the water level is used to detect the absence of water in the system, or to check if the level is within permissible limits. The distance sensor used in this system has a measuring range from 0.3 m to 4 m, and a response time of ~ 32 ms. Two resistors, 1kΩ and 2kΩ, are added for connecting the distance sensor to the microcontroller. The value for second resistor is approximately 2kΩ but the system uses two 1kΩ resistors in serial connection.

The schematic diagram of the solution is given in Fig. 2.

B. *Software Description*

The whole application is divided into three parts: Raspberry Pi, Front-End, and Back-End.

a) Raspberry Pi

Operates on Linux and the Python scripting language is used for programming. The communication between the user and the Pi is done using the terminal, known as the command line interface. All hardware-related codes are written in scripts.

b) Front-End.

HTML, CSS, and JS are used for creating a user-friendly application.

c) Back-End.

PHP, specifically the Laravel PHP framework, is implemented for the back-end programming. The reason for choosing Laravel is that it provides all the features required to build a modern web application, including: routing, validation, caching, queues, file storage, and middleware. It also allows for easy communication with the database.

d) Database.

Two databases are used for the purposes of the project: MySQL and Firebase. Firebase is used for interacting with the sensors, and the data is displayed on the user dashboard through Laravel. The MySQL database is used for storing the registered users and posted news.

IV. DEVELOPMENT OF A SYSTEM

The proposed sensor system is designed to provide real-time water quality monitoring. This system comprises measuring instruments, communication between the

microcontroller and database, and communication between the database and Laravel application. Fig. 3 illustrates the data delivery process to the user.

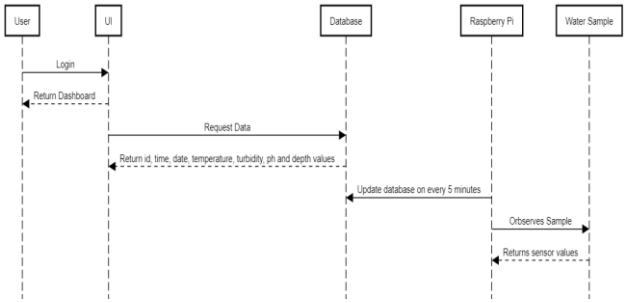


Figure 3. Sequence Flow Diagram

The starting point of this system is an application developed in Laravel, providing users with all relevant information related to the system. Any user, registered or not, can access the site and utilize the items in the navigation menu. Requests are made when the user performs an action, such as accessing a page, submitting a form, posting news, or updating news. Depending on the action, the route supports different methods.

Users of the system can access their dashboard, which provides a visual representation of the measured data. The data is retrieved from the database and organized in a table.

Users may be unregistered (visitors) or registered (admin and authenticated user). The application also provides constant support to its users, managed by the admin. The admin is responsible for ensuring that changes made to the system are reported to the users immediately. However, there is no limitation on what is published, and the admin can post any article that users may find useful and interesting. Furthermore, the admin has a mailbox for messages from users.

When the user logs in, they are redirected to the parameters dashboard. This page provides a menu for the user to choose from the parameters, and there are links to a full guide on using the system and relevant news. Fig. 4 shows the home page, which contains cards displaying the current values and notifications based on the parameter value, as well as a table with the measured data starting

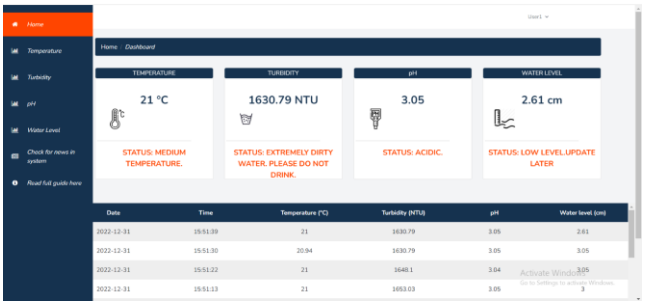


Figure 4. User Dashboard

from the latest added record. The data is organized in six columns: date, time, temperature, turbidity, pH, and depth.

The menu items and cards are clickable, directing the user to the desired parameter. Fig. 5 displays what appears after the temperature card or temperature sidebar item is clicked. There is also a message that shows the status of the particular parameter, the current temperature, minimum, maximum, and their difference. The same page with the corresponding parameter appears when clicking the rest of the items.

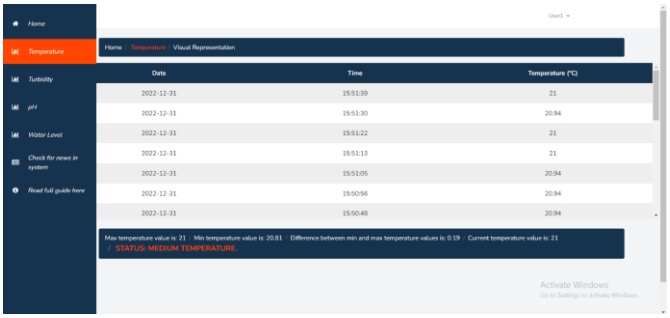


Figure 5. Temperature page

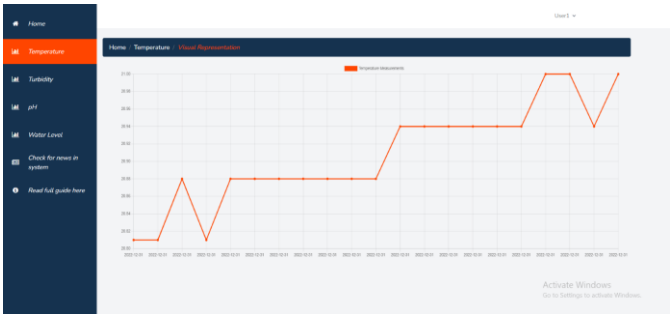


Figure 6. Temperature Graph

Proposed system is dependent on a constant current flow. The system can be adapted to work without constant flow by powering Raspberry Pi with batteries, power banks or other power sources. Moreover, it is an alternative which ensures stability in cases of defects.

The application provides graphs for a better visual representation of the measured parameters, available for

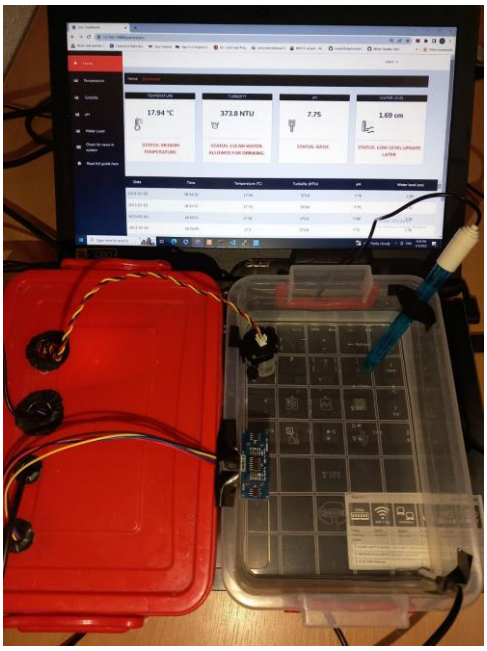


Figure 7. System prototype

all parameters. Fig. 6 shows an example of the temperature graph.

Every water source can be exposed to contamination due to agricultural, industrial, or domestic activities. Water can be harmed by heavy metals, pesticides, chemicals, oils, and other pollutants. The developed application determines the status of the water based on the values of the temperature, turbidity, pH, and level parameters using the classification given in [17], which are potable, palatable, contaminated, and infected.

Fig. 7 presents the system prototype, both the hardware and the software.

V. TESTING AND RESULTS

The development of the sensor system for real-time water quality monitoring necessitated extensive testing to adjust the formulas used to calculate the final values, as well as to determine the correct interval of measurement since the sensors had different response times. To assess the accuracy and reliability of the system, three phases of testing were conducted: unit testing, integration testing, and system testing. To validate the results obtained by each unit, different samples of water were used to ensure that the system works in different environments.

A. *Unit Testing*

Unit testing was the first step in evaluating the performance of each sensor. At this stage every unit is tested one by one to check accuracy and reliability of results. The unit testing phase validates that each unit of the software performs as designed.

a) pH Unit Testing

For assessing the correctness of the system measurements, the results were compared to the reference values of each sample. Fig. 8 presents examples of the samples used, while Table I shows the measurements

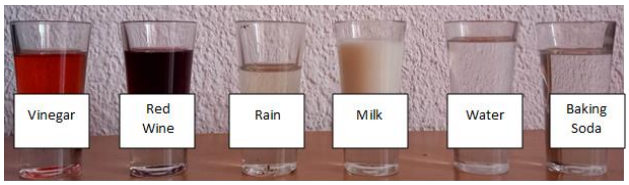


Figure 8. Liquid Samples

TABLE I. PH COMPARISION OF EXPECTED AND OBTAINED VALUES

Substance	***Vinegar***	***Wine***	***Rain***	***Milk***	***Pure Water***	***Baking Soda***
Expected Value	2.8	4	6.5	6.8	7	8.3
Expected Category	Acidic	Acidic	Acidic	Acidic	Neutral	Basic
System Value	2.76	4.09	6.38	6.69	7.24	8.19
System Category	Acidic	Acidic	Acidic	Acidic	Basic	Basic
Error Rate	1.42%	2.25%	1.88%	1.61%	2.25 %	1.32

obtained and their corresponding reference values [18]. At this stage, the value of each sample was tested and categorized as acidic, neutral, or basic depending on the measured value.

b) Temperature Unit Testing

Two samples were used to evaluate the temperature readings. The results obtained by the sensors are displayed in Fig. 9.

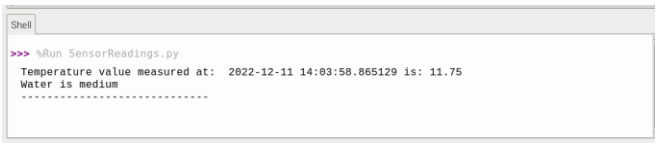


Figure 9. Temperature Sensor Readings

c) Turbidity Unit Testing

Three samples were used to test the turbidity sensor. Fig. 10 and Fig. 11 show the results for two of them.



Figure 10. Turbidity Sample 1



Figure 11. Turbidity Sample 2

TABLE II DISTANCE COMPARISION OF EXPECTED AND OBTAINED VALUES

	Measurement 1	***Measurement 2***
Expected Value	10.00 cm	4.00 cm
System Value	10.27 cm	3.98 cm
Error Rate	2.7%	0.5%

Note that turbidity sensor reacts on particles in water, not on water colour.

d) Water Level Unit Testing

To compare the measurements with the ruler and the sensor, two different distances were used and the results are summarized in Table II.

B. *Integration Testing*

This phase of testing was used to prove that all sensors worked together and were able to measure the water parameters accurately. Three samples were tested and the results are displayed in Figures 12 (Sample 1) and 13 (Sample 3). Sample 1 was clear water with a pH that was within the expected range and a turbidity that was expected to be between 0 and 1. Sample 2 was a mixture of milk and water, resulting in a reduction of the pH from 7.7 to 6.95, and an increase of turbidity due to the milk.

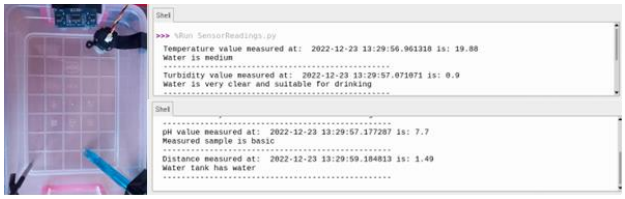


Figure 12. Integration Testing, Sample 1

Sample 3 was a mixture of water, milk, sand, and baking soda, which affected the pH of the previous solution from 6.95 to 9.01, and drastically increased the turbidity due to the sand.

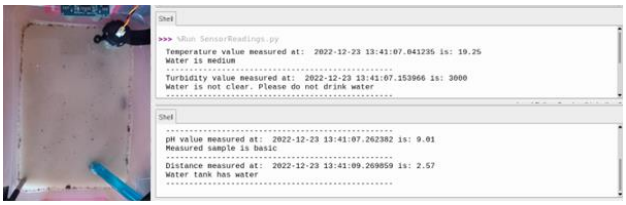


Figure 13. Integration Testing, Sample 2

C. *System Testing*

The aim of this phase of testing was to prove that the sensors, the communication with the database, and the Laravel application were all functioning properly. First, the parameters were measured, then sent to the database, and displayed on the user dashboard (Fig. 14). Fig. 15 demonstrates that the data inserted into the database was successfully retrieved on the dashboard. The table also starts with the most recent record, and the time, date, and

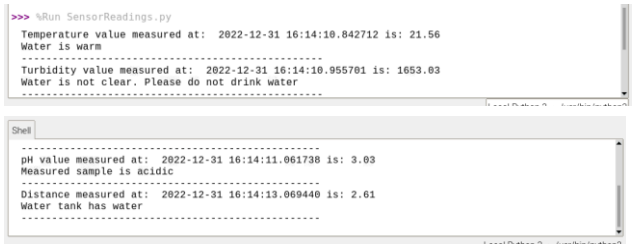


Figure 14. Parameter values obtained

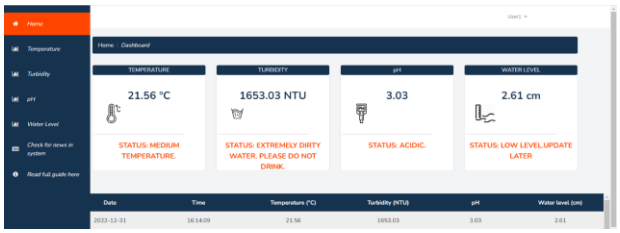


Figure 15. Parameter in dashboard

values all correspond to those in the Shell.

VI. CONLUSION

In conclusion, the proposed smart and real-time water quality monitoring system helps in detecting water parameters and issues in quality in early stages. It is an effective, robust and scalable platform which can be expanded with more sensors to ensure more precise insight into water quality and to address the growing demand for environmental monitoring and protection. The system can be used for a variety of purposes, with the potential to help protect the environment and living organisms. Through continuous monitoring of water parameters, harmful consequences can be avoided. This system has the potential to benefit many industries and organizations, such as healthcare, agriculture, industry, and ecological societies, and could greatly enhance our everyday lives. Additionally, the availability of both web and mobile versions makes it accessible to a wide range of users, providing them with direct insight into the results.

REFERENCES

[1] "UN General Assembly, The human right to water and sanitation: resolution adopted by the General Assembly", A/RES/64/292, 2010.[Online].Available: https://www.refworld.org/docid/4cc926b02.html. [Accessed: 19-Oct - 2022].

[2] "Why is Clean Water Important to Communities?", Healing Waters Media, 2021. [Online]. Available: https://healingwaters.org/why-do-communities-need-clean-water/#:~:text=Unsafe%20water%20causes%20water-borne,%2C%20and%20hygiene%20(WASH). [Accessed: 22- Oct - 2022].

[3] V. Lakshmikantha, A. Hiriyannagowda, A. Manjunath, A. Patted, J. Basavaiah, and A. A. Anthony, "IoT based smart water quality monitoring system", Global Transitions Proceedings, vol. 2, Issue 2, 2021, pp. 181-186, ISSN 2666-285X, https://doi.org/10.1016/j.gltp.2021.08.062.

[4] N. Cloete, R. Malekian, and L. Nair, "Design of Smart Sensors for Real-Time Water Quality Monitoring", IEEE Access, Vol. 4, pp. 3975 - 3990, 2016.

[5] World Health Organization (WHO), "Drinking water", 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/drinking-water

[6] D. Dunea, "Water Quality: Factors and Impacts", Intechopen, 2022. DOI 10.5772/intechopen.96701. ISBN 978-1-83969-548-3

[7] S. Geetha, and S. Gouthami, "Internet of things enabled real time water quality monitoring system". Smart Water 2, 1, 2016. https://doi.org/10.1186/s40713-017-0005-y

[8] S. Pasika, and S. T. Gandla, "Smart water quality monitoring system with cost-effective using IoT". Heliyon, vol.6, issue7, E04096, 2020. DOI:https://doi.org/10.1016/j.heliyon.2020.e04096

[9] N. Hassan Omer, "Water Quality Parameters," Water Quality - Science, Assessments and Policy, Jul. 2020, doi: 10.5772/intechopen.89657.

[10] G. Halfacree, "The Official Raspberry Pi Beginner's Guide (The Official Raspberry Pi Beginner's Guide: How to use your new computer)", Raspberry Pi Press; 4th edition, 2020.

[11] A.K. Dennis, "Raspberry pi computer architecture essentials: Explore raspberry pi's architecture through innovative and Fun Projects", Birmingham, UK: Packt Publishing, 2016.

[12] W. Gay, "Advanced Raspberry Pi: Raspbian Linux and GPIO Integration", Apress, 2nd edition. 2018.

[13] B. Charan Patel, G. R. Sinha, and N. Goel, "Introduction to sensors", Book chapter in book: Advances in Modern Sensors, 2020. DOI: 10.1088/978-0-7503-2707-7ch1

[14] J. Fraden, "Handbook of Modern Sensors: Physics, Designs, and Applications" (5th ed. 2016), Springer, 2015.

[15] D. Perlstein, "Classifying Different IoT Sensors & Their Uses – Axonize Types of IoT Sensors", Axonize, 2019.

[16] A. Azzuni, "Design, implementation, and evaluation of an online water quality monitoring system in Lake Saimaa, Finland", 2014. DOI: 10.13140/RG.2.1.3420.6561

[17] N. Hassan Omer, "Water Quality Parameters," Water Quality - Science, Assessments and Policy, Jul. 2020, doi: 10.5772/intechopen.8965

[18] Periodic Table of the Elements - pH. (n.d.). [Online]. Available: http://coolperiodictable.com/resources/acids-and-bases/pH-of-some-common-substances.php

Proposal of Prediction Model for Smart Agriculture Based on IoT Sensor Data

Jakup Fondaj*, Mentor Hamiti * , Samedin Krrabaj**, Besnik Selimi* and Ermira Idrizi*
* South East European University, Tetovo, North Macedonia
** University "Ukshin Hoti", Prizren, Kosovo
jf13459@seeu.edu.mk
m.hamiti@seeu.edu.mk
samedin.krrabaj@uni-prizren.com
b.selimi@seeu.edu.mk
e.idrizi@seeu.edu.mk

Abstract **- Prediction is a mechanism to prevent and make changes in different sectors. Climate changes are affecting different sectors but the most damaged is the agriculture sector because it affects the quality and quantity of agricultural products. Farmers who produce grapes in recent years face many challenges due to climate change, this is evident also in our region Kosovo. Prizren region is known for grape products and many families depend on grape production. For this reason, in this paper, we are going to show recent research in this area and present the most appropriate algorithms used for the prediction of smart IoT sensor data for the prediction of the quality and quantity of agricultural products. In this paper, we propose a model for the prediction of the quantity and quality of grapes in order to recommend what is the right time to do grape processing and grape sterilization because those are the crucial fact that affects the quality and quantity of the grape.**

Keywords - Internet of Things; Artificial intelligence; data mining; IoT sensor data; smart agriculture;

I. INTRODUCTION

The quality and quantity of agricultural products nowadays are affected a lot because of climate change. Using IoT smart sensors to collect data and predict what will happen in the future to prevent possible damage to agricultural products. By using smart agriculture sensor data, we predict the temperature for next year in order to help farmers in grape cultivation by specifying the period of time when the grape harvest is to be done which has a direct impact on the quality and quantity of grapes.

Human history can be described through the history of agriculture. Agricultural landscapes represent the very essence of the agricultural and cultural heritage of the population living and working in it. Food, land, water, environment, and population growth are challenges to which governments and institutions must respond. In the past, the combination of measures and instruments was used to support agriculture that had insufficient results due to climate variability and market changes, but today with new technology we can have sustainable solutions.

The authors in this paper [1] elaborate on the term smart farming and show different examples of why it is important to use technology in order to improve the production of different agricultural products.

The objective of smart agriculture is to provide solutions that can be applied to all farmers, regardless of farm size, region, or sector, exploiting scale effects and keeping the cost of technology low. In this context, policies are needed to ensure high-speed data transmission and harmonized interoperability, the adoption of standards at the European level to promote connections in the rural area, and adequate infrastructure and data processing services.

So far, the expected benefits from the integration of technological processes in agriculture are high production efficiency and quality, cost reduction, optimization of inputs, minimization of medical impacts, also the creation of jobs for specialized personnel. Yet thanks to support for agriculture, policies and growing awareness of the economic and ecological benefits that new technologies in agriculture can produce real the potential to spark a process of radical transformation in the sector. According to analysts [2], the global smart agriculture market which was worth almost $ 10 billion in 2017, will reach $ 23 billion by the end of 2022.

Having a strategy for smart agriculture will help farmers to make a decision and improve food production [3]. FAO [3] estimated that if the present production and consumption rates continue, agricultural production should increase 60% by 2050 to meet the needs of food of the world's population. Because the climate has changed, it is important to integrate Information and communication technologies (ICT) to protect food production.

The main idea is to propose a model for predicting the grape yield and its quality based on weather conditions parameters and past year's grapes yield.

The aim of this paper is to analyze the existing research in this field and based on the analysis we will give information to farmers for a potential year what the

quality of the grape will be and recommend when is the right time to do the grape processing and grape sterilization because these are crucial factors to the quality and quantity of the grape.

The following in the second section is related work on this particular topic compared to what we have done in this research. The third section defines what smart agriculture is and its relation to IoT. The fourth section describes the prediction model for Smart agriculture. Then in the fifth section, the fourth phase model for prediction in smart agriculture is presented. And last is the conclusion and future works.

II. Related Work

Statistical models predict that the world's population will reach 8.5 billion by the end of 2030 [4]. This represents a real threat to our food security and puts the current food production system under pressure. Efficient use of Earth's natural resources is the only solution to facing future challenges such as global hunger. The implementation of precision agriculture using new technologies such as artificial intelligence, big data, IoT, and remote sensing is the first step toward this goal [5], [6].

S. S. Rathore et al. [7] proposed a prediction model for smart agriculture using IoT sensor data. The model incorporates machine learning algorithms and data analytics to predict crop yields, plant health, and soil moisture. The authors also discuss the importance of using IoT sensor data to optimize agricultural practices, reduce waste, and increase efficiency.

Y. Liu et al. [8] developed a predictive model for precision agriculture using IoT sensor data. The authors used a decision tree algorithm to analyze data from soil moisture and temperature sensors and made predictions about crop growth and yield. The authors note that their model can be used to optimize irrigation and fertilization practices, leading to increased crop yields and reduced water usage.

M. Tariq et al. [9] proposed a crop yield prediction model for smart agriculture using IoT sensor data. The authors used data from various sensors, including temperature, humidity, and light sensors, to predict crop yields. The authors also discuss the potential benefits of using IoT sensor data for smart agriculture, including reduced labor costs and increased crop yields.

D. Kumar et al. [10] developed a predictive model for crop disease detection and prevention using IoT sensor data. The authors used machine learning algorithms to analyze data from soil moisture and temperature sensors, as well as image data from cameras, to identify crop diseases and make recommendations for prevention and treatment. The authors note that their model can help farmers reduce crop losses due to diseases and increase crop yields.

S. H. Jung et al. [11] proposed a prediction model for grape quality using IoT sensor data. The authors used data from various sensors, including temperature, humidity, and leaf wetness sensors, to predict grape quality. The authors also discuss the potential benefits of using IoT sensor data for smart agriculture, including improved crop quality and reduced labor costs.

These studies highlight the potential benefits of using IoT sensor data for smart agriculture, including increased efficiency, reduced labor costs, and improved crop yields and quality. The studies also demonstrate various predictive models that can be developed using IoT sensor data, including models for crop yield prediction, precision agriculture, crop disease detection and prevention, and grape quality prediction.

Agriculture has always been an area that has required a lot of work and frequent maintenance even for the smallest work. It also requires careful supervision so that the person is able to know when the plants are ready to be harvested, or even to see the diseases of the harvest as early as possible, and to definitely notice the insects in the plants because otherwise they are damaged. In a word, we can say that a lot of work is needed to increase productivity in this area, and here begins to include smart agriculture.

Based on the source [12] by 2050 the world's population will increase from 7.3 billion to 9.7 billion inhabitants on the planet, which means that food production must increase by 70%, where the only way to achieve growth is with the intervention of new technology.

In recent years, grape cultivation has become difficult for farmers due to frequent climate change. The goal of mining and predicting the grape quality and yield of grape is to help viticulture farmers to take decisions related to their farms and business. We have 5 years of data from weather hydrometeorological institution where the air pressure, air temperature, humidity, wind direction, and wind speed is measured which are very important parameters to be analyzed related to our topic. From IoT sensors for smart agriculture, we also collect data for land humidity where we are going to analyze how it affects the grape quality and yield.

The main idea is to propose a model for predicting the grape yield and its quality based on weather conditions parameters and past year’s grape yield.

Based on the analyses we will give information to farmers for the potential year on what the quality of the grape will be and recommend what is the right time to do the grape processing and grape sterilization because are crucial factors that affect the quality and quantity of the grape.

There are different studies in this field but no such study is done yet for our region.

III. Smart Agriculture and IoT

Human history can be described through the history of agriculture. Agricultural landscapes represent the very essence of the agricultural and cultural heritage of the population living and working in it. Food, land, water, environment, and population growth are challenges to which governments and institutions must respond. In the past, the combination of measures and instruments was used to support agriculture that had insufficient results due

to climate variability and market changes, but today with new technology we can have sustainable solutions.

By smart agriculture [1], [13], [3] we mean the use of digital techniques to recreate, govern, and optimize production processes. Digital transformation encourages human intervention in agriculture and helps reduce the workload, carry out specific measures, to calibrate the use of chemical products of soils and crops, in addition to guaranteeing and increasing crops. It also helps manage all those processes that enable or support agricultural production, including economic administrative processes.

The authors in this paper [15] elaborate on the term of smart farming and show different examples of why is important to use technology to improve the production of different agricultural products.

The objective of smart agriculture is to provide solutions that can be applied to all farmers, regardless of farm size, region, or sector, exploiting scale effects and keeping the cost of technology low. In this context, policies are needed to ensure high-speed data transmission and harmonized interoperability, the adoption of standards at the European level to promote connections in the rural area, coupled with adequate infrastructure and data processing services.

So far, the expected benefits from the integration of technological processes in agriculture are high production efficiency and quality, cost reduction, optimization of inputs and minimization of medical impacts, and also the creation of jobs for specialized personnel. Yet thanks to support for agriculture, policies and growing awareness of the economic and ecological benefits that new technologies in agriculture can produce really have the potential to spark a process of radical transformation in the sector.

IV. A Prediction Model of Smart Agriculture Based on IoT sensor Data

Smart agriculture in recent years is a very interesting area to research and are doing a lot of studies in this area. A prediction is a tool for decision-making for farmers before they predict the yield of certain crops through their own practice and observed weather parameters. For prediction, the main field of research is Data Mining because it can predict crop yield by discovering unseen information from the influence of the existing climatic dataset [14].

Here we are going to show the researches which uses prediction algorithms to analyze data collected from IoT sensors.

The aim of this paper [15] is to show the impact that has climate changes on the cultivation of grapes in northeast Italy. The paper aims to verify if the grape yield and quality are affected by planting options and climate conditions. Six weather stations are placed in the study area in Italy where a new dataset is created and data as a mean of daily rainfall, temperature, and relative humidity among the six stations are collected. They analyzed 10-year data where weather conditions are analyzed. Also, geo-data are analyzed to see how the land position affects the quality and quantity of grapes. They measured sugar concentration and tartaric acid. There was a positive linear regression between climate conditions and the quality and quantity of grapes.

The authors [12] in this study propose a prediction model of Apple disease in the apple orchards of Kashmir valley using data analytics and Machine learning in IoT systems. To collect data a survey is conducted with local farmers in order to know if technologies will help them in decision-making. The paper also shows the challenges that are faced when technologies for smart agriculture are incorporated into the traditional environment. The paper proposes a prediction model based on linear regression and data analyses.

The other study [17] also proposed a model for smart agriculture based on mobile which integrates IoT and cloud-based big data analytics. The aim of the proposed model is to predict the crop yield and decide the better crop sequence based on the previous crop sequence in the same farmland with the soil nutrient current information. The proposed model estimates also the total production and for crop region-wise estimates the total fertilization required. This will help farmers to manage the costs of agricultural products.

The authors in this paper [14] present a model for classifying the data of sugarcane yield by implementing various decision tree algorithms where the classes are categorized as low, moderate, and high. They use Hadoop and Elasticsearch for Big Data mining. The proposed model is divided into three phases: Relief feature selection where training data are given as input after that is attribute ranking and then the filtering of the attribute is done. The second phase is called classification model implementation where is done MapReduce decision tree classification and comparison of C4.5, C5.0, and random forest decision tree is done. Last is the third phase named accuracy prediction where is done computation of classification accuracy. Comparison of these three algorithms C4.5, C5.0, and Random Forest is done by some form of measurement [14] as Accuracy, Error, Precision, Sensitivity, and Specificity. The result of this study shows that there is a correlation between climate conditions and sugarcane productivity. They propose MapReduce-based classification algorithms for classifying the sugarcane yield classes from accessible observed climatic parameters [14].

The authors in this paper [17] show a model for finding the correlation between weather conditions and Grape berry quality. The study is conducted in Portugal. They measured the temperature as average, min, and max for three months Jun, July, and August, and how the temperature affects the grape berry quality and quantity. They found that high temperature decreases berry weight, titratable acidity, anthocyanins, and TPI, and increases pH and potential alcohol. They use different algorithms such as correlation analysis, Pearson product-moment correlation coefficients, Student's t-test, and ANOVA analysis.

The authors in this paper [18] present the impact of climate change on viticultural products in Europe. They found that Increased future warming and dryness will probably result in an eventual overall loss of viticultural

suitability in the Mediterranean-like climatic areas of southern Europe, while in central and northern Europe, warming conditions will potentially benefit grapevine cultivation.

The authors [19] in this paper propose a model for grape quality prediction by combining open-source AutoML techniques and Normalized Difference Vegetation Index (NDVI) data for vineyards obtained from four different platforms-two proximal vehicle-mounted canopy reflectance sensors, orthomosaics from UAV images and Sentinel-2 remote sensing imagery-during the 2019 and 2020 growing seasons. The author of this paper investigates the application of open-source AutoML and multi-platform multi-temporal NDVI data for the evaluation of quality attributes of grapes. They compare it with the method proposed before [20].

V. PROPOSED MODEL FOR SMART AGRICULTURE SENSOR DATA

The recommendations that emerge from the model will depend on the specific predictions made by the algorithm. For example, if the model predicts that grapes will be of low quality due to high humidity, a recommendation might be to apply protective measures to prevent mold or other diseases that could affect grape quality. Alternatively, if the model predicts that grapes will be of high quality due to favorable weather conditions, a recommendation may be to increase vine irrigation to ensure optimal growth and yield.

Overall, the model focuses on predicting the quality and quantity of grapes based on environmental factors, and the resulting recommendations depend on the specific predictions made by the algorithm.

The proposed model now predicts the weather temperature, and humidity based on the data collected from IoT sensor data and the hydrometeorological institute based on it we analyze the data for grape yield for the last five year and we analyze the time frame when the grape sterilization is done.

Prediction of the quality and quantity of grapes is based on four essential elements:

- IoT sensor data which are used for training the model
- Data collected by interviews with farmers on how to cultivate grape
- Data from the Kosovo Hydrometeorological Institute and Institute of viticulture
- One or more prediction algorithms.

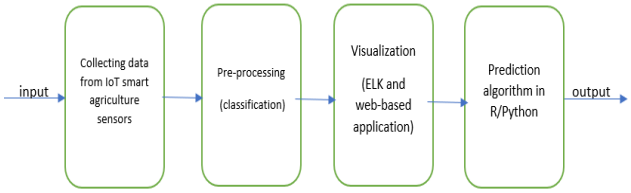


Figure 1. Proposed model for prediction of temperature

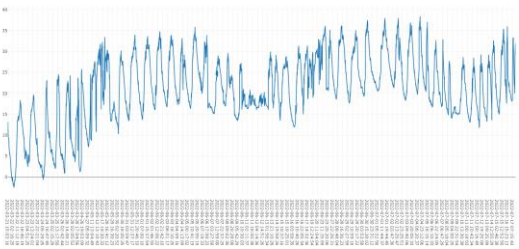


Figure 2. Temperature data collected from IoT smart agriculture sensors

A four-phase framework is proposed, consisting of IoT sensor data collection in real-time, preprocessing of data, real-time visualization of data, and at last prediction of quality and quantity of grapes. The proposed model is shown in figure 1.

The first step is to install IoT smart city and smart agriculture sensors and configure them to collect data in real-time and send them to Mashlium server where automatically the data are saved then in MySQL. The data are collected by sensors and sent throw a Wi-Fi connection to the main server where a MySQL database is installed. Is important to mention that the sensors can be located at a distance of 5 km from the center where the data are saved. Libelium sensors that we use are Smart City Pro and Smart Agriculture. The programing language for configuring them is like Arduino, based on C and C++. Table 1 below presents the data that are collected from Libelium IoT SmartAgriculture and SmartCity sensors. There are six parameters collected every minute (the time can be defined by us). The data are registered every five minutes and we have a dataset from march 2022 that continues. The second phase is pre-processing of data, the data are seasonal time-series data as shown in figure 2 below.

This phase includes the classification of data and removing unnecessary data. Here we use classification methods to remove unnecessary data and prepare them for visualization.

The third phase is visualization of data by ELK [21] and analyzing them in real time. ELK is composed from Elasticsearch, Kibana and Logstash where we have configured the data from the datacenter where we collect IoT data automatically in real time to send to ELK and visualize there. In Logstash we can configure and do pre-processing of data before visualizing them in Kibana. Kibana offer the possibility of visualization and also analyzing and the possibility of adding additional algorithms for visualization and prediction of data. Also, we have developed a web-based platform for monitoring the data collected from IoT sensors mentioned above.

TABLE I. DATA COLLECTED FROM IOT LIBELIUM SENSOR DATA FOR SMARTAGRICULTURE AND SMARTCITY

id	*id_wasp*	*id_secret*	*frame_type*	*Frame number*	*sensor*	*Value*	*timestamp*	*sync*	*raw*	*parser_type*	*MeshliumID*
8049	SmartCityPro	3E381CE819623C3D	134	0	CO2	-1	6/8/2021 14:55	0	noraw	1	2.11E+13
8050	SmartCityPro	3E381CE819623C3D	134	0	PRES	96821.63	6/8/2021 14:55	0	noraw	1	2.11E+13
8051	SmartCityPro	3E381CE819623C3D	134	0	HUM	41.6	6/8/2021 14:55	0	noraw	1	2.11E+13
8052	SmartCityPro	3E381CE819623C3D	134	0	TC	27.23	6/8/2021 14:55	0	noraw	1	2.11E+13
8053	SmartCityPro	3E381CE819623C3D	134	0	US	270	6/8/2021 14:55	0	noraw	1	2.11E+13

Last is the prediction phase. In this phase different algorithms are compared and chosen the best one or more to be used for the prediction of the quality and quantity of grapes, the data mining techniques will be used in this phase. Algorithms that have been used to predict temperature are NeuralProphet [22], Random Forest Regression [23], SARIMA[24], and ANN by KERAS[25].

Samples (which from now on we are going to refer to as datasets) contain hourly temperature values from the 1st of January 2017 up to the 31st of December 2021. Using these values, we are going to try and predict temperatures for the whole of 2022 year. There are some periods during which sensors haven't been measuring temperature. Those values have been replaced with a specific value that doesn't leave holes in our dataset.

Forecasting temperature is done in a straightforward way based on assumption. There are many factors affecting air temperatures, such as pressure, humidity, wind speed, and direction, but seasonality allows us to predict daily temperature with a small margin of error. Since temperatures are seasonal (going from winter into summer, temperatures start to get higher, and vice versa from summer into winter), today's predicted temperature will be based on yesterday's temperature, and yesterday's temperature will be based on the day before yesterday, and so on.

Grabbing and formatting our datasets is going to be handled by Pandas, a Python library that can read and manipulate CSV Files. Plotly [26], another Python library, will be used to show plots and visualize our datasets and predicted values.

VI. CONCLUSION AND FUTURE WORKS

Smart Agriculture is a fairly new term and most farmers are not aware of what exactly lies behind it. The term smart agriculture refers to the use of technologies such as the Internet of Things, sensors, location systems, robots, and artificial intelligence on farms [1], [6]. The ultimate goal is to increase the quality and quantity of the harvest.

Some examples of technology used in smart agriculture are:

- Precise watering and proper nutrition of plants;
- Management and control of temperature in the greenhouse;
- Sensors for soil, water, light, humidity, and temperature management;
- Software platform;
- Location systems – GPS, satellites, etc.;
- Communication systems – based on mobile connection;
- Robots;
- Analytics and optimization platforms.

The connection between all these things is the Internet of Things, which is a mechanism for the connection between sensors and machines resulting in a complex system that manages the farm based on the received data. Thanks to this system, farmers can monitor the processes on their farms and make strategic decisions remotely - tablet, phone, or any other mobile device without being present in the fields.

This paper proposes a model for the prediction of the quality and quantity of grapes based on IoT sensor data from agriculture. This proposal will help farmers to take decisions and improve the process of grape processing and grape sterilization.

Because the world is facing climate change on one side and population growth on the other side [5], [6] is very important to use data mining techniques in agriculture to predict the quality and quantity of agriproducts. We use sensor data for the past five years as temperature, air humidity, air pressure, and land humidity which is a very important factors related to grape quality and based on them the proposed model will be able to recommend an adequate period for grape processing and grape sterilization. The quality and quantity for the past five years will be analyzed and based on them we do predictions for next year.

The proposed prediction model in this paper is focused on predicting grape quality and quantity based on environmental factors, and the recommendations that arise depend on the specific predictions made by the algorithm.

The next step of our research is a selection of the prediction models and algorithms used by other authors, and to compare them and select the best to compose the new model which is better than the existing one.

Also, a web-based platform for visualizing the data in real-time will be proposed in the future where the farmers will be able in real-time to monitor the weather parameters such as temperature, air humidity, air pressure, and land humidity.

REFERENCES

[1] Kasimati, A., Espejo-García, B., Darra, N., and Fountas, S. (2022). Predicting Grape Sugar Content under Quality Attributes Using Normalized Difference Vegetation Index Data and Automated Machine Learning. Sensors (Basel, Switzerland), 22(9), 3249. https://doi.org/10.3390/s22093249

[2] I. Global Industry Analysts, "prnewswire," Journalists, 23 03 2022. [Online]. Available: https://www.prnewswire.com/news-releases/global-smart-agriculture-market-to-reach-17-1-billion-by-2026--301507420.html. [Accessed 26 06 2022].

[3] Eitfood.eu, Sustainably feeding the world in 2050: Are efficiency and equity the answer? Accessed on: 25.06.2022, link: https://www.eitfood.eu/blog/sustainably-feeding-the-world-in-2050-are-efficiency-and-equity-the-answer,

[4] Jian Y., and Zhongyou L. Smart Agriculture and IoT Technology. Asian Agricultural Research. 2022;14(2):10-13. doi:10.19601/j.cnki.issn1943-9903.2022.02.002

[5] Alaa A.,Mai M., Nada M. M., Loa'a A. S.,Nada A., Yomna T. A. All., M. Saeed D., Mohamed A. F. and Hassan M. Smart IoT Monitoring System for Agriculture with Predictive Analysis. 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), 2019, pp. 1-4, doi: 10.1109/MOCAST.2019.8741794.

[6] Almalki FA, Soufiene BO, Alsamhi SH and Sakli H. A Low-Cost Platform for Environmental Smart Farming Monitoring System Based on IoT and UAVs. Sustainability (2071-1050). 2021;13(11):5908. doi:10.3390/su13115908

[7] Rathore, S. S., Kumar, S., and Soni, A. (2018). A prediction model for smart agriculture using IoT. International Journal of Engineering and Technology, 7(3.25), 107-110.

[8] Liu, Y., Kang, J., and Wang, G. (2019). Prediction model for precision agriculture based on IoT sensors. Journal of Ambient Intelligence and Humanized Computing, 10(8), 3139-3147.

[9] Tariq, M., Qayyum, A., and Baig, I. (2020). Crop yield prediction model for smart agriculture using IoT. In Proceedings of the International Conference on Artificial Intelligence and Data Engineering (pp. 91-98). Springer.

[10] Kumar, D., Biswas, P., and Singh, R. K. (2020). Crop disease detection and prevention model using IoT sensor data. Journal of Ambient Intelligence and Humanized Computing, 11(2), 557-568.

[11] Jung, S. H., Joo, G. J., and Choi, Y. J. (2020). A prediction model for grape quality using IoT sensor data. Computers and Electronics in Agriculture, 170, 105276.

[12] Droulia F. and Charalampopoulos I. (2021). Future Climate Change Impacts on European Viticulture: A Review on Recent Scientific Advances. Atmosphere. 12. 10.3390/atmos12040495.

[13] Ballesteros R., Intrigliolo D.S., Ortega J.F., Ramírez-Cuesta J.M., Buesa I. and Moreno M.A. Vineyard Yield Estimation by Combining Remote Sensing, Computer Vision and Artificial Neural Network Techniques. Precis. Agric. 2020, 21, 1242–1262.

[14] Alessia C., Franco M., Francesco P., Alberto C. and Francesco M. Analysis and impact of recent climate trends on grape composition in north-east Italy. BIO Web Conf. 13 04014 (2019). DOI: 10.1051/bioconf/20191304014

[15] Cogato A., Meggio F., Pirotti F., Cristante A. and Marinello F. (2019). Analysis and impact of recent climate trends on grape composition in north-east Italy. BIO Web of Conferences. 13. 10.1051/bioconf/20191304014.

[16] Kasimati, A., Espejo-García, B., Darra, N., and Fountas, S. (2022). Predicting Grape Sugar Content under Quality Attributes Using Normalized Difference Vegetation Index Data and Automated Machine Learning. Sensors (Basel, Switzerland), 22(9), 3249. https://doi.org/10.3390/s22093249

[17] Ravesa A. and Shabir A. Sof. Precision agriculture using IoT data analytics and machine learning. Journal of King Saud University Computer and Information Sciences. 2021.

[18] Rajeswari S., Kannan S. and Rajakumar K.. (2018). A smart agricultural model by integrating IoT, mobile, and cloud-based big data analytics. International Journal of Pure and Applied Mathematics. 118. 365-369.

[19] Revathy, R., Murali P. and Saminathan B. “HADOOP BIG DATA MINING: AN EFFECTIVE MAPREDUCE TOOL FOR CLASSIFYING SUGARCANE YIELD DATA.” (2020).

[20] Costa C., Graça A., Fontes N., Teixeira M., Gerós H. and Santos J. (2020). The Interplay between Atmospheric Conditions and Grape Berry Quality Parameters in Portugal. Applied Sciences. 10. 4943. 10.3390/app10144943.

[21] Ferehan N, Haqiq A. and Ahmad MW. Smart Farming System Based on Intelligent Internet of Things and Predictive Analytics. Journal of Food Quality. May 2022:1-8. doi:10.1155/2022/7484088

[22] Neural Prophet, Accessed on: 01.12.2022, Link:https://www.section.io/engineering-education/building-a-time-series-weather-forecasting-application-in-python/

[23] Random Forest Regression, Accessed on: 20.12.2022, Link:https://machinelearningmastery.com/random-forest-for-time-series-forecasting/

[24] SARIMA, Accessed on: 11.10.2022, link:https://github.com/imkhoa99/Time-Series-Analysis-and-Weather-Forecast-/blob/master/Time_Series_Project.ipynb

[25] KERAS, Accessed on: 21.11.2022, Link:https://github.com/MTxSouza/TemperaturePrediction-TimeSeries

[26] Python plotly, Accessed on: 21.11.2022, link https://plotly.com/python/

BIM Based Information Management in Renewable Energy Projects

Marin Ljuban*, Marin Curavić**, Luka Budin***, Ivan Duilo***, and Marko Delimar***
* University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana, Slovenia
** IT Sistemi, Split, Croatia
*** University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
marin.ljuban@gmail.com, marin.curavic@itsistemi.com, luka.budin@fer.hr, ivan.duilo@fer.hr, marko.delimar@fer.hr

***Abstract* – Proper implementation of Building Information Modelling (BIM) principles leads to better information management, and consecutively higher productivity in construction projects. BIM is an emerging methodology in the world of the built environment. It is a set of processes and standards, supported by digital technologies that aim to enhance information management in every phase of a project. From design and construction, through operations and maintenance, and finally the demolishing phase, BIM has shown advantages over traditional information management processes. Renewable energy is one of the areas which has significant potential, but a low adoption rate of BIM. The objectives of this paper are twofold. Firstly, to introduce the key concepts of BIM methodology to the reader and secondly, to examine them through the lens of renewable energy projects, more specifically, through the framework proposed in the FieldWork 4 RES platform.**

Keywords - BIM; renewable energy; information management; asset management; CDE

I. INTRODUCTION

Driven by policies and the market, as well as the negative climate change activities, renewable energy capacity is expected to increase more than twice as much in the next five years as it did in the past five years [1]. The sheer amount of projects that need to be developed on one side and the increasing costs of building materials and workforce on the other make it a necessity to optimize the construction process as much as possible. In that matter, optimized information management makes a key difference [2]. When dealing with optimized information management, as described in ISO 19650, two main fields are recognized: Building Information Modeling (BIM) and Common Data Environment (CDE).

The BIM approach takes its place as a response to the previously stated need for the optimization of the project's processes. As stated in [3], BIM represents a digital information management approach, using 3-D modelling with attached data that can be distributed between all project members, at all project stages from planning and design through maintenance and demolition. The main result of using the BIM approach is the more efficient use of resources, as well as more effective communication which enables greater flexibility between partners and improved long-term planning of the project itself. In short, Building Information Modeling provides a methodology that enables structuring the information so that technology can process it.

This paper is based on research developed through the project "Integrated solution for asset management and support of investment processes of design, planning, and implementation of renewable energy sources construction" supported by the European Union within the European Regional Development Fund (KK.01.2.1.02.0146).

Design stage researchers are exploring the potentials of BIM, generative design, and machine learning for early-stage optimization, both for structural design [4] as well as occupant comfort [5]. Additionally, notable efforts are being done in the field of digital building permitting [6]. In the construction phase, a combination of BIM data and lean principles can ensure better project management [7]. The operations and maintenance phase has also seen an uptake in research in the last years [8], with research being done in the integration of BIM, BMS, and IoT for Facility Management [9], among many others. This research is being done both for new buildings, as well as existing buildings, with the special attention being on maintaining buildings of special cultural value, called Historic or Heritage BIM (HBIM) [10]. From the computer science point of view, the new trend in research is to explore the move from traditional, object-oriented data models towards more interoperable semantic web-based data models [11]. In addition to academic research, a strong link with the industry in the EU is fueled by public funding such as the Horizon Europe program, with many projects regarding digital building permitting, construction safety, energy consumption optimization, etc. currently being developed [12].

CDE, as stated in [13], presents a combination of technical solutions and process workflows represented by software (or another form of a tool) that allows all project participants to share important information about the project (including BIM models). The CDE workflow defines the processes used for collecting, managing, and disseminating structured and unstructured information, while the CDE solution is the technology that enables these processes. The introduction of CDE is a significant part of the realization of the ISO 19650 standard. The ISO 19650 standard, as shown in Figure 1, presents a subset of the complete ISO 55000 standard for asset management.

Furthermore, the surveys, carried out in [14], show that most of the surveyed companies are aware of the benefits of asset management, but lack the time, resources, and knowledge to properly implement the processes and tools required. Qualitative asset data is vital to risk management, but most companies still cope with it, often opting for the traditional paper-based workflows in more than 50% of cases. Also, 50% of companies don't use an IT system to handle their assets. It is stated that digitization of information management according to ISO 19650 could bring benefits to the renewable energy industry, both itself, as well as an intermediate step towards overall standardization of asset management per ISO 55000.

FieldWork 4 RES (FW4RES) is a platform specialized in supporting Renewable Energy Sources (RES) projects in all project phases, from planning and preparation, throughout the construction of the RES projects, as well as the maintenance of the RES assets, including a full range of functionalities to satisfy the requirements of all end users – investors, planners, contractors, and managers. The FW4RES unites all the previously mentioned barriers and uses CDE and BIM for optimized information management. The FW4RES also includes a various number of functionalities for RES projects, which are listed and described in [15].

The objective of this paper is to present the implementation of key principles of information management using BIM methodology, as introduced in ISO 19650, on RES projects. This approach could utilize the data produced in the design stage to optimize the construction process, making it quicker and more efficient. At the end of the construction process, data would be verified and passed over to the asset managers, who could use it as input in their information management systems, extending it with new information about project maintenance as it is created.

This paper is presented in the following order; firstly, an overview of key concepts, terminology, and standardization in information management in construction is presented to introduce the theoretical framework of the application in RES projects; then, the BIM terminology is presented. Both areas are followed by application in the specific FW4RES use case.

II. STANDARDIZATION OF INFORMATION MANAGEMENT

The productivity of the construction industry grew by only 1% over the last two decades, in comparison with an average growth of 2.8%. The key way of actually improving productivity is adopting a collaborative approach across the industry [16]. To support collaboration efforts concerning new technology development, a standard was developed in the United Kingdom (BS 1192), which formed the base of the later published international standard ISO 19650 – „Organization and digitization of information about buildings and civil engineering works, including building information modelling“. The standard defines collaborative production and management of information across the lifecycle of a built asset. It is aimed at all people and organizations involved in the whole life cycle of an asset. In the context of asset management, it is also important that ISO 19650 can be interpreted as a subset of the international standards that define organizational and asset management.

A. *Concepts and principles of ISO 19650*

The key objects for decision-making throughout the life cycle of an asset are information models. Depending on the phase, they are named project information models (PIM), or asset information models (AIM). These information models can include structured information containers, such as video clips or unstructured information containers such as BIM models, or databases [18].

These information models are of key importance to all the participants involved in the project. To enable the



Figure 1. Project and asset information management lifecycle [17]

successful delivery of these information models, ISO 19650 introduces many different terms regarding actor roles (e.g. appointing party, appointed party), teams (e.g. delivery team, task team), and document types (e.g. organizational information requirements, exchange information requirements), as well as the workflows that describe interactions between these actors, teams, and documents. All the aforementioned is described in detail in the standard itself and, for the sake of length, will not be discussed in this paper. However, one term, named common data environment (CDE) is of particular importance and will be examined more thoroughly.

B. *Common Data Environment (CDE) workflow*

Common Data Environment is an “agreed source of information for any given project or asset for collection, managing, and dissemination of each information container through a managed process”, where the information container is a “named persistent set of information retrievable from a file, system or application hierarchy” [17]. The CDE is a centralized repository for the management of data in the entire life cycle of a built asset. The ISO 19650 standard makes a clear distinction between the CDE workflow and a CDE solution, where a CDE workflow defines the processes to be used, while a certain CDE solution provides the technology that makes the implementation of these processes easier [17]. Several key concepts that make up a standardized CDE workflow are explained in the rest of this chapter.

Classification system - classification is a method of organizing entities logically, so they can be differentiated one from another. Many classification systems were derived from the core ISO 12006 standard, and the one commonly used in Europe is called Uniclass 2015. Uniclass is a faceted system that allows the classification of all building elements according to several tables (e.g. Products, Entities, Systems). Besides building elements, Uniclass has several additional tables that allow the

classification of other important entities in the project, such as actors, and the activities they need to carry out.

Naming standard - information is, as mentioned, kept in the CDE in form of a structured or unstructured information container. To ease the navigation of the CDE user between all the different containers, a naming convention should be adopted following conventions set in the ISO 19650 national annex.

The information states – each of these containers can exist in various states. There are four states, shown in Figure 2. with their important characteristics. The presence of information states ensures that proper information is shared with exactly those actors that should receive the information, thus reducing the risk of information overload.

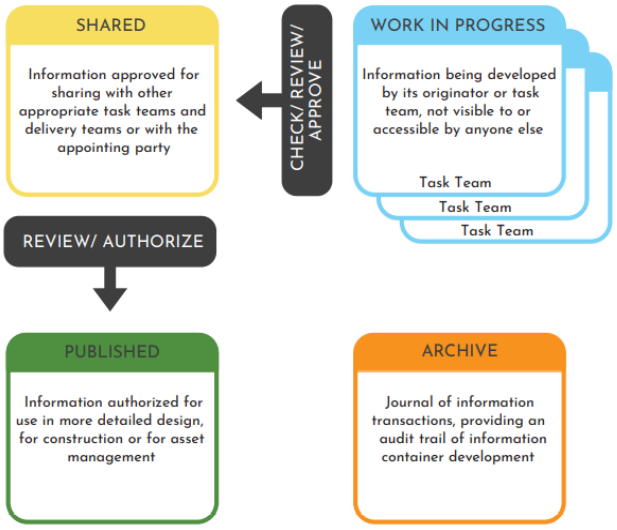


Figure 2. Common Data Environment information states [17]

Roles – similarly, all the project actors should be organized according to their roles and project teams (e.g., surveying company, civil engineering company, electrical engineering company) in the project to ensure that only the information needed is communicated to them.

- Classification code

Approval workflows – to put all the mentioned terms in a useful system, approval workflows are developed. These approval workflows take an information container as input, process it by sending it to different actors higher up the decision chain depending on the workflow, and after gaining approval, automatically change the information state of the specific container, making it visible to other actors.

C. Common Data Environment (CDE) solution

It is important to differentiate the workflow from a technological solution. While not practical, even a non–digital solution such as postal service can be classified as a CDE solution while CDE workflows are respected [15]. On the other hand, a cloud file repository (e.g. Dropbox, Google Drive) is not a CDE by itself because it does not have some of the features of the CDE workflow. The most common solution is the integration of CDE workflows in electronic document management systems (EDMS), which is also the situation in the FW4RES platform. Due to the complexity of the construction industry, CDEs are still primarily file-based opposed to most business process management solutions. It is useful for FW4RES to utilize both the database and file-based management of information.

FW4RES supports both horizontal and vertical hierarchical organizational arrangement at multiple levels, which adapts to any actual organizational structure of the business environment it serves as a CDE solution. Organizational units and corresponding project teams can be internal and external. Internal organizational units are departments or project teams within the business entity of the owner of the application that is specialized in performing a particular type of work. External organizational units or project teams are all organizational units that do not appertain to the business entity of the owner of the CDE solution. External organizational units are typically contractors or subcontractors as executors of the work.

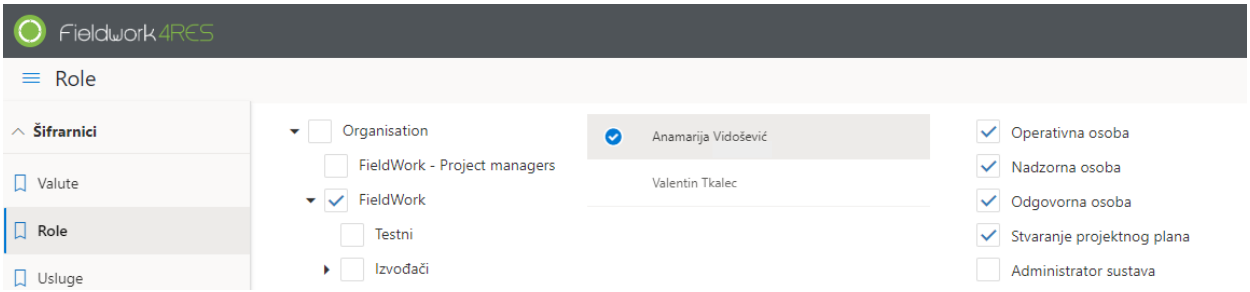


Figure 3. Screenshot from FW4RES – role management within an organizational unit

Metadata – metadata is formally defined as a set of data that describes information about other data. Container metadata is the information that describes this particular container more thoroughly. The ISO 19650 recommends that the following metadata is included in each information container:

- A revision code
- Status code

Also, FW4RES users can be assigned one or more roles within an organizational unit to work with documents, as shown in Figure 3. Through roles, the user is assigned a process role (action) on documents. Usually, three levels (operational, supervisor, and responsible) are considered quite sufficient and an optimum between administrative effort and the potential benefit. If necessary, however, the anticipated Uniclass system defines more than two

hundred roles that could be used to describe actors in finer detail. For description purposes, Uniclass roles can be used as additional metadata for system users with the assigned solution role. The Uniclass roles table is also extendable and customizable according to specific business demands so it allows the additional definition of an arbitrary number of roles. The naming convention recommended for information containers is the one from the ISO 19650 UK Annex but can be customized as needed.

Considering project execution, the operational person as a first level should have the possibility of entering bills of quantities, distributing works, accepting the project plan, changing contractors, entering calculations, creating receipts, and manually opening production orders. The supervisory person as a second level of document approval supervises the operational person and accepts or rejects the report on the works carried out through the return for supplementation (correction). The responsible person as a third level of approval supervises the operational person and accepts or rejects the report on the works carried out through the return for supplementation (correction). Performs financial control and approves settlements. All these changes can go through predefined approval workflows and result in changing the database or moving the information container to another state in the CDE structure (e.g., from shared to published folder).

FW4RES's core function as a CDE solution is the planning and execution of construction or operation and maintenance projects. Hence, the project plan is the first document in the development process and as such is a prerequisite for the creation of all other documents because all types of later-created documents in the process are linked to the approved project plan and derived subprojects. Through the project plan (creation of the new project is shown in Figure 4.), the primary division of the work in the planning phase is done. This work-in-progress information state has multiple substates (new project plan, approved supervisor, approved responsible, accepted). Approved and accepted project plan is translated into shared or in realization state creating a detailed design plan with detailed subprojects and situations whose structure is identical to the structure of the project plan – subprojects. With the creation of a detailed design plan with detailed sub-projects and situations, the planning phase is completed, and the implementation phase begins with the project being published. After it stops being useful, each information container changes its state to the archive information state to keep its content and audit trail available.

III. Building Information Modelling

A building information model should be the main container in the information management process aligned to ISO 19650, and all the other information containers should refer to it [18]. A building information model is the consequence of the realization of BIM methodology in a project. Due to its key role in information management, it is of key significance to understand what BIM represents.

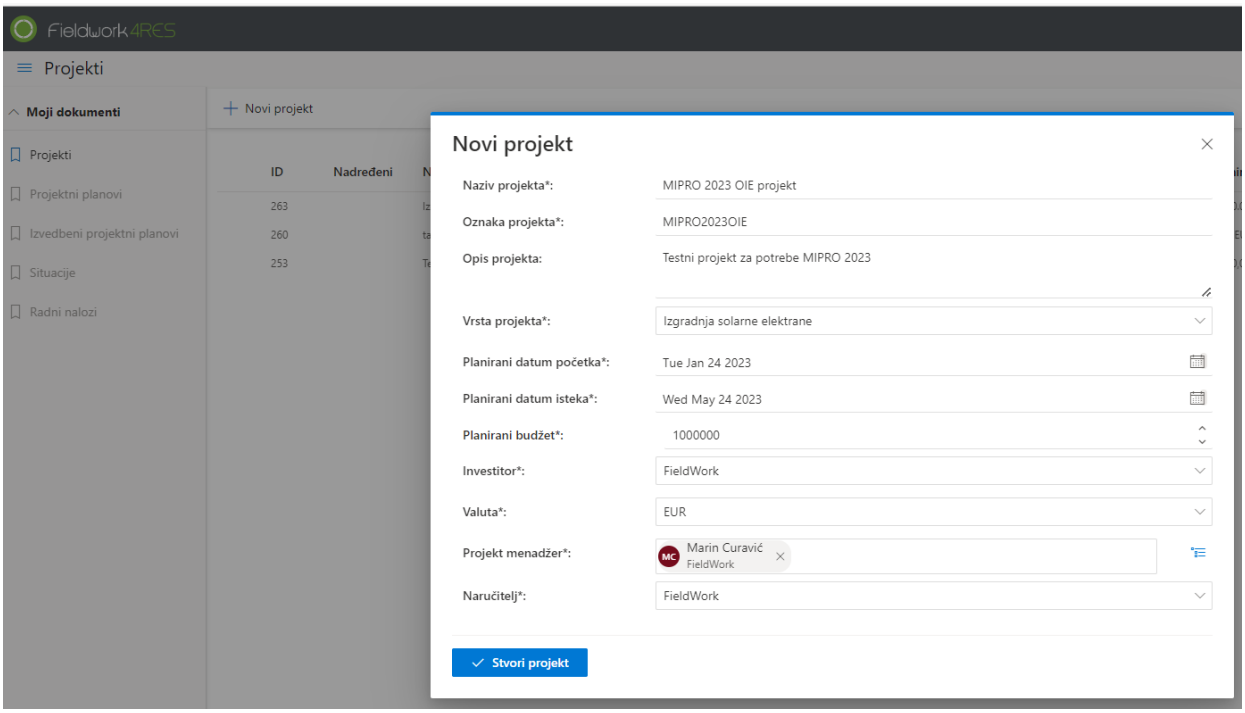


Figure 4. Screenshot from FW4RES – creation of the new project using input data

BIM is most easily explained as a set of standardized digital construction processes. These processes are supported by technology and result in an accurate virtual model of the project being designed. The model is built by placing instances of BIM objects in the model, and each of these objects contains alphanumeric metadata that describes it. By building the project digitally before it is built, BIM brings several important benefits to the involved actors. This model is used as the key part of information exchange in all the project stages. Benefits of BIM consist of, however, are not limited to [19]:

- Design:
  - earlier and improved collaboration of different disciplines,
  - earlier and more accurate visualizations,

- increased consistency of data and design drawings,
- improved cost estimation

- Construction:
  - synchronization of design and construction planning,
  - better implementation of lean construction techniques,
  - synchronization of procurement with design and construction
- Operations and maintenance:
  - improved commissioning and handover,
  - improved visualization of operations,
  - easier communication with field workers.

Furthermore, it is essential to highlight that although a 3D model (like the one shown in Figure 5.a) is the most prominent deliverable of the process, 3D models that do not contain metadata cannot be categorized as BIM models. BIM models are data-rich and parametric, and very often alphanumeric data contained within elements proves to be more useful than the geometry itself, especially in the O&M stage. In the example, each element of the geometrical model (Figure 5.a) can be accessed individually and contains information about the manufacturer, installation date, and warranty, among others (Figure 5.b). BIM can also be used for better information management of already existing projects using the scan-to-BIM method, as shown in figure 6.

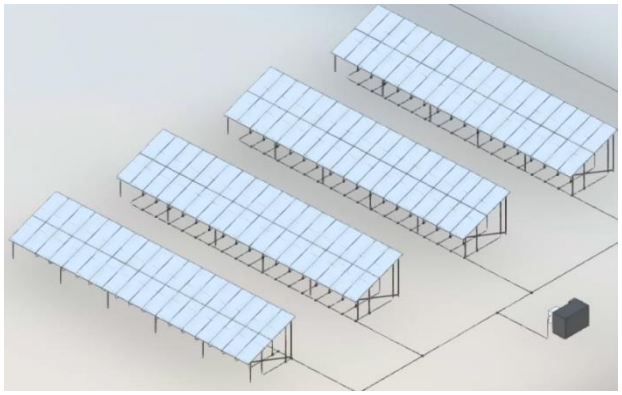


Figure 5.a Solar power plant IFC model – SE VIS Solis

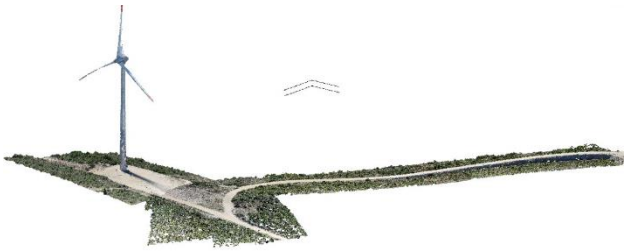


Figure 6. Photogrammetry model used as a starting point for scan-to-BIM method for creating wind power plant model – VE Danilo

successful efforts were formalized in the form of ISO standards. Others are not formalized in the ISO standard form, but their wide use makes them the *de facto* standard. The ones that are key to understanding the workflow are explained below.

IFC – to standardize information exchange on the software implementation level, a common language needed to be created. This language is called Industry Foundation Classes, often abbreviated as IFC. The IFC is an open, non-proprietary, and standardized digital representation of built assets. It is an EXPRESS-based, object-oriented data model that describes the identity, characteristics, and relationships of objects. Objects can be built objects, such as columns, slabs, or energy conversion devices, but also processes (installation, maintenance), people (owners), and abstract concepts (performance, costs). A such data model is encoded in different file formats (XML, JSON, STEP), and used to exchange information between actors [20]. As the core interoperability concept in construction, IFC is currently undergoing a transformation that will enable it to move from a file-based sharing to being used in a transactional environment [21], making it possible to integrate it completely with the business process management solution in FW4RES.

BCF – while IFC enables interoperability between software tools, it does not solve the issue of communication between actors that work in different software. For this purpose, the BIM Collaboration Format (BCF) was created. It can be described as an additional layer on top of the IFC. After selecting an IFC object, the user can create issues and assign them metadata such as

TagNumber	IfcGUID	Accessibili	AssetT	ClassificationCategory	Colour	Description	DurationUnit	Ex	Manufacturer	ModelNumber
01_01_01	3Di2VlmD99f8MjR0Bn8ral	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_02	3Di2VlmD99f8MjR0Bn8rbY	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_03	3Di2VlmD99f8MjR0Bn8rbo	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_04	3Di2VlmD99f8MjR0Bn8rb2	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_05	3Di2VlmD99f8MjR0Bn8rbl	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_06	3Di2VlmD99f8MjR0Bn8rwY	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_07	3Di2VlmD99f8MjR0Bn8rwo	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450
01_01_08	3Di2VlmD99f8MjR0Bn8rw2	NoIssues	Fixed	Pr_60_70_65_63	Aluminium	Solar Panel	Years	25	Longi	450

Figure 5.b Part of the non – graphical data for the solar power plant IFC model

A. Terminology

Since BIM aims to manage information of a project in all of its stages, as well as the information of all the disciplines (e.g., architecture, mechanical, and electrical engineering), it has to solve many interoperability issues to be able to communicate effectively. To do so, many standardization efforts have been undertaken, most of them led by an international non-profit organization buildingSMART. After spending some time in the development, the most

description, responsibility, and deadline. These issues can then be delivered to other actors via the cloud (through BCF-API) or encoded in a file format (BCF-XML).

B. BIM in FieldWork 4 RES (FW4RES)

Generally, the utilization of BIM models in tools downstream from the authoring tools depends on the BIM uses that model is designed for. Since the model is always a simplified representation of reality, it is important to

establish these model uses before a specific project, according to one of the established models, such as the Penn State or BIMe Initiative model framework.

FW4RES aims to utilize the BIM model throughout the complete lifecycle of the built asset, from the design and planning phase until the decommission, but like any other tool, is not powerful enough to have all the functionalities needed in a project. The most prominent uses it is aimed for can be grouped into two model use series per BIMe Initiative [22]. The first series is capturing and representing, and the second one is operating and maintaining.

Depending on specific project workflows, the IFC model can be designed in an authoring tool (e.g. Revit), and later enriched with data regarding quantity takeoff, cost estimation, or project planning in another tool (e.g. Bexel Manager). Such a model can be imported to the FW4RES platform at any stage. The platform has an integrated IFC parser that can obtain all the information about the model such as a spatial tree, non-graphical properties, and metadata, among others, and display it to the user. It can also render the IFC model, showing it in 3D view, where all the elements and their properties can be accessed individually. There are also possibilities to extend the platform with specific tools, that would parse the data from the model and then show it in specific views, such as cost breakdowns and Gantt charts, among others. This is an important benefit that is aimed at supervising and responsible roles in the project. They can access the latest model at any stage of the project, and without knowledge of specific software and access to often expensive licenses, they can examine and report possible issues back to the operating personnel.

The second important aspect is operation and maintenance. After the project close-out and handover of the verified as-built model, the need for specific tools does not exist anymore, and the FW4RES platform can be utilized as the only source of information for all the actors in the project. The supervising and responsible actors can access the web platform and input the data about the planned maintenance through the BCF format. Depending on the specifics of each project and other tools used (e.g., ERP), integrations via BCF-API can be made to ensure relevant data streams. After doing the planned maintenance activities, the operating actors can report the successful completion of these activities, as well as possible issues, again via BCF.

IV. Conclusion

Most stakeholders involved in renewable energy projects are conscious of the significance of proper asset management, but most still do not do it in a standardized way. Information management is an important subset of asset management, and the digitalization of information management is an important intermediate step towards efficient asset management throughout the life cycle of a renewable energy project. This paper investigates the possibilities of implementation of standardized information management principles from construction projects, as stated in ISO 19650.

Starting as an attempt to organize the novel methods of project delivery revolving around BIM methodology, ISO 19650 introduced several important concepts in information management, the most notable of which is the notion of a Common Data Environment. The CDE is a central repository for the collection, management, and dissemination of information in the project, where ISO 19650 separates the CDE workflow from a CDE solution. The paper presents universal concepts of CDE workflow and also presents how these should be implemented in the CDE solution of the FW4RES platform. An additional facilitator of information management is the introduction of the most important principles of BIM and their implementation in the FW4RES platform.

The future work will be focused on the complete implementation of the concepts presented in a functional FW4RES solution. Besides the implementation, future work will also keep track of the latest developments in the interoperability domain (IFC 5, openCDE API) to make the FW4RES interoperable with further solutions, both more granular (e.g., contractor CDEs), as well as more consolidated (e.g., municipality digital twin platform).

References

[1] International Energy Agency, “Renewables2022 - Analysis and forecast to 2027,” 2022. [Online]. Available: link

[2] P. Adekunle, C. Aigbavboa, O. Akinradewo, A. Oke, and D. Aghimien, “Construction Information Management: Benefits to the Construction Industry,” *Sustainability (Switzerland)*, vol. 14, no. 18, Sep. 2022

[3] “BIM and ISO 19650 from a project management perspective.” [Online]. Available: link

[4] L. Gradišar, R. Klinc, Ž. Turk, and M. Dolenc, “Generative Design Methodology and Framework Exploiting Designer-Algorithm Synergies,” Buildings, vol. 12, no. 12, Dec. 2022

[5] H. H. Hosamo, M. S. Tingstveit, H. K. Nielsen, P. R. Svennevig, and K. Svidt, “Multiobjective optimization of building energy consumption and thermal comfort based on integrated BIM framework with machine learning-NSGA II,” Energy Build, vol. 277, Dec. 2022

[6] F. Noardo et al., “Unveiling the actual progress of Digital Building Permit: Getting awareness through a critical state of the art review,” Building and Environment, vol. 213. Elsevier Ltd, Apr. 01, 2022.

[7] K. el Mounla, D. Beladjine, K. Beddiar, and B. Mazari, “Lean-BIM Approach for Improving the Performance of a Construction Project in the Design Phase,” Buildings, vol. 13, no. 3, p. 654, Feb. 2023

[8] O. Adetayo, D. Onatayo, “A Scientometric Review of BIM in Facility Management Research,” 2023, [Online]. Available: link

[9] L. Chamari, E. Petrova, and P. Pauwels, “A web-based approach to BMS, BIM and IoT integration: a case study Holistic BIM-Based Sustainable Building Design and Performance Assessment View project Flemish Cities in Transition-A Framework for a Web-based Built Heritage Platform Using BIM, GIS and Linked Data View project A web-based approach to BMS, BIM and IoT integration: a case study”

[10] M. Bonduel, “A Framework for a Linked Data-based Heritage BIM”, 2021, [Online]. Available: link

[11] K. Jeon and G. Lee, “The Status Quo of Graph Databases in Construction Research”, 2022, [Online]. Available: link

[12] SmartBuilt EU, “Smart Buildings EU-funded Innovations – January 2023” 2023.

[13] J. Ford and G. Try, “Facilitating the common data environment (workflow and technical solutions).” [Online]. Available: link

[14] Deloitte, “Asset Management: A Risk-Based ApproachEnergy & Resources Benchmark Survey”. [Online]. Available: link

[15] M. Curavić, Z. Mavretić, I. Duilo, L. Budin and M. Delimar, "Blockchain Application in Digital Platform FieldWork 4 RES used for Planning and Realization of Renewable Energy Sources Projects," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 65-70

[16] McKinsey Global Institute, “Reinventing Construction: A Route to Higher Productivity”. [Online]. Available: link
[17] International Organization for Standardization, *ISO 19650-1:2018(en) Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles*. [Online]. Available: link
[18] Sphere, “How to move digital twin environments to the AECOO sector.” [Online]. Available: link
[19] C. Eastman, P. Teicholz, R. Sacks, and K. Liston, “BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors.” Wiley, 2018.
[20] T. Pazlar and Ž. Turk, “Interoperability in practice: Geometric data exchange using the IFC standard,” *Electronic Journal of Information Technology in Construction*, vol. 13, pp. 362–380, Jan. 2008.
[21] L. van Berlo, T. Krijnen, H. Tauscher, A. van Kranenburg, P. Paasiala, and solibricom Solibri, “Future of the Industry Foundation Classes: towards IFC 5.” [Online]. Available: link
[22] B. Succar, N. Saleeb, W. Sher, “Model uses: foundations for a modular requirements clarification language” Australasian Universities Building Education (AUBEA2016), Cairns, Australia, July 6-8, 2016

Case Management Model Design for Obtaining Documentation Using FieldWork 4 RES Application in Renewable Energy Sources Projects

Zrinka Tolvajčić*, Tanja Ugrina*, Marin Curavić*, Ivan Duilo**, Luka Budin**, and Marko Delimar**
* IT Sistemi, Split, Croatia
** University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
zrinka.tolvajcic@itsistemi.com, tanja.ugrina@itsistemi.com, marin.curavic@itsistemi.com, ivan.duilo@fer.hr, luka.budin@fer.hr, marko.delimar@fer.hr

***Abstract* – FieldWork 4 RES is an application in the final stage of production as part of the Development of an integrated asset management solution and support for investment processes in the design, planning, and implementation of renewable energy construction projects. The purpose of the FieldWork 4 RES system is to provide application support for the activities of planning and realization of projects ranging from the construction of infrastructure projects to property maintenance projects. During this process, part of the activity relates to the collection and recording of project documentation, depending on the type of project and legal regulations. This paper deals with the module design for obtaining project documentation. It is based on the communication between FieldWork 4 RES, Camunda, and OfficePoint. The case management model is presented and all its parts are defined.**

Keywords – renewable energy; CMMN; fieldwork4res; camunda;

I. INTRODUCTION

FieldWork 4 RES is an application that will be used for the planning, preparation, realization, and maintenance of renewable energy sources assets. It will also apply to other business processes which deal with investment and project planning, managing contractors and contracts, financial management, ensuring profitability, procurement of needed documentation, material and inventory supervision, Building Information Management (BIM) application, digitization, and automation of business processes in general, etc.

The European Commission states that a large amount of energy from renewable sources has been successfully implemented in the EU Member States, however, there is still a significant number of obstacles that slow down further development of renewable energy sources projects, which could affect the EU's 2030 and 2050 climate target. As stated in [1], almost 25% of all existing barriers are related to administrative and grid connection processes, where these obstacles appear in all EU markets. Observing the mentioned barriers, renewable energy installations are affected by a share of 83%, the heat sector by 13%, and the transport sector by only 4%. Renewable energy installations are affected the most by administrative barriers, especially wind power (44%) and solar power (23%), as shown in Figure 1. Thus, document management systems (DMS) come to their full potential.

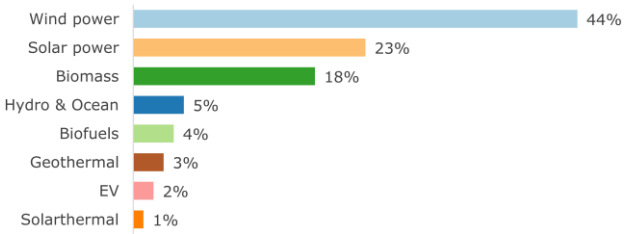


Figure 1. The impact of administrative barriers on RES project development [1]

Issues in regulatory processes and document management systems are very common in today's business, energy, traffic, and other sectors. This paper presents the case management model used together with applications FieldWork 4 RES and OfficePoint (both products of IT Sistemi). It is a unified solution for optimized documentation obtainment mainly used for renewable energy construction projects, but it is also applicable to many others asset management solutions.

The paper deals with the part of the application that oversees obtaining documentation for RES projects. Through the system, the user will be informed about the current phase of the project, documents that need to be obtained, submitted requests whose processing is in progress, the expiration of already issued permits, etc. The focus is on the necessary and obtained documentation, and not on the order in which the documents are obtained within a particular phase. The system is flexible so it can support different types of projects and at the same time enables the user to easily enter the necessary changes even in the event of a change in legislation.

This paper is based on research developed through the project "Integrated solution for asset management and support of investment processes of design, planning, and implementation of renewable energy sources construction" supported by the European Union within the European Regional Development Fund (KK.01.2.1.02.0146).

II. CASE MANAGEMENT AND DOCUMENT MANAGEMENT SYSTEM

A. Case Management Model and Notation (CMMN)

Case Management Model and Notation (CMMN) is a graphical approach used in work methods based on solving cases that require different activities that can be achieved in an unpredictable arrangement in response to changing situations. Using an event-driven method and the case file concept, CMMN expands the limits of what can be developed with Business Process Modeling and Notation (BPMN), including less designed work efforts and those determined by expert workers. BPMN represents a visual "flowchart-like" methodology used to map out the various steps involved in a business process from start to finish. As a critical component of Business Process Management, it provides a visual representation of the entire sequence of business activities and information flows required to execute the process. The primary objective of BPMN is to devise ways to enhance efficiency, accommodate changes in circumstances, and gain a competitive edge by creating models that identify process improvements. Applying a combination of BPMN and CMMN allows clients to deal with a wider selection of work techniques [2][3].

The main goal of CMMN is to define a general meta-model and representation for modeling and graphical expression of the case. A case primary includes tasks performed in relation to the subject in a certain situation to accomplish the preferred result. The situation usually includes data that informs and drives actions recorded in the case [3].

CMMN deals with defining tasks whether they are applicable or what are the subsequent assignments necessary, considering the condition of the case. Decisions and flow can be controlled by events or new details that constantly occur through the duration of the case, for instance the delivery of new documents, accomplishment of certain barrels, or achievement of certain milestones. The individual tasks that are planned and carried out in the context of a case may themselves be predefined procedural processes, but the total case can't be coordinated by a predefined sequence of tasks. Model and notation are used to express case models in the usual notation for a certain type of case, and the resulting model can subsequently be instantiated for handling a specific case instance [3].

B. Camunda

Camunda is an open-source and free Java-based framework used for various workflow, case management modeling, and decision-making automation systems. It allows the user to create BPMN processes and decision tables. It also allows the execution of process models as so-called process instances. During the execution of the model, a rule engine can automatically calculate decisions.

A CMMN consists of the following basic components:

- Case plan
- Roles
- Case plan items
- Tasks
- Stages
- Case file
- Case file items
- Criteria
- Event listeners [4].

C. Document Management System

A document management system refers to systems that manage certain processes with documents to organize them. A basic part of digital document management is the conversion of paper documents into electronic files on a computer. After digitization, every document is available on the computer, through the computer network, document management system, and to all other users. Many organizations around the world use document management systems on a daily basis instead of paper archives. Documents entering the DMS are scanned (as a rule, documents created outside the parent organization) or uploaded as electronic documents (documents created and available in electronic form). The system saves these files where the user indexes the document. Indexing is describing a document with terms that will later make these documents searchable. To access documents, a hierarchical structure of document organization or a search engine built into the system is used. Which documents can be read and which actions can be performed on them depends on the rights assigned to the user or group of users by the DMS administrator [5]–[7].

DMS consists of basic 5 components [8]–[10]:

- Entry: Scanning documents in paper form.
- Storage: the storage system contains documents in electronic form and provides additional security mechanisms to protect documents from loss.
- Indexing: provides a good description of documents, quick and easy future access, and search.
- Retrieval: the system keeps track of the arrangement of files on the media and enables quick retrieval using an index.
- Access: documents are available to those who need them, and access is limited to authorized users.

D. Communication Architecture

The basic concept used for obtaining documentation is based on the communication protocols between frontend application, case management, and content storage.

The main components, shown in Figure 2, are:

- FieldWork 4 RES (referred to as FW4RES in the following text) – frontend application
- Case management module – Camunda (CMMN + BPMN)
- OfficePoint (referred to as OP in the following text) – content storage, DMS

The communication protocols are as follows:

- FieldWork – Case management (user actions): in the FW user interface, the user is provided with a series of actions necessary to be able to fully manage the project and its parts. The actions themselves are implemented in the case management (CM) module.
- FieldWork – OfficePoint (create/update content): the user creates or edits documents through the FieldWork interface, and the content is stored in OfficePoint.
- FieldWork – OfficePoint (search data): FieldWork retrieves and searches data from OP to display to the user. For example, when it is necessary to display all phases of a project or all documents within a particular phase.
- Case management – FieldWork (get codebook): through the FW representational state transfer (REST) application, various codebooks are retrieved that are used for calculations and validations carried out in the CM module. (Codebook represents a list of mandatory documents that are created in various phases of the project).
- Case management – FieldWork (update project state): if necessary, the CM updates the project status or some project metadata in FW using the FW REST API.
- Case management – OfficePoint (create/update content): the CM module communicates with the OfficePoint application to create the necessary items. Based on the actions implemented in CM, the documentation stored in OP can be managed.
- OfficePoint – Case management (content update and notifications): OfficePoint triggers notifications to the CM module in moments when content is updated. CM internally decides which events to react to and which actions are necessary at those moments.

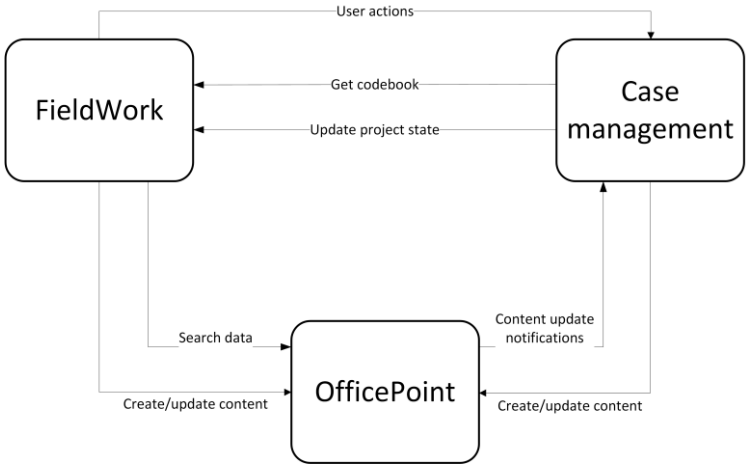


Figure 2. Communication protocols concept for data retrieval

The FieldWork 4 RES interface is shown in Figure 3.

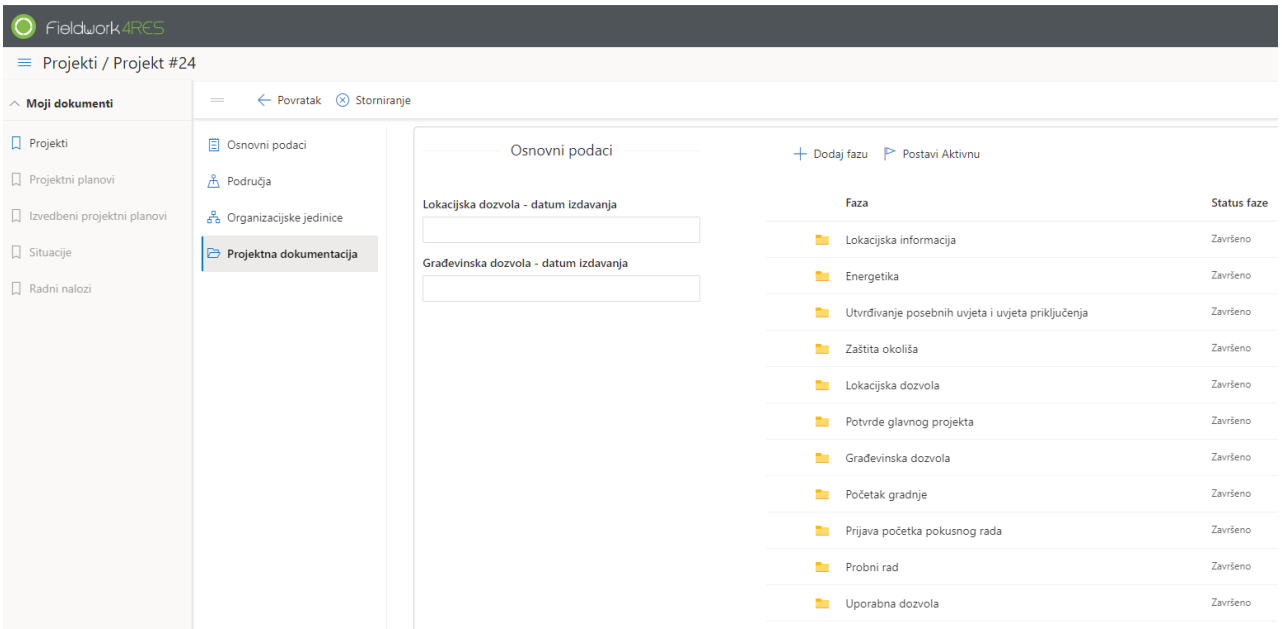


Figure 3. Screenshot from the FieldWork 4 RES interface

III. PROJECT DOCUMENTATION

Obtaining the project documentation is based on communication between FW4RES, OfficePoint, and Camunda. The necessary requirements for instantiating the Camunda case plan are as follows:

- List of Mandatory Documents – the codebook contains a list of all mandatory documents that can be created in various phases of the projects. Certain criteria are used to define under which conditions which document is mandatory. For example, the document is mandatory for the phase of obtaining a building permit, or the document is mandatory if it is a 100 MW (or higher) power plant, etc. It is necessary to enable simple and quick changes in the rules to meet the needs of the client or changes in legislation, without the need to create and deploy a new version of the system or individual components.
- Creating a new CMMN model in Camunda – a new project created in Camunda represents a graphical notation of a CMMN model or a case plan. It consists of:
  - Tasks (human tasks, process tasks, case tasks, discretionary tasks)
  - Event listeners (timer event listener, user even listener)
  - Milestone(s)
  - Stage and Discretionary Stage
  - Criteria
  - Entry criteria
  - Exit criteria
  - Case file
  - Planning table

An example of a CMMN model, made in Camunda, is shown in Figure 4.

After all the necessary input requirements are met, the Camunda case plan instance can be initiated. The structure of the case plan is as follows:

- Starting the Camunda case instance – when a case management instance is created, the "Project creation" action is started, which creates all the necessary files in OP, where all the documents are later created by the users. The mandatory project phases and documents are calculated using complex logic implemented in various codebooks.
- Calculation of the project phase – depending on the type of project (e.g., construction of a wind farm, construction of a residential building, etc.), required metadata is stored in Camunda variables

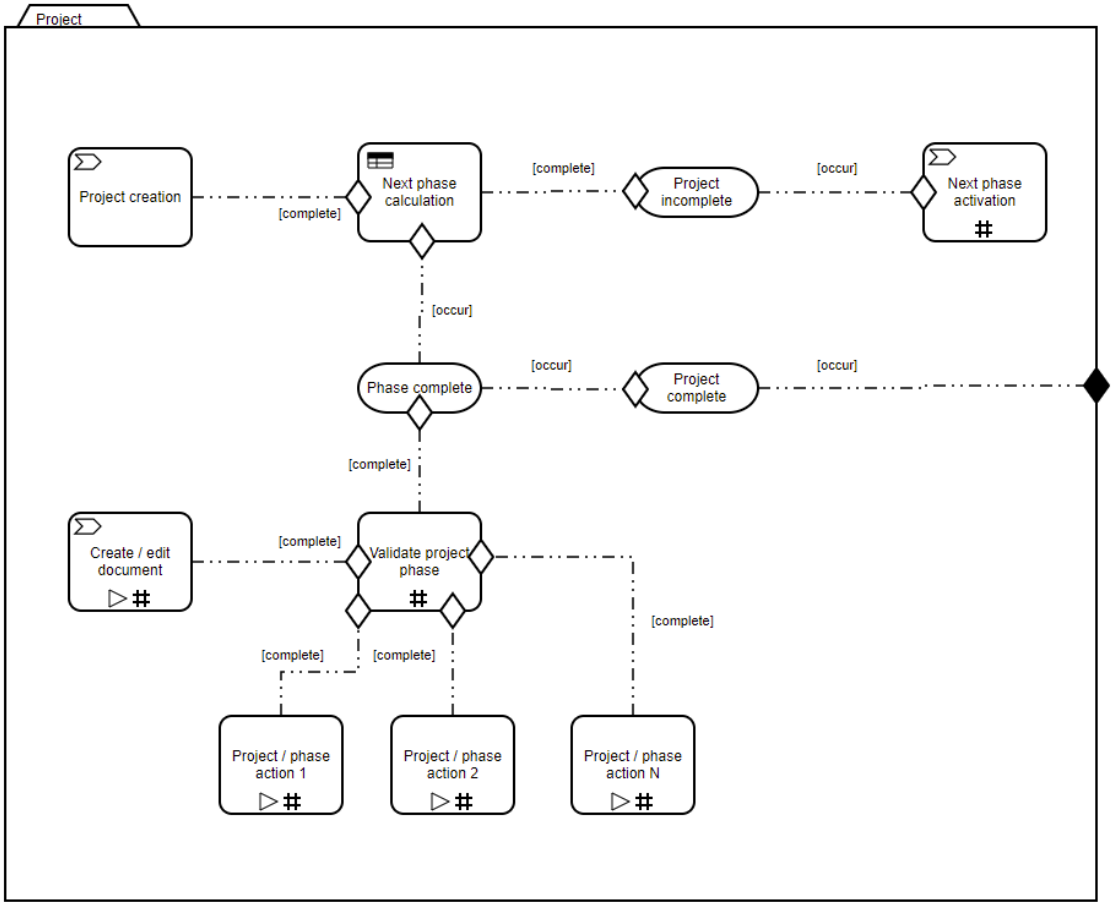


Figure 4. Screenshot from Camunda representing the CMMN model

to enable decisions to be made as the process progresses. An integral part of the project is a Decision Model and Notation (DMN) calculation table, which decides on the phase sequence depending on all current inputs.

- Submitting requests and adding documents – each created document corresponds to a specific document type predefined in OP. Based on the document type and the internal business logic, various rules are applied, and the required processes are started. Some of the processes are fully automated, but others may require additional user actions to be completed.
- Project completion and reactivation – the project may be considered complete when all phases have been completed or when a major obstacle has occurred that has prevented full completion. One of the impediments may be the inability to obtain some of the permits, or perhaps a lack of funding. In either case, it is not necessary for a case management instance to "live", there for it is set to be closed. If circumstances change at some point in the future, there is always a way to reactivate the old case instance or create a new one.

Furthermore, the system of notifications will be available in the FW4RES for the user to be informed of certain events. There will be two types of notifications:

- Type 1 – after the expiration of the legal deadline for obtaining a solution, the user can be notified to check if there are any changes.
- Type 2 – expiration of expired licenses - N days before the expiration or on the day of the expiration, inform the user so that he can submit a request for the extension of the corresponding license.

Last of all, each stage of the project documentation obtainment process is secured since it depends on specific criteria before moving forward. If the criteria are not met, documentation obtainment is paused, canceled, or restarted, depending on the criteria type. The whole process is ready to be updated and improved further (e.g., integration with external services).

IV. CONCLUSION

Despite the digitization and automation of business processes, obtaining the necessary documentation for various construction projects is still relatively slow. The administrative processes in which the documentation is obtained, such as numerous permits, remain a very critical and time-consuming issue. This is a significant constraint, especially in the construction of RES power plants such as wind farms and solar power plants, but also in other large projects that are being stalled because of administrative issues. The paper describes the basics of CMMN and DMS, which, in communication with FieldWork 4 RES, greatly facilitate the previously mentioned issues.

In the future, it is planned to implement a system of notifications and integration with external systems such as e-Permits. The user will have full up-to-date documentation in one place with no need of going to external services and websites. Also, the notification system will allow the user to stay alert for important events such as the expiration date of a certain document, various deadlines, and other upcoming related to project events.

REFERENCES

[1] European Commission, Directorate-General for Energy, Tallat-Kelpšaitė, J., Brückmann, R., Banasiak, J., et al., Technical support for RES policy development and implementation: simplification of permission and administrative procedures for RES installations (RES simplify): interim report, Publications Office of the European Union, 2022

[2] M. A. Marin "Introduction to the Case Management Model and Notation (CMMN)." ArXiv, 2016.

[3] I. Routis, M. Nikolaidou, N. Alexopoulou and D. Anagnostopoulos, "Empowering Knowledge Workers with CMMN: The Concept of Case Learning," 2018 IEEE 22nd International Enterprise Distributed Object Computing Workshop (EDOCW), Stockholm, Sweden, 2018, pp. 33-36

[4] Š. Sielskaitė and D. Kalibatienė, "On CMMN Model for Software System Project Simulation," 2021 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2021, pp. 1-6

[5] C. Corea, P. Delfmann, "A Tool to Monitor Consistent Decision-Making in Business Process Execution", 16th International Conference on Business Process Management, Sydney, Australia, 2018

[6] A. Zensen and J. Küster, "A Comparison of Flexible BPMN and CMMN in Practice: A Case Study on Component Release Processes," 2018 IEEE 22nd International Enterprise Distributed Object Computing Conference (EDOC), Stockholm, Sweden, 2018, pp. 105-114

[7] K. Abaci, I. T. Medeni, "Efficiency of electronic document management systems: a case study," Science, Education and Innovations in the context of modern problems. Vol. 5. 2022

[8] D. Etinger, S. D. Simić and L. Buljubašić, "Automated decision-making with DMN: from decision trees to decision tables," 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2019, pp. 1309-1313

[9] M. K. Ugale, S. J. Patil and V. B. Musande, "Document management system: A notion towards paperless office," 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM), Aurangabad, India, 2017, pp. 217-224

[10] Shanmugaraja. P, S. K. Susmitha, S. Swadha, R. Vijay, G. Naveen, "Document Management System using Blockchain," 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), Trichy, India, 2022, pp. 888-891

Influence of Personal Recommendations towards Electricity Consumption on Consumer Behavior

J. Héjjová, J. Bucko
Technical University of Košice, Faculty of Economics, Department of Applied Mathematics and business Informatics, Košice, Slovakia
janahejjova@centrum.sk, jozef.bucko@tuke.sk

***Abstract* – Reducing electricity consumption belongs to significant prerequisites of mitigating the negative impact of the current energy crisis in Europe. At the same time, it is a necessary condition for slowing down the environmental deterioration of the Earth. Various studies have shown that the Bidgely technology, which provides consumers with disaggregated data on individual electric appliance consumption, has proven to be a successful way of achieving savings and more responsible consumer behavior. Bidgely enables sending personal recommendations to an individual consumer, when their consumption compared to consumers within a close radius, is higher and above the benchmark. Our study examines the impact of personal recommendations by Bidgely on consumer behavior. It compares the consumption of individual appliances in a household before and after a consumer has received the recommendations.**

***Keywords* - energy efficiency, disaggregation, behavioral changes, nudge, consumer behavior**

I. INTRODUCTION

Energy is a two faces type of commodity with its bright and dark side, both of noteworthy impact on humans. Brightly, it has enabled the game changing inventions such as: internet, new ways of medical treatments, options towards universe exploration and many more. As a result, humans have become heavily dependent on energy consumption in order to maintain sustainable economic development, which leads to increase of the world's demand for energy (Benromdhane 2015; Mahi et al. 2020). The second and darker side of energy is linked to its production, which belongs to the major factors causing the Earth's deterioration as 62 % of the global energy supply origins from fossil fuels (EMBER, 2022).

Followingly, there are 2 motivators increasing the importance of changes in consumption from the long-term perspective:

1. Energy's remarkable role within preserving current society's functions and its further evolution.

2. Limited volumes of clean energy, as production from fossil fuels needs to be eliminated to slow down the climate change, but sustainable sources can be unpredictable as shown in during the year 2021 resulting into energy crisis.

70 % of carbon dioxide emissions into the environment origin from households (Wang et al., 2017). This fact linked to the knowledge that energy production is one of the major factors of Earth's deterioration is a reason strong enough to pursue investigation into ways how household consumers use electricity and to search for the moments, when it can be more efficient. Households' energy consumption has a large saving potential estimated up to 27% of current households' energy use, which can be saved through more efficient energy use (Zhou, Yang, 2016).

This study examines, if technology providing disaggregated data on household's electricity consumption has a potential to be included into set of tools nudging electricity consumers towards the sustainable change in their behavior within electricity consumption. It focuses on behavior within individual categories of electricity consumption with the aim to reveal the areas providing the best potential for remarkable and sustainable decrease in consumption as an outcome of behavioral change.

This paper is structured as follows:

- Section 2 "Current state overview" introduces various findings on the topic of sustainable behavioral change of habits. Also, it briefly describes the Bidgely technology, which currently belongs to the major disaggregation techniques and data analyzed in this study origin from there.
- Section 3 "Data and methodology" describes the dataset and approach to examining the data related to energy consumption.
- Section 4 "Results and limitations" presents our findings within behavior related to energy efficiency.
- Section 5 "Conclusion and Discussion" summarizes our study and provides the review of implementation options for the findings.

II. THE CURRENT STATE OVERVIEW

Electricity production is among major factors generating carbon footprint. It has been estimated that through the realistic implementation of already known changes in consumer behavior, the European Union (EU) could reduce its carbon footprint by about 25% (Moran et al., 2020). The most impactful are changes in the consumption pattern (28% of the total), reduced consumption (26%), switching to goods with a lower carbon footprint in production (17%) and to goods with less carbon emission during use (19%).

Electricity usage is mostly connected to daily routines or habits. As habits are one of the strongest impediments to lifestyle change, acting to 'lock in' behaviour (Marechal, Lazaric, 2011), this reality makes achievement of lasting sustainable change in consumers' behavior a complex issue with wide spectrum of opinions and suggested approaches.

Appropriate timing of change initiation

A growing amount of literature points to the importance of not only how to intervene to achieve social and lifestyle change, but also when. Many interventions (e.g., information campaigns) are ineffective because they are not strong enough to disrupt habits (Verplanken et al., 1997). But, because habits are cued by stable contexts as e.g.: the same time, place and/or social group, (Wood et al., 2005), change in context disrupts habits (Verplanken et al., 2008). For example, low-carbon behavior, such as bus use, energy efficiency and waste reduction measures, have been shown to be more effectively changed using low-cost interventions in the 12 weeks after relocation (Verplanken, Roy, 2016; Ralph, Brown, 2017), as well as at other moments of change (Nicolson et al., 2017).

Also, the times of significant change or transition (Thompson et al., 2011) have been identified as key opportunities for reconfiguring lifestyles (Capstick et al., 2014; Graham et al., 2011) and identities (Devine et al., 2020).

A number of researches reveal, that disruptions either life-course (e.g. moving home) or structural events (e.g. economic downturn, extreme weather events, the COVID-19 pandemic) provide opportunities to more effectively change behavior (Birkman et al., 2010; Carroll, Conboyt, 2020; Verplanken et al., 2018; Marsden et al., 2020).

This knowledge increases actuality and relevancy of the topic of changing consumer behavior within electricity consumption due to turbulent market situation triggered by energy crisis starting in 2021 and being deepened by impact of war in Ukraine on the energy market. That only confirms humans' dependency on energy and its importance.

We prefer to be nudged in simple and encouraging manner

The implementation of policies aimed at encouraging and preserving the change in behavior includes a set of uneasy topics to be sold. Cutting down consumption in general, no matter if related to food and obesity impact on healthcare system or related to environmental goals, is unpopular among citizens as cutting down means restriction or change in a habit.

There have been a numerous studies providing various outcomes and opinions on effective way of behavioral change.

One of the strongest factors shaping behavior is social influence, yet rarely recognized by individuals themselves (Nolan, Schults, Cialdini, 2008). Adoption of low-carbon innovations, such as electric cars and solar panels, is significantly shaped by social norms and neighborhood effects (Bollinger, Gillingham, 2012; Grazziano, Gillingham 2015; Pettifor et al., 2017). According to study by Bogueva et al. (2017), social norms have a significant and positive relationship with both consumer attitudes and intention towards reducing their meat consumption related to environmental protection.

Waste separation is more simple and less financially demanding example of social influence impact. This type of behavioral change is compatible with the tendency of environmental psychologists to focus on low-impact, incremental behavior changes that are 'simple and painless' (Thøgersen, Crompton, 2009) rather than higher impact, more transformative behavior changes, such as purchasing energy-efficient or renewable energy equipment (Nielsen et al., 2021), which are necessary for lifestyle change that is in line with effective climate change mitigation (Whitmarsh, 2021).

Within area of environment protection, a positive effect on the willingness to change behavior appears to be conditioned by the link between our collective action and climate change salient. On the contrary, making salient how one's own behavior harms the climate decreases the chance of desired behavioral change (Lavallee, Di Giusto, 2019). Peifer et al. (2020) find that voluntary simplicity is positively related to perceived consequences for others and that making the link between consumption and climate change salient significantly reduces intentions to buy a new pair of shoes that is desired, but not needed.

Households usually have an economic incentive to save heating and electricity, and utilities sometimes provide additional incentives. A Swizz study found that electricity consumers generally prefer positive incentives for reducing consumption to negative incentives against increasing consumption (Mahmoodi et al., 2020).

The demographic factors have been generally seen to have small effects on acceptability. Among the stable factors, ideology seems to be a consistent predictor, while the influence of generalized and political trust is more limited and varied. Values, specifically self-transcendence or bio-spheric values, show stable positive effects on acceptability, but the association is typically quite weak. (Ejelov, Nilsson, 2020)

Most of the studies agree, that desired behavior should be presented with more of encouragement to do something with benefiting result, and less as a necessity to change something to prevent undesired outcome. They also reveal that consumer acceptance of incentive schemes depends on loss and risk-aversion as well as on how optimistic consumers are regarding their ability to meet the goal of their behavioral change.

Working ways of delivering the motivation

Often used informational approaches appear to be generally less effective than other types of intervention (Abrahamse, Matthies, 2012). Informational campaigns may raise awareness and concern but do not always produce behavioral change (Staats et al., 1996). Their effectiveness could be supported by respecting a few principles as e.g.: tailor messages to audience values and beliefs (Whitmarsh, Corner, 2017); target times and locations of decision-making (WRAP, 2020; Kaiser et al. 2020); promoting self-efficacy instead of, or in addition to appealing to fear (Hunter, Röös, 2016), encouraging and

setting specific and realistic goals to motivate action (Abrahamse et al., 2005).

Disaggregated data on electricity consumption presented in a user friendly and simplified way meet the above listed preconditions, which makes it a potential tool to support the behavioral change within electricity consumption. Therefore, it has been selected to be a subject of this study. One of the current major disaggregation technologies is Bidgely technology, which provides household electricity consumers with regular reports on household's consumption; and recommendations on the household management areas such as: space heating, refrigeration, lighting, always on appliances.

Disaggregated information itself has two major impacts (Gupta, Chakravarty, 2013):

- increased consumer engagement,
- reduced energy usage.

The technology addresses a combination of particular behavioral principles towards nudging consumers to change their behavior within energy consumption. It can help effectively shape energy efficiency policies and campaigns (Dawnay, E., Shah, H.; 2005):

a) **Habits**. It is more difficult to change daily routines requiring little or no cognitive efforts, which is of significant relevancy in case of electricity use.

b) **Rewards**. The change is conditioned by reward, ideally shortly after the action. Energy is not a sugar. We cannot see, how much have we saved by implementing some change in usual behavior. Therefore, in case of energy efficiency an example is an easy access to the data on consumption, which is now available thanks to disaggregation technologies analyzing the aggregate household main power measurement in the house, and disaggregating this into individual appliances (Devlin, Hayes, 2019).

c) **Consequences**. The consequences of the change are important to us. In this field, the No. 1 is an electricity bill. The first-hand output from smart meters is disaggregated data presented in the kilowatt-hours (kWh) consumed.

d) **Involvement**. People need to feel involved and effective to make a change. In case of electricity consumption, a level of accessibility of information and its complexity can lead to feelings of helplessness and inaction. Currently, there have been applications linked with disaggregation technology, which are able to produce the user-friendly regular report on consumption.

e) **Approval**. Humans' behavior is strongly influenced by other people's behavior. An example of the tool enabling common natural influence among consumers is benchmarking to similar consumers to present the effective level of consumption and support the feasibility of recommendations.

III. DATASET DESCRIPTION AND METHODOLOGY

A. Dataset description and processing and hypothesis

The database examined consists of two datasets containing electricity consumption data of approximately 30 000 consumers within the household segment in Slovakia. The data contains consumers meeting the legislative criteria for the obligatory installment of smart meters, which has been proceeding sequentially in Slovakia since 2014. These consumers are primarily households using electricity for heating or equipped with electric water boilers. Free application Bidgely has been offered to consumers with smart meters and the dataset includes the ones, who have activated it.

The first dataset "peer.insight" contained more than 3 760 million rows in 14 columns with monthly electricity consumption of consumers during 3 years: from 2018 until 2021. Electricity monthly usage (in kWh) of consumers is disaggregated according to following appliances (areas):

- Air conditioning
- Always on
- Lighting
- Pool
- Refrigeration
- Space heating
- Water heating
- Z-Other
- EV
- Cooking

The second dataset named "bidgely_recommendations" has more than 11 091 million rows in 10 columns with messages (text recommendations – tips or advice), which consumers receive to reduce the consumption in a specific area or as an advice for a particular appliance. Customers have been receiving notifications from January 2020. Some examples of recommendations are presented in Table 1.

TABLE I. SAMPLE OF DATASET BIDGELY_RECOMMENDATIONS

user_id	start_date	end_date	id_recommendation	title	Description
5000047068	2020-02-01	2020-03-01	R180	Look for ENERGY STAR when shopping for audio and video equipment	ENERGY STAR certified equipment is up to 50% more efficient than conventional models.
5000047068	2020-02-01	2020-03-01	R160	Don't want to keep unplugging?	Use a Smart Plug! These compact devices allow remote control and scheduling to avoid unnecessary energy use.
5000047068	2020-02-01	2020-03-01	R340	Improve your insulation	Repairing or replacing your insulation and weatherstripping can save you 20 percent or more on heating and cooling costs.

5000504793	2020-01-01	2020-02-01	R360	Avoid the night sweats	Set your thermostat between 65 and 68 degrees before going to bed to reduce heating costs by up to 10 percent.
5000901823	2020-02-01	2020-03-01	R590	Consider task lighting	Why light an entire room? Light only what you need.
5000841845	2020-05-01	2020-06-01	R700A	Adjust your fridge temperature settings	An ideal fridge temperature is 33°F to 39°F.
5000185210	2020-04-01	2020-05-01	R710A	Adjust your freezer temperature settings	An ideal freezer temperature is 0°F to 5°F.
...	...	...	...	...	...

Source: Own processing

The first step in data pre-processing was to aggregate recommendations to ten groups according to a type of area consumption (appliances).

TABLE II. NUMBER OF RECOMMENDATIONS IN PARTICULAR AREAS OF CONSUMPTION

IS_appliance	Name (area)	Number of recommendations
1	Air conditioning	272025
2	Always on	2893344
3	Lighting	1475221
4	Pool	8213
5	Refrigeration	2084255
6	Space heating	1164120
7	Water heating	1819720
8	Z-Other	373313
9	EV	4817
10	Cooking	996351

Source: Own processing

R-Studio was used in following work with datasets. Some unnecessary columns were deleted from both datasets. Our research focused on five consumptions areas: Space heating, Refrigeration, Lighting, Always on, and Air conditioning. There were filtered the unique consumers, who obtained relevant recommendations from the dataset "bidgely_recommendations" for every area of consumption. The resulting sub-dataset was joined with the dataset "peer_insight". A new dataset with monthly electricity usage of consumers was obtained from June 2018 to August 2021 in the filtered area (appliance). Finally, the monthly customers' average electricity consumption was calculated. Each of the resulting five datasets consists of the monthly average usage (kWh) of consumers from June 2018 to August 2021. Values before January 2020 represent the average consumers' electricity consumption before receiving recommendations, and values after January 2020 represent the average consumers' electricity consumption after they have started to receive recommendations.

The study examines 5 hypotheses:

HYPOTHESIS 1: The average electricity consumption of customers in the area of Space heating will not statistically significantly change after sending recommendations for its decrease.

HYPOTHESIS 2: The average electricity consumption of customers in the area of Refrigeration will not statistically significantly change after sending recommendations for its decrease.

HYPOTHESIS 3: The average electricity consumption of customers in the area of Lighting will not statistically significantly change after sending recommendations for its decrease.

HYPOTHESIS 4: The average electricity consumption of customers in the area of Always on will not statistically significantly change after sending recommendations for its decrease.

HYPOTHESIS 5: The average electricity consumption of customers in the area of Air conditioning will not statistically significantly change after sending recommendations for its decrease.

B. Methodology of analyzing trend in electricity consumption

The analysis compares consumers' average electricity consumption before and after receiving recommendations. Graphical presentation was used to show the trend of average consumers' electricity consumption for each use case. To confirm the Hypotheses, the data was divided into two sets - before and after January 2020. The Shapiro-Wilk test at 0.01 level of significance was used to test the normality of the data. After the confirmation, that the data is normally distributed, the F-test was used to test whether two samples have equal variances. If the data samples have equal variances, the t-test was performed to determine if both population means are equal. (Hudec, Liptáková, 2018)

IV. RESULTS AND LIMITATIONS

A. Results

This research has focused on influence analysis of personal recommendations to customer behavior related to electricity use in individual areas of electricity consumption: space heating, refrigeration, lighting, always on, and air conditioning. Five hypotheses were examined. Each hypothesis was set to each case of electricity consumption and supposed that consumption would not change statistically significantly after sending recommendations.

TABLE III. SHAPIRO-WILK NORMALITY TEST

	Before recommendations	During recommendations

	W	p-value	W	p-value
Space Heating	0.88031	0.02174	0.90898	0.0709
Refrigeration	0.92582	0.145	093441	0.2087
Lighting	0.93875	0.2503	0.9667	0.7092
Always On	0.96794	0.7345	0.96794	0.7345
Air Conditioning	0.92803	0.3913	0.96796	0.8834

Source: Own processing

Firstly, the data were divided into two datasets - before and during receiving recommendations within each category of electricity consumption. Shapiro-Wilk normality test was used to test data for each sub-dataset at the significance level $\alpha = 0.01$. The results shown in Table 3 determine, that data within all sub-datasets are normally distributed, that is the presumption to test if two samples within each category have equal variances by using F-test and therefore, if variances of average consumer consumption differ between periods before and during receiving recommendations.

TABLE IV. F-TEST RESULTS

	p-value	results
Space heating	0.461	p value > 0,05 , two samples have equal variances
Refrigeration	0.218	p value > 0,05 , two samples have equal variances
Lighting	0.241	p value > 0,05 , two samples have equal variances
Always on	0.479	p value > 0,05 , two samples have equal variances
Air conditioning	0.251	p value > 0,05 , two samples have equal variances

Source: Own processing

We have used F-test to confirm that the two samples have equal variances. Calculations of the results are shown in Table 4. As by the F-test results, the variances between datasets within all the categories are equal, T-test is used to test if a means of average consumer consumption differ between periods before and during receiving recommendations.

TABLE V. T-TEST RESULTS

	p-value	results
Space Heating	0.0029	$H1_0$ was rejected. There exists difference between values before and after at the significance level $p<0.05$
Refrigeration	0.9876	$H2_0$ was confirmed. There exists no difference between values before and after at the significance level $p<0.05$
Lighting	0.5593	$H3_0$ was confirmed. There exists no difference between values before and after at the significance level $p<0.05$
Always On	0.1123	$H4_0$ was confirmed. There exists no difference between values before and after at the significance level $p<0.05$
Air Conditioning	0.0015	$H5_0$ was rejected. There exists difference between values before and after at the significance level $p<0.05$

Source: Own processing

The results of the T-test are shown in Table 5. For categories Space heating and Air conditioning means of average consumer consumption differ between periods before and during receiving recommendations. For the categories Refrigeration, Lighting, and Always on there are no differences between periods before and during receiving recommendations.

Space heating and Air Conditioning

Due to the results, there are statistically significant differences between consumption before and during receiving recommendations in the categories Space heating and Air conditioning.

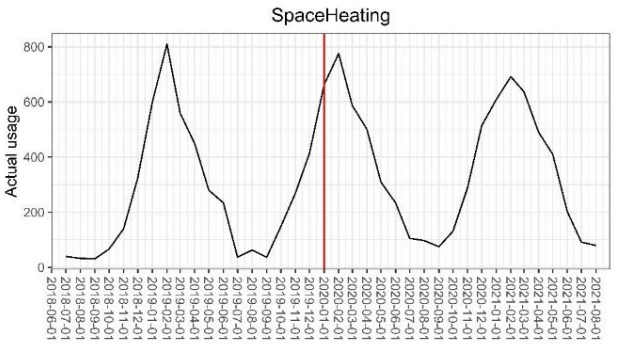


Figure 1. Monthly average electricity consumption for Space heating

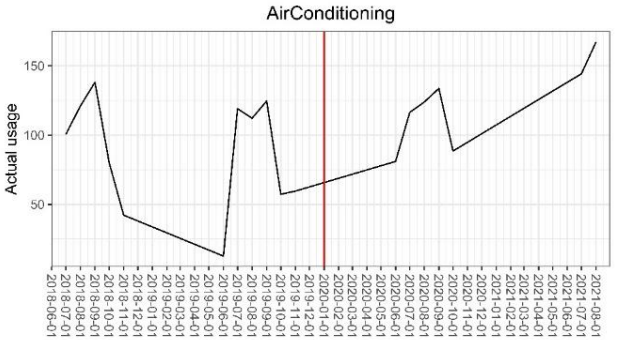


Figure 2. Monthly average electricity consumption for Air Conditioning

Figure 1 and Figure 2 represents a graphical view of the monthly average consumption for the categories. The periods before and during receiving recommendations are split by red lines. The results present, that in case of heating and air conditioning the average consumption of a household increases. As the electricity consumption within these categories are strongly related to weather situation, further research needs to reflect this aspect to specify the impact of disaggregation in these two areas. The same pattern applies to air conditioning and number of days with summer temperatures. The positive impact of disaggregation should not be rejected as abovementioned specific conditions require further research.

Refrigeration, Lighting and Always on

Based on the results, refrigeration consumer electricity consumption has not changed after they started receiving the recommendations. The reason links to the way this appliance works within household and its consumption is permanent. In this area the usefulness of monitoring its consumption links to the malfunctions of refrigerator,

when significantly increased consumption suggest that the appliance should be changed or repaired.

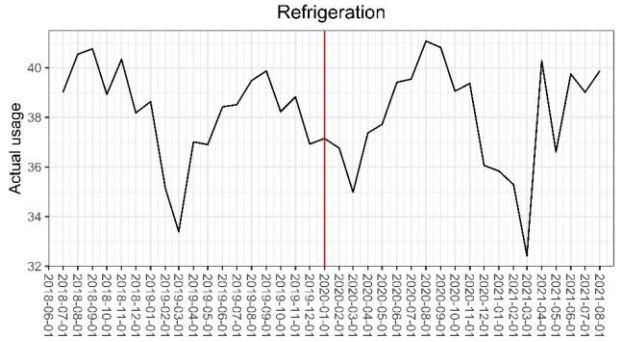


Figure 3. Monthly average electricity consumption for Refrigeration

Based on the results, Lighting consumer electricity consumption has not changed after they started receiving the recommendations. The reason should be, that lightning is area with the lowest electricity consumption. Quantified financially, this area doesn't provide sufficient motivation to change habits.

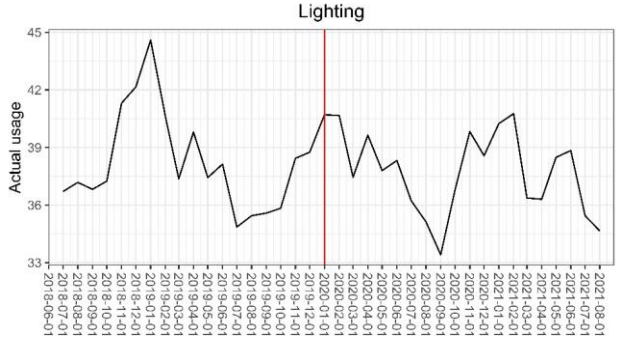


Figure 4. Monthly average electricity consumption for Lighting

Based on the results, Always on consumer electricity consumption has not changed after they started receiving the recommendations. The reason would be similar to lightning category, providing not attractive trigger to change a consumer behavior.

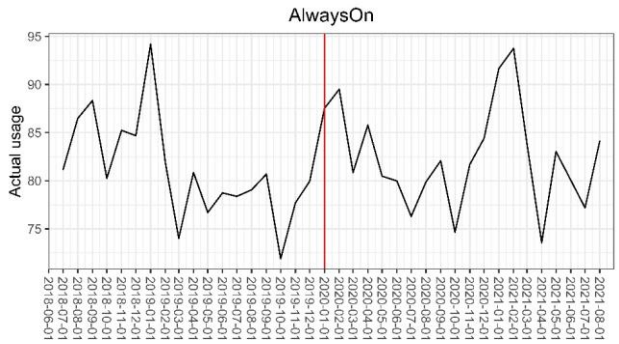


Figure 5. Monthly average electricity consumption for AlwaysOn

B. Limitations

Our current research has been accompanied by a few limitations, that need to be taken into account within interpretation of the results:

- For the purpose of the further use of these results, we have worked with average values, which do not reflect the seasonality of energy consumption.
- The level of energy consumption and efficiency is a complex topic influenced by the group of factors. We focused on the impact of an individual tool based on behavioral principles. Its impact could be intensified when combined e.g. with campaigns by state authorities.
- The personalized report on a particular household consumption has, within this database, been send through only one tool, an e-mail message without reminder notification. Therefore, deliverability of the message would have an impact on effectivity as well and way how to increase it, is another topic with the potential of improvement the total outcome.

V. CONCLUSION AND DISCUSSION

Behavior change is often narrowly conceived as individual-level consumer action (e.g. buying a low carbon product, recycling, reducing meat-eating), but is more appropriately understood as extending across the many roles and contexts humans occupy: as members of communities, participants in organizations, and as citizens who can influence policies (Nielsen et al., 2021).

The aim of this research is to study human motivations to proceed long term and sustainable change in behavior and apply them when creating policies to reach a globally beneficial outcome. As some studies show, favorable conditions towards change in habits are provided by the times of significant social change or transition, such as current tough global political and economic situation shaped by recent Covid pandemic as well as ongoing energy crisis and war in Ukraine.

The research has pointed at households, but they are persons, who can be decisions makers in their professional lives. When they change their behavior towards electricity consumption at home, it will become natural to them to transfer behavioral pattern it into professional life and behave same towards the decision at work, which expands the impact of households' behavioral change into company sector.

Disaggregation of electricity consumption and simple approach to user friendly and easily understandable data on consumption is among the tools with potential to support the behavioral change thanks to reflecting a number of behavioral principles such as: habits, rewards, consequences, involvement and approval (Dawnay, Shah, 2005)

This study has pointed at examining the particular areas of electricity consumption within households. The results show, that areas with low consumption levels and therefore also low costs on electricity appear to provide insufficient motivation to change behavior towards changing habits and consequently the electricity consumption.

The results provide an outcome, that in case of heating and air-conditioning the average consumption of a

household increases. But the amount of electricity used in these areas is significantly influenced by the weather. Therefore, further methods need to be applied to reflect the weather conditions in the period examined. In other words, average consumption of a consumer related to heating would be higher during a year when winter temperatures appear from December till January in comparison to for example a period when winter temperatures start in December but last till end of March. The same pattern applies to air conditioning and number of days with summer temperatures. The positive impact of disaggregation should not be rejected as abovementioned specific conditions require further research.

The study investigates changes in consumers' behavior triggered by comfortable access to user friendly reports containing disaggregated data related to their energy consumption. Consumers included in the dataset receive the report on their e-mail therefore the impact of recommendations on behavioral changes is strongly influenced by delivery of the message to a consumer. Therefore, open rates of the e-mails containing reports need to be reflected in more detail within further research.

Lowering energy consumption through permanent and lasting behavioral change is a complex topic requiring involvement of different subjects such as policy makers, schools, professionals and other authorities. It is also conditioned by combining a set of tools and choice of appropriate and inspiring contents. As explained in the study, disaggregation technology delivered to a consumer in a simplified and easily understandable way meets a number of preconditions to support the desired behavioral change.

REFERENCES

[1] Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. J Environ Psychol 2005, 25(3):273–291.

[2] Abrahamse, W., Matthies, E. (2012). Informational strategies to promote pro-environmental behaviours: changing knowledge, awareness, and attitudes. In Environmental Psychology: an Introduction. Edited by Steg L, van den Berg AE, de Groot JIM, Stafford: Wiley- Blackwell; 2012:224–231.

[3] Bain, P., Milfont, T., Kashima, Y., et al. (2016). Co-benefits of addressing climate change can motivate action around the world. Nature Clim Change 2016, 6:154–157.

[4] Balderjahn, I., Lee, M.S.W, Seegebarth, B., Peyer, M. (2020). A sustainable pathway to consumer wellbeing. The role of anticonsumption and consumer empowerment. J Consum Aff 2020, 54:456-488.

[5] Balderjahn, I., Seegebarth, B., Lee, M.S.W. (2020). Less is more! The rationale behind the decision-making style of voluntary simplifiers. J Clean Prod 2020, 284:124802.

[6] Beattie, G., McGuire, L. (2020). The modifiability of implicit attitudes to carbon footprint and its implications for carbon choice. Environ Behav 2020, 52:467-494.

[7] Birkmann, J., Buckle, P., Jaeger, J., et al. (2010). Extreme events and disasters: a window of opportunity for change? Analysis of organizational, institutional and political changes, formal and informal responses after mega-disasters. Nat Hazards 2010, 55:637–655.

[8] Bollinger, B., Gillingham, K. (2012). Peer effects in the diffusion of solar photovoltaic panels. Market Sci 2012, 31(6):900–912.

[9] Bogueva, D., Marinova, D., & Raphaely, T. (2017). Reducing meat consumption: The case for social marketing. Asia Pacific Journal of Marketing & Logistics, 29(3), 477–500.

[10] Capstick, S., Lorenzoni, I., Corner, S., Whitmarsh, L. (2014) Social science prospects for radical emissions reduction. Carbon Manag 2014, 4(5):429–445.

[11] Carroll, N., Conboy, K. (2020). Normalising the "new normal": changing tech-driven work practices under pandemic time pressure. Int J Inf Manag 2020, 55:102186.

[12] Dawnay, E., Shah, H. (2005). Behavioural economics: seven principles for policy-makers. Published by New economics foundation. Based on the report Extending the "rational man" model of human behaviour: seven key principles commissioned by The Environment Agency. ISBN 1 904882 03 X.

[13] Devine-Wright, P., Pinto de Carvalho, L., Di Masso, A., Lewicka, M., Manzo, L., Williams, D.R. (2020). "Re-placed" - reconsidering relationships with place and lessons from a pandemic. J Environ Psychol 2020, 72:101514.

[14] Ejelöv, E., Nilsson, A. (2020). Individual Factors Influencing Acceptability for Environmental Policies: A Review and Research Agenda. Sustainability 2020, 12, 2404; doi:10.3390/su12062404 www.mdpi.com/journal/sustainability

[15] EMBER (2022). https://ember-climate.org/insights/research/global-electricity-review-2022/

[16] Graham-Rowe, E., Skippon, S., Gardner, B., Abraham, C. (2011). Can we reduce car use and, if so, how? A review of available evidence. Transport Res Pol Pract 2011, 45(5):401–418.

[17] Graziano, M., Gillingham, K. (2015). Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. J Econ Geogr 2015, 15(4):815–839.

[18] Gupta, A., Chakravarty, P. (2013). Impact of Energy Disaggregation on Consumer Behavior. 2013 BECC Conference Sacramento CA

[19] Hudec, O., Liptáková, E. (2018): Pravdepodobnosť a metódy induktívnej štatistiky, 1. vyd., Košice, Technická univerzita v Košiciach, 198 p. [print]. - ISBN 978-80-8086-270-1.

[20] Hunter, E., Röös, E. (2016). Fear of climate change consequences and predictors of intentions to alter meat consumption. Food Pol 2016, 62:151–160. [56].

[21] Kaiser, M., Bernauer, M., Sunstein, C.R., Reisch, L.A. (2020). The power of green defaults: the impact of regional variation of opt-out tariffs on green energy demand in Germany. Ecol Econ 2020, 174:106685.

[22] Lavallee, J.P., Di Giusto, B., Yu, T.Y. (2019). Collective responsibility framing also leads to mitigation behavior in East Asia: a replication study in Taiwan. Clim Change 2019, 153:423-438.

[23] Mahmoodi, J., Hille, S., Patel, M.K., Brosch, T. (2020). Using rewards and penalties to promote sustainability: who chooses incentivebased electricity products and why? J Consum Behav 2020 http://dx.doi.org/10.1002/cb.1870.

[24] Maibach, E.W., Nisbet, M., Baldwin, P., et al. 2010). Reframing climate change as a public health issue: an exploratory study of public reactions. BMC Publ Health 2010, 10(299).

[25] Marechal, K., Lazaric, N. (2011). Overcoming inertia: insights from evolutionary economics into improved energy and climate policies. Clim Pol 2011, 10:103–119.

[26] Marsden, G., Anable, J., Chatterton, T., Docherty, I., Faulconbridge, J., Murray, L., Shires, J. (2020). Studying disruptive events: innovations in behaviour, opportunities for lower carbon transport policy? Transport Pol. 2020, 94:89–101, https://doi.org/ 10.1016/j.tranpol.2020.04.008.

[27] Moran, D., Wood, R., Hertwich, E., Mattson, K., Rodriguez, J.F.D, Schanes, K., Barrett, J. (2020). Quantifying the potential for consumer oriented policy to reduce European and foreign carbon emissions. Clim Policy 2020, 20:S28-S38.

[28] Nielsen, K.S., Clayton, S., Stern, P.C., Dietz, T., Capstick, S., Whitmarsh, L. (2021). How psychology can help limit climate change. Am Psychol 2021, 76(1):130–144.

[29] Nicolson M, Huebner GM, Shipworth D, Elam S: Tailored emails prompt electric vehicle owners to engage with tariff switching information. Nature Energy 2017, 2:17073.

[30] Nolan, J.M., Schultz, P.W., Cialdini, R.B., Goldstein, N.J., Griskevicius, V. (2008). Normative social influence is underdetected. Pers Soc Psychol Bull 2008, 34:913–923.

[31] Peifer, J.L, Chugani, S., Roos, J.M. (2020). The ethical underpinnings of nonmaterialistic values and voluntary simplicity behavior in the United States. Psychol Market 2020, 37:232-249.

[32] Pettifor, H., Wilson, C., McCollum, D., Edelenbosch, O.Y. (2017). Modelling social influence and cultural variation in global low-carbon vehicle transitions. Global Environ Change 2017, 47:76–87.

[33] Ralph, K.M., Brown, A.E. (2017). The role of habit and residential location in travel behavior change programs, a field experiment. Transportation 2017, https://doi.org/10.1007/s11116-017- 9842-7.

[34] Sorrell, S., Gatersleben, B., Druckman, A. (2020). The limits of energy sufficiency: a review of the evidence for rebound effects and negative spillovers from behavioural change. Energy Res Soc Sci 2020, 64:101439.

[35] Staats, H.J., Wit, A.P., Midden, C.J.H. (1996). Communicating the greenhouse effect to the public : evaluation of a mass media campaign from a social dilemma perspective. J Environ Manag 1996, 46(2):189–203.

[36] Thompson, S., Michaelson, J., Abdallah, S., Johnson, V., Morris, D., Riley, K., Simms, A. (2011). Moments of Change as Opportunities for Influencing Behaviour: a Report to the Department for Environment, Food and Rural Affairs. London: The New Economics Foundation. Defra; 2011.

[37] Thøgersen, J., Crompton, T. (2009). Simple and painless? The limitations of spillover in environmental campaigning. J Consum Pol 2009, 32(2):141–163.

[38] Verplanken, B., Aarts, H., van Knippenberg, A. (1997). Habit, information acquisition, and the process of making travel mode choices. Eur J Soc Psychol 1997, 27:539–560.

[39] Verplanken, B., Walker, I., Davis, A., Jurasek, M. (2008). Context change and travel mode choice: combining the habit discontinuity and self-activation hypotheses. J Environ Psychol 2008, 28: 121–127.

[40] Verplanken, B., Roy, D. (2016). Empowering interventions to promote sustainable lifestyles: testing the habit discontinuity hypothesis in a field experiment. J Environ Psychol 2016, 45:127–134.

[41] Verplanken, B., Roy, D., Whitmarsh, L. (2018). Cracks in the wall: habit discontinuities as vehicles for behavior change. In The Psychology of Habit. Edited by Verplanken B, Springer; 2018. This chapter reviews the evidence of how habits are disrupted during periods of transition, such as relocation and retirement, and how these periods represent opportunities to more effectively intervene to change behaviour (e.g., towards low-carbon action).

[42] Videbæk, P.N., Grunert K.G. (2020). Disgusting or delicious? Examining attitudinal ambivalence towards entomophagy among Danish consumers. Food Qual Prefer 2020, 83:103913.

[43] Wang, Z., Wang, X., & Guo, D. (2017). Policy implications of the purchasing intentions towards energy-efficient appliances among China's urban residents: Do subsidies work? Energy Policy, 102, 430–439.

[44] Whitmarsh, L., Corner, A. (2017). Tools for a new climate conversation: a mixed-methods study of language for public engagement across the political spectrum. Global Environ Change 2017, 42:122–135.

[45] Whitmarsh, L., Poortinga, W., Capstick, S. (2021). Behaviour change to address climate change Current Opinion in Psychology 2021, 42:76–81.

[46] Wood, W., Tam, L., Guerrero, Witt. M. (2005). Changing circumstances, disrupting habits. J Pers Soc Psychol 2005, 88:918–933.

[47] WRAP: The Effectiveness of Providing Pre-Purchase Factual Information in Encouraging More Environmentally Sustainable Product Purchase Decisions. 2019. Downloaded on 28 Nov 2020 from, https://www.wrap.org.uk/providing-pre-purchaseinformation.

[48] Zhou, K., Yang, S. (2016). Understanding household energy consumption behavior: The contribution of energy big data analytics. Renewable and Sustainable Energy Reviews 56 (2016) 810–819

Advanced and Smart Maintenance of Induction Motors Based on an Algorithm for Motor Faults Recognition

S. Tvorić, T. Đuran, V. Jančić and D. Sente
Zagreb University of Applied Sciences/Department of Electrical Engineering, Zagreb, Croatia
stvoric@tvz.hr, tduran@tvz.hr, vjancic@tvz.hr, dario.sente@tvz.hr

Abstract **- The paper describes the algorithm and the methodology to recognize rotor faults of the squirrel-cage motor based on the induced voltages in the measuring coils spatially spaced for the pole pitch. Based on the FEM calculations and measurements, the functional algorithm for early fault detection and reliable recognition of rotor faults was designed. In the paper is presented and described flowchart of the proposed algorithm. Furthermore, the two special algorithms were designed, based on the obtained research results. The input parameters for the proposed algorithm are induced voltages in four measuring coils spatially spaced for pole pitch. The first algorithm is intended for early and reliable detection of the broken rotor bars and the broken ring of the squirrel-cage induction motor. The second algorithm is proposed for early and reliable detection of the rotor static eccentricity. The proposed algorithm uses induced voltage in four measuring coils. In the case of the broken rotor bars and ring, the algorithm uses voltage difference between two induced voltages in measuring coils spatially spaced for $2\tau_p$. In the case of static eccentricity, the difference between the RMS value of each measuring coil installed in the air gap of the induction motor is analyzed. The proposed algorithms are aimed to improve and contribute to the advanced and smart maintenance of induction motors.**

Keywords – FEM; broken rotor bar; broken rotor ring; squirrel-cage rotor; static eccentricity; magnetic field; air gap; measuring coils; faults detection

I. INTRODUCTION

Induction motors are one of the most common types of electrical machines, and they play a key role in various industrial plants and processes (power plants, oil refineries, chemical plants, metal foundries, pumping stations, coal mills, paper industry, etc.). Because of their importance in the process, these machines must achieve reliable and safe operation. To ensure this, all key machines contain some types of protection, supervision, and monitoring. The main goal is early machine fault detection and the possibility to recognize the type and nature of a fault. The cause of induction machines faults lies in electrical and mechanical conditions together with operational environment. According to the literature [1 – 4], the most common faults of induction machines are bearing faults, following stator and rotor faults and other faults such as different types of eccentricity (static, dynamic or both). In order to detect and recognize induction machine faults early, various types of fault detection methods have been developed [5, 6, 7, 8]. These methods are a result of long-term testing, measurement, and monitoring of the induction machine operation. All these methods with more or less success enable fault detection of induction motor.

Online magnetic field monitoring via permanently installed measuring coils inside air gap is a well-established methodology which enables winding fault detection [1, 9, 10, 11, 12]. In industry and according to the literature [3, 4, 13, 14], rotor winding faults are very common problem. Most commonly used method for detection of broken rotor bars is method based on current signature analysis (MCSA) [8, 15, 16, 17, 18, 19, 20, 21].

The next very common motor problem is the occurrence of rotor static eccentricity that exceeds the allowable limit [22]. There are generally two kinds of rotor eccentricities: static or dynamic (showed on the figure 1). In the case of static eccentricity, the rotor turns around its center of gravity, but this centre is displaced of the stator centre line. From the stator point of view, the place of minimal air gap occurs at unchanged circular position. In the dynamic eccentricity case, the central axis of the rotor is not fixed but during rotation changes its position relative to the stator centre line. In this case the position of minimal airgap does not stay at the same stator circular position, but rather changes in time. Within this paper only the phenomenon of static eccentricity was analyzed. The occurrence of rotor static eccentricity in induction machine can cause some of the following undesirable effects: vibrations, unbalanced magnetic pull (UMP), additional losses in the rotor winding and in parallel circuits of the stator windings, occurrence of the shaft voltages and currents, additional losses in the stator or rotor core, and finally in the worst-case scenario physical contact of the rotor with the stator [23]. These undesirable effects either reduce the machine performance and efficiency or make secure operation of the machine impossible and therefore lead to the complete machine shutdown.

In this paper, a method based on the magnetic field measurement in the air gap using the measuring coils embedded in the machine is presented. This methodology is used for reliable and unambiguous detection of squirrel-cage rotor faults such as: broken rotor bar, broken rotor ring, and static eccentricity. Within this paper, the issues

regarding the rotor faults detection are discussed. Also, the corresponding algorithm for rotor faults recognition has been designed. The aim of the paper is to contribute to the early and reliable fault detection of electrical rotating machines based on the magnetic field analysis in the air gap, with application in advanced and smart maintenance of induction motors.

II. THEORETICAL APPROACH

A. *Theory*

The novelty of this methodology is measurement of the magnetic field in a rotating machine with four measuring coils, installed on the places in the machine which have, by absolute value, equal magnetic vector potential. The distance between the measuring coils is $n \cdot \tau_p$, where τ_p is a pole pitch, and n=1, 2, 3, 4, … is a multiple of the pole pitch. The Fig.1. shows a presentation of the magnetic field line distribution in a four-pole induction machine along with locations for installation of measuring coils.

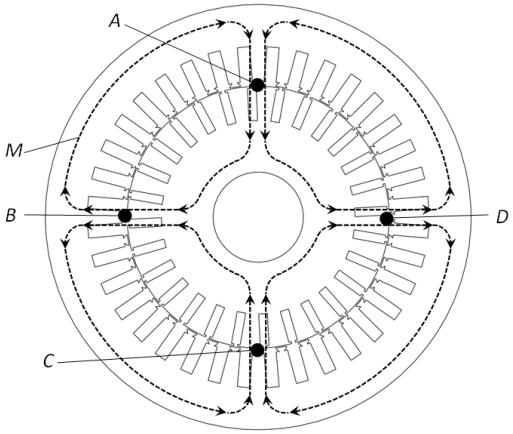


Figure 1. Presentation of the magnetic field line distribution along with locations for installation of measuring coils for magnetic filed measurement [24].

Positions for installation of measuring coils in a four-pole machine, at a mutual distance of multiple pole pitch $n \cdot \tau_p$, where n=1, 2, 3, 4, 5, 6, 7…, marked by positions A, B, C and D as presented in Fig. 1, have by absolute value equal magnetic potential for one selected time point. However, the direction of the magnetic field lines (pos. M Fig. 1) on the observed positions is not the same, but it differs on the positions A and C in relation to positions B and D. At positions A and C the magnetic field lines are in the direction from the periphery of the machine toward the center, whereas at positions B and D they are in the direction from the center toward the periphery. This spatial arrangement enables the detection of squirrel-cage induction motor rotor failures depending on the type and location of the failure. This new methodology is based on the magnetic field analysis obtained by measuring coils installed in the airgap of induction machine. The main advantage of this methodology is that it avoids complex signal processing and the need for harmonic analysis of the measured signal. The presence of rotor faults, such as static eccentricity is determined by detection of changes in induced voltage waveform and RMS, and detection of broken rotor bar or broken rotor ring is detected based on the differential measurement of the induced voltages in the two measuring coils. These statements are proven by numerous FEM calculations and by experimental measurements in the laboratory [24, 25].

B. *FEM Analysis*

The FEM analysis was performed in the commercial software *Simcenter Infolytica Magnet*. A two-dimensional FEM model of the squirrel-cage induction motor is shown on Fig.2. The Fig.3. shows distribution of the magnetic field together with the position of the measuring coils. These measuring coils are in the airgap, installed on the stator teeth and circularly spaced for a pole pitch, τ_p. To achieve reliable static eccentricity or broken rotor bar detection, it is necessary to ensure exactly this kind of measuring coils layout. Based on the waveform and RMS value of the induced voltage in the measuring coils, or based on their voltage difference it is possible to detect rotor eccentricity or rotor winding damages (such as broken rotor bars and ring) of the of the induction machine.

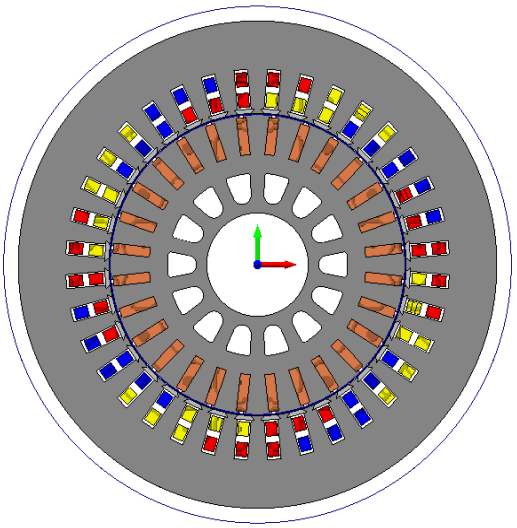


Figure 2. 2D FEM model of induction motor.

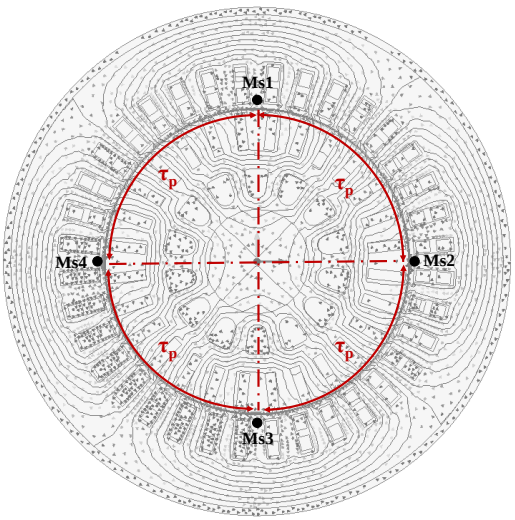


Figure 3. Distribution of the magnetic field with position of the measuring coils.

To determine changes of these voltage values, numerical simulations on the FEM model have been performed. Table 1 lists regular and simulated faulty conditions of the analyzed induction motor. In the case of static eccentricity simulation, percentage of eccentricity is given in relation to the minimal air gap width. First, FEM calculation for a regular machine operation was performed. Then, using the same model with corresponding modifications the simulations of static eccentricity and rotor winding damages (broken rotor bars and broken rotor ring) were performed. In the FEM model, the machine rotor was shifted in the direction of ±y axes for achieving changes of static eccentricity.

TABLE I. CALCULATED NORMAL AND FAULTY CONDITIONS OF THE INDUCTION MACHINE

Operating state	*Normal condition*	*Faulty condition*	*Model variant*
Rated load of the induction motor	•	static eccentricity	15 %
		static eccentricity	20 %
		static eccentricity	25 %
		broken rotor bar	1 bar
		broken rotor bar	2 bars
		broken rotor ring	1 ring

The influence on the induced voltage in measuring coils circularly spaced for a pole pitch was observed and analyzed with the aim to detect the static eccentricity presence and presence of broken rotor bars and ring. Based on the RMS value change of the induced voltage in the coils, the presence of rotor displacement and static eccentricity is determined. The rotor winding damages are detected based on the voltage difference between two measuring coils circularly spaced for a pole pitch, τ_p. Fig. 4 shows the voltage waveform induced in the measuring coils mutually spaced for the pole pitch, τ_p obtained for normal operation condition. For correct voltage waveform interpretation, it is necessary to have certain information regarding machine active part elements and understand their impact on the measuring coil voltage waveform. Fig. 5 presents the impact of the squirrel-cage rotor bars on the induced voltage waveform in the measuring coils. Numbers on this figure mark the impact of every rotor bar on the waveform. In one rotor turn, the waveform contains the number of rises and drops of the voltage that is exactly equal to the number of rotor bars. In the analyzed case the induction machine has 28 rotor bars. Following the regular machine operation, FEM simulations of static eccentricity listed in the table I were performed. Fig. 6 shows induced voltage waveform in the measuring coils when rotor is displaced for 25 % of the air gap width in the + y axis direction. Fig. 6 shows comparison of the induced voltage in the measuring coils Ms1 and Ms3, and Fig. 7. comparison of the voltage in the measuring coils Ms2 and Ms4. Through observation of these figures, it can be noticed that the absolute value of voltage induced in the coil Ms1 is greater than the voltage induced in the coil Ms3. This result is expected, as the chosen rotor displacement direction places squirrel-cage rotor nearest to the coil Ms1, and farthest apart from the coil Ms3.

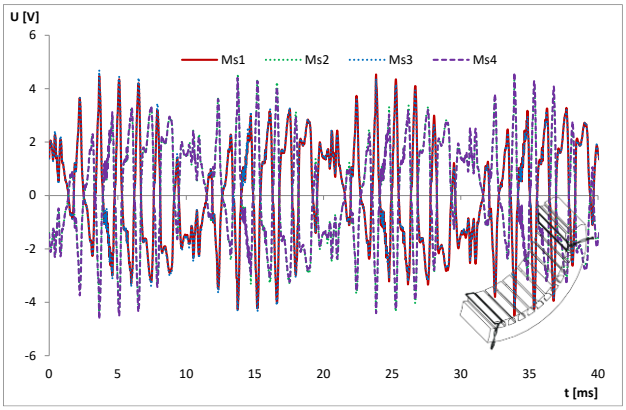


Figure 4. Induced voltage waveform in the measuring coils Ms1, Ms2, Ms3 and Ms4 – normal condition.

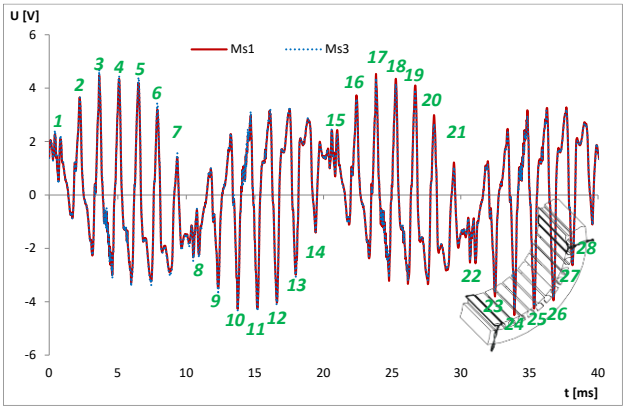


Figure 5. Impact of the squirrel-cage rotor bars on the induced voltage waveform in measuring coils.

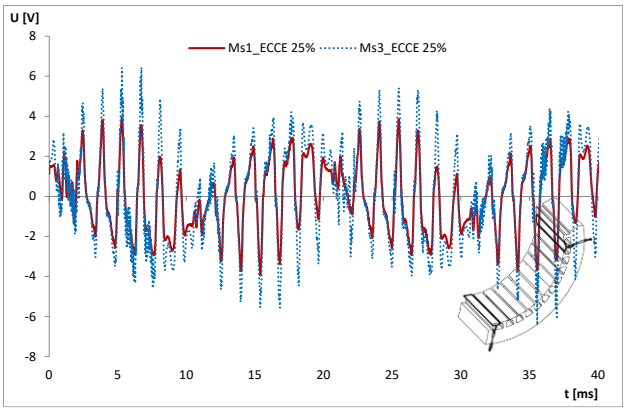


Figure 6. Comparison of the induced voltage waveform in the measuring coils for the static eccentricity amount 25% of the air gap (a shift in the + y axis): Ms1 and Ms3 diametrically spaced.

The comparison of the induced voltage in the measuring coils Ms2 and Ms4 is shown on Fig. 7. From this figure it can be concluded that the induced voltage in the mentioned measuring coils is equal. This result is expected as, in the case of the chosen rotor displacement direction, the relative position of the rotor to these two coils stays the same.

On Fig. 8 the voltage difference between two measuring coils, Ms1 and Ms3 spatially spaced for 2 τ_p is shown. The voltage difference is shown for the normal (blue waveform) and for the faulty (red waveform)

condition. The observed fault is a broken rotor bar of the squirrel-cage induction motor.

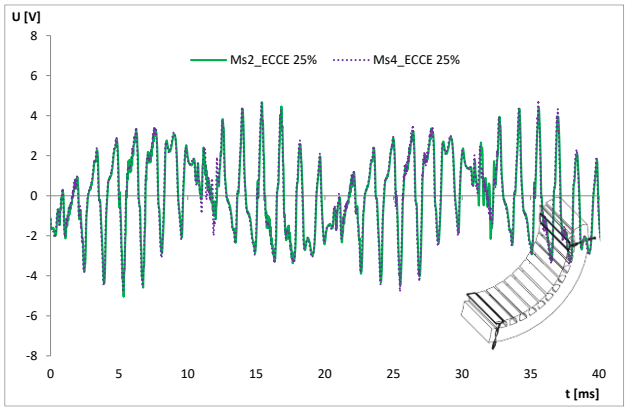


Figure 7. Comparison of the induced voltage waveform in the measuring coils for the static eccentricity amount 25% of the air gap (a shift in the + y axis): Ms2 and Ms4 diametrically spaced.

By observing the waveform of voltage difference for two measuring coils, presented in Fig. 8, the rotor winding damage is easy to detect. The measuring voltage is sensitive only to machine faults. In Fig. 8, winding damage is marked with voltage peek, which is repeated two times during a full turn in the case of one broken rotor bar. When the broken rotor bar comes across the measuring coil, one voltage peak appears. Therefore, in the measured waveform during one machine turn two extremely high voltage peaks will appear. The waveform is presented for the two machine turns, which in this case is 80 ms, because the analyzed induction motor is four-pole.

Accordingly, the number of such measured peaks of the measured voltage during one full machine turn, divided by two, gives the number of broken rotor bars.

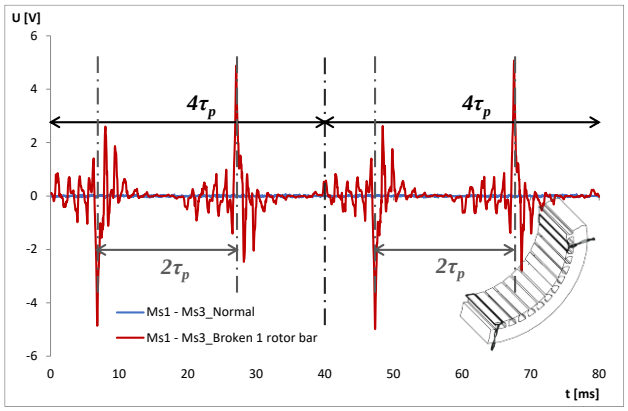


Figure 8. Voltage difference between measuring coils Ms1 and Ms3 for normal (blue waveform) and for one broken rotor bar (red waveform).

Fig. 9 presents waveform of the voltage difference for two measuring coils Ms1 and Ms3, installed around stator tooth without fault (blue waveform) and with two broken rotor bars (red waveform). By observing the voltage waveform obtained from two measuring coils and presented in Fig. 9, four extremely high voltage peaks can be seen for one machine turn. Therefore, the machine with this voltage waveform has two broken rotor bars.

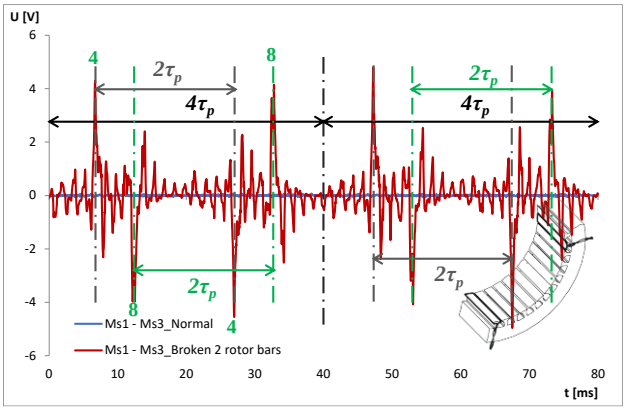


Figure 9. Voltage difference between measuring coils Ms1 and Ms3 for normal (blue curve) and for two broken rotor bar (red curve).

Furthermore, Fig. 10 presents waveform of the voltage difference for two measuring coils Ms1 and Ms3, installed around stator tooth without fault (blue waveform) and with broken rotor ring (red waveform). By observing the voltage waveform obtained from two measuring coils and presented in Fig. 10, extremely high voltage peaks can be seen for one machine turn. If you look in more detail, it can be seen that obtained waveform for the observed failure is different than in the case of broken rotor bars. The waveform is very similar to the waveform obtained for two broken bars. But, if you look closer it can be seen that the peaks are closer to each other and more pronounced (the higher voltage value). Therefore, it is possible to recognize the type and nature of the rotor fault. Based on these waveforms, obtained for different types of induction motor rotor failures, it was possible to design an algorithm for reliable and early detection and mutual recognition of induction motor faults.

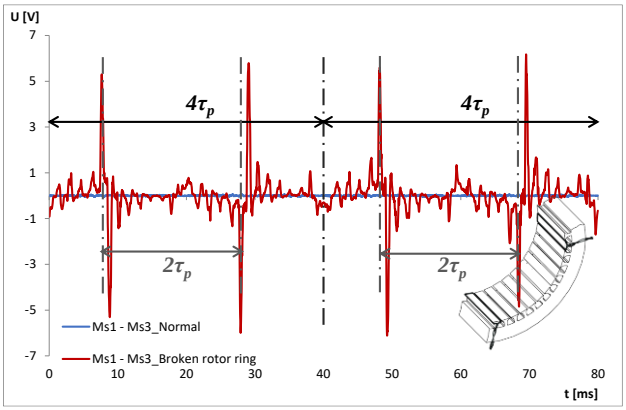


Figure 10. Voltage difference between measuring coils Ms1 and Ms3 for normal (blue curve) and for broken rotor ring (red curve).

This methodology is verified experimentally via series of laboratory tests performed on the real machines specially designed for fault study of broken rotor bars, broken ring and inter-coil short circuit in a rotor winding. Because of the complexity and extensiveness of the observed failures, measurement results are not shown within this paper.

III. ALGORITHM FOR MOTOR FAULTS RECOGNITION

Based on the FEM calculations and measurements, the functional algorithm for early fault detection and reliable recognition of rotor faults was designed. Fig. 11 presents and describes the flowchart of the proposed algorithm. Based on the obtained research results. Two special algorithms were designed. The input parameters for the proposed algorithm are induced voltages in four measuring coils spatially spaced for pole pitch. The first algorithm is intended for early and reliable detection of the broken rotor bars and the broken ring of the squirrel-cage induction motor. Based on the voltage difference obtained from induced voltages in measuring coils spatially spaced for $2\tau_p$, it is possible to detect and reliably recognize the fault of the squirrel-cage rotor. The algorithm mutually recognizes the broken bar and the broken ring of the rotor with a high degree of reliability. The second algorithm is proposed for early and reliable detection of rotor static eccentricity. The proposed algorithm uses induced voltage in four measuring coils. In the case of broken rotor bars and ring, the algorithm uses voltage difference between two induced voltages in measuring coils spatially spaced for $2\tau_p$. In the case of static eccentricity, the difference between the RMS value of each measuring coil installed in the air gap of the induction motor is analyzed. If there is a deviation between RMS values greater than the permissible value, the algorithm alarms the end user to the occurrence of a static eccentricity. Both algorithms for output results give graphical signalization with the corresponding message of the occurrence of each fault.

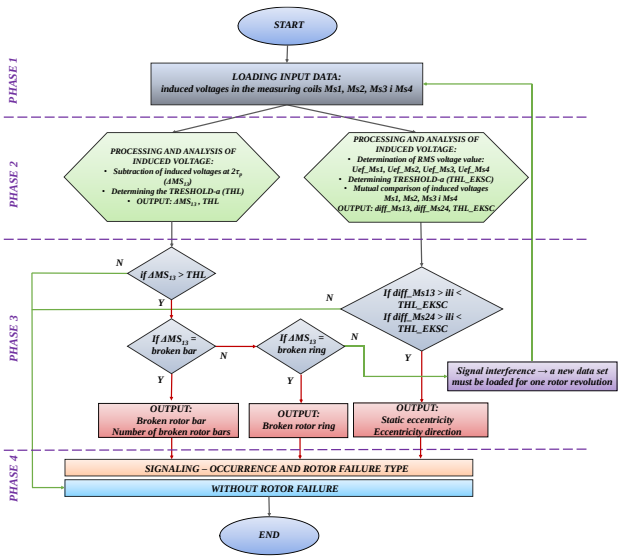


Figure 11. Flowchart of the smart and advanced algorithm for reliable faults detection and their mutual recognition.

Based on the input data (induced voltage in the measuring coils), using the function for determining characteristic peaks in the voltage waveform and voltage threshold, the algorithm detects and recognizes rotor faults of the induction motor. Respectively, Figures 12, 13, 14 and 15 shows algorithm output message for various conditions of the squirrel-cage rotor. In the case of static eccentricity, algorithms output message is type of fault and occurrence of static eccentricity in the direction of one of measuring coils. The output message in the case of static eccentricity occurrence is shown on Fig. 16 and 17, respectively.

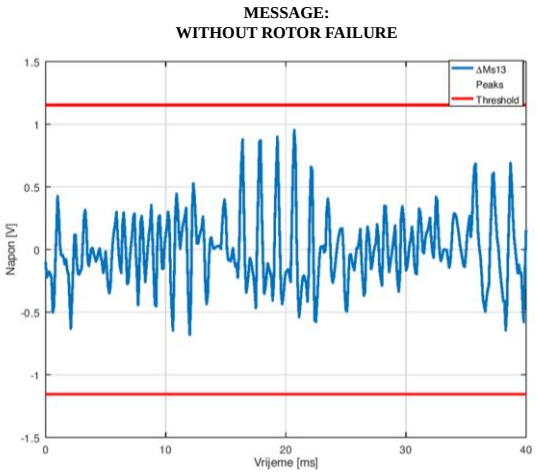


Figure 12. Algorithm output message – without rotor failure

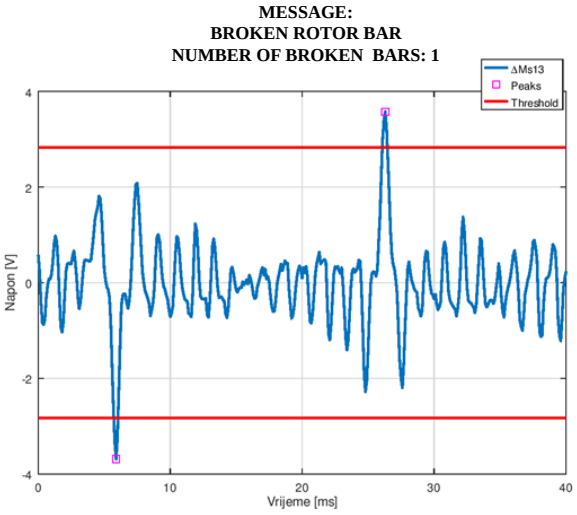


Figure 13. Algorithm output message – broken one rotor bar

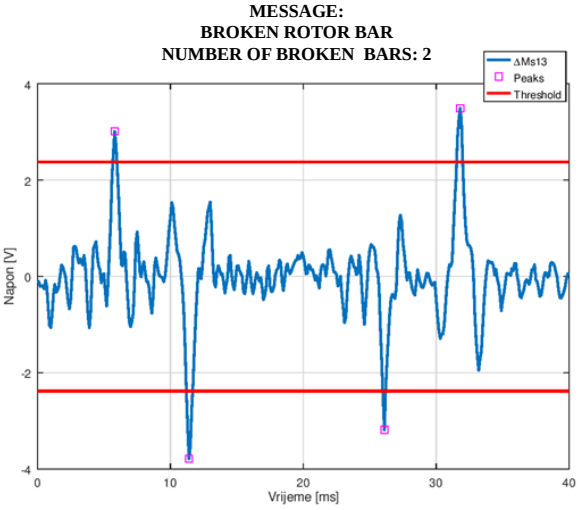


Figure 14. Algorithm output message – broken two rotor bars

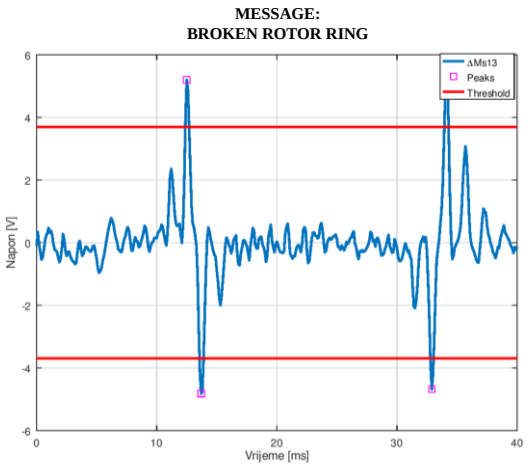


Figure 15. Algorithm output message – broken rotor ring

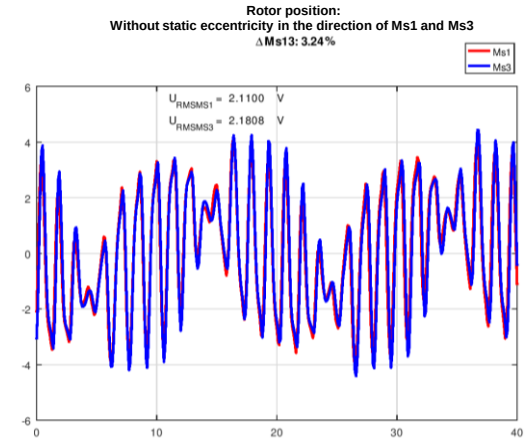


Figure 16. Algorithm output message – without the presence of static eccentricity

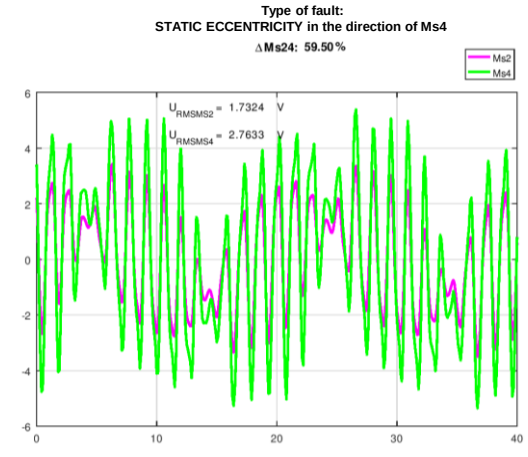


Figure 17. Algorithm output message – occurrence of static eccentriciy in the direction of the measuring coil Ms4

IV. CONCLUSION

The aim of this paper is to contribute to the early and reliable rotor eccentricity and rotor winding damages detection of induction machines based on the magnetic field analysis in the air gap. The two new and innovative approaches for rotor static eccentricity and rotor winding damages are presented. The novelty of this new methodology is magnetic field analysis obtained by measuring coils installed in the airgap of induction machine, and spatially spaced for pole pitch, τ_p. By FEM calculations, it is proven that by using the measuring coils on the stator teeth it is possible to detect static eccentricity occurrence or broken rotor bars, in an early stage. One of the main advantages of this methodology is that it avoids complex signal processing and the need for harmonic analysis of the measured signals. Determination of rotor static eccentricity is based on the changes in induced voltage waveform and RMS value of the measuring coil voltage. Determination of broken rotor bar and rotor ring is based on the voltage difference between two measuring coils spatially spaced for pole pitch, τ_p. The result of these simulations is an algorithm that is used for reliable and early rotor fault detection and recognition of the squirrel-cage induction motor and can be used for smart and advanced maintenance of induction motors.

REFERENCES

[1] S. Nandi, H.A. Toliyat, X. Li: "Condition Monitoring and Fault Diagnosis of Electrical Motors", IEEE transactions on energy conversion, Vol. 20, No. 4, December 2005.

[2] P. Tavner, L. Ran, J. Penman, H. Sedding: "Condition Monitoring of Rotating Electrical Machines", The Institution of Engineering and Technology, London, United Kingdom, 2008.

[3] O.V. Thorsen,M. Dalva: "A Survey of Faults on Induction Motors in Offshore Oil Industry, Petrochemical Industry, Gas Terminals, and Oil Refineries", IEEE transactions on industry applications, Vol. 31, No.5, September/October 1995.

[4] O.V. Thorsen, M. Dalva: "Failure Identification and Analysis for High Voltage Induction Motors in Petrochemical Industry", IEEE, 1998.

[5] M.L. Sin, W.L. Soong, N. Ertugrul: "Induction machine on-line monitoring and fault diagnosis – a survey", AUPEC 2003, Australasian Universities Power Engineering Conference, Christchurch, New Zealand, October 2003.T. Seppa "Fried Wire?" (Public Utilities Fortnightly, December 2003, pages 39-41)

[6] A. Siddique, G.S. Yadava, B. Singh: "A Review of Stator Fault Monitoring Techniques of Induction Motors", IEEE transactions on energy conversion, Vol. 20, No. 1, March 2005.Prospectiva del Sector Eléctrico 2002-2011. (Secretaría de Energía. Mexico, 2002).

[7] M.J. Picazo-Rodenas, R. Royo, J. Antonino-Daviu, J. Roger-Folch: "Use of Infrared thermography for computation of heating curves and preliminary failure detection in induction motors", ICEM 2012, XXth International Conference on Electrical Machines, Marseille, France, September 2012.

[8] M.E.H. Benbouzid: "A Review of Induction Motors Signature Analysis as a Medium for Faults Detection", IEEE transactions on industrial electronics, Vol. 47, No.5, October 2000.

[9] D.R. Albright, "Inter turn Short-circuit Detector for Turbine-Generator Rotor Windings", IEEE trans. on power apparatus and systems, Vol. PAS-90 Number 2, March/April 1971.

[10] M. Sasic, B.A. Lloyd, S.R. Campbell, "Flux Monitoring Improvement", IEEE industry applications magazine, Sept/Oct 2011.

[11] M. Šašić, L. Blake, A. Elez, "Finite Element Analysis of Turbine Generator Rotor Winding Shorted Turns", IEEE trans. on energy conversion. 27 (2012), 4; 930-937.

[12] A. Elez, S. Car, Z. Maljković, "A new method for inter-coil short circuit detection in synchronous machine armature winding", The international review of electrical engineering (IREE). Vol. 7, N. 6 (2012).

[13] H. C. Karmaker; - "Broken Damper Bar Detection Studies Using Flux Probe Measurements and Time-Stepping Finite Element Analysis For Salient – Pole Synchronous Machines", SDEMPED

2003, Symposium for Electric Machines, Atlanta USA August 2003.

[14] P. O'Donell: "Report of Large Motor Reliability Survey of Industrial and Commercial Installations Parts I, II and III", IEEE Transactions on Industry Applications, Motor Reliability Working Group, IEEE Industry Applications Society, July/August 1985, January/February 1987.

[15] B.M. Ebrahimi, J. Faiz, S. Lotfi-fard, P. Pillay: "Novel indices for broken rotor bars fault diagnosis in induction motors using wavelet transform", Elsevier ScienceDirect, Journal of Mechanical Systems and Signal Processing 30, 131-145, February 2012.

[16] C.C. Yeh, G. Sizov, A. Sayed-Ahmed, N.A.O. Demerdash, R.J. Povinelli, E.E. Yaz, D.M. Ionel: "A Reconfigurable Motor for Experimental Emulation of Stator Winding Interturn and Broken Bar Faults in Polyphase Induction Machines", IEEE transactions on energy conversion, Vol. 23, No. 4, December 2008.

[17] N.M. Elkasabgy, A.R. Eastham, G.E. Dawson: "Detection of Broken Bars in the Cage Rotor on an Induction Machine", IEEE transactions on industry applications, Vol. 28, No.1, January/February 1992.

[18] G. Didier, E. Ternisien, O. Caspary, H. Razika: "A new approach to detect broken rotor bars in induction machines by current spectrum analysis", Elsevier, Mechanical Systems and Signal Processing, Vol.21, 2007.

[19] Z. Ye, B. Wu,A. Sadeghian: "Current Signature Analysis of Induction Motor Mechanical Faults by Wavelet Packet Decomposition", IEEE transactions on industrial electronics, Vol. 50, No.6, December 2003.

[20] H. Douglas, P. Pillay, A. Ziarani: "Detection of broken rotor bars in induction motors using wavelet analysis", IEEE International Electric Machines and Drives Conference, IEMDC'03, June 2003.

[21] V.V.Thomas, K. Vasudevan, V.J. Kumar: "Online Cage Rotor Fault Detection Using Air-Gap Torque Spectra", IEEE transactions on energy conversion, Vol.18, No. 2, June 2003.

[22] W.T. Thomson, R.J. Gilmore: "Motor current signature analysis to detect faults in induction drives – fundamentals, data interpretation and industrial case histories", Proceedings of the thirty-second turbomachinery symposium, 2003.

[23] M.Y. Kaikaa, M. Hadjami: "Effects of the Simultaneous Presence of Static Eccentricity and Broken Rotor Bars on the Stator Current of Induction Machine", IEEE Transactions on Industrial Electronics, Vol.61, No.5, May 2014.

[24] S. Tvorić: "Rotor Faults Detection in Squirrel-Cage Induction Motor based on Magnetic Field Analysis in Air-Gap", doctoral thesis, Faculty of Electrical Engineering and Computing, Zagreb 2019.

[25] A. Elez, S. Car, S. Tvorić, B. Vaseghi: "Rotor cage and winding fault detection based on machine differential magnetic field measurement (DMFM)", IEEE Transactions on Industry Applications, 53(3):3156–3163, 2016.

Analytics Use Cases for Landside Traffic Optimization in the Catchment Area of the Airport: Case Study of Zagreb Airport

Marin Tica*, Igor Štimac**, Krešimir Vidović***, Saša Vojvodić*** and Irina Stipanović****
*International Zagreb Airport Jsc., Zagreb, Croatia
**Zagreb Airport Ltd., Zagreb, Croatia
***Ericsson Nikola Tesla, Zagreb, Croatia
****Infraplan, Zagreb, Croatia
kresimir.vidovic@ericsson.com

Abstract **- Airports and their facilities are general attractors that significantly influence the supply and demand of landside traffic in the adjacent and larger catchment areas of the airport. This paper will draft a concept of a decision support solution for strategic planning that will enhance, optimize and significantly improve the management of land transportation-related activities for catchment areas of the airport. Following the principles of FRAME Next ITS architecture creation process, data-driven analytical use cases will be identified and proposed that utilize commonly available data sources in the airport environment (e.g. anonymized telecom network data, public transport data, flight information data, airport parking data, transport on-demand data, etc.) and identify and prioritize real analytical use cases of importance to key stakeholders (e.g. airlines, local community, public transport operators, etc.). Proposed analytical use cases are demonstrated for Franjo Tuđman Airport in Zagreb, Croatia.**

Keywords**:** ***airport landside traffic, optimization, data analytics***

I. INTRODUCTION

Airports are in general known for having a significant impact on both the environment and public health and are, therefore, subject of improvement so that they can meet nominated Sustainable Development Goals of the United Nations regarding environmental efficiency. Specific actions are required with the purpose of increasing performance from these perspectives, and one of the most significant ones is related to the improvement of ground transport in catchment areas of the airport. The goal is to foster the use of sustainable and green multimodal solutions for all journeys related to the airport (employees, passengers, supporting services, etc.) in order to decrease the direct environmental footprint of the airport in the area of ground transportation. Therefore, this research will propose a decision-making system concept that will utilize available common (traffic counter, airline data, etc.) and advanced data sources (anonymized mobile network data) that will be used as input for defined analytical use cases. Analytical use cases will be identified in order to target the real problems and needs from relevant stakeholders influenced by ground transport.

According to literature review, the existing researches have identified that both conventional and emerging data sources can be used as a base for quality data-driven decision-making in the area of urban mobility management in small/medium-sized cities. Also, a software-based decision-making system for urban mobility management can be used as a tool in Intelligent Transport System (ITS) applications, where it, besides data requirements, usually consists of a data aggregation platform and relevant applications as a key building blocks. Functional requirements on those subsystems are identified especially in priority areas which are related to passenger information, weather and environmental monitoring, traffic management and operations, people's safety in the transport system, public urban transport and electronic transport-related payments. [1]–[3].

Besides the introduction in first chapter, the paper will address the process of identification and core use cases in Chapter 2. Chapter 3 will provide overview of data sources used in this research, Chapter 4 will present the decision support system architecture, Chapter 5 will present the proposed methodology, and Chapter 6 will include conclusion remarks.

II. USE CASES IDENTIFICATION

Since the goal of the research is to develop a decision support solution for strategic planning that will enhance, optimize and significantly improve the management of activities related to land transportation for the catchment area of the airport, relevant stakeholders were involved in order to identify targeted areas of improvements. Since optimization measures will be proposed in the area of Intelligent Transport Systems (ITS), the methodology for use case development is based on the combination of Delphi method and utilization of FrameNext ITS architecture (FRAME). An ITS architecture is a conceptual design that defines the structure and/or behavior of an integrated Intelligent Transport System (ITS). The Delphi method is a process used to arrive at a

group opinion or decision by surveying a panel of experts [4]. The FRAME Architecture comprises top level requirements and functionality, or use cases, for almost all ITS applications and services that have been considered for implementation somewhere in the European Union. It is at a "level" such that it can be used as a reference by all ITS architects and is intended to be the foundation for building other types of architecture that will be necessary. It will enable them to guarantee compliance at the interfaces of other systems so that seamless services can be provided to cross-border travelers, and an open European market of compatible components can be established. [5], [6] Therefore, the Delphi methodology has been used to extract knowledge from relevant experts nominated from key stakeholders, and FRAME architecture was used to shape the user requirements in the matter to fit the standardized requirements. During a series of workshops, a number of stakeholders (12) were interviewed so as to identify key issues of landside transportation in the catchment areas of the airport. Relevant stakeholders involved representatives of neighboring cities, representatives of the airport and connected companies, representatives of public transport companies operating in the area, and they were nominated in order to fullfill the FRAME next principles, according to which, all relevant stakeholders' groups have to be represented (stakeholders that want ITS, use ITS, rule ITS, make ITS and service providers). The entire process consisted of five workshops, which resulted in the identification of analytical use cases proposed for land transport optimization in narrower airport area. Following the identification of analytical use cases that will be addressed in this research (user aspirations), research team has proposed a research methodology, required data sources and infrastructure, and has consolidated them in a consistent manner that is suitable for the next stage in the process (user needs). The consolidated requirements where then presented to stakeholders, and after a common agreement has been achieved, use cases have been prioritized and completed. These use cases are set up as key requirements for the decision-making system. The use cases are the following:

The first use case is public transport optimization for airport users (passengers) with the goal of optimization of existing public transport lines or introduction of new public transport lines in a narrower catchment area of the airport (Zagreb, Velika Gorica) serving airport passengers.

The second use case is public transport optimization for airport users (employees) with the goal of optimization of existing public transport lines or introduction of new public transport lines in a narrower airport catchment area (City of Zagreb, City of Velika Gorica) serving airport employees.

The third use case is related to airport strategic planning and includes an analysis of catchment areas of the airport based on the analysis of users migrations where either trip origins and/or destinations are at the airport, with the goal of identification of real catchment areas, enhancement of sustainable transport connections for these areas, analysis of potential areas which are not, but are expected to be, served by the analyzed airport.

The fourth use case is related to airline strategic planning (analysis of airline users). It consists of an analysis of airline users based on the information of the initial country of origin/destination in order to identify a passenger's country of origin based on telecom data and to identify countries that have a significant travel demand, but have not initially been served by airline companies.

The fifth use case is based on the analysis of migration and retention habits of passengers that include analysis of migration and retention habits of passengers, identification of passenger type based on its migration patterns (tourists, business users) and optimization of service/offering for targeted users.

The sixth use case is based on the analysis of transport demand prediction based on the information from airport arrivals/departures and expected utilization of public transport services serving the airport area, including providers of on-demand transport. Based on historical data, migration patterns of airport passengers following the arrival of an aircraft (e.g. the number of passengers leaving airport with a provider of on-demand transport or public transport), the demand will be predicted for a specific transport mode/service (e.g. on-demand transport, rent-a-car, etc.)

The next step in the research is to propose a system architecture and data sources that will enable realization of identified use cases.

III. IDENTIFIED DATA AND CORRESPODING DATA SOURCES USED FOR DECISION-MAKING

The system relies on a number of data sources used as input data for decision-making. These data sources include transport data, public transport data, anonymized mobile network data, airline data and other data. Since anonymized mobile operator data have a great potential and are still not widely used, most of this chapter will be devoted to them.

A. Transport data

Traffic data includes data on traffic flow parameters as well as data on traffic situations. The Datex standard is used to exchange traffic data between traffic control, information centers, traffic operators and other stakeholders. Datex allows the collection of data on traffic accidents, current roadwork and other traffic-related events. The data is presented in XML format and modeled with a UML diagram. The structure of the message includes a publication from the table of measurement locations and the publication of measured data. The data source includes the following data: vehicle type and vehicle flow rate. Individual Vehicle Data includes the data on the arrival and exit road lane, vehicle speed, direction. Moreover, situation data is available as well. Situation data includes information on incidents, like improper turn, a vehicle went through red light, a vehicle moving in the wrong way, a vehicle improperly stopped or parked, and other events, such as accidents. In this research, vehicle flow data will be mostly utilized.

B. Public transport data

Public transport data is exchanged via NeTEx/GTFS or Siri (excluded from this research due to the unavailability of data in this format). NeTEx provides a means to exchange data for passenger information, such as stops, routes, timetables and fares, among different computer systems, together with related operational data. It can be used to collect and integrate data from many different stakeholders, and to reintegrate it as it evolves through successive versions. NeTEx provides a breakdown of transport modes by types: scheduled (air, rail including high-speed rail, conventional rail, light rail, long-distance coach, maritime including ferry, metro, tram, bus, trolleybus), demand-responsive (shuttle bus, shuttle ferry, taxi, car-sharing, car-pooling, car-hire, bike-sharing, bike-hire) and personal (car, motorcycle, bicycle.) In this research, data is obtained using NeTEx protocol from National Access Point for Multimodal data and the following is used: information on public transport lines, stops and timetables.

C. Anonymized mobile network data

Commonly used traffic sensors within the road network are very precise in performing traditional measurements (traffic flow characteristic on a road section). However, since the requirements from the road authorities and operators might surpass that level of data or be oriented on those road sections where the sensor infrastructure does not exist, alternative methods have been studied for years [7]–[11]. One of the most researched data sources, with the highest potential, is mobile network operator data, with an increased rate of popularity as a result of increased penetration of mobile phone users (market penetration in developed countries is more than 100%), and a result of introduction of new generations of mobile networks (3G, 4G, 5G). Every new generation brings significant improvements in telecommunication service and, as a result, a more useful accompanying databases generated because of network operations. Furthermore, every new generation brings improvement in user positioning as well, and since it is still not as precise as Global Navigation Satellite System (GNSS) systems, this data can still be very useful in transport engineering.

The most commonly used data sources from mobile network data include Call/Charging Data Records (CDR) data, mobile network signalization data logs (CTUM, CTR, etc.), mobile network subscriber database data, and performance counters data.

Call/Charging Data Records (CDR) data represents a dataset that is collected because of user telecommunication activities, since every time a user performs a telecommunication activity, the activity is logged in this database for future charging purposes. Mobile network signalization data consists in majority of cell trace data, where communication between mobile network terminal and corresponding antenna is logged. An advantage of these data sources is that the logs are populated even if mobile phone is not being actively used (in call, data usage, etc.). These two data sources are user-based, and required processes, organization measures and anonymization techniques have to be applied which ensure that data is handled in accordance with the General Data Protection Regulation (EU) 2016/679 (GDPR), which regulates data protection and privacy in the European Union (EU) and the European Economic Area (EEA). The use of this data enables anonymized reconstruction of user migrations, and therefore user trips with additional parameters (origins, destinations, duration, length, speed) can be extracted. If fused with anonymized user database data (CRM – Customer Relationship Management), additional data can be applied to extracted trip database, and this data can include information on user sex and age range. That enables various kinds of analytics for identification of habits: identification of travel patterns for age group and sex, identification of average speed, identification of the type of users which obey speed limits, etc. As opposed to previously mentioned data sources, performance counters data are not user-oriented; rather, they are location-oriented. LTE/E-UTRA performance management (PM) counters are aggregated over 15-minute intervals. These counters are inherently privacy-preserving, since they can represent the number of terminals being served by a corresponding cell and can be used as virtual traffic counters.

Telecom data sources can be utilized for various transport-related analyses. Signaling data, CDR data and subscriber data can be used for both strategic and operational purposes for the identification of traffic flow characteristics (traffic volume, speed), for driving distraction use cases (analysis of mobile phone usage by drivers), for the identification of transport demand (origin and destination data). Performance counters data can be utilized for highly efficient operational virtual traffic counters, for supplementing locations where traditional traffic counters do not exist.

In this research, this data source will be used to identify user types and migration patterns (more details in the Methodology chapter).

D. Airport data

Airport data contains information on aircraft arrivals and departures. The source of data is the airport management system. This data source provides information on real-time aircraft arrivals and departures, including information on air carrier data, scheduled time data of arrival and departure, estimated time data, stopover data, country data, airport data, flight code – airline code and flight number, gate data/exit data, aircraft stand data and status data (arrived/departed). In this research, all this data will be used in order to match landside transport demand and its correlation with aircraft arrivals and departures.

E. Parking data

Parking data contains information on the occupancy and utilization of the airport parking areas. The data source is the parking management system and due to the lack of standardization in this area, the data is not standardized and is available in a proprietary format. The data contains the following information: entry and exit time, user type (for taxi vehicles), entry and exit point, etc. This data source will be used to identify specific types of airport users (e.g. taxi) to validate transport mode data obtained from anonymized mobile network data.

F. *Other*

Other data sources might include general and well-known data, such as statistical data on the number of employees at the airport, census data on user migrations, data from transport models, etc. This data will be used to validate results of analyses from different data sources (like anonymized mobile network data).

IV. DECISION SUPPORT SYSTEM ARCHITECTURE

The proposed decision support system architecture is presented on Figure 1. It consists of a four-level architecture. The first layer consists of available and required data sources. These data sources are integrated with the second layer—the data aggregation platform using a set of application programming interfaces. Where applicable, standardized protocols should be used to ensure compatibility of the solution. Therefore, transport and public transport data will be supported by using standardized protocols (Datex, Netex, GTFS, SIRI). Other data sources are not covered with standardization, so anonymized telecom data, parking data, airport data and other data sources will be integrated using proprietary data formats.

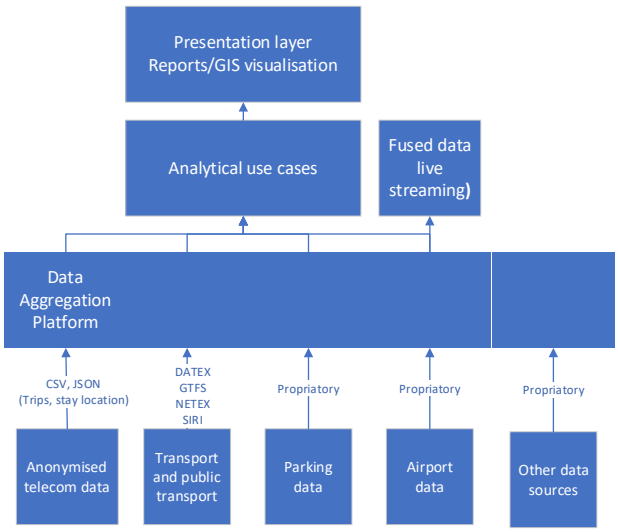


Figure 1. System architecture

The data is then stored in a data aggregation platform. It includes a three-layer architecture, where the southern interface is used for acquisition of data (input interface), the middle layer is used for operations with data, and the third layer—the northern interface—is used for publication of data toward users/applications (output interface).[12]

In the application layer, applications will be used to perform analytical use cases identified in this project. Besides an analytical engine, an interface will be ensured that enables live streaming of fused data entering into the platform. Above these layers, data presentation layers will be established that will enable generation of reports and visual presentation of data on an interactive web GIS tool.

V. PROPOSED METHODOLOGY FOR IDENTIFIED ANALYTICAL USE CASES

Following the identification of available data sources and the decision support system architecture, a methodology has to be proposed in order to fulfill the use case requirements using available data sources. The methodology is presented on Fig. 2.

Anonymized telecom data is used for identification of all users stay locations, which are then transformed into user migrations (trips). The output is a table of stay locations per anonymized user with the information on time spent in a specific geographical area, together with a list of trips of all users that will include information on the trip start and end location, time of the beginning and end of a trip and information about distance between the starting and the last point of the trip. Based on this data, the first step includes identification of different user types. The following categories are identified: airport employee, airport user—passenger— arriving by aircraft airport user—passenger—departing by aircraft airport user (greeter/byer), airport user taxi driver, person passing near the airport—local residents—and commuters. All these user types can be identified based on their characteristic behavior following criteria such as: identification of Home/Work/Other sector, retention time at airport sector, trip characteristics, user type (domestic or foreign), time of the day, etc.

For example, a candidate for an airport employee might be a person whose estimated home sector is not the airport but spends significant time in their day at the airport sector. An airport employee should have a pair of trips originating and departing the airport. These trips should occur during airport working hours. On the other hand, a taxi driver should have a significant number of trips originating and departing at the airport and retention time at the airport should not be significant. A passenger arriving at the airport should have first occurrence at the airport sector, passenger leaving the airport should have last occurrence at the airport sector, etc. After identification, all these user types can be validated using various data sources, as presented on Fig. 2. Following this step, all trips in the trip database can be assigned to the appropriate user type, and therefore, for each use case only trips relevant to a specific user type can be considered.

As previously mentioned, based on trip characteristics (origin, destination, duration, speed and distance), transport mode can be identified. For the purpose of this research, a methodology has been applied that has differentiated two most common transport modes used to arrive to the airport—personal vehicles and public transport. The proposed methodology is based on a set of rules and application of Bayes statistical modelling, as presented in [13]. Following this step, all trips that include identified user types are then assigned with an appropriate transport mode. Transport mode data can be validated using supporting datasets from the airport parking, and data from traffic counters on road network in the airport area.

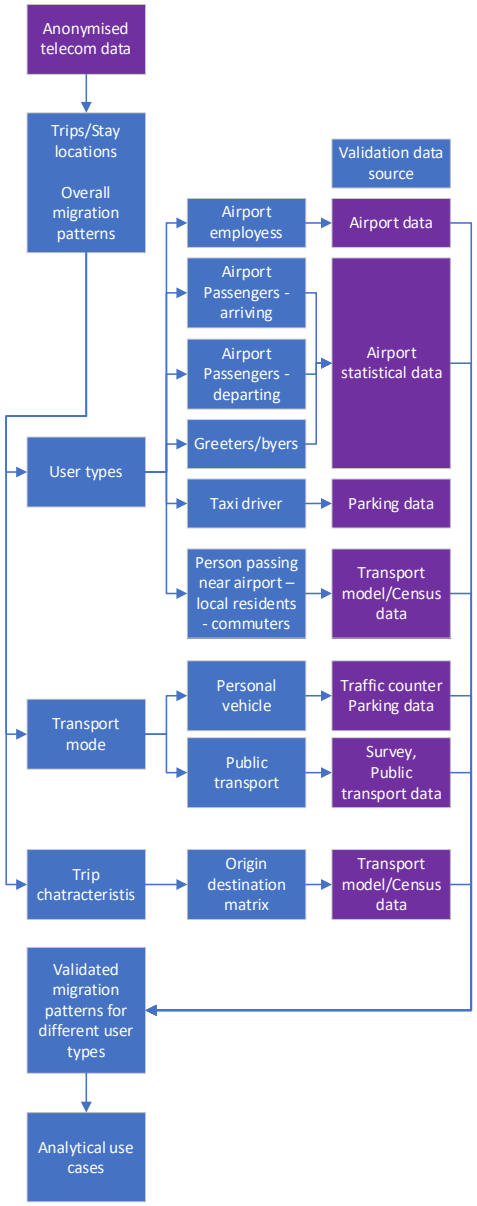


Figure 2. Methodology

At the end of this process, validated migration patterns have been generated and stored into an appropriate database, and can, therefore, serve previously identified use cases. Validation data sources have been proposed and considered as a ground truth data for each data source and each measured/identified parameter. At the end of analytical process, the analytical results will be validated using ground truth data in order to identify results inaccuracy.

VI. Conclusion

In order to meet health and environmental goals for the improvement of airport-related ground transportation activities, a decision support system concept has been introduced in this paper. Relevant stakeholders have identified most common ground transport-related issues, and analytical use cases have been proposed in order to quantify these values and become the base for optimization activities. Available relevant data sources have been utilized to identify target airport users (employees, passengers, supporting services operator, etc.) and to identify what their real travel needs and migration patterns are and what real airport catchment area is. These objectively measured values represent a real snapshot of airport ground transportation reach using origin destination matrixes for these users, together with the transport mode that they use when performing these travels. The proposed methodology can be used for other analytical use cases in various research domains besides transport (tourism, economy…) that are based on spatio temporal analysis of population migration. Next steps of the research will focus on the bare results of this proposed concept and will focus on optimization measures that will be introduced in order to fulfill the sustainability related goals.

Acknowledgment

This paper presents the results of activities performed within the OLGA project. OLGA project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement № 101036871.

References

[1] Medved, D., D. Blažinić, V. Galijan, & N. Antolović, Evolution of data sources for integrated data-driven urban mobility management, Transp. Res. Procedia, vol. 64, pp. 68–75, 2022.

[2] Draganić, M. et al., Requirements on applications within the decision-making system for urban mobility management, Transp. Res. Procedia, vol. 64, pp. 60–67, 2022.

[3] Radović, R., I. Marasović, V. Čačković, D. Pleština, D. Keresteny, & Z. Anić, The concept of a data aggregation platform in the function of a decision-making system for urban mobility management, Transp. Res. Procedia, vol. 64, pp. 53–59, 2022.

[4] Pérez, V. L. & R. Schüler, The Delphi Method as a tool for information requirements specification, Inf. Manag., vol. 5, no. 3, pp. 157–167, 1982.

[5] Bossom, R. & P. Jesty, Extend FRAMEwork architecture for

cooperative systems - D15 – FRAME Architecture – Part 1: Overview, 2011.

[6] Jesty, P. H. & R. A. P. Bossom, Using the FRAME Architecture for planning integrated Intelligent Transport Systems, 2011 IEEE Forum Integr. Sustain. Transp. Syst. FISTS 2011, pp. 370–375, 2011.

[7] Galloni, A., B. Horváth, & T. Horváth, Real-time Monitoring of Hungarian Highway Traffic from Cell Phone Network Data, in ITAT 2018 Proceedings, 2018, vol. 2203, pp. 108–115.

[8] Kos, G., P. Brlek, & K. Vidovic, SHARED SPACE CONCEPT IN LOCAL COMMUNITIES: CASE STUDY, in 8TH INTERNATIONAL CONFERENCE ON ROAD SAFETY IN LOCAL COMMUNITY, 2013, pp. 1–8.

[9] Goulding, J., Best Practices and Methodology for OD Matrix Creation from CDR Data, NLAB, Univ. Nottingham, vol. 44, no. 0, pp. 1–37, 2018.

[10] Furletti, B. et al., Use of mobile phone data to estimate mobility flows. Measuring urban population and inter-city mobility using big data in an integrated approach, Sis, pp. 1–10, 2014.

[11] Yaghoubi, F., A. Catovic, A. Gusmao, J. Pieczkowski, & P. Boros, Traffic Flow Estimation using Machine Learning and 4G/5G Radio Frequency Counters, vol. 2022-June, no. 1. Association for Computing Machinery, 2022.

[12] Radović, R., I. Marasović, V. Čačković, D. Pleština, D. Keresteny, & Z. Anić, The concept of a data aggregation platform in the function of a decision-making system for urban mobility management, in Transportation Research Procedia, 2022, vol. 64, pp. 53–59.

[13] Vidović, K., P. Čolić, S. Vojvodić, & A. Blavicki, Methodology for public transport mode detection using telecom big data sets: case study in Croatia, Transp. Res. Procedia, vol. 64, pp. 76–83, 2022.

MEET

International Conference on

MICROELECTRONICS, ELECTRONICS AND ELECTRONIC TECHNOLOGY

Steering Committee

Chairs: Marko Koričić, University of Zagreb, Croatia
Željko Butković, University of Zagreb, Croatia

Members: Slavko Amon, University of Ljubljana, Slovenia
Dubravko Babić, University of Zagreb, Croatia
Maurizio Ferrari, CNR-IFN, Povo-Trento, Italy
Mile Ivanda, Ruđer Bošković Institute, Zagreb, Croatia
Tihomir Knežević, Ruđer Bošković Institute, Zagreb, Croatia
Branimir Pejčinović, Portland State University, United States
Mirko Poljak, University of Zagreb, Croatia
Jörg Schulze, Fraunhofer Institute for Integrated Systems and Device Technologies, Erlangen, Germany
Tomislav Suligoj, University of Zagreb, Croatia
Aleksandar Szabo, IEEE Croatia Section, Zagreb, Croatia
Davor Vinko, University of Osijek, Croatia

PAPERS

Back-Hopping in Ultra-Scaled MRAM Cells

M. Bendra*,**, S. Fiorentini*,**, J. Ender*,**, R.L. de Orio**, T. Hadámek*,
N.P. Jørstad*, B. Pruckner*, S. Selberherr**, W. Goes†, and V. Sverdlov*,**
* Christian Doppler Laboratory for Nonvolatile Magnetoresistive Memory and Logic at the
** Institute for Microelectronics, TU Wien, Gußhausstraße 27-29, A-1040 Wien, Austria
† Silvaco Europe Ltd., Compass Point, St Ives, Cambridge, PE27 5JL, United Kingdom
e-mail: bendra@iue.tuwien.ac.at

***Abstract*—The development of advanced magnetic tunnel junctions with a footprint in the single-digit nanometer range can be achieved using an elongated multilayer ferromagnetic free layer structure. Using the spin drift-diffusion model, we investigated the back-hopping effect in ultra-scaled STT-MRAM devices. Unwanted switching of the middle-layer structure has been identified as a possible cause of the back-hopping effect, which leads to a writing error in the magnetization state of the free layers. To understand the switching of the free layer, the torque acting on both parts of the composite free layer is studied in detail. A reduction in the size of MRAM components to increase the memory density leads to lower anisotropies and thus increases the likelihood of back-hopping due to the presented mechanism. A possible solution to avoid erroneous switching is to increase the magnetic anisotropy of the layers.**

Keywords*—*Back-Hopping, Spin Transfer Torques, Ultra-Scaled MRAM Cells, Perpendicular Magnetic Anisotropy, Writing Error

I. Introduction

Emerging nonvolatile spin-transfer torque (STT) magnetoresistive random access memories (MRAM) have generated significant interest because of their high speed, and endurance. Thus, they are attractive for Computing-in-Memory concepts [1], embedded automotive [2], low-latency industrial applications [3], last-level cache [4], and high density MRAM arrays [5]. An STT-MRAM cell is based on a magnetic tunnel junction (MTJ) with perpendicular magnetic anisotropy (PMA) and is composed of several layers, including a CoFeB reference layer (RL) and a free magnetic layer (FL) separated by an MgO tunnel barrier (TB), which form the MTJ. To increase the PMA, the FL, typically composed of two CoFeB layers and a thin metal buffer, is interfaced with a second MgO layer [6]. Introducing more MgO layers [7] and elongating the FL allows to boost the perpendicular anisotropy while reducing the FL diameter [8]. A further reduction in the size of MRAM components to increase the memory density leads to lower anisotropies and thus increases the likelihood of erroneous switching. At high currents running through a cell the erroneous switching can occur due to an unwanted flip of the RL, or back hopping [9]. For the accurate design of ultra-scaled MRAM cells (Fig. 1) it is paramount to understand the magnetization dynamics of magnetic materials related to the performance of MTJs.

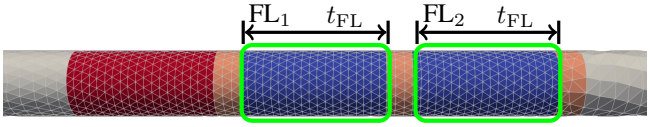


Figure 1: Simplified mesh of the simulated ultra-scaled MRAM cell. This composite structure consists of a CoFeB (5)|MgO (0.9)|CoFeB (5)|MgO (0.9)|CoFeB (5)|MgO (0.9) MTJ connected to normal metal contacts (50), where the numbers in parentheses indicate the thickness of each layer in nanometer. The diameter is $2.3\,\mathrm{nm}$. A bias of $1.5\,\mathrm{V}$ is applied across the structure. The color coding represents the following: red for the RL, blue for the FL, orange for the TB, and gray for the non-magnetic contacts. The green framed section denotes the FL segments for FL_1 and FL_2, with the length t_{FL}.

II. Micromagnetics Model

We developed a fully three-dimensional finite element method (FEM) based modeling and simulation approach which includes all relevant physical phenomena, (described in equations (1)-(6)) [10], responsible for proper ultra-scaled MRAM operation. We numerically solve the magnetization dynamics, described by the Landau-Lifshitz-Gilbert equation (LLG) (1).

$$\frac{\partial \mathbf{m}}{\partial t} = -\gamma \mathbf{m} \times \mathbf{H}_{\text{eff}} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \frac{1}{M_S} \mathbf{T}_{\mathbf{S}} \quad (1)$$

The effective field $\mathbf{H}_{\text{eff}}$ accounts for the magnetic anisotropy field, the exchange field, as well as the demagnetization field. The latter contribution to $\mathbf{H}_{\text{eff}}$ is evaluated only on the disconnected magnetic domain by using a hybrid approach combining the boundary element method and the finite element method [11]. To appropriately model the switching of ultra-scaled MRAM cells, the torque $\mathbf{T}_{\mathbf{S}}$ is described by the following equation:

$$\mathbf{T}_{\mathbf{S}} = -\frac{D_e}{\lambda_J^2} \mathbf{m} \times \mathbf{S} - \frac{D_e}{\lambda_\varphi^2} \mathbf{m} \times (\mathbf{m} \times \mathbf{S}). \quad (2)$$

The coupled spin and charge drift-diffusion method, as defined by equations (3)-(5), accurately describes the charge and the spin transport through a nanometer-sized magnetic valve [12].

$$D_e \left(\frac{\mathbf{S}}{\lambda_{sf}^2} + \frac{\mathbf{S} \times \mathbf{m}}{\lambda_J^2} + \frac{\mathbf{m} \times (\mathbf{S} \times \mathbf{m})}{\lambda_\varphi^2} \right) = -\boldsymbol{\nabla} \cdot \mathbf{J}_{\mathbf{S}} \quad (3)$$

$$\mathbf{J}_{\mathbf{S}} = -\frac{\mu_B}{e} \beta_\sigma \big(\mathbf{J}_{\mathbf{C}} \otimes \mathbf{m} + \\ + \beta_D D_e \frac{e}{\mu_B} [(\boldsymbol{\nabla} \mathbf{S}) \mathbf{m}] \otimes \mathbf{m} \big) - D_e \boldsymbol{\nabla} \mathbf{S} \quad (4)$$

Table I: Simulation Parameters.

Parameter	Value
Gilbert damping, α	0.015
Gyromagnetic ratio, γ	$1.76 \cdot 10^{11}\ \frac{\mathrm{rad}}{\mathrm{s \cdot T}}$
Saturation magnetization, M_S	$1.2 \cdot 10^{6}\ \frac{\mathrm{A}}{\mathrm{m}}$
Exchange constant, A	$1 \cdot 10^{-11}\ \frac{\mathrm{J}}{\mathrm{m}}$
Free layer length, t_{FL}	X nm
Interface anisotropy, K_i	$1.53 \cdot 10^{-3}\ \frac{\mathrm{J}}{\mathrm{m}^2}$
Uniaxial anisotropy constant, K	$\mathrm{K}_i / t_{\mathrm{FL}}$
Current spin polarization, β_σ	0.7
Diffusion spin polarization, β_D	1.0
Electron diffusion coefficient, D_e	$2 \cdot 10^{-2}\ \frac{\mathrm{m}^2}{\mathrm{s}}$
Spin-flip length, λ_{sf}	$10\,\mathrm{nm}$
Exchange length, λ_J	$1\,\mathrm{nm}$
Spin dephasing length, λ_φ	$0.4\,\mathrm{nm}$
Electrical resistance in the anti-parallel state, $\mathrm{R_{AP}}$	$750\,\mathrm{k\Omega}$
Electrical resistance in the parallel state, $\mathrm{R_P}$	$410\,\mathrm{k\Omega}$

$$\mathbf{J_C} = \sigma \mathbf{E} - \beta_D D_e \frac{e}{\mu_B} [(\boldsymbol{\nabla} \mathbf{S}) \mathbf{m}] \tag{5}$$

We extended this method to MTJs by modeling the TB as a poor conductor (6) with a local resistance dependent on the relative orientation of the FL magnetization [10].

$$\sigma(\theta) = \frac{\sigma_P + \sigma_{AP}}{2} \left(1 + \left(\frac{\mathrm{TMR}}{2 + \mathrm{TMR}} \right) \cos\theta \right) \tag{6}$$

$\sigma_{P(AP)}$ is the conductivity in the parallel (anti-parallel) state, θ is the angle of the FL direction with respect to the RL one. The tunneling magnetoresistance $\mathrm{TMR} = {G_P - G_{AP}}/{G_{AP}}$, where $G_{P(AP)}$ is the conductance in the parallel (anti-parallel) state. The tunneling spin current polarization taken from [10] is introduced by applying appropriate magnetization dependent boundary conditions at the left and right interface of the tunneling layers.

III. Results

The following sections report the results of switching simulations performed in the structure depicted in Fig. 1. The parameters employed are presented in Table I. They are consistent with CoFeB and MgO for the FM layers and TB layers, respectively. The low values of λ_J and λ_φ are employed to have complete absorption of the transverse spin accumulation components in the FL near the TB interface [10]. The tunneling spin current polarization creates a jump between the values of the spin accumulation at the left and right interface of the TB. This is the manifestation of the MTJ polarization effects on the spin current. The transverse spin accumulation is quickly absorbed, resulting in torques that are acting near the TB|FL interface.

In Fig. 2 the magnetization trajectories for the switching from anti-parallel (AP) to parallel (P) and from P to AP of an ultra-scaled MRAM with different FL configurations are displayed. The FL configuration is either symmetric or antisymmetric. The parts of the FL can be equal or different. In Fig. 2 bold lines represent a symmetric configuration, for dashed lines with triangles, the first FL_1 length is fixed with $5\,\mathrm{nm}$ and the second FL_2 is varied to be $2\,\mathrm{nm}$, $3\,\mathrm{nm}$, and $4\,\mathrm{nm}$, respectively. For dotted lines with crosses, the reverse is the case: FL_2 is fixed with $5\,\mathrm{nm}$, and FL_1 is varied. The applied bias is $1.5\,\mathrm{V}$ or $-1.5\,\mathrm{V}$ depending on the magnetization switching direction, P to AP, or AP to P. Fig. 2(a) reports, that the successful switching of the FL from AP to P is always obtained. Fig. 2(b) shows that successful switching from P to AP is only obtained for the symmetric configuration and a larger length t_{FL} of $5\,\mathrm{nm}$. The difference between the intermediate states of the dashed and dotted lines comes from the different lengths, and therefore, the magnetizations of the FL_1 and FL_2 resulting in the distinct total magnetization of the FL in the final state.

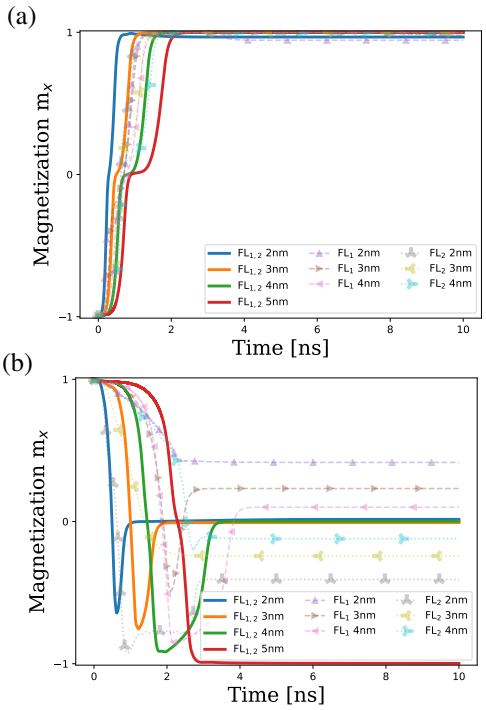


Figure 2: Magnetization trajectories for the switching from (a) AP to P and (b) P to AP for different combinations of FL lengths. The solid line represents a symmetric configuration where both FL have the same length. The dotted line and the dashed line correspond to an asymmetric structure. In the former, the first FL layer is fixed, the second is varied, and in the latter, the second FL layer is fixed, and the first FL layer is varied. The polarization of the TB used is 0.6, 0.5, and 0.2, in order from left to right in Fig. 1.

The differences in the response times of the cells are due to the different uniaxial anisotropies, which depend on the length of the layer. A shorter layer possesses a reduced energy barrier separating the two magnetization configurations, so that the speed of the response times is improved in the case of the shorter layer.

In Fig. 3 we report switching realizations where we modeled the polarization of the TBs, such that the TB in the middle of the FL is the higher one. Higher TB polarization leads to higher spin current so that the torque acting on the two FL segments is stronger than the torque acting on RL and FL_1. In Fig. 3(a), all structures switch successfully except for the symmetric configuration with an FL length of $2\,\mathrm{nm}$. Here, the magnetization develops a single oscillation before reaching a stable state. This is due to the weak anisotropy, and the torques from FL_2 acting on FL_1 being more substantial than from RL, leading

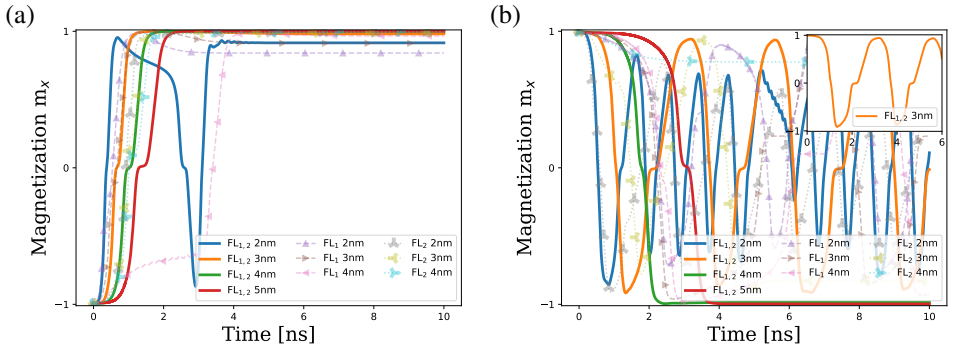


Figure 3: Magnetization trajectories for the switching from (a) AP to P and (b) P to AP for different combinations of FL lengths, were the same methodology as in Fig. 2 was applied, but with increasing polarization of TB in the middle of the FL. In addition, the symmetrical structure with 3 nm thick FLs is highlighted in the inset. The polarization of the TB used is 0.5, 0.9, and 0.2, in order from left to right in Fig. 1

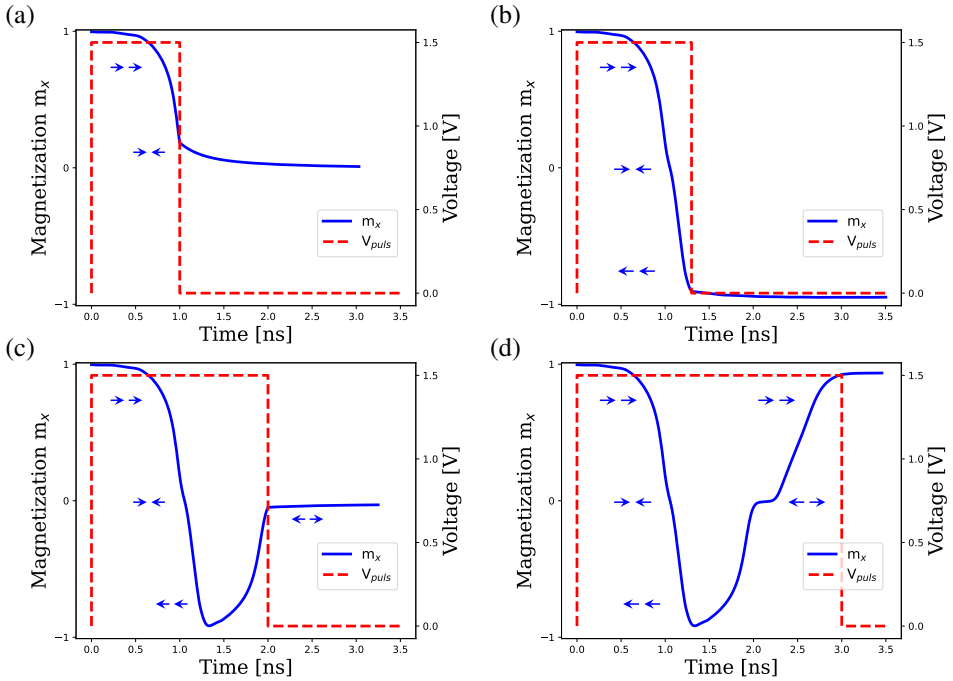


Figure 4: Magnetization trajectories for a symmetric structure with 3 nm thick FLs, switching from P to AP for different duration of the applied bias. The solid blue line represents the x-component of the magnetization and the red dashed line the duration and amplitude of the bias pulse. The arrows along the magnetization represent the state of FL_1 and FL_2.

to an initial back-and-forth switching before reaching the final state. Fig. 3(b) illustrates the back-hopping effect. As reported in the inset, the structure with 3 nm shows continuous transitions between the sequence of states as long as the current is applied. The sequential switching procedure of the composite FL is as follows: The torque acting from FL_2 to FL_1 is opposite to that of the RL, so FL_2 switches first. After FL_2 switches, the torque contributions from FL_2 and the RL act on FL_1 in the same sense flipping FL_1 and completing the switching.

As the bias is maintained, the magnetization of FL_2 reverses due to the higher polarization in the middle TB and the back-hopping effect of FL_2 occurs. The torques from the RL favor the anti-parallel state and those from FL_2 favor the parallel state. However, due to the higher polarization in the middle TB, the torques from FL_2 are stronger than those from the RL, so FL_1 flips, and the whole FL returns to the initial configuration. This process repeats as long as the bias is maintained.

In Fig. 4 we report switching realizations with different durations of the applied bias. If we adjust and vary the duration of the bias pulse, we can realize a multi-level cell with four different states. These states of the FL are illustrated by the blue arrows and follow the same switching procedure as mentioned earlier. By properly modulating the duration of the pulse, all four states can be addressed. A distinction between the two intermediate states can be recognized by the fact that the first state FL_1 parallel to the RL and anti-parallel to FL_2 has a lower resistance than the state with FL_1 anti-parallel to the RL and FL_2. Fig. 5 demonstrates the snapshots of the important stages of the switching sequence for the AP

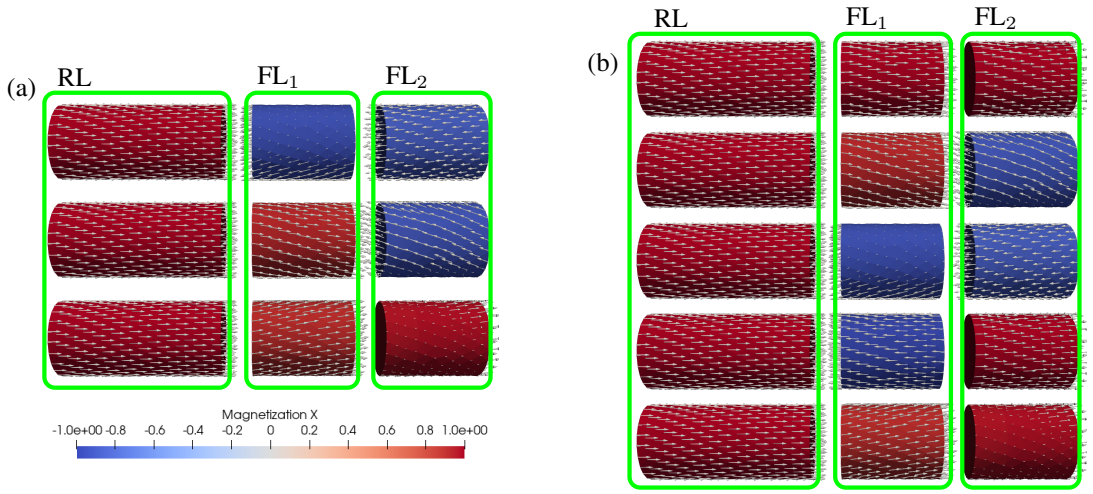


Figure 5: Switching snapshot of MRAM cell's in the following order: Initial, intermediate and final state, for the configuration (a) AP to P and (b) P to AP. Arrows indicate the magnetization directions, and color-coding represents the average magnetization in x-direction. Depending on the switching direction, a negative bias for (a) or a positive bias for (b) is applied. To achieve the respective states, the pulse duration is modulated as shown in the previous Fig. 4.

to P and P to AP process. Fig. 5(b) displays the arrow configuration of Fig. 4. To get from the configuration in Fig. 3(a) or Fig. 3(c) back to the desired AP or P configuration, only a short bias pulse is necessary to reset the structure into the desired state.

IV. Conclusion

The drift-diffusion approach is supplemented with appropriate boundary conditions for the spin current $\mathbf{J_S}$ at the TB interfaces to account for the dependence of the torque on the relative magnetization orientation specific to MTJs. The boundary condition is key to describe the spin current and the spin accumulations in the RL and FL. This gives us the opportunity to describe the spin and charge transport coupled to the magnetization in arbitrary stacks of MTJs. We demonstrated an appearance of the back-hopping effect in ultra-scaled STT-MRAM devices with a composite free layer. By proper analysis of the back-hopping effect, we demonstrated the potential of employing this effect to engineer devices as a multi-bit cells with four different states.

Acknowledgment

The financial support by the Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

References

[1] S. Jung, H. Lee, S. Myung, H. Kim, S. K. Yoon, S.-W. Kwon, Y. Ju, M. Kim, W. Yi, S. Han, and et al., "A crossbar array of magnetoresistive memory devices for in-memory computing," *Nature*, vol. 601, no. 7892, p. 211–216, 2022.

[2] V. B. Naik, K. Yamane, T. Lee, J. Kwon, R. Chao *et al.*, "Jedec-qualified highly reliable 22nm FD-SOI embedded MRAM for low-power industrial-grade, and extended performance towards automotive-grade-1 applications," in *IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 11.3.1–11.3.4.

[3] S. Ikegawa, K. Nagel, F. B. Mancoff, S. M. Alam, M. Arora, M. DeHerrera, H. K. Lee, S. Mukherjee, G. Shimon, J. J. Sun, I. Rahman, F. Neumeyer, H. Y. Chou, C. Tan, A. Shah, and S. Aggarwal, "High-speed (400MB/s) and low-BER STT-MRAM technology for industrial applications," in *International Electron Devices Meeting (IEDM)*, 2022, pp. 10.4.1–10.4.4.

[4] G. Hu, C. Safranski, J. Z. Sun, P. Hashemi, S. L. Brown, J. Bruley, L. Buzi, C. P. D'Emic, E. Galligan, M. G. Gottwald, O. Gunawan, J. Lee, S. Karimeddiny, P. L. Trouilloud, and D. C. Worledge, "Double spin-torque magnetic tunnel junction devices for last-level cache applications," in *International Electron Devices Meeting (IEDM)*, 2022, pp. 10.2.1–10.2.4.

[5] S. M. Seo, H. Aikawa, S. G. Kim, T. Nagase, Y. Ito, T. J. Ha, K. Yoshino, B. K. Jung, T. Oikawa, K. Y. Jung, H. I. Moon, B. S. Kim, F. Matsuoka, K. Hatsuda, K. Hoya, S. Kim, S.-H. Lee, M.-H. Na, and S. Y. Cha, "First demonstration of full integration and characterization of 4f² 1s1m cells with 45 nm of pitch and 20 nm of mtj size," in *International Electron Devices Meeting (IEDM)*, 2022, pp. 10.1.1–10.1.4.

[6] H. Sato, M. Yamanouchi, S. Ikeda, S. Fukami, F. Matsukura, and H. Ohno, "MgO/CoFeB/Ta/CoFeB/MgO recording structure in magnetic tunnel junctions with perpendicular easy axis," *IEEE Transactions on Magnetics*, vol. 49, no. 7, pp. 4437–4440, 2013.

[7] K. Nishioka, H. Honjo, S. Ikeda, T. Watanabe, S. Miura *et al.*, "Novel quad interface MTJ technology and its first demonstration with high thermal stability and switching efficiency for STT-MRAM beyond 2Xnm," in *Symposium on VLSI Technology*, 2019, pp. T120–T121.

[8] B. Jinnai, J. Igarashi, K. Watanabe, T. Funatsu, H. Sato *et al.*, "High-performance shape-anisotropy magnetic tunnel junctions down to 2.3 nm," in *IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 24.6.1–24.6.4.

[9] C. Abert, H. Sepehri-Amin, F. Bruckner, C. Vogler, M. Hayashi, and D. Suess, "Back-hopping in spin-transfer-torque devices: Possible origin and countermeasures," *Physical Review Applied*, vol. 9, p. 054010, 2017.

[10] S. Fiorentini, M. Bendra, J. Ender, R. L. de Orio, W. Goes, S. Selberherr, and V. Sverdlov, "Spin and charge drift-diffusion in ultra-scaled mram cells," *Sci. Rep.*, vol. 12, no. 1, 2022.

[11] J. Ender, M. Mohamedou, S. Fiorentini, R. Orio, S. Selberherr, W. Goes, and V. Sverdlov, "Efficient demagnetizing field calculation for disconnected complex geometries in STT-MRAM cells," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2020, pp. 213–216.

[12] C. Petitjean, D. Luc, and X. Waintal, "Unified drift-diffusion theory for transverse spin currents in spin valves, domain walls, and other textured magnets," *Phys. Rev. Lett.*, vol. 109, p. 117204, 2012.

Tunneling Attenuation and Leakage Current in MoS_2 Nanoribbon MOSFETs

I. Prevarić, M. Matić, and M. Poljak*
Computational Nanoelectronics Group
Micro and Nano Electronics Laboratory, Faculty of Electrical Engineering and Computing
University of Zagreb, HR-10000 Zagreb, Croatia
*E-mail: mirko.poljak@fer.hr

***Abstract*—We study the OFF-state leakage current in quasi-one-dimensional MoS_2 nanoribbon (MoS_2NR) FETs using *ab initio* Hamiltonians and quantum transport simulations based on Green's functions. Complex band structure is computed for these devices and the energy-dependent tunneling attenuation inside the bandgap is obtained. We investigate the tunneling component of the OFF-state leakage for sub-20 nm long and sub-3 nm wide MoS_2NR FETs, using the under-the-barrier (UTB) and top-of-the-barrier (ToB) ballistic models. We report that using the parabolically-approximated attenuation overestimates the OFF-state leakage significantly. Furthermore, we demonstrate that all MoS_2NR FETs show good tunneling suppression due to high attenuation even for the shortest devices where the OFF-state leakage is under 16.5 nA/μm for nFETs and lower than 22 nA/μm for pFETs.**

***Keywords*—MoS_2, molybdenum disulfide, leakage current, complex bandstructure, tunneling, quasi-one-dimensional, nanoribbon, quantum transport, NEGF, ab initio, DFT**

I. Introduction

Monolayer molybdenum disulfide (MoS_2), alongside graphene, is one of the most promising two-dimensional (2D) materials for future field-effect transistors (FETs) due to high stability and very high ON-OFF current ratio [1]. The 2D MoS_2 is a transition metal dichalcogenide (TMD) that exhibits a sandwiched S-Mo-S structure as illustrated in Fig. 1a and b. The electronic, transport, and device properties of devices with 2D MoS_2 monolayer were comprehensively studied previously [2]–[4], however, patterning MoS_2 into quasi-one-dimensional (quasi-1D) structures such as nanoribbons provides an additional degree of freedom. This nanostructure enables the tuning of electron and transport properties via quantum confinement effects and allows a strong influence of edge atoms on the characteristics of MoS_2 nanoribbons (MoS_2NRs). Consequently, a high number of possible MoS_2NR configurations exists and many of them have been studied previously in both armchair and zigzag directions [5]–[7]. Nevertheless, there are some interesting configurations, such as MoS_2NRs with hybrid OH-passivated edges shown in Fig. 1a and b, whose transport and device properties have not yet been explored. Electronic properties of hybrid OH-passivated MoS_2NRs were studied in [8] where this configurations was proven to be one of the most stable edge passivation configurations for future MoS_2-based nanodevices.

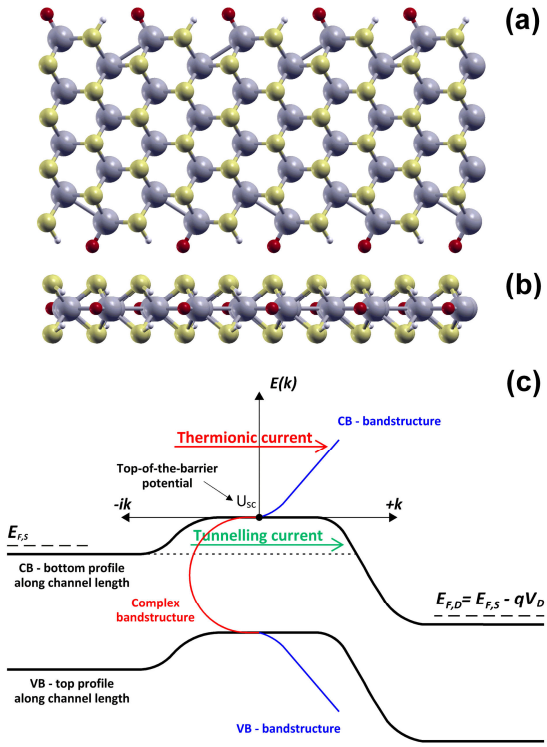


Fig. 1. (a) Top and (b) side view of MoS_2NR with OH passivation (explained in the text). (c) Conduction and valence band profiles along the channel length with an illustration of real and complex bandstructure used in UTB and TOB calculations.

In addition to stability, low OFF-state leakage current (I_{OFF}) is one of the critical requirements for future low power/high density devices. However, I_{OFF} has not yet been studied in ultra-scaled FETs with channels made of MoS_2NRs with hybrid OH passivation. In this paper, we present a study on I_{OFF} of MoS_2NR FETs using our in-house non-equilibrium Green's function (NEGF) formalism based quantum transport solver [9], [10]. The NEGF calculations are employed in conjunction with top-of-the-barrier (ToB) model [11] for the ballistic thermionic current and with the under-the-barrier (UTB) model [12],[13] that accounts for intra-band tunneling, as indicated in Fig. 1c. Highly accurate space-localized

Hamiltonians are obtained from *ab initio* plane-wave density functional theory (DFT) calculations using maximally-localized Wannier functions (MLWFs). We report strong suppression of tunneling for all the studied MoS_2NRs even in transistors with 5 nm-long channels, which indicates a robust performance and low I_{OFF} in ultra-scaled MoS_2NR FETs.

II. Methodology

Armchair MoS_2 nanoribbons passivated with OH of various widths (W) ranging from ~0.8 nm to ~3.0 nm are constructed for *ab initio* calculations. For example, the 1.11 nm-wide MoS_2NR is shown in the top (Fig. 1a) and side view (Fig. 1b) with the transport direction along the horizontal orientation. The hybrid OH passivation is achieved by terminating the edge Mo atoms with O atoms and S atoms with H atoms. Plane-wave DFT is employed in combination with transformation to tight-binding-like sparse matrices using MLWFs [14] that are suitable for NEGF simulations. Quantum Espresso [15] is used for DFT calculations, where a vacuum region of 20 Å is added in the confined directions due to the nature of DFT to assume periodicity in all three directions. The Brillouin zone (BZ) is sampled using an 1 × 15 × 1 Monkhorst-Pack grid [16], where 15 points are used in the transport direction. Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [17] exchange-correlation (XC) functional is used along projector augmented wave (PAW) pseudopotentials. The plane-wave cutoff energy is set to 1360 eV, whereas the convergence threshold is set to 10^{-3} eV/Å for the ionic force and to 10^{-4} eV for energy. Wannier90 program [18] is employed to transform the DFT Hamiltonians into the space-localized basis while preserving the accuracy of the DFT. The accuracy of transformation using MLWFs greatly depends on the trial projections on the Bloch manifold. In this work, for OH-passivated MoS_2NRs, d orbitals are chosen as trial orbitals for Mo atoms and p orbitals for S and O atoms. During wannierization, we obtain a Wannier spread lower than 2.5 Å^2 per atom for all MoS_2NR widths, which indicates successful transformation. After MLWF Hamiltonians of a MoS_2NR unit-cell are obtained, where the unit-cell designates a super-cell along the entire nanoribbon width, the cell is scaled in the transport direction to construct the total MLWF Hamiltonian of the given MoS_2NR with lengths (L) ranging from ~5 nm to ~20 nm.

The main topic of this paper is the study of I_{OFF} of n- and p- type MoS_2NR MOSFET devices with ideal contacts. The I_{OFF} consists of the thermionic (above barrier) and tunneling (under barrier) current components, as shown in Fig. 1c. The thermionic current is calculated using the ToB model that depends on the density of states (DOS) and transmission spectra obtained using the NEGF simulations [19]–[21], as implemented in our in-house code [9], [10], [22], [23]. Furthermore, UTB model [13] is employed to calculate the tunneling current through the bandgap of the S/D potential barrier. For the UTB model, tunneling attenuation is needed, which can be obtained from the complex bandstructure. While using real wave-vector k leads to dispersion of propagating electron waves, setting imaginary values of the wave-vector k and calculating the dispersion results in tunneling attenuation $\kappa(E)$ that describes evanescent electron waves. Transmission probability is calculated according to the Wentzel-Kramers-Brillouin (WKB) approximation

$$T_{WKB} = \sum_n \exp\left[-2\, l(E)\, \kappa(E)\right] \tag{1}$$

where n represents the number of complex subbands included in the calculation, while $l(E)$ is an energy-dependent tunneling barrier length. For simplicity, the potential barrier between the source and drain (S/D) is approximated with a parabolic function. Tunneling current is calculated using the Landauer formula that depends on WKB transmission, barrier height and Fermi-window, i.e. difference between S/D Fermi functions.

In order to provide meaningful comparison between various MoS_2NR devices we set a common ToB I_{OFF} of 10 nA/μm as defined in the International Roadmap for Devices and Systems (IRDS) for future high-performance (HP) CMOS nodes [24]. The I_{OFF} is set by a proper gate work function, which modifies the flat-band voltage, and accordingly lowers the OFF-state ToB potential that is used in the UTB model. Gate equivalent oxide thickness (EOT) is set to 1 nm and S/D doping is set at 0.01 molar fraction of the MoS_2NR areal atom density. Device OFF-state performance is studied when the tunneling attenuation $\kappa(E)$ is extracted directly from the complex bandstructure, and when $\kappa(E)$ is approximated with a parabolic function as is originally done in the literature [13]. The parabolic $\kappa(E)$ is defined by the maximum attenuation, occurring usually near the mid-gap, and the bandgap value. Furthermore, nanoribbon width and length scaling effects on the tunneling component of I_{OFF} are studied because the thermionic current is ballistic and, hence, independent of nanoribbon length.

III. Results and Discussion

Figure 2 reports the bandstructure consisting of the complex part for imaginary k as negative values, and the real part for real k as positive values, for $W = 0.81$ nm (Fig. 2a) and $W = 3.04$ nm (Fig. 2b). Scaling down the nanoribbon width decreases the number of subbands and, consequently, the number of complex subbands inside the bandgap also decreases. While scaling down the width generally decreases the bandgap, i.e. from 1.5 eV for $W = 3.04$ nm to 1.25 eV for the 0.81 nm-wide nanoribbon, the attenuation curves of the dominant subband do not change as much. We observe that the two attenuation characteristics exhibit qualitatively similar behavior with maximum attenuation of ~2 nm^{-1} being equal for both MoS_2NRs in Fig. 2.

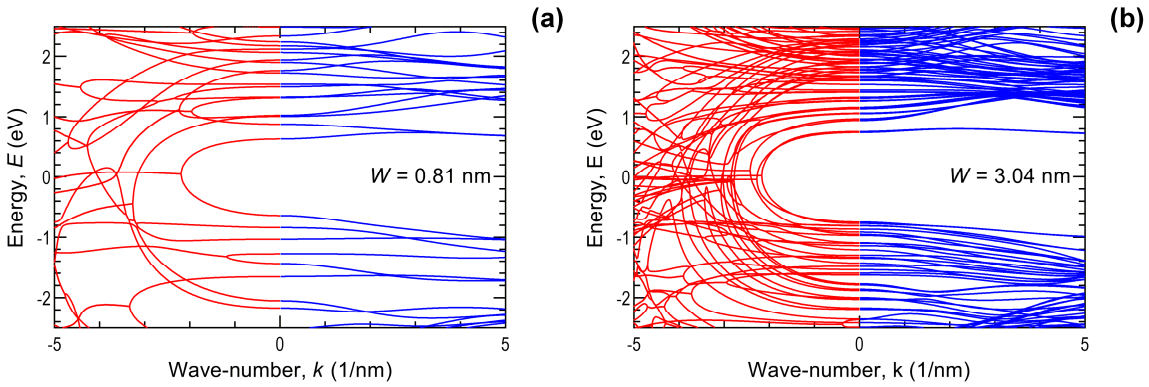


Fig. 2. Complex and real bandstructure of MoS_2NRs with the widths of (a) $W = 0.81$ nm, and (b) $W = 3.04$ nm. Real part is on the right-hand-side for positive k, and the complex part (attenuation) is on the left-hand-side for negative k.

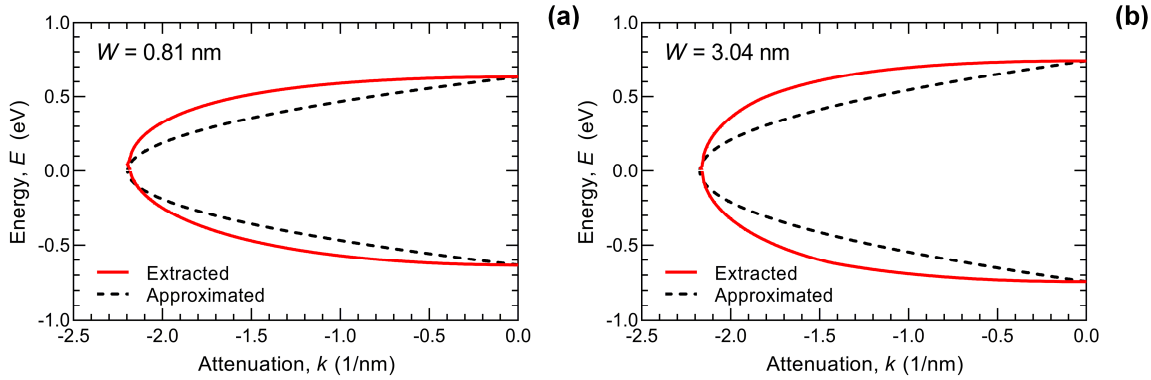


Fig. 3. Extracted and approximated attenuation characteristics of the dominant (with the lowest attenuation) complex subband for (a) 0.81 nm- and (b) 3.04 nm-wide MoS_2NRs.

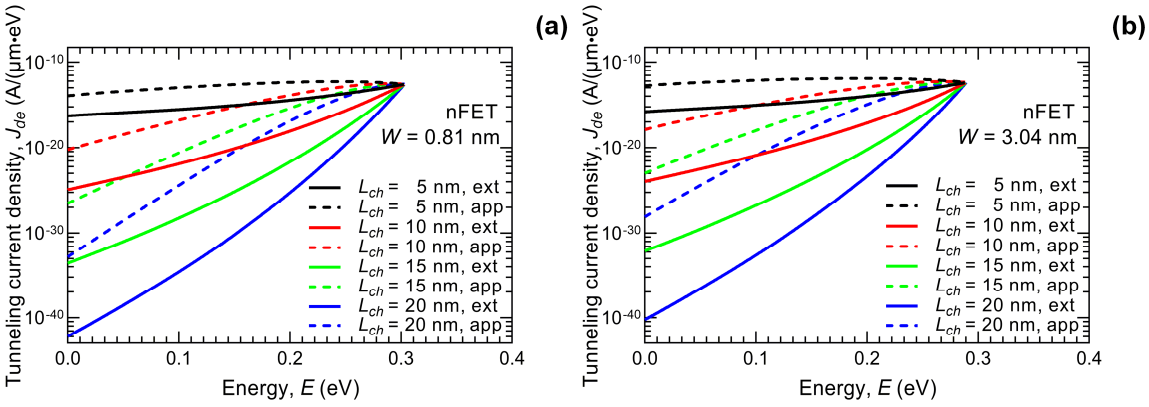


Fig. 4. Tunneling current energy density for MoS_2NR nFETs calculated with the exact extracted (ext) and approximated (app) attenuation for (a) $W = 0.81$ nm, and (b) $W = 3.04$ nm, and for four channel lengths equal to 5 nm, 10 nm, 15 nm and 20 nm.

The first subband belonging to evanescent waves, which connects the highest valence band with the lowest conduction band, can be extracted and then approximated with a parabola. Although using the approximation is much simpler to implement in the simulation environment, a large difference in attenuation is observed when comparing the extracted and approximated attenuation. The difference is reported in Fig. 3a for the 0.81 nm-wide nanoribbon, and an even larger mismatch is observed for $W = 3.04$ nm in Fig. 3b. Attenuation is the lowest near the conduction band minimum (CBM) and valence band maximum (VBM) which means that the tunneling OFF-state current is dominantly determined in the energy regions near the CBM for nFETs, and near the VBM for pFETs. Going towards mid-gap, the attenuation increases in both MoS_2NRs and reaches ~2.2 nm^{-1}, which turns off significant tunneling in this energy range.

In order to assess the differences in using the extracted and approximated attenuation curves, the tunneling current energy density (J_{de}), hereafter current density, is shown in Fig. 4 for various MoS_2NRs nFETs, with four different

channel lengths in the range from ~5 nm to ~20 nm, and the widths of $W = 0.81$ nm (Fig. 4a) and $W = 3.04$ nm (Fig. 4b). The tunneling transport occurs between $E = 0$ eV (mid-gap) and $E \sim 0.3$ eV (CBM). The current density calculated with the approximated attenuation is up to 12 orders of magnitude higher than the current density calculated using the exact extracted attenuation, which demonstrates the inadequacy of the parabolic approximation for quantitative prediction of the tunneling component of I_{OFF}. In both attenuation models, J_{de} increases while scaling down MoS_2NR channel length from 20 nm to 5 nm due to decreasing tunneling length. When the extracted $\kappa(E)$ is used, the maximum J_{de} occurs at the top of the barrier for all channel lengths. On the other hand, J_{de} calculated using the parabolic $\kappa(E)$ exhibits a maximum at the top of the barrier only for channels longer than ~15 nm. As shown in Fig. 4, for shorter channels the J_{de} maximum shifts towards the mid-gap, which is clearly unphysical. Significant difference in J_{de} magnitude and the shift of the maximum show that the approximated parabolic $\kappa(E)$ model is not appropriate to assess the OFF-state performance and, therefore, in further calculations only the extracted $\kappa(E)$ is used.

The extracted attenuation exhibits differences between electron and hole attenuation characteristics, i.e. $\kappa(E)$ curves from the mid-gap towards either CBM or VBM, as shown in Fig. 5a. In turn, nFET and pFET devices with MoS_2NR channels are expected to have different tunneling characteristics. Scaling down MoS_2NR width increases the mismatch between the electron and hole attenuations, with hole $\kappa(E)$ being lower than the electron $\kappa(E)$. This difference points to a higher OFF-state current for pFET devices generally, and the difference is expected to increase as W decreases. Maximum attenuation equals ~2.2 nm^{-1} for all MoS_2NRs with the exception for $W = 1.11$ nm where the maximum $\kappa(E)$ is slightly higher at ~2.5 nm^{-1}, as reported in Fig. 5b.

The total OFF-state current per nanoribbon width is calculated for various MoS_2NR widths and channel lengths in the range from 20 nm to 5 nm and is shown in Fig. 6. The total I_{OFF} is a sum of the tunneling current, calculated by integrating the current density J_{de} over the Fermi window, and of the L-independent thermionic current, set to 10 nA/μm for all devices in ToB simulations. Scaling down channel length results in a significant increase of I_{OFF} in both nFETs (Fig. 6a) and pFETs (Fig. 6b), especially in 5 nm-long devices with the shortest S/D tunnel barrier. The total I_{OFF} is of the same magnitude for all n- and p- type MoS_2NR FETs with $L \geq 10$ nm, while a considerable difference is observed only for $L = 5$ nm. For nFETs with the 5 nm-long channel, maximum I_{OFF} of 16.5 nA/μm is observed for $W = 0.81$ nm and minimum

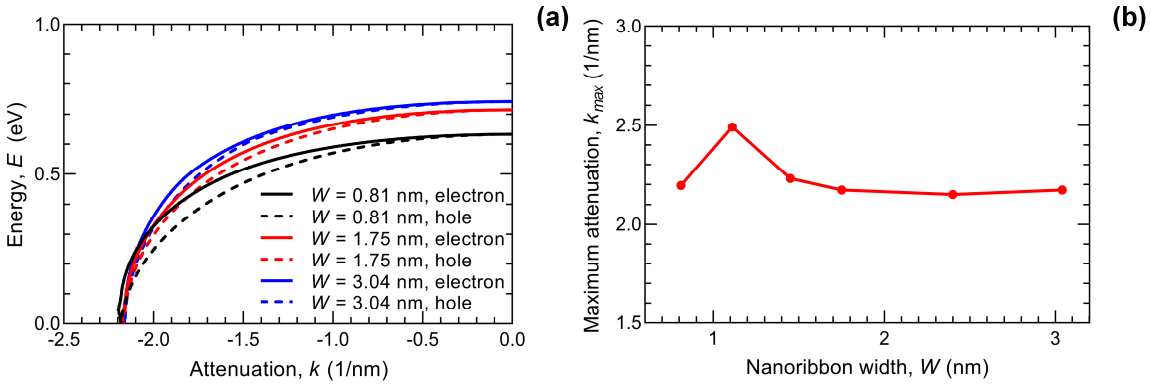


Fig. 5. (a) Extracted attenuation characteristics for electrons and holes for various widths of the MoS_2NR. The energy range spans from mid-gap at 0 eV to band minimum or maximum. (b) Dependence of the maximum attenuation of the dominant complex subband on the MoS_2NR width.

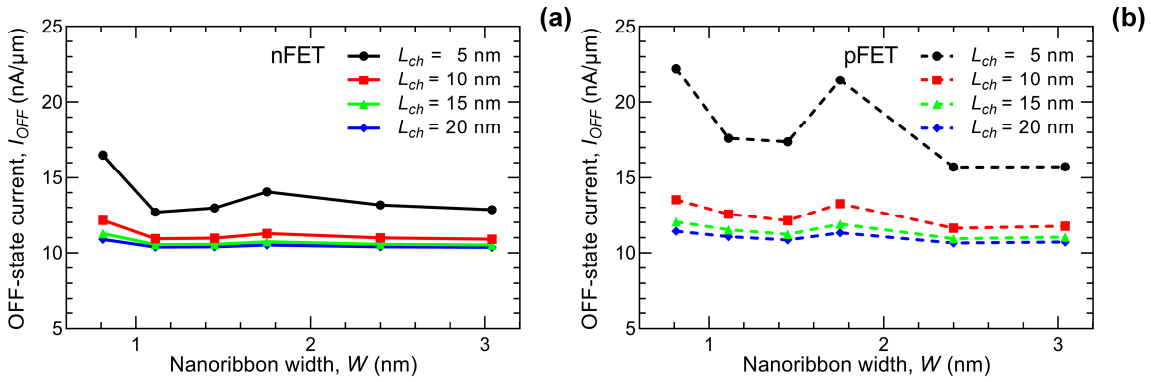


Fig. 6. Impact of MoS_2NR width scaling on total OFF-state current for (a) nFET and (b) pFET for various channel lengths ranging from 5 nm to 20 nm.

I_{OFF} of 12.7 nA/μm for the nFET with the 1.11 nm-wide MoS_2NR channel. On the other hand, pFET with the channel length of 5 nm shows the maximum I_{OFF} of 22 nA/μm for $W = 0.81$ nm and the minimum I_{OFF} of 15.7 nA/μm for $W = 3.04$ nm. These results show that a direct consequence of the lower hole attenuation compared to electron $\kappa(E)$ is the higher tunneling current in pFETs that ranges from 5.7 to 12 nA/μm in pFETs, in comparison to 2.7 to 6.5 nA/μm in nFETs. Finally, we note that the general increase and non-monotonic dependence of I_{OFF} on W reported in Fig. 6 is a consequence of non-monotonic features of $\kappa(E)$ (see Fig. 5b) and width-normalization of the drain current.

IV. Conclusions

Some features of NEGF calculations combined with the UTB and ToB models are used to calculate the tunneling and thermionic OFF-state currents in ultra-scaled MoS_2NR FETs with n- and p-type channels. Regarding the UTB model, we report a significant overestimation of tunneling leakage when a simple parabolic approximation of $\kappa(E)$ is used, rendering it useless for quantitative predictions about tunneling in ultra-scaled MoS_2 nanostructures. We report that scaling down channel length results in the expected growth of the tunneling current that is especially evident for sub-10 nm-long MoS_2NRs n- and pFETs. Due to weaker attenuation pFETs exhibit larger tunneling leakage, i.e. for $L = 5$ nm the total I_{OFF} is lower than 16.5 nA/μm for nFETs and under 22 nA/μm for pFETs. Generally, all MoS_2NRs show strong attenuation that ensures an appropriate suppression of tunneling in sub-3 nm-wide and sub-20 nm-long MoS_2NR FETs.

Acknowledgments

This work was supported by the Croatian Science Foundation under the project CONAN2D (Grant No. UIP-2019-04-3493).

References

[1] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS2 transistors," *Nature Nanotechnology*, vol. 6, no. 3, pp. 147–150, 2011, doi: 10.1038/nnano.2010.279.

[2] W. Cao, J. Kang, D. Sarkar, W. Liu, and K. Banerjee, "Performance evaluation and design considerations of 2D semiconductor based FETs for sub-10 nm VLSI," in *2014 IEEE International Electron Devices Meeting*, Dec. 2014, p. 30.5.1-30.5.4. doi: 10.1109/IEDM.2014.7047143.

[3] A. Nourbakhsh *et al.*, "MoS2 Field-Effect Transistor with Sub-10 nm Channel Length," *Nano Lett.*, vol. 16, no. 12, pp. 7798–7806, Dec. 2016, doi: 10.1021/acs.nanolett.6b03999.

[4] Y. Lee, S. Fiore, and M. Luisier, "Ab initio mobility of single-layer MoS2 and WS2: comparison to experiments and impact on the device characteristics," in *2019 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2019, p. 24.4.1-24.4.4. doi: 10.1109/IEDM19573.2019.8993477.

[5] C. Ataca, H. Şahin, E. Aktürk, and S. Ciraci, "Mechanical and Electronic Properties of MoS2 Nanoribbons and Their Defects," *J. Phys. Chem. C*, vol. 115, no. 10, pp. 3934–3941, Mar. 2011, doi: 10.1021/jp1115146.

[6] K. Dolui, C. D. Pemmaraju, and S. Sanvito, "Electric Field Effects on Armchair MoS2 Nanoribbons," *ACS Nano*, vol. 6, no. 6, pp. 4823–4834, Jun. 2012, doi: 10.1021/nn301505x.

[7] Q. Yue *et al.*, "Bandgap tuning in armchair MoS2 nanoribbon," *J. Phys.: Condens. Matter*, vol. 24, no. 33, p. 335501, Jul. 2012, doi: 10.1088/0953-8984/24/33/335501.

[8] L. Zhang *et al.*, "Modulation of Electronic Structure of Armchair MoS_2 Nanoribbon," *J. Phys. Chem. C*, vol. 119, no. 38, pp. 22164–22171, Sep. 2015, doi: 10.1021/acs.jpcc.5b04747.

[9] M. Matić;, T. Župančić, and M. Poljak, "Parallelized Ab Initio Quantum Transport Simulation of Nanoscale Bismuthene Devices," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 118–123. doi: 10.23919/MIPRO55190.2022.9803335.

[10] M. Matić and M. Poljak, "Ab initio quantum transport simulations of monolayer GeS nanoribbons," *Solid-State Electronics*, vol. 197, p. 108460, Nov. 2022, doi: 10.1016/j.sse.2022.108460.

[11] A. Rahman, J. Guo, S. Datta, and M. S. Lundstrom, "Theory of ballistic nanotransistors," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1853–1864, Sep. 2003, doi: 10.1109/TED.2003.815366.

[12] M. Matic, M. Leljak, and M. Poljak, "Estimating OFF-state Leakage in Silicene Nanoribbon MOSFETs from Complex Bandstructure," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia: IEEE, Sep. 2021, pp. 85–89. doi: 10.23919/MIPRO52101.2021.9596876.

[13] A. Szabo and M. Luisier, "Under-the-Barrier Model: An Extension of the Top-of-the-Barrier Model to Efficiently and Accurately Simulate Ultrascaled Nanowire Transistors," *IEEE Trans. Electron Devices*, vol. 60, no. 7, pp. 2353–2360, Jul. 2013, doi: 10.1109/TED.2013.2263386.

[14] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," *Phys. Rev. B*, vol. 56, no. 20, Art. no. 20, Nov. 1997, doi: 10.1103/PhysRevB.56.12847.

[15] P. Giannozzi *et al.*, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials," *J. Phys.: Condens. Matter*, vol. 21, no. 39, Art. no. 39, Sep. 2009, doi: 10.1088/0953-8984/21/39/395502.

[16] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," *Phys. Rev. B*, vol. 13, no. 12, Art. no. 12, Jun. 1976, doi: 10.1103/PhysRevB.13.5188.

[17] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Phys. Rev. Lett.*, vol. 77, no. 18, Art. no. 18, Oct. 1996, doi: 10.1103/PhysRevLett.77.3865.

[18] G. Pizzi *et al.*, "Wannier90 as a community code: new features and applications," *J. Phys.: Condens. Matter*, vol. 32, no. 16, p. 165902, Apr. 2020, doi: 10.1088/1361-648X/ab51ff.

[19] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1997.

[20] M. Pourfath, *The Non-Equilibrium Green's Function Method for Nanoscale Device Simulation*. in Computational Microelectronics. Wien: Springer-Verlag, 2014. doi: 10.1007/978-3-7091-1800-9.

[21] Y. He, T. Kubis, M. Povolotskyi, J. Fonseca, and G. Klimeck, "Quantum transport in NEMO5: Algorithm improvements and high performance implementation," in *2014 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sep. 2014, pp. 361–364. doi: 10.1109/SISPAD.2014.6931638.

[22] M. Poljak, "Electron Mobility in Defective Nanoribbons of Monoelemental 2D Materials," *IEEE Electron Dev. Lett.*, vol. 41, no. 1, pp. 151–154, Jan. 2020, doi: 10.1109/LED.2019.2952661.

[23] M. Poljak, M. Matić, and A. Zeljko, "Minimum Contact Resistance in Monoelemental 2D Material Nanodevices With Edge-Contacts," *IEEE Electron Device Letters*, vol. 42, no. 8, pp. 1240–1243, Aug. 2021, doi: 10.1109/LED.2021.3087908.

[24] "'IEEE Intl. Roadmap for Devices and Systems (IRDS), 2019 Update.' https://irds.ieee.org/."

Transport Properties and Device Performance of Quasi-One-Dimensional MoS_2 FETs

M. Matić and M. Poljak*
Computational Nanoelectronics Group
Micro and Nano Electronics Laboratory, Faculty of Electrical Engineering and Computing
University of Zagreb, HR-10000 Zagreb, Croatia
*E-mail: mirko.poljak@fer.hr

***Abstract*—We investigated the bandstructure, transport and device properties of semiconducting MoS_2 nanoribbons (MoS_2NR) with hybrid OH-passivated armchair edges using orbitally-resolved *ab initio* Hamiltonians and quantum transport simulations based on Green's functions. The impact of MoS_2NR width scaling on the bandstructure, transmission, bandgap, injection velocity, charge density and ON-state current are analyzed in detail using the ballistic FET model. We find that sub-3 nm-wide and ~15 nm-long MoS_2NR FETs offer low driving currents under 0.43 mA/μm for nFETs and under 0.6 mA/μm for pFETs. Moreover, the current is only weakly modulated by nanoribbon width downscaling due to immunity of the MoS_2NR bandstructure to quantum confinement effects.**

***Keywords*—MoS_2, molybdenum disulfide, quasi-one-dimensional, nanoribbon, quantum transport, Green's function, NEGF, ab initio, DFT, MLWF**

I. Introduction

Since the discovery of graphene in 2004, research focus on two-dimensional (2D) materials as potential candidates to replace silicon in future electronic devices is growing exponentially [1]–[3]. Atomically-thin and dangling-bond-free surfaces along with near-ballistic transport properties of some 2D materials are ideal for future field-effect transistors (FETs). After graphene, monolayer molybdenum disulfide (MoS_2) is one of the most studied 2D materials which showed promise due to high stability, compatibility with graphene, and exceptionally high ON/OFF current ratio > 10^8 [4]. The MoS_2 monolayer is a transition metal disulfide (TMD) with a sandwiched S-Mo-S structure, as shown in Fig. 1. Electronic, transport, and device properties of MoS_2 monolayer devices have been studied extensively and it is often used as a benchmark for other 2D materials [5]–[7]. On the other hand, patterning MoS_2 into quasi-one-dimensional (quasi-1D) nanoribbons, which are illustrated in Fig. 1, enables the tuning of electronic and transport properties, with a potential for beneficial adjustment of device properties as well. Although MoS_2 nanoribbons (MoS_2NRs) in armchair and zigzag directions have been studied previously [8]–[10], due to high influence of edge atoms on the device performance and wide range of possible edge passivation atoms, there are still interesting configurations that have not been explored such as MoS_2NRs with hybrid OH-passivated edges. Electronic properties study in [11] of the hybrid OH-passivated MoS_2NRs showed that this edge configuration is the most stable among all studied edge passivation configurations and, therefore, presents an interesting choice for future MoS_2 based nanodevices once fabrication becomes mature enough. Therefore, in this work we study the device performance of the hybrid OH-passivated MoS_2NRs.

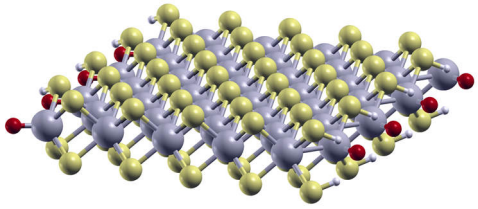


Figure 1. Illustration of an MoS_2 nanoribbon with OH-passivated armchair edges (Mo, S, O and H atoms are represented by grey, yellow, red and white balls, respectively).

Advanced modeling and simulations are used in this study to investigate MoS_2 nanoribbons and MoS_2NR FETs at an atomic level, which is necessary due to inherently strong quantum effects in 2D material nanostructures. Electronic band structure is calculated using the *ab initio* density functional theory (DFT) and, afterwards, maximally-localized Wannier functions (MLWFs) are employed to transform DFT Hamiltonians into a localized basis. The localized MLWF Hamiltonians are much sparser, thus enabling the simulations of realistically sized nanodevice. Quantum transport in such nanostructures is calculated by using the non-equilibrium Greens function (NEGF) formalism. In this paper, we analyze the electronic, transport and device properties of MoS_2NR and MoS_2NR FETs with OH-passivated edges by employing our in-house DFT-NEGF-MLWF solver [12], [13]. We report only a weak impact of MoS_2NR width scaling from ~3.0 nm to ~0.8 nm on the ballistic performance, which results in small degradation and robust performance of ultra-scaled MoS_2NR FETs.

II. Methodology

The MoS_2NR structure of various widths ranging from ~0.8 nm to ~3.0 nm is constructed along the armchair

direction from a MoS_2 unit-cell obtained from Materials Cloud [14] and passivated with OH, as shown in Fig. 1. The Mo edge atoms are passivated with O atoms, while S atoms are passivated with H atoms. Plane-wave DFT is used to obtain highly accurate Hamiltonians of MoS_2 nanostructures investigated in this work. Since DFT calculations assumes periodicity in all three directions, we add a vacuum region of 20 Å in confined directions to exclude any interaction between adjacent layers. For DFT calculations we employ Quantum Espresso [15] with Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [16] exchange-correlation (XC) functional and projector augmented wave (PAW) pseudopotentials . The Brillouin zone (BZ) is sampled using an 15-point equally-spaced Monkhorst-Pack grid [17] in the transport direction and a single point in the confined directions. The plane-wave cutoff energy is set to 100 Ry, whereas the convergence threshold is set to 10^{-3} eV/Å for the ionic force and to 10^{-4} eV for energy.

Dense DFT Hamiltonians are localized in energy and are, therefore, not suitable for NEGF quantum transport simulations which prefer sparser matrices, i.e. space-localized Hamiltonians. Transformation into a space-localized basis is performed using the maximally-localized Wannier functions (MLWFs) [18] which results in tight-binding-like sparse matrices. To obtain MLWF Hamiltonians, we use the Wannier90 [19] program. The main inputs of Wannier90 are the trial orbital projections on Bloch manifold, and in this work the trial orbitals for MoS_2NRs are *d* orbitals for Mo atoms and *p* orbitals for O and S atoms. For all MoS_2NR widths the Wannier spread smaller than 2.5 $Å^2$ per atom is obtained. For each nanoribbon width, MLWF Hamiltonians of a MoS_2NR unit-cell are scaled in the transport direction to construct ~15 nm-long MoS_2NRs.

For quantum transport simulations we use the NEGF formalism [20]–[22], as implemented in our in-house code [12], [13], [23], [24]. Within NEGF, device is modeled with ideal contacts i.e. semi-infinite regions of the same material as the channel using the recursive Sancho-Rubio method [25]. Top-of-the-barrier (ToB) model [26] is used in combination with NEGF to assess ballistic device performance of ultra-scaled MoS_2NR FETs with n- and p-type channels. In ToB ballistic FET simulations, gate equivalent oxide thickness (EOT) is set to 1 nm, while source/drain (S/D) region doping levels are set at 0.01 molar fraction of the MoS_2NR areal atomic density. In all devices we set a common OFF-state current (I_{OFF}) of 10 nA/μm as projected in the International Roadmap for Devices and Systems (IRDS) for high-performance (HP) devices in future CMOS technology nodes [27]. Gate control over the channel in the ToB model devices is assumed ideal and, since no tunneling is included, ideal 60 mV/dec subthreshold slope is obtained in all devices. The ballistic ON-state current (I_{ON}) is extracted when both gate and drain are biased at the supply voltage, i.e. $V_{GS} = V_{DS} = 0.7$ V.

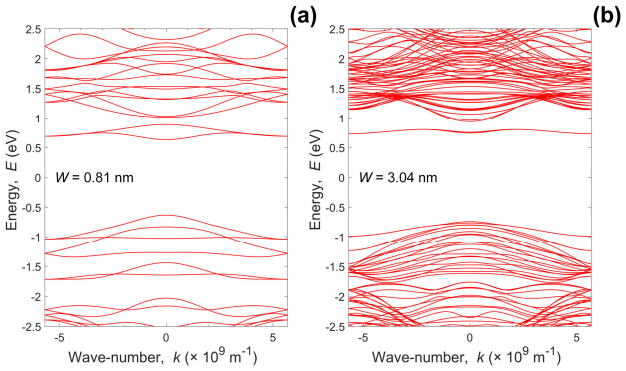


Figure 2. Bandstructure of armchair MoS_2NRs with the widths of (a) 0.81 nm and (b) 3.04 nm.

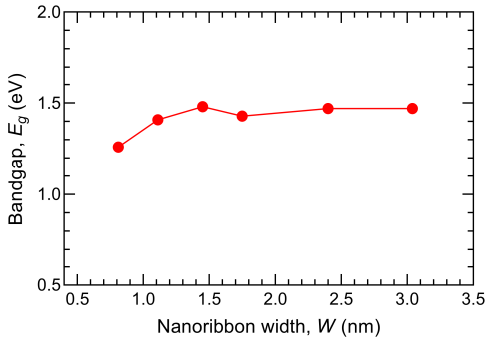


Figure 3. Impact of nanoribbon width scaling on the bandgap of armchair MoS_2NRs.

III. Results and Discussion

In Figure 2 we plot the bandstructure of the widest (Fig. 2a) and the narrowest (Fig. 2b) analyzed MoS_2NR with the widths of $W = 3.04$ nm and $W = 0.81$ nm, respectively. Bandstructure plots contain two subbands that are isolated from the rest of the conduction band (CB). These subbands are isolated due to the influence of edge atoms, which also shift the bandgap from indirect to direct with no conclusive pattern. Scaling down nanoribbon width from 3.04 nm to 0.81 nm results in the separation of these two subbands in CB near the Γ point, which also results in higher curvature, i.e. lower electron effective mass, near CB minimum (CBM) for $W = 0.81$ nm. On the other hand, in the valence band (VB) a higher number of available bands in wider MoS_2NR results in higher number of bands near the VB maximum (VBM), but the narrowest MoS_2NR exhibits a higher curvature and lower hole effective mass in the dominant subband near VBM.

Width-dependence of the bandgap (E_g), extracted from the bandstructure of MoS_2NRs of various widths, is reported in Fig. 3. The plot show a nearly constant bandgap of ~1.47 eV for MoS_2NRs with $W > 1.4$ nm, while scaling down the width below 1.4 nm results in a sharp E_g decrease to 1.26 eV in the 0.81 nm-wide MoS_2NR. The OH-passivated MoS_2NRs exhibit a similar bandgap compared to 2D monolayer MoS_2 for which we obtain the bandgap of 1.7 eV. This value is in line with

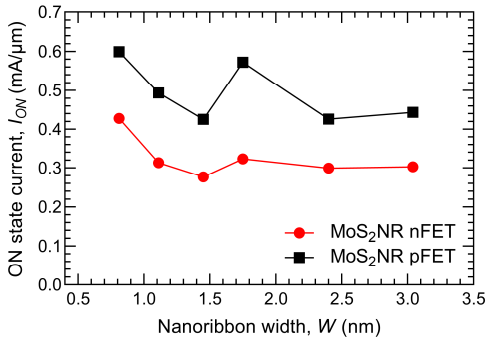


Figure 4. Width-dependence of the ON-state current in n- and p-type MoS_2NR FETs. I_{ON} extracted at $V_{GS} = V_{DS} = V_{DD} = 0.7$ V, with a common I_{OFF} = 10nA set for all devices.

$E_g = 1.67$ eV reported in [11] with the main difference coming from the isolated edge-state bands in the CB.

In order to assess the scaling and confinement effects on device performance, we plot the ON-state current in Fig. 4 for different widths and both device types. Both n- and pFETs exhibit similar scaling laws of I_{ON} with width-downscaling. Qualitatively, we can separate the plot into two groups or two width ranges. The first group includes MoS_2NR FETs with channel widths from $W = 3.04$ nm to $W = 1.75$ nm, which shows increase of I_{ON} while scaling down MoS_2NR width. The second group starts with $W = 1.45$ nm where the I_{ON} shows a sharp decrease compared to the $W = 1.75$ nm case, while further width downscaling results in an increase of I_{ON} down to the 0.81 nm-wide MoS_2NR FET. The maximum I_{ON} for both n- and pFETs is obtained for 0.81 nm-wide MoS_2NR transistor with $I_{ON} = 0.43$ mA/μm for the nFET and $I_{ON} = 0.60$ mA/μm for the pFET. Comparing the obtained I_{ON} results for MoS_2NR FETs to IRDS requirements at the "3 nm" and "2.1 nm" nodes, we conclude that the I_{ON} goal set at 1.9 mA/μm is fulfilled by none of the MoS_2NR devices explored in this paper.

To further understand the behavior of the ON-state current we plot density of states (DOS), transmission, the ON-state charge density (Q_s) and injection velocity (v_{inj}) at ToB, and the ON-state current energy-density. The DOS and transmission characteristics, plotted separately for electrons and holes i.e. in the CB and VB, are shown in Fig. 5. The DOS near the CBM is similar for all MoS_2NRs up to ~0.08 eV after which, due to the existence of isolated edge-state bands, there are no free states and DOS goes to zero. There is an exception for $W = 0.81$ nm where we previously in Fig. 2a observed the broadening of the isolated band near the Γ point. This band broadening results in a nonzero DOS in a broader energy range in comparison to wider nanoribbons. Transmission exhibits similar characteristics as DOS in the CB. At the CBM, transmission equals 1 for the widths of 0.81 nm and 1.75 nm, and sharply increases to 2 at an energy ~20 meV away from the CBM. Interestingly, these MoS_2NRs have a direct bandgap and show highest drain current of all the simulated MoS_2NR nFETs. On the other hand, the 1.45 nm-wide and 3.04 nm-wide MoS_2NRs have a higher transmission equal to 2 at the CBM. This is because these MoS_2NRs have an indirect bandgap with a lower curvature of the CB dispersion, which results in lower I_{ON}. On the other hand, in VB we observe that scaling down nanoribbon width decreases the number of Van Hove

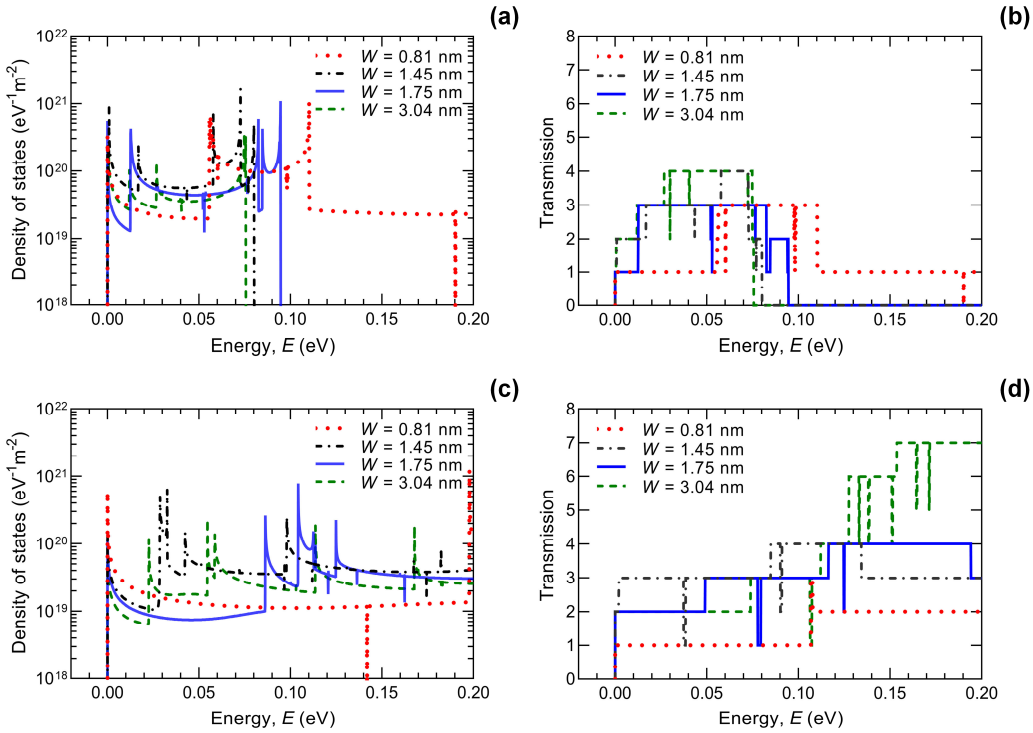


Figure 5. Comparison of DOS (panels on the left) and transmission (panels on the right) in (a, b) CB and (c, d) VB of MoS_2NRs of various widths.

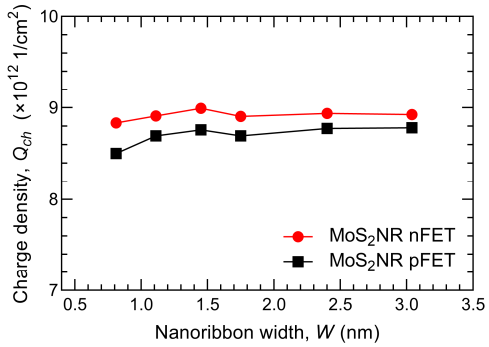


Figure 6. Width-dependence of the channel charge density at ToB in n- and p-type MoS_2NR FETs. The Q_s is extracted in the ON-state, i.e. $V_{GS} = V_{DS} = V_{DD} = 0.7$ V.

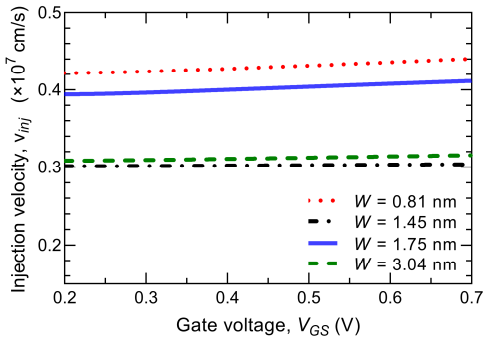


Figure 7. Hole injection velocity dependence on gate bias in MoS_2NR pFETs with various channel widths. $V_{DS} = V_{DD} = 0.7$ V.

singularities (VHS) in the hole DOS near the VBM due to the lower number of available subbands in narrower MoS_2NRs. For the same reason, transmission near the VBM decreases while scaling down MoS_2NR width. When comparing the transport of electrons and holes, we see that holes generally exhibit a higher transmission probability than electrons over the 200 meV energy range away from the VBM or CBM, respectively, which is partly responsible for the higher I_{ON} in p-type than in n-type MoS_2NR FETs reported in Fig. 4.

The DOS and transmission exhibit complex characteristics and are unable to provide a clear insight into FET performance. The influence of width scaling on the ON-state charge density at the ToB is plotted in Fig. 6. Despite considerable variations in the DOS, similar values of Q_s are obtained for all MoS_2NR nFETs and pFETs. On average over all nanoribbon widths, electron density is ~8.9×10^{12} cm^{-2}, while hole density is somewhat lower and equals ~8.7×10^{12} cm^{-2}. The impact of scaling and confinement is visible only in the 0.81 nm-wide MoS_2NRs that show slightly lower electron and hole density at 8.8×10^{12} cm^{-2} and 8.5×10^{12} cm^{-2}, respectively. This decrease is due to setting the common I_{OFF} value, because MoS_2NR with the width of 0.81 nm has the smallest bandgap and, therefore, the lowest applied bias is needed to set I_{OFF} which results in slightly lower Q_s at ToB.

Within the ballistic ToB model, we have $I_{ON} = Q_s \cdot v_{inj}$ at ToB in the ON-state and, hence, the peculiar I_{ON} - W characteristic in Fig. 4 is a consequence of the injection velocity behavior [13], [26]. In Fig. 7 we plot the gate bias dependence of v_{inj} in pFETs, with gate voltage ranging from the threshold voltage (~0.2 V) to supply voltage (0.7 V). Only the v_{inj} in pFETs is plotted because nFETs do not show any noticeable velocity modulation with the applied gate voltage. On the other hand, hole injection velocity in MoS_2NR pFETs with the widths of 0.81 nm and 1.75 nm show a slight increase of v_{inj} with increasing gate bias. The modulation is rather weak, equaling around 5% increase with maximum value obtained in the ON-state of 0.41×10^7 cm/s and 0.44×10^7 cm/s, respectively. With the proportional relationship of the ON-state current to v_{inj} and Q_s, and v_{inj} showing similar scaling laws as I_{ON} while Q_s being of the similar value for all MoS_2NR widths we determine that ON-state performance is dominantly determined by v_{inj}. In comparison to other contending 2D materials, MoS_2NR FETs exhibit significantly lower v_{inj} that deteriorates up to ~9× in comparison to graphene nanoribbons, and ~1.5× when compared against conventional Si MOSFETs [28]. Hence, it is no surprise that the I_{ON} of MoS_2NR FETs is lower than in other 2D material devices [28]–[31].

Finally, the ON-state current energy-density (J_{de}) for MoS_2NR nFET (Fig. 8a) and pFET (Fig. 8b) shows that the current flows mostly in the ~100 meV energy range above/under the CBM/VBM. This range coincides with the energy range of isolated bands in the CB and, therefore, only the isolated edge-state bands in the CB determine the performance of nFETs. On the other hand, VB offers higher number of available bands in the same energy

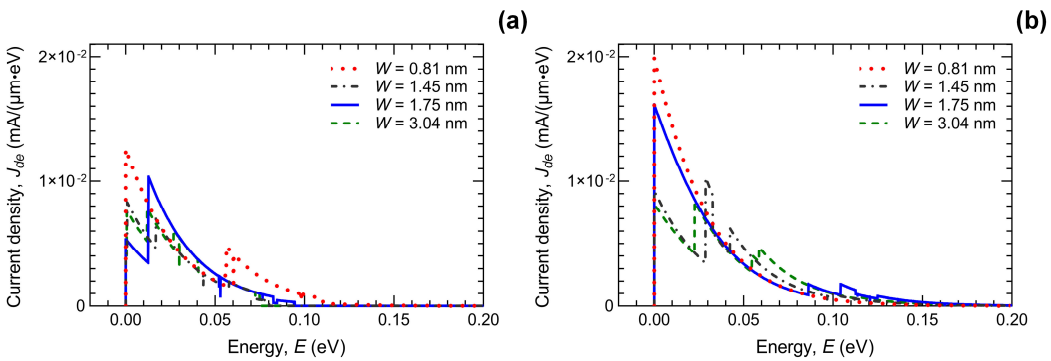


Figure 8. Current energy-density at ToB in the ON-state for MoS_2NR (a) nFETs and (b) pFETs with different nanoribbon channel widths.

range, which results in higher J_{de}, i.e. maxima of 1.2×10^{-2} mA/(eV·μm) for nFETs and 2×10^{-2} mA/(eV·μm) for pFETs, with both values reached at the CBM and VBM, respectively. Hence, any improvement of MoS_2NR nFET performance depends on the effective turn-off of the isolated band contribution.

IV. Conclusions

We studied the electronic, transport and ballistic device characteristics of sub-3 nm-wide and ~15 nm-long quasi-1D MoS_2 nanostructures with OH-passivated edges. We employed DFT and MLWFs to obtain the electronic structure, and NEGF simulations with the ToB model to assess the ballistic performance of armchair MoS_2NR FETs. We found that nFET is limited by the two isolated bands in the CB, which exist due to edge configuration or edge states. The analysis revealed that Q_s is similar for all MoS_2NR widths, and that I_{ON} - W dependence is mainly determined by the features of v_{inj}. The maximum ON-state current of 0.43 mA/μm is obtained for nFETs and 0.6 mA/μm for pFETs. Therefore, MoS_2NR FETs passivated with OH do not meet IRDS requirements for advanced nodes unless a method of eliminating edge states and associated isolated bands in the CB is found.

Acknowledgments

This work was supported by the Croatian Science Foundation under the project CONAN2D (Grant No. UIP-2019-04-3493).

References

[1] K. S. Novoselov *et al.*, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, no. 5696, pp. 666–669, Oct. 2004, doi: 10.1126/science.1102896.

[2] G. Fiori *et al.*, "Electronics based on two-dimensional materials," *Nat. Nanotechnol.*, vol. 9, no. 10, pp. 768–779, Oct. 2014, doi: 10.1038/nnano.2014.207.

[3] N. Briggs *et al.*, "A roadmap for electronic grade 2D materials," *2D Mater.*, vol. 6, no. 2, p. 022001, Jan. 2019, doi: 10.1088/2053-1583/aaf836.

[4] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS2 transistors," *Nature Nanotechnology*, vol. 6, no. 3, pp. 147–150, 2011, doi: 10.1038/nnano.2010.279.

[5] W. Cao, J. Kang, D. Sarkar, W. Liu, and K. Banerjee, "Performance evaluation and design considerations of 2D semiconductor based FETs for sub-10 nm VLSI," in *2014 IEEE International Electron Devices Meeting*, Dec. 2014, p. 30.5.1-30.5.4. doi: 10.1109/IEDM.2014.7047143.

[6] A. Nourbakhsh *et al.*, "MoS2 Field-Effect Transistor with Sub-10 nm Channel Length," *Nano Lett.*, vol. 16, no. 12, pp. 7798–7806, Dec. 2016, doi: 10.1021/acs.nanolett.6b03999.

[7] Y. Lee, S. Fiore, and M. Luisier, "Ab initio mobility of single-layer MoS2 and WS2: comparison to experiments and impact on the device characteristics," in *2019 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2019, p. 24.4.1-24.4.4. doi: 10.1109/IEDM19573.2019.8993477.

[8] C. Ataca, H. Şahin, E. Aktürk, and S. Ciraci, "Mechanical and Electronic Properties of MoS2 Nanoribbons and Their Defects," *J. Phys. Chem. C*, vol. 115, no. 10, pp. 3934–3941, Mar. 2011, doi: 10.1021/jp1115146.

[9] K. Dolui, C. D. Pemmaraju, and S. Sanvito, "Electric Field Effects on Armchair MoS2 Nanoribbons," *ACS Nano*, vol. 6, no. 6, pp. 4823–4834, Jun. 2012, doi: 10.1021/nn301505x.

[10] Q. Yue *et al.*, "Bandgap tuning in armchair MoS2 nanoribbon," *J. Phys.: Condens. Matter*, vol. 24, no. 33, p. 335501, Jul. 2012, doi: 10.1088/0953-8984/24/33/335501.

[11] L. Zhang *et al.*, "Modulation of Electronic Structure of Armchair MoS_2 Nanoribbon," *J. Phys. Chem. C*, vol. 119, no. 38, pp. 22164–22171, Sep. 2015, doi: 10.1021/acs.jpcc.5b04747.

[12] M. Matić;, T. Župančić, and M. Poljak, "Parallelized Ab Initio Quantum Transport Simulation of Nanoscale Bismuthene Devices," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 118–123. doi: 10.23919/MIPRO55190.2022.9803335.

[13] M. Matić and M. Poljak, "Ab initio quantum transport simulations of monolayer GeS nanoribbons," *Solid-State Electronics*, vol. 197, p. 108460, Nov. 2022, doi: 10.1016/j.sse.2022.108460.

[14] L. Talirz *et al.*, "Materials Cloud, a platform for open computational science," *Sci Data*, vol. 7, no. 1, Art. no. 1, Dec. 2020, doi: 10.1038/s41597-020-00637-5.

[15] P. Giannozzi *et al.*, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials," *J. Phys.: Condens. Matter*, vol. 21, no. 39, Art. no. 39, Sep. 2009, doi: 10.1088/0953-8984/21/39/395502.

[16] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Phys. Rev. Lett.*, vol. 77, no. 18, Art. no. 18, Oct. 1996, doi: 10.1103/PhysRevLett.77.3865.

[17] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," *Phys. Rev. B*, vol. 13, no. 12, Art. no. 12, Jun. 1976, doi: 10.1103/PhysRevB.13.5188.

[18] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," *Phys. Rev. B*, vol. 56, no. 20, Art. no. 20, Nov. 1997, doi: 10.1103/PhysRevB.56.12847.

[19] G. Pizzi *et al.*, "Wannier90 as a community code: new features and applications," *J. Phys.: Condens. Matter*, vol. 32, no. 16, p. 165902, Apr. 2020, doi: 10.1088/1361-648X/ab51ff.

[20] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1997.

[21] M. Pourfath, *The Non-Equilibrium Green's Function Method for Nanoscale Device Simulation*. Wien: Springer-Verlag, 2014. doi: 10.1007/978-3-7091-1800-9.

[22] Y. He, T. Kubis, M. Povolotskyi, J. Fonseca, and G. Klimeck, "Quantum transport in NEMO5: Algorithm improvements and high performance implementation," in *2014 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sep. 2014, pp. 361–364. doi: 10.1109/SISPAD.2014.6931638.

[23] M. Poljak, "Electron Mobility in Defective Nanoribbons of Monoelemental 2D Materials," *IEEE Electron Dev. Lett.*, vol. 41, no. 1, pp. 151–154, Jan. 2020, doi: 10.1109/LED.2019.2952661.

[24] M. Poljak, M. Matić, and A. Zeljko, "Minimum Contact Resistance in Monoelemental 2D Material Nanodevices With Edge-Contacts," *IEEE Electron Device Letters*, vol. 42, no. 8, pp. 1240–1243, Aug. 2021, doi: 10.1109/LED.2021.3087908.

[25] M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, "Quick iterative scheme for the calculation of transfer matrices: application to Mo (100)," *J. Phys. F: Met. Phys.*, vol. 14, no. 5, Art. no. 5, May 1984, doi: 10.1088/0305-4608/14/5/016.

[26] A. Rahman, Jing Guo, S. Datta, and M. S. Lundstrom, "Theory of ballistic nanotransistors," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1853–1864, Sep. 2003, doi: 10.1109/TED.2003.815366.

[27] "'IEEE Intl. Roadmap for Devices and Systems (IRDS), 2019 Update.' https://irds.ieee.org/."

[28] G. Liang, N. Neophytou, D. E. Nikonov, and M. S. Lundstrom, "Performance Projections for Ballistic Graphene Nanoribbon Field-Effect Transistors," *IEEE Transactions on Electron Devices*, vol. 54, no. 4, pp. 677–682, Apr. 2007, doi: 10.1109/TED.2007.891872.

[29] S. Kaneko, H. Tsuchiya, Y. Kamakura, N. Mori, and M. Ogawa, "Theoretical performance estimation of silicene, germanene, and graphene nanoribbon field-effect transistors under ballistic transport," *Appl. Phys. Express*, vol. 7, no. 3, p. 035102, Mar. 2014, doi: 10.7567/APEX.7.035102.

[30] X. Cao and J. Guo, "Simulation of Phosphorene Field-Effect Transistor at the Scaling Limit," *IEEE Trans. Electron Devices*, vol. 62, no. 2, pp. 659–665, Feb. 2015, doi: 10.1109/TED.2014.2377632.

[31] M. Poljak and M. Matić, "Bandstructure and Size-Scaling Effects in the Performance of Monolayer Black Phosphorus Nanodevices," *Materials*, vol. 15, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/ma15010243.

Bandgap Narrowing in Silicene Nanoribbons with Metal Edge Contacts

M. Poljak* and M. Matić
Computational Nanoelectronics Group
Micro and Nano Electronics Laboratory, Faculty of Electrical Engineering and Computing
University of Zagreb, HR-10000 Zagreb, Croatia
*E-mail: mirko.poljak@fer.hr

***Abstract*—By employing atomistic quantum transport simulations we investigate the impact of metal edge contacts on the transport gap (E_{TG}) of silicene nanoribbons (SiNR). Transmission and E_{TG} are investigated for sub-5 nm-wide and sub-15 nm-long SiNRs for various metal-device interaction strengths. We find that metallization occurs in certain cases, especially in wider and shorter devices, which sets fundamental limits to device scaling of potential SiNR-based field effect transistors (FET). The findings are elaborated through analytical and numerical examples by discussing transmission and eigenvalue evolution with increasing metal-device interaction.**

Keywords*—*silicene, nanoribbon, quasi-one-dimensional, bandgap, transport gap, bandgap narrowing, metallization, quantum transport, NEGF, wide-band limit, edge contact

I. Introduction

Silicene is a monolayer and monoelemental two-dimensional (2D) material similar to graphene with its Dirac linear dispersion in K points and zero bandgap, but with a somewhat buckled crystal lattice in contrast to the flat lattice of graphene [1]–[4]. Cutting silicene into nanoscale quasi-one-dimensional (1D) silicene nanoribbons (SiNRs) potentially allows denser device integration on chip and possibilities for adjusting material and device characteristics by quantum confinement [5], [6]. Material and device properties of SiNRs and SiNR field-effect transistors (FETs) have been investigated previously, although mainly for the case of ideal contacts [7]–[10]. Ideal contacts result in perfect step-like transmission characteristics because contact regions and the device are assumed to be made of identical materials, which turns off any scattering of electron waves at electrode-device interfaces [11]. Moreover, ideal contacts preserve the bandgap (E_G) while attaching realistic metal contacts changes the transmission function and induces a transport gap (E_{TG}) [12], [13]. Generally, $E_{TG} \neq E_G$ and it is crucial to study the effects of metal contacts on E_{TG} since the transport gap determines several fundamental figures-of-merit of FETs.

In this paper, we employ atomistic Hamiltonians coupled with the non-equilibrium Green's function (NEGF) formalism to study the electronic and transport properties, including E_{TG} behavior, of semiconducting SiNRs of various sizes. We investigate the impact of metal edge contacts since edge contacts are foreseen as a solution for better integration and lower contact resistance in ultra-scaled FETs based on 2D materials [14], [15]. An illustration of edge contacts for a SiNR FET is given in Fig. 1. Transmission and E_{TG} are analyzed for SiNR widths ranging from ~5 nm to ~0.5 nm, for nanoribbon lengths from ~15 nm to ~4.5 nm, and for different electrode metals modeled via a metal-device interaction strength parameter. We find that in certain cases SiNRs undergo metallization, i.e. transition from semiconductor to conductor, which limits the practical applicability of SiNRs of certain dimensions in extremely-scaled FETs.

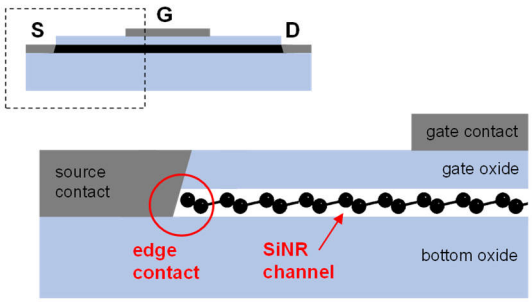


Fig. 1. Illustration of a SiNR FET (top) and a zoomed-in region on the source side showing a metal edge contact (bottom).

II. Methodology

In order to construct atomistic Hamiltonians for armchair-edge SiNRs we use the tight-binding (TB) model from the literature with nearest-neighbor interactions only where the Si-Si coupling or hopping parameter (t) for silicene is $t = 1.03$ eV [16]. Edge relaxation effect is accounted for by using a larger $t = 1.15$ eV for edge Si-Si bonds [7]. A 4-atom unit cell is constructed and repeated NW times along the nanoribbon width (W), which results in a column super-cell. This super-cell is then repeated NL times along the SiNR length (L), resulting in the total nanoribbon Hamiltonian matrix that contains $(4 \cdot NW \cdot NL)^2$ elements.

The focus of this study is on the bandgap, which can be extracted from the bandstructure, and more importantly on the transport gap, which is obtained from the energy-resolved quantum transmission function, $T(E)$, as the energy range where $T(E) < 0.001$. The E_{TG} is related to transport properties of a finite device and is therefore a more relevant figure-of-merit for potential device applications of ultra-scaled SiNRs. Transmission function of SiNRs is calculated within our existing NEGF code that deals with quasi-1D nanostructures such as nanoribbons of 2D materials [17]–[20]. The NEGF method starts by finding the retarded Green's function of the device, defined as

$$G^R(E) = \left[\left(E + i0^+\right) I - H - \Sigma_{C1}^R(E) - \Sigma_{C2}^R(E)\right]^{-1} \quad (1)$$

for a two-terminal nanodevice described by the TB Hamiltonian matrix H, while $\Sigma_C{}^R$ matrices are the retarded contact self-energies [21], [22]. The $\Sigma_C{}^R$ matrices account for open boundary conditions between the nanodevice, i.e. SiNR of a given size, and two electrodes. After calculating the retarded Green's function of the device, the transmission is found from

$$T(E) = Tr\left[\Gamma_{C1}(E)\, G^R(E)\, \Gamma_{C2}(E)\, G^A(E)\right] \quad (2)$$

where $G^A(E)$ is the advanced Green's function of the device, i.e. $G^A(E) = G^R(E)^\dagger$, and $\Gamma_C(E) = -2\,\mathrm{Im}\Sigma_C(E)$ is the contact-induced broadening for each electrode.

In the case of ideal contacts, the resulting transmission is a step-like curve, and $\Sigma_C{}^R$ matrices need to be calculated with an iterative algorithm, e.g. Sancho-Rubio method [22], [23]. On the other hand, in this work we are interested in how E_{TG} changes in SiNRs with metal edge contacts (MECs). The MECs are included by using the wide-band limit (WBL) approximation for reservoirs, which amounts to assuming metals with a constant density of states (DOS) and a constant metal-device coupling or hopping parameter (t_{MD}) [22]. The resulting $\Sigma_C{}^R$ matrices contain a single negative purely complex parameter (marked hereafter as $-\mathrm{Im}\Sigma_C{}^R$) since we neglect the real part and any energy shifts caused by attaching the metal electrodes. In order to assess a realistic magnitude of $-\mathrm{Im}\Sigma_C{}^R$ for an initial description of WBL contacts we use $\mathrm{DOS}(E = E_F) \sim 0.25\ \mathrm{eV}^{-1}$ similar to DOS of Au at Fermi level [24], and assume t_{MD} of 3 eV close to the coupling parameter for a graphene TB model. Hence, we obtain $-\mathrm{Im}\Sigma_C{}^R = t_{MD}{}^2 \cdot \mathrm{DOS}(E_F)/2 \approx 1.1$ eV and use $-\mathrm{Im}\Sigma_C{}^R = 1$ eV for initial simulations.

Later in the paper we examine the impact of reducing $-\mathrm{Im}\Sigma_C{}^R$ down to 0.1 eV, thus modeling a less interacting metal, and increasing $-\mathrm{Im}\Sigma_C{}^R$ up to 10 eV, which mimics metal electrodes with strong interaction towards the SiNR structure. The wide range of $-\mathrm{Im}\Sigma_C{}^R$ values that are explored comes from a study of graphene-metal and carbon nanotube-metal contacts that reported WBL interaction parameters ranging from ~0.01 eV for Pd electrodes to ~20 eV for Ni contacts [25], [26]. The wide range is a consequence of various DOS values at Fermi level in different metals, and of various coupling parameters between metals and graphene. In this work concerning SiNRs, the effects of attaching MECs with different interaction strengths $-\mathrm{Im}\Sigma_C{}^R$ are studied for various nanoribbon widths and lengths to assess the potential design space, i.e. allowed geometries of SiNRs in terms of the transport gap.

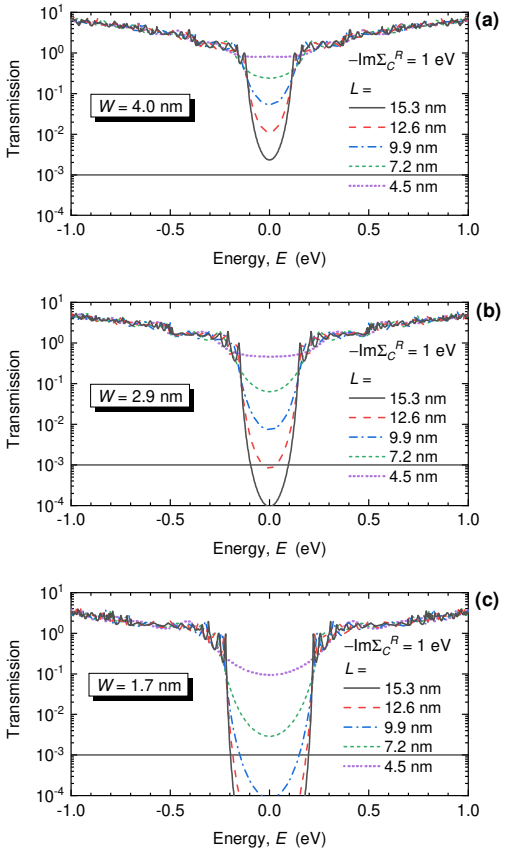


Fig. 2. Transmission vs. energy for (a) 4.0 nm-, (b) 2.9 nm-, and (c) 1.7 nm-wide SiNRs of different lengths with the metal-device interaction parameter of 1 eV.

III. Results and Discussion

First, we investigate the impact of attaching WBL metal edge contacts, with $-\mathrm{Im}\Sigma_C{}^R = 1$ eV as interaction parameter, on the transmission of 15.3 nm to 4.5 nm long SiNRs. Transmission characteristics for SiNR widths of 4.0 nm, 2.9 nm and 1.7 nm are plotted in Fig. 2a, b and c, respectively. Downscaling of SiNR width increases the bandgap and, consequently, the transport gap, which is easily visible for 15.3 nm-long structures in the central energy range where the transmission is suppressed considerably. However, the transmission does not cross the defined E_{TG} limit of $T(E) < 0.001$ in all cases. For $W = 1.7$ nm, E_{TG} exists for SiNRs that are at least 9.9 nm long. When L equals 4.5 nm and 7.2 nm the transmission is ~0.01 and ~0.1 at mid-gap, i.e. transport gap effectively closes and these 1.7 nm-wide SiNRs become metallized.

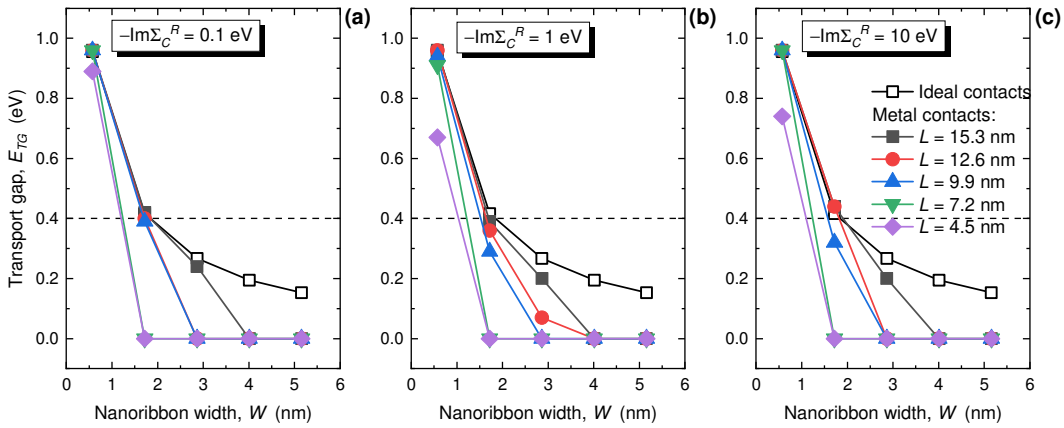


Fig. 3. Impact of SiNR width and length downscaling on the transport gap in SiNRs with metal edge contacts for various metal-device interaction strengths, i.e. with $-\mathrm{Im}\Sigma_C^R$ values of (a) 0.1 eV, (b) 1 eV, and (c) 10 eV. The characteristic of a SiNR with ideal contacts is inserted for comparison.

For $W = 2.9$ nm, metallization occurs even for longer nanoribbons so only for $L = 15.3$ nm and $L = 12.6$ nm a clear E_{TG} can be extracted. Strikingly, a clear transport gap completely disappears in the widest SiNR, with $W = 4.0$ nm shown in Fig. 2a, irrespective of the length. The widest structures with the smallest bandgap are, hence, more susceptible to metallization effects than the narrower ones, as reported previously for graphene and other 2D material nanoribbons [11], [13], [27]. In the shortest ($L = 4.5$ nm) 2.9 nm- and 4.0 nm-wide SiNRs the transmission at mid-gap is on the order of ~1, which indicates unacceptable tunneling and very poor switching between the ON and OFF states in SiNR FETs for logic applications. Therefore, attaching metal edge contacts with a realistic interaction parameter of $-\mathrm{Im}\Sigma_C^R = 1$ eV causes significant changes in the transmission, can close the transport gap, and sets limits to acceptable nanoribbon sizes for applications in nanodevices.

The extracted E_{TG} is reported in Fig. 3 for various SiNR widths and lengths, and for different metal-device interaction parameters to enable an easier comparison later in the text. The analyzed values of $-\mathrm{Im}\Sigma_C^R$ are 0.1, 1 and 10 eV and the corresponding results are presented in Fig. 3a, b and c, respectively. First, we focus on the initial value of the interaction parameter, i.e. $-\mathrm{Im}\Sigma_C^R = 1$ eV as shown in Fig. 3b. When SiNR width is scaled from 5.2 nm down to 0.6 nm, E_G increases from 0.15 eV to 0.96 eV in SiNRs with ideal contacts where $E_{TG} = E_G$. However, when MECs are attached, E_{TG} generally decreases with decreasing length, irrespective of the width. The 5.2 nm- and 4.0 nm-wide SiNRs with MECs are metallized regardless of the length. For $W = 2.9$ nm, the SiNR is semiconducting only for $L \geq 12.6$ nm, end even for these lengths the E_{TG} decreases from 0.27 eV (ideal contacts), over 0.2 eV ($L = 15.3$ nm), to 0.07 eV ($L = 12.6$ nm). The metallization effects become weaker for the two narrowest SiNRs with the widths of 1.7 nm and 0.6 nm. For $W = 1.7$ nm, only the shortest ($L = 4.5$ nm) nanoribbon has $E_{TG} = 0$ eV, whereas for $W = 0.6$ nm all SiNRs exhibit a transport gap irrespective of nanoribbon length. For the narrowest device the only significant effect of attaching MECs is E_{TG} decrease, or bandgap narrowing, from 0.96 eV to 0.67 eV when $L = 4.5$ nm (i.e. change of −30%). We can summarize the observed effects as metallization and bandgap narrowing being the strongest in wider and shorter nanodevices. The wide SiNRs have a small bandgap that is more susceptible to metallization-induced narrowing, whereas in short SiNRs metal electrodes are close and their interaction induces high local density of states and closes the gap [11]. The plot in Fig. 3 also contains a horizontal line showing $E_G = 0.4$ eV, which is the minimum acceptable bandgap for future extremely-scaled FETs [28]. Clearly, for technology nodes where channel lengths under ~10 nm are needed, only the narrowest ~0.6 nm-wide SiNR offers an appropriate bandgap. Although SiNRs have been experimentally reported [5], [29], [30], fabricating sub-1 nm-wide SiNRs on a wafer scale would present a serious challenge.

Next, different metals are studied and their impact on E_{TG} is shown in Fig. 3a for $-\mathrm{Im}\Sigma_C^R = 0.1$ eV, and in Fig. 3c for $-\mathrm{Im}\Sigma_C^R = 10$ eV. The former value describes a weakly-interacting metal or a moderately-interacting contacts with defects and other nonidealities on the metal-device interface that decrease the effective $-\mathrm{Im}\Sigma_C^R$. Surprisingly, Fig. 3a shows that an order of magnitude weaker interaction does not lead to significantly weaker metallization effects in comparison to Fig. 3b. In comparison to the case when $-\mathrm{Im}\Sigma_C^R = 1$ eV, for $-\mathrm{Im}\Sigma_C^R = 0.1$ eV the transport gap vanishes in one additional device, in the 2.9 nm-wide and 12.6 nm-long SiNR. For other devices with a preserved energy gap the impact of MECs is somewhat weaker for $-\mathrm{Im}\Sigma_C^R = 0.1$ eV than for the initial case, but E_{TG} is still lower than in ideal devices. Increasing the metal-device interaction parameter to 10 eV (Fig. 3c), which models a strongly-interacting metal, leads to similar effects as discussed previously. In comparison to the case of $-\mathrm{Im}\Sigma_C^R = 1$ eV, E_{TG} goes to zero for the SiNR with $W = 2.9$ nm and $L = 12.6$ nm. Other nanoribbons experience similar E_{TG} changes as in

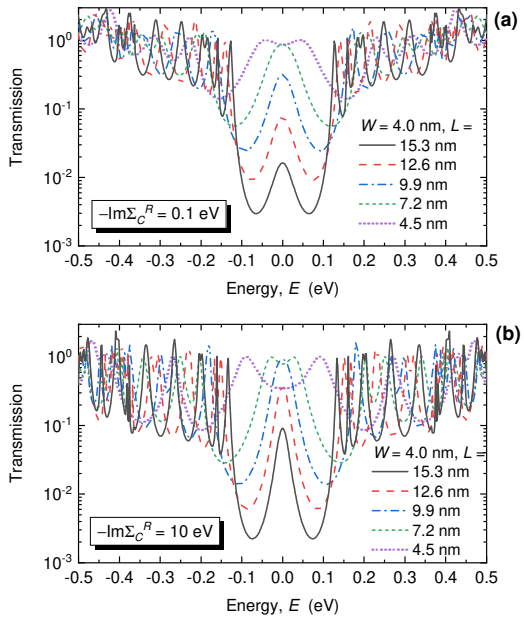


Fig. 4. Transmission of a 4.0 nm-wide SiNR of different lengths with the metal-device interaction parameter being (a) 0.1 eV, and (b) 10 eV.

the initial case of moderately-interacting metal electrodes defined by $-\mathrm{Im}\Sigma_C{}^R = 1$ eV.

The properties of E_{TG} in SiNR with weaker and stronger interactions are unexpected since one would expect a lesser E_{TG} change in Fig. 3a, and a larger impact of MECs in Fig. 3c, when compared to data presented in Fig. 3b for $-\mathrm{Im}\Sigma_C{}^R = 1$ eV. In order to clarify this issue, we plot the transmission for the 4.0 nm-wide SiNR with different lengths in the case of $-\mathrm{Im}\Sigma_C{}^R = 0.1$ eV in Fig. 4a and $-\mathrm{Im}\Sigma_C{}^R = 10$ eV in Fig. 4b. Both plots show that the transmission is greatly suppressed in the conduction and valence bands, with sharp transmission peaks occurring at specific energies. This result stands in stark contrast to step-like transmission characteristics of nanoribbons with ideal contacts, which is a signature of 1D and quasi-1D nanostructures [12], [26], [27]. In addition, the transmission increases inside the bandgap with decreasing nanoribbon length, in both cases. A similar effect was reported for graphene nanoribbons, with a pronounced peak at mid-gap that belongs to a strongly localized zero-energy state, i.e. metal-induced gap states (MIGS) are formed [12], [27]. This mid-gap increase due to MIGS is especially detrimental for SiNR FET operation as it would enhance direct tunneling through the channel and make the SiNR FET unable to turn off adequately.

As shown in supporting information of [13], assuming constant imaginary contact self-energies leads to Lorentzian-shaped transmission peaks at nanosystem eigenenergies. For example, the transmission of a simple one-level nanostructure equals

$$T(E) = \frac{1}{1+\left(E/\Gamma\right)^2} \qquad (3)$$

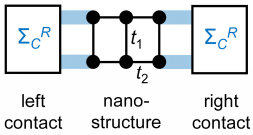


Fig. 5. Illustration of a quasi-1D nanostructure used for transmission analysis with respect to $-\mathrm{Im}\Sigma_C{}^R$. $t_1 = -3$ eV, $t_2 = -1$ eV.

for zero local energy, and with WBL contacts with $\Sigma_{C1} = \Sigma_{C2} = -\mathrm{i}\,\Gamma/2$. Hence, in realistic nanosystems with many eigenenergies, such as SiNRs, one can expect a transmission characteristic with a series of Lorentzians if WBL MECs are attached, which is exactly what is seen in Fig. 4. In devices such as SiNRs the individual Lorentzians are closely spaced due to a large number of states. Consequently, this crowding leads to overlaps and the local constructive interference causes transmission increase at certain energies.

Since SiNR is a relatively complicated system, it is easier to understand the evolution of transmission and E_{TG} with increasing $-\mathrm{Im}\Sigma_C{}^R$ using a 1D toy example. In the following analysis we assume a rectangular nanostructure that consists of two parallel 1D atomic chains, as illustrated in Fig. 5. The structure is a 2-by-3-atom nanoribbon with two WBL MECs that interact only with the two edge atoms on either side. For this system we increase the interaction parameter from 0.01 eV to 10 eV and plot the resulting transmissions in Fig. 6. For ideal contacts, the results assert that the nanostructure is a semiconductor with a 2 eV-wide bandgap and 4 eV-wide conduction and valence bands. However, with MECs the transmission changes significantly, with its shape evolving from delta-functions (Fig. 6a), over multiple Lorentzians in the bands (e.g. Fig. 6c), to a single Lorentzian in each band (e.g. Fig. 6h). The delta-functions, obtained for very weakly interacting MECs with $-\mathrm{Im}\Sigma_C{}^R = 0.01$ eV, indicate the localization or resonant regime with negligible transport probability, which means that the nanostructure is nearly isolated. Increasing $-\mathrm{Im}\Sigma_C{}^R$ to at least 0.1 eV increases the transmission to non-negligible levels. The six peaks that are visible in Fig. 6a-d are due to six states in the nanostructure. From Fig. 6e onwards, i.e. for $-\mathrm{Im}\Sigma_C{}^R$ equal to and higher than 1.3 eV, the separate peaks merge into one peak per band. This merging occurs because the contribution of contacts to system eigenvalues becomes dominant, which is seen in Fig. 7 that reports the real and imaginary parts of eigenvalues of the effective device Hamiltonian, $H_{eff} = H_0 + \Sigma$. The real part indicates where the transmission peaks occur, whereas the imaginary part defines the lifetime of a state via $\tau \sim \hbar / \mathrm{Im}(\varepsilon)$, where ε is a complex eigenvalue. The single peaks from Fig. 6e onwards become narrower as the interaction strength increases, in accordance with Eq. (3). For $-\mathrm{Im}\Sigma_C{}^R = 10$ eV (Fig. 6i) the Lorentzians are similar in shape to those obtained for $-\mathrm{Im}\Sigma_C{}^R = 0.1$ eV (Fig. 6b), but with a decreased number of peaks in the former case also comes an increased ETG of ~4 eV, in comparison to

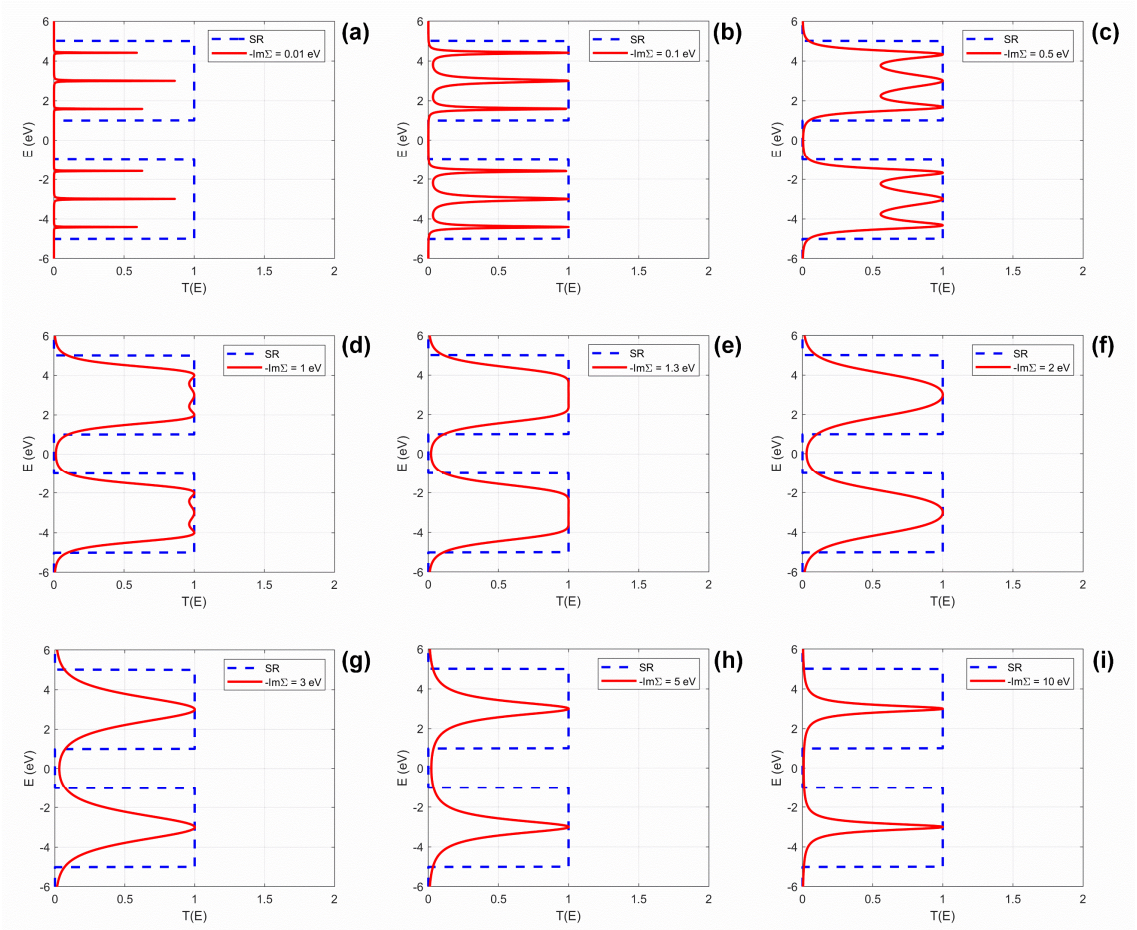


Fig. 6. Comparison of transmission functions of the device illustrated in Fig. 5 with ideal contacts (dashed line, obtained by the Sancho-Rubio (SR) method) and with metal edge contacts (full line). The nine plots show the impact of increasing $-\mathrm{Im}\Sigma_C^R$ from 0.01 eV to 10 eV.

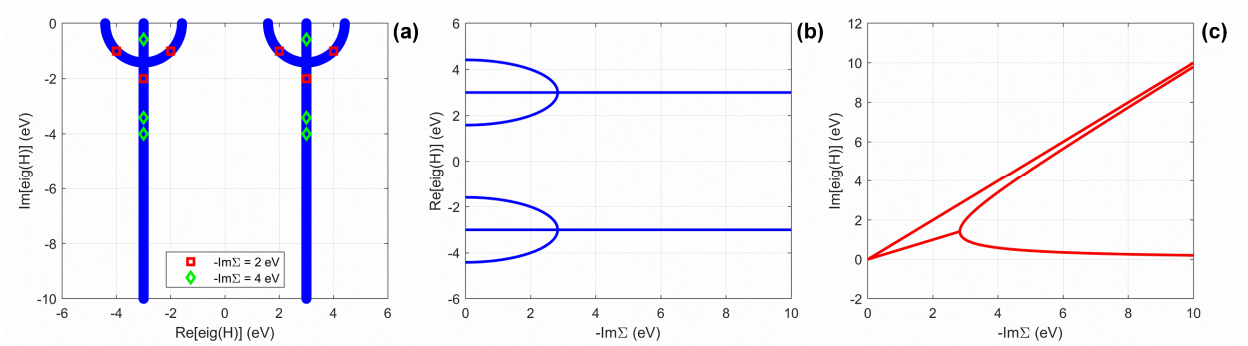


Fig. 7. (a) Eigenvalues in the complex plane for the effective Hamiltonian of the structure illustrated in Fig. 5 for various metal-device interaction strengths. Dependence of the (b) real and (c) imaginary part of the eigenvalues on $-\mathrm{Im}\Sigma_C^R$.

~3 eV when $-\mathrm{Im}\Sigma_C^R = 0.1$ eV, and 2 eV in SiNRs with ideal contacts. Therefore, this simplified analysis agrees with the findings on differences between Fig. 3a and c, i.e. stronger ETG narrowing is observed in the case of strongly-interacting metal edge contacts.

IV. Conlusions

Quantum transport simulations are employed to study metallization effects and bandgap narrowing in SiNRs with metal edge contacts. The analysis is done for various metal-device interaction strengths, and for technologically relevant SiNR dimensions, i.e. W under ~5 nm to reach relevant bandgap values, and L under ~15 nm for channels of future ultra-scaled FETs. We show that even moderately-interacting MECs ($-\mathrm{Im}\Sigma_C^R = 1$ eV) significantly suppress the transmission in the conduction and valence bands, increase the transmission inside the bandgap, cause considerable bandgap narrowing, and even metallization ($E_{TG} = 0$ eV). These MEC-induced effects set severe limits to scaling of SiNRs as channel material for FETs, e.g. SiNRs with $W > 3$ nm become metallized even when their length is ~15 nm. In addition, we demonstrate that weakly-interacting metals

($-\mathrm{Im}\Sigma_C^R << 1$ eV) could lead to localization regime with very poor transport properties and semiconductor-to-insulator transition. On the other hand, strongly-interacting metals ($-\mathrm{Im}\Sigma_C^R >> 1$ eV) can either decrease or further widen the transport gap, depending on the nanostructure and interaction strength. Overall, and specifically for SiNRs, only the sub-1 nm-wide devices are capable for length scaling down to ~5 nm due to the strong bandgap narrowing. These findings can be extended to other nanostructures with similar electronic properties and, therefore, could prove to be important for practical applications of 2D materials in nanoscale electron devices.

ACKNOWLEDGMENTS

This work was supported by the Croatian Science Foundation under the project CONAN2D (Grant No. UIP-2019-04-3493).

REFERENCES

[1] P. Vogt *et al.*, "Silicene: Compelling Experimental Evidence for Graphenelike Two-Dimensional Silicon," *Phys. Rev. Lett.*, vol. 108, no. 15, p. 155501, Apr. 2012, doi: 10.1103/PhysRevLett.108.155501.

[2] A. Kara *et al.*, "A review on silicene — New candidate for electronics," *Surface Science Reports*, vol. 67, no. 1, pp. 1–18, Jan. 2012, doi: 10.1016/j.surfrep.2011.10.001.

[3] K. S. Novoselov *et al.*, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, no. 5696, pp. 666–669, Oct. 2004, doi: 10.1126/science.1102896.

[4] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene," *Rev. Mod. Phys.*, vol. 81, no. 1, pp. 109–162, Jan. 2009, doi: 10.1103/RevModPhys.81.109.

[5] P. De Padova *et al.*, "Evidence of graphene-like electronic signature in silicene nanoribbons," *Applied Physics Letters*, vol. 96, no. 26, pp. 261905-261905–3, Jun. 2010, doi: doi:10.1063/1.3459143.

[6] B. van den Broek, M. Houssa, A. Lu, G. Pourtois, V. Afanas'ev, and A. Stesmans, "Silicene nanoribbons on transition metal dichalcogenide substrates: Effects on electronic structure and ballistic transport," *Nano Res.*, vol. 9, no. 11, pp. 3394–3406, Nov. 2016, doi: 10.1007/s12274-016-1217-4.

[7] S. Kaneko, H. Tsuchiya, Y. Kamakura, N. Mori, and M. Ogawa, "Theoretical performance estimation of silicene, germanene, and graphene nanoribbon field-effect transistors under ballistic transport," *Appl. Phys. Express*, vol. 7, no. 3, p. 035102, Mar. 2014, doi: 10.7567/APEX.7.035102.

[8] M. A. Kharadi, G. F. A. Malik, K. A. Shah, and F. A. Khanday, "Sub-10-nm Silicene Nanoribbon Field Effect Transistor," *IEEE Transactions on Electron Devices*, vol. 66, no. 11, pp. 4976–4981, Nov. 2019, doi: 10.1109/TED.2019.2942396.

[9] M. Poljak, "Length-Dependent Electron Transport Properties of Defective Silicene Nanoribbons," *IEEE Transactions on Nanotechnology*, vol. 19, pp. 315–321, 2020, doi: 10.1109/TNANO.2020.2982071.

[10] M. Poljak, "Impact of Width Scaling and Parasitic Series Resistance on the Performance of Silicene Nanoribbon MOSFETs," *IEEE Trans. Electron Devices*, vol. 67, no. 11, pp. 4705–4708, Nov. 2020, doi: 10.1109/TED.2020.3017465.

[11] M. Poljak and M. Matić, "Metallization-Induced Quantum Limits of Contact Resistance in Graphene Nanoribbons with One-Dimensional Contacts," *Materials*, vol. 14, no. 13, p. 3670, Jan. 2021, doi: 10.3390/ma14133670.

[12] G. Liang, N. Neophytou, M. S. Lundstrom, and D. E. Nikonov, "Ballistic graphene nanoribbon metal-oxide-semiconductor field-effect transistors: A full real-space quantum transport simulation," *Journal of Applied Physics*, vol. 102, no. 5, p. 054307, Sep. 2007, doi: 10.1063/1.2775917.

[13] M. Poljak, M. Matić, T. Župančić, and A. Zeljko, "Lower Limits of Contact Resistance in Phosphorene Nanodevices with Edge Contacts," *Nanomaterials*, vol. 12, no. 4, Art. no. 4, Jan. 2022, doi: 10.3390/nano12040656.

[14] L. Wang *et al.*, "One-Dimensional Electrical Contact to a Two-Dimensional Material," *Science*, vol. 342, no. 6158, pp. 614–617, Nov. 2013, doi: 10.1126/science.1244358.

[15] A. Jain *et al.*, "One-Dimensional Edge Contacts to a Monolayer Semiconductor," *Nano Lett.*, vol. 19, no. 10, pp. 6914–6923, Oct. 2019, doi: 10.1021/acs.nanolett.9b02166.

[16] S. Cahangirov, M. Topsakal, and S. Ciraci, "Armchair nanoribbons of silicon and germanium honeycomb structures," *Phys. Rev. B*, vol. 81, no. 19, p. 195120, May 2010, doi: 10.1103/PhysRevB.81.195120.

[17] M. Poljak, E. B. Song, M. Wang, T. Suligoj, and K. L. Wang, "Influence of Edge Defects, Vacancies, and Potential Fluctuations on Transport Properties of Extremely Scaled Graphene Nanoribbons," *IEEE Trans. Electron Devices*, vol. 59, no. 12, pp. 3231–3238, Dec. 2012, doi: 10.1109/TED.2012.2217969.

[18] M. Poljak and T. Suligoj, "Immunity of electronic and transport properties of phosphorene nanoribbons to edge defects," *Nano Res.*, vol. 9, no. 6, pp. 1723–1734, Jun. 2016, doi: 10.1007/s12274-016-1066-1.

[19] M. Poljak, M. Glavan, and S. Kuzmić, "Accelerating simulation of nanodevices based on 2D materials by hybrid CPU-GPU parallel computing," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, May 2019, pp. 47–52. doi: 10.23919/MIPRO.2019.8756964.

[20] M. Matić, T. Župančić, and M. Poljak, "Parallelized Ab Initio Quantum Transport Simulation of Nanoscale Bismuthene Devices," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, May 2022, pp. 118–123. doi: 10.23919/MIPRO55190.2022.9803335.

[21] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1997.

[22] M. Pourfath, *The Non-Equilibrium Green's Function Method for Nanoscale Device Simulation*. Wien: Springer-Verlag, 2014. doi: 10.1007/978-3-7091-1800-9.

[23] M. P. L. Sancho, J. M. L. Sancho, J. M. L. Sancho, and J. Rubio, "Highly convergent schemes for the calculation of bulk and surface Green functions," *J. Phys. F: Met. Phys.*, vol. 15, no. 4, p. 851, 1985, doi: 10.1088/0305-4608/15/4/009.

[24] K. Schwarz and P. Blaha, "DFT calculations of solids in the ground state," in *DFT calculations of solids in the ground state*, De Gruyter, 2018, pp. 67–100. doi: 10.1515/9783110433920-003.

[25] A. Fediai, D. A. Ryndyk, and G. Cuniberti, "The modular approach enables a fully ab initio simulation of the contacts between 3D and 2D materials," *J. Phys. Condens. Matter*, vol. 28, no. 39, p. 395303, 05 2016, doi: 10.1088/0953-8984/28/39/395303.

[26] A. Fediai *et al.*, "Towards an optimal contact metal for CNTFETs," *Nanoscale*, vol. 8, no. 19, pp. 10240–10251, May 2016, doi: 10.1039/C6NR01012A.

[27] G. Liang, N. Neophytou, M. S. Lundstrom, and D. E. Nikonov, "Contact Effects in Graphene Nanoribbon Transistors," *Nano Lett.*, vol. 8, no. 7, pp. 1819–1824, Jul. 2008, doi: 10.1021/nl080255r.

[28] F. Schwierz, J. Pezoldt, and R. Granzner, "Two-dimensional materials and their prospects in transistor electronics," *Nanoscale*, vol. 7, no. 18, pp. 8261–8283, Apr. 2015, doi: 10.1039/C5NR01052G.

[29] N. F. Kleimeier *et al.*, "Unoccupied electronic band structure of pentagonal Si nanoribbons on Ag(110)," *Phys. Chem. Chem. Phys.*, vol. 21, no. 32, pp. 17811–17820, Aug. 2019, doi: 10.1039/C9CP02239B.

[30] S. Yue *et al.*, "Observation of One-Dimensional Dirac Fermions in Silicon Nanoribbons," *Nano Lett.*, vol. 22, no. 2, pp. 695–701, Jan. 2022, doi: 10.1021/acs.nanolett.1c03862.

Determining Graphene and Substrate Quality from the Coupled Hall Mobility Measurements and Theoretical Modeling

K. Japec[1], M. Matić[1], R. Lukose[2], M. Lisker[2,3], M. Lukosius[2,**] and M. Poljak[1,*]

[1]Computational Nanoelectronics Group, Micro and Nano Electronics Laboratory, Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia
[2]2D Materials Group, IHP–Leibnitz Institute for High Performance Microelectronics, 15236 Frankfurt (Oder), Germany
[3]Technical University of Applied Science Wildau, 15745 Wildau, Germany
[*]E-mail: mirko.poljak@fer.hr, [**]E-mail: lukosius@ihp-microelectronics.com

***Abstract*—An increase of mobility up to ~2600 cm²/Vs is observed in graphene by Hall bar characterization within the temperature range from 40 K to 300 K. The increasing trend is attributed to Coulomb scattering by employing theoretical modeling based on the momentum relaxation time approximation of the Boltzmann transport equation. We also find that at room temperature and for higher charge densities additional mechanisms such as lattice defect and/or substrate corrugation scattering become important and restrict the mobility down to only ~200 cm²/Vs and carrier mean free paths well under ~35 nm.**

***Keywords*—graphene, mobility, scattering rate, Coulomb impurity, corrugations, defects, surface roughness, momentum relaxation time (MRT) approximation**

I. Introduction

The high carrier mobility and its atomically-thin crystal structure make graphene a promising two-dimensional (2D) material for applications in electronics, photonics and other engineering fields [1]–[3]. However, successful integration of graphene monolayers into CMOS is limited by the lack of cost-efficient growth methods on CMOS compatible substrates. Recently, an appropriate substrate for growing graphene was found in Ge, mainly due to very low solubility of carbon [4]. While Ge(110) is superior in terms of graphene grain formation and mobility [4], Ge(001) is more compatible with the standard Si(001) wafers and process flow [5], [6]. For graphene growth on Ge(001) substrates, carrier mobility ranging from a few ~100 cm²/Vs [7] to ~10,000 cm²/Vs [8] at room temperature and low carrier density is reported. However, in the latter case it is unclear what limits the mobility at low densities, and transport features at high carrier densities remained unknown.

Complementary to experiments, theoretical modeling and numerical simulations of carrier transport can provide physical insights into material and device properties [9]–[11], which allows an optimization of processing parameters. In this work, we employ the momentum relaxation time (MRT) approximation of the Boltzmann transport equation to calculate electron/hole mobility for all relevant scattering mechanisms ranging from phonon, over Coulomb to substrate roughness scattering. Calibration on temperature-dependent measurements allows the extraction of graphene and substrate quality parameters such as lattice defect or charged impurity density. We report that the mobility increases with increasing temperature from 40 K to 300 K, which is attributed to the dominant impact of Coulomb scattering, and that it reaches ~2600 cm²/Vs at $T = 300$ K, being limited by substrate roughness/corrugations and/or defect scattering within graphene.

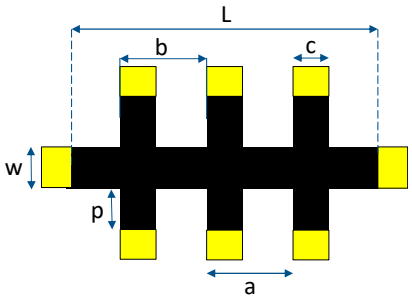


Fig. 1. Schematic illustration of 8-contact Hall bar devices. $L = 20$ µm, $a = 5$ µm, $b = 5$ µm, $c = 2$ µm, $w = 6$ µm, $p = 2$ µm.

II. Methodology

A. Fabrication and Characterization

The fabrication process of 8-contact Hall bar devices (Fig. 1) from chemical vapor deposition (CVD) grown graphene involves several steps. Firstly, a layer of graphene is grown on a Ge/Si(001) substrate using the CVD technique at 885 °C. The graphene layer is then transferred onto a target substrate (SiO_2 in this work) using a polymer-based electrochemical delamination transfer method. Next, lithography is used to define the device geometry, followed by etching of the unwanted graphene using reactive ion etching (RIE) process. After graphene is patterned, Au metal contacts are deposited using a thermal evaporation technique, followed by lift-off procedure of the polymer resist. Finally, the devices were annealed at

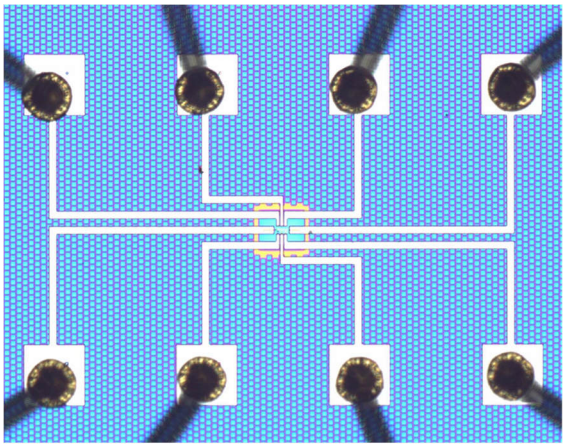


Fig. 2. Photograph under an optical microscope of the 8 contact Hall bar device under study.

250 °C temperature to improve the contact resistance between the metal contacts and the graphene monolayer.

Hall bar measurements in the range of 40 – 300 K in the magnetic field up to 1 Tesla were used to investigate the electronic and transport properties of graphene Hall bar devices (Fig. 2). The Hall effect occurs when a magnetic field is applied perpendicularly to the direction of current flow in a conductor. This creates a transverse voltage across the conductor, which is proportional to the strength of the magnetic field and the carrier density of the material. By measuring the Hall voltage as a function of magnetic field strength and temperature, information about the carrier mobility, carrier sheet density, and carrier type was extracted. The range of 40 – 300 K is chosen to study the material behavior at both room temperature and lower temperatures where quantum mechanical effects related to Dirac transport become more pronounced.

B. Theory and Numerical Simulations

The main goal of this paper is to find the mechanisms responsible for the increase of mobility with temperature in the fabricated graphene samples, and to estimate quality parameters of graphene and substrate upon which it is deposited. We calculate carrier mobility by employing the MRT approximation of the Boltzmann transport equation [9], [10], [12], thus building on our previous work [13], especially concerning quasi-one-dimensional graphene nanoribbons [14], [15]. Mobility in graphene is limited by various scattering mechanisms that are more or less significant for the overall transport depending on e.g. temperature and carrier density. In this work, we implement all relevant mechanisms such as scattering on intrinsic acoustic phonons (AP) [16] and optical phonons (OP) [10] in graphene, surface optical phonons (SOP) [17] that originate from the dielectric substrate, Coulomb (CO) scattering originating from charged impurities on the substrate [11], atomic-scale lattice defect (DEF) scattering [18] from missing carbon atoms, and surface roughness or substrate corrugation (COR) scattering [11].

The expressions for MRT spectra are given in the literature, and here we explicitly define only the two most relevant ones. For Coulomb impurity scattering the energy-dependent CO scattering rate is given as in [11], i.e.

$$\frac{1}{\tau_{CO}} = \frac{N_{imp}}{4\pi\hbar E}\left(\frac{Ze^2}{\varepsilon}\right)^2 \int_0^{\pi/2} d\theta \frac{\sin^2(\theta)\, e^{-k\sin(\theta)d_s}}{\left[\sin(\theta)+\frac{q_S}{2k}\right]^2} \quad (1)$$

where N_{imp} represents the aerial density of charged impurities, q_S is the screening wave vector, while d_S is the distance between impurities, i.e. the substrate, and the graphene sheet. The screening function considered in this work is static and equals

$$q_s = \frac{e^2 E(k)}{2\pi\varepsilon(\hbar v_F)^2} \quad (2)$$

where ε is graphene dielectric constant, and v_F is the Fermi velocity in graphene. Screening is taken into account for the CO and SOP MRT calculations. Another important scattering mechanism defined here is the substrate corrugation scattering which is assumed to arise after graphene sheet is deposited on the (oxide) substrate that is not perfectly flat. The energy-dependent COR scattering rate is given as in [11], i.e.

$$\frac{1}{\tau_{COR}(E)} = \frac{2k}{v_F}\left(\frac{\Delta L e^2 n_S}{2\hbar\varepsilon}\right)\frac{1}{\sqrt{1+k^2L^2}}E\left(\frac{kL}{\sqrt{1+k^2L^2}}\right) \quad (3)$$

where n_S is the gate-induced sheet carrier density, Δ is the rms height and L is the correlation length for the corrugations or roughness of the substrate. In the experiments, the n_S is found from the current, magnetic field and Hall voltage, whereas in the simulations the n_S is defined by Fermi level shift in graphene.

After each MRT spectrum is known, the mobility limited by each of the different mechanisms is calculated by energy averaging [10], [14], [19]. For example, the COR-limited mobility is designated as μ_{COR}, the SOP-limited mobility is designated as μ_{SOP}, and so on. After calculating these components, the total mobility is calculated using the Matthiessen's rule defined as

$$\frac{1}{\mu_{TOT}} = \sum_m \frac{1}{\mu_m}. \quad (4)$$

where m presents individual scattering mechanisms. Unless stated otherwise, the constants and parameters used in the simulations are those defined in the cited literature.

III. Results and Discussion

Figure 3 shows the fitting between the measured and simulated mobility in the temperature range from 40 K to 300 K. The simulated curve is obtained for the charge carrier density $n = 1.867 \times 10^{12}$ cm^{-2}, close to carrier density maintained in Hall measurements. Other simulation parameters set to achieve the fitting in Fig. 3 are as follows: density of impurities $N_{imp} = 2.18 \times 10^{11}$ cm^{-2} (for CO scattering), defect concentration $N_{def} = 0.18 \times 10^{10}$ cm^{-2} (for DEF scattering), and rms height $\Delta = 1.3$ nm and the correlation length $L = 2$ nm (for COR scattering). The simulated mobility matches the experimental data well, especially for $T > 100$ K, while for lower temperatures the mismatch is smaller than 8.5%. We attribute this difference visible for $T < 100$ K to simplifications in scattering modeling such as equipartition approximation for acoustic phonons, absence of an incomplete ionization model, using a very simple screening model etc.

Generally, mobility increases with temperature with the values of the measured mobility ranging from 2226 cm^2/Vs to 2626 cm^2/Vs, while the simulated mobility being in the range from 2390 cm^2/Vs at $T = 40$ K to 2640 cm^2/Vs at $T = 300$ K. Given this increase with temperature, only two mechanisms could be responsible for the overall mobility behavior, i.e. CO and/or COR scattering. In the case of CO scattering, increasing T increases the average kinetic energy of carriers, which in turn decreases CO-MRT rate as given in Eq. (1), increases screening according to Eq. (2), with both effects resulting in μ_{CO} increase with temperature. In order to clarify this issue, in Fig. 4 we plot the impact of temperature increase on the individual mechanisms and we can see how they contribute to the total mobility according to the Matthiessen's rule. Clearly, CO scattering, followed by COR scattering, exhibit the strongest impact on the total mobility, and are responsible for the overall increase of μ_{TOT} with increasing temperature. In contrast, other mechanisms such as AP, OP, SOP and DEF scattering become stronger at higher temperatures. If, for example, SOP or DEF scattering become dominant in certain graphene samples, that would be seen from a moderately decreasing μ_{TOT} with increasing temperature. The μ_{OP} is several orders of magnitude higher than mobilities limited by other sources of scattering and is, therefore, of little significance for electron transport in our samples.

It is worth looking into MRT spectra to investigate the relative influence of different mechanisms of scattering at room temperature, as reported in Fig. 5 for $T = 300$ K. We observe that at low (kinetic) energies the strongest mechanism is the CO scattering with rates as high as ~10^{14} s^{-1}, whereas at energies $E > 0.17$ eV the COR scattering dominates over all other scattering mechanisms. For $E > 0.5$ eV the CO scattering becomes the weakest

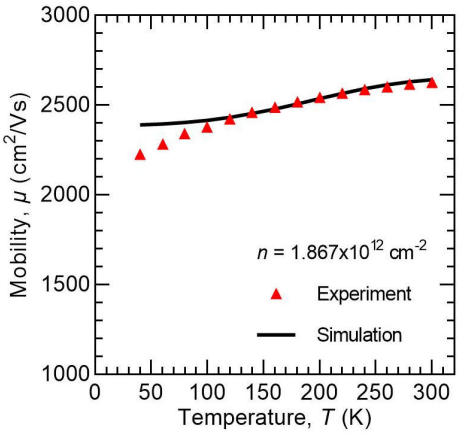


Fig. 3. Comparison between the measured and simulated temperature-dependent electron mobility in the range 40 – 300 K.

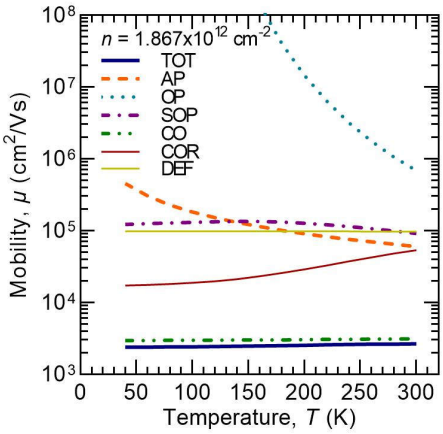


Fig. 4. Temperature-dependence of individual mobilities limited by different scattering mechanisms.

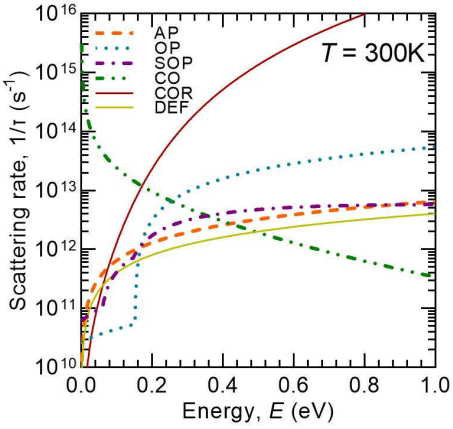


Fig. 5. Scattering rate spectra at 300 K for all the included scattering mechanisms.

mechanism, with the lowest rate of ~4×10^{11} s^{-1} at $E = 1$ eV, and the weakest influence on μ_{TOT}. We note that for all spectra, except for the CO mechanism, the scattering rate increases with energy and most mechanisms reach the

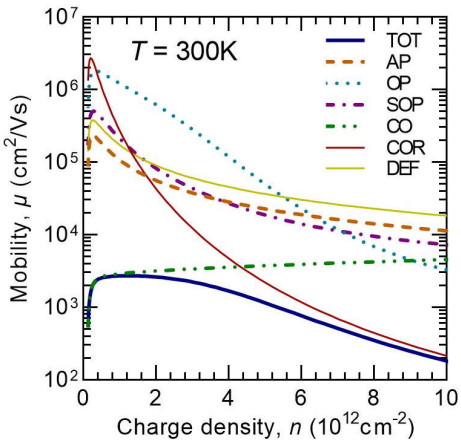


Fig. 6. Mobility in graphene vs. charge density at 300 K for all the included scattering mechanisms.

rate of ~5 × 10^{12} s^{-1} at high energies. In turn, the increasing scattering rate results in mobility decrease with increasing kinetic energy, except for the CO scattering that behaves oppositely.

The previous assertions about the scattering spectra are confirmed in Fig. 6 where we show the mobility as a function of carrier charge density for T = 300 K. When n increases, the total mobility first increases up to 2708 cm^2/Vs at n = 1.26 × 10^{12} cm^{-2} and after that point it starts to decrease. In this point of the local mobility maximum the mean free path (MFP) equals $\hbar\mu\sqrt{\pi n}/q \approx 35.5$ nm. As shown in Fig. 6, at low carrier concentrations the limiting factor is the CO scattering and for the higher concentrations the COR scattering is dominant. Hence, increasing μ_{CO} for low n, and decreasing μ_{COR} for high n determines the μ_{TOT}-n characteristic, at least in the current constellation of simulation parameters. The only other way that could provide a decreasing μ_{TOT} with increasing n, as reported in all experiments so far [7], [8], [20], [21], is to assume COR scattering to be negligible, and DEF scattering strong at high kinetic energies or carrier densities, well above the initially assumed value of N_{def} = 0.18 × 10^{10} cm^{-2}, e.g at about ~10^{11} cm^{-2}. The increased N_{def} would infer a rather short average distance between individual defects of ~32 nm, assuming their distribution in a perfectly square arrangement. Hence, N_{def} = 10^{11} cm^{-2} seems unrealistically high when compared to previously fabricated graphene by the same group at IHP [6]–[8], and this should be further analyzed by a detailed Raman spectroscopy study.

Finally, while the Hall measurements are done at different temperatures, but at the same carrier density, the well-calibrated transport model presented here allows us to extract the mobility at technologically more relevant densities up to n ~ 10^{13} cm^{-2}. At these values corresponding to e.g. strong inversion in field-effect transistors (FETs), the simulated mobility equals 200 cm^2/Vs. While this value is orders of magnitude lower than in suspended graphene samples at low temperatures [21], it is much higher than the field-effect mobility measured or simulated in ultra-thin semiconductor layers of comparable thickness [22], [23], [13], [24]. Additionally, the MFP in this case equals $\hbar\mu\sqrt{\pi n}/q \approx 7.4$ nm, which is comparable to channel length in modern CMOS FETs and should allow for at least quasi-ballistic carrier transport.

IV. Conclusions

The quality of graphene and substrate upon which it is grown are analyzed by a coupled experimental and theoretical study of carrier mobility. The Hall bar structures were characterized in the temperature range from 40 K to 300 K at moderate charge density, whereas MRT approximation of the Boltzmann transport equation is used to theoretically calculate scattering spectra of all relevant scattering mechanisms. The experiments reveal an increasing μ_{TOT} with increasing temperature, which is reproduced by simulations and attributed to CO scattering with $N_{imp} \approx 2.2 \times 10^{11}$ cm^{-2} over the entire temperature range. Regarding charge density dependence and room temperature characteristics, the theory predicts either strong COR scattering with high roughness/corrugation amplitudes (Δ = 1.3 nm, L = 2 nm) or a high density of lattice defects within the graphene monolayer ($N_{def} > 10^{11}$ cm^{-2}) to achieve a good match with experimental data. Therefore, additional improvements are needed in the transfer process to reduce defect density and/or graphene corrugations if mobilities greater than 2700 cm^2/Vs and MFPs above 35 nm are needed.

Acknowledgments

This work was supported by the Croatian Ministry of Science and Education and the German Academic Exchange Service (DAAD) through the Croatia-Germany Bilateral Project GRONGER. M. P. also acknowledges partial support by the Croatian Science Foundation under the project CONAN2D (Grant No. UIP-2019-04-3493).

References

[1] K. S. Novoselov *et al.*, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, no. 5696, pp. 666–669, Oct. 2004, doi: 10.1126/science.1102896.

[2] F. Bonaccorso, Z. Sun, T. Hasan, and A. C. Ferrari, "Graphene photonics and optoelectronics," *Nature Photonics*, vol. 4, no. 9, pp. 611–622, Aug. 2010, doi: 10.1038/nphoton.2010.186.

[3] G. Fiori *et al.*, "Electronics based on two-dimensional materials," *Nat. Nanotechnol.*, vol. 9, no. 10, pp. 768–779, Oct. 2014, doi: 10.1038/nnano.2014.207.

[4] J.-H. Lee *et al.*, "Wafer-Scale Growth of Single-Crystal Monolayer Graphene on Reusable Hydrogen-Terminated Germanium," *Science*, vol. 344, no. 6181, pp. 286–289, Apr. 2014, doi: 10.1126/science.1252268.

[5] I. Pasternak *et al.*, "Large-area high-quality graphene on Ge(001)/Si(001) substrates," *Nanoscale*, vol. 8, no. 21, pp. 11241–11247, May 2016, doi: 10.1039/C6NR01329E.

[6] M. Lukosius *et al.*, "Metal-Free CVD Graphene Synthesis on 200 mm Ge/Si(001) Substrates," *ACS Appl. Mater. Interfaces*, vol. 8, no. 49, pp. 33786–33793, Dec. 2016, doi: 10.1021/acsami.6b11397.

[7] M. Lisker *et al.*, "Processing and integration of graphene in a 200 mm wafer Si technology environment," *Microelectronic Engineering*, vol. 205, pp. 44–52, Jan. 2019, doi: 10.1016/j.mee.2018.11.007.

[8] J. Aprojanz *et al.*, "High-Mobility Epitaxial Graphene on Ge/Si(100) Substrates," *ACS Appl. Mater. Interfaces*, vol. 12, no. 38, pp. 43065–43072, Sep. 2020, doi: 10.1021/acsami.0c10725.

[9] M. Lundstrom, *Fundamentals of Carrier Transport*, 2nd ed. New York, NY, USA: Cambridge University Press, 2000.

[10] R. S. Shishir and D. K. Ferry, "Intrinsic mobility in graphene," *J. Phys.: Condens. Matter*, vol. 21, no. 23, p. 232204, Jun. 2009, doi: 10.1088/0953-8984/21/23/232204.

[11] R. S. Shishir, F. Chen, J. Xia, N. J. Tao, and D. K. Ferry, "Room temperature carrier transport in graphene," *J Comput Electron*, vol. 8, no. 2, pp. 43–50, Jun. 2009, doi: 10.1007/s10825-009-0278-y.

[12] M. Bresciani, P. Palestri, D. Esseni, and L. Selmi, "Simple and efficient modeling of the E–k relationship and low-field mobility in Graphene Nano-Ribbons," *Solid-State Electronics*, vol. 54, no. 9, pp. 1015–1021, Sep. 2010, doi: 10.1016/j.sse.2010.04.038.

[13] M. Poljak, V. Jovanovic, D. Grgec, and T. Suligoj, "Assessment of Electron Mobility in Ultrathin-Body InGaAs-on-Insulator MOSFETs Using Physics-Based Modeling," *IEEE Transactions on Electron Devices*, vol. 59, no. 6, pp. 1636–1643, Jun. 2012, doi: 10.1109/TED.2012.2189217.

[14] M. Poljak, T. Suligoj, and K. L. Wang, "Influence of substrate type and quality on carrier mobility in graphene nanoribbons," *Journal of Applied Physics*, vol. 114, no. 5, p. 053701, Aug. 2013, doi: 10.1063/1.4817077.

[15] M. Poljak, K. L. Wang, and T. Suligoj, "Variability of bandgap and carrier mobility caused by edge defects in ultra-narrow graphene nanoribbons," *Solid-State Electronics*, vol. 108, pp. 67–74, Jun. 2015, doi: 10.1016/j.sse.2014.12.012.

[16] E. H. Hwang and S. Das Sarma, "Acoustic phonon scattering limited carrier mobility in two-dimensional extrinsic graphene," *Phys. Rev. B*, vol. 77, no. 11, p. 115449, Mar. 2008, doi: 10.1103/PhysRevB.77.115449.

[17] I.-T. Lin and J.-M. Liu, "Surface polar optical phonon scattering of carriers in graphene on various substrates," *Appl. Phys. Lett.*, vol. 103, no. 8, p. 081606, Aug. 2013, doi: 10.1063/1.4819395.

[18] D. K. Ferry, "Short-range potential scattering and its effect on graphene mobility," *J Comput Electron*, vol. 12, no. 2, pp. 76–84, Jun. 2013, doi: 10.1007/s10825-012-0431-x.

[19] D. K. Ferry, S. M. Goodnick, and J. Bird, *Transport in Nanostructures*, 2nd ed. New York, NY, USA: Cambridge University Press, 2009.

[20] V. E. Dorgan, M.-H. Bae, and E. Pop, "Mobility and saturation velocity in graphene on SiO2," *Applied Physics Letters*, vol. 97, no. 8, pp. 082112-082112–3, Aug. 2010, doi: doi:10.1063/1.3483130.

[21] K. I. Bolotin *et al.*, "Ultrahigh electron mobility in suspended graphene," *Solid State Communications*, vol. 146, no. 9–10, pp. 351–355, Jun. 2008, doi: 10.1016/j.ssc.2008.02.024.

[22] K. Uchida, H. Watanabe, A. Kinoshita, J. Koga, T. Numata, and S. Takagi, "Experimental study on carrier transport mechanism in ultrathin-body SOI nand p-MOSFETs with SOI thickness less than 5 nm," in *IEDM Tech. Dig.*, IEEE, 2002, pp. 47–50. doi: 10.1109/IEDM.2002.1175776.

[23] D. Esseni, A. Abramo, L. Selmi, and E. Sangiorgi, "Physically based modeling of low field electron mobility in ultrathin single- and double-gate SOI n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 50, no. 12, pp. 2445–2455, Dec. 2003, doi: 10.1109/TED.2003.819256.

[24] S. Krivec, M. Poljak, and T. Suligoj, "Electron mobility in ultra-thin InGaAs channels: Impact of surface orientation and different gate oxide materials," *Solid-State Electronics*, vol. 115, no. Part B, pp. 109–119, Jan. 2016, doi: 10.1016/j.sse.2015.08.009.

Stress-dependent MOSFET Model for use in Circuit Simulations

Andro Žamboki*, Leo Gočan*, Josip Mikulić†, Niko Bako†, Gregor Schatzberger†, Tomislav Marković* and Adrijan Barić*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
† ams-OSRAM AG, Premstaetten, Austria
andro.zamboki@fer.hr

***Abstract*—This paper presents a stress-dependent MOSFET model based on existing transistor models. The encapsulation of chips is often made from polymers (such as epoxy) which are poured over the silicon chips. When the encapsulation cools, it generates a significant amount of static mechanical stress in the ICs which changes their performance. To compensate for the offsets caused by the stress, chip designers need to be able to simulate the effects of stress on their circuits and design stress sensors and active stress-compensating circuits. In this paper, a stress-dependent MOSFET model is implemented by combining existing transistor models and a Verilog-A cell. In this way, temperature, corner and other transistor variations can be simulated in combination with stress. Simulations show that this model can be used to simulate the stress effects on circuits and to design and optimize stress-compensating circuits.**

***Keywords*—stress, MOSFET model, SPICE simulation, optimization, stress compensation, encapsulation**

I. INTRODUCTION

THE stress caused by integrated circuit (IC) packaging has been thoroughly studied in literature [1] [2] [3]. During the IC packaging process, a significant amount of static stress that varies with temperature can be introduced by the encapsulation to the surface of the silicon. These stresses can have values higher than 100 MPa [3]. It has been shown extensively that this induced stress has a detrimental impact on circuits, especially biasing sources and oscillators [2] [4].

To analyze and compensate these stress-induced effects, researchers have designed stress sensors, complex stress sensing circuits and active stress offset-canceling subcircuits but, often, circuit simulators do not have an easy way of simulating the stress-induced effects on circuits. In [5] authors propose a Verilog-A stress upgrade for the PSP MOSFET model with the focus on highly flexible thin silicon chips. In [4] authors propose a statistical model focused on reference circuits. A more generalized model is needed that can be applied to different types of MOS silicon technologies that circuit designers use and which can be easily implemented, customized and improved.

In Section II the basics of the semiconductor stress theory are shown. In Section III a model calibration technique is laid out. In Section IV we propose stress-enhanced add-ons which can be attached to the existing MOSFET or resistor models. Simulations in Section V show that these models perform well in designing and testing stress sensors and identifying the most stress-sensitive parts of the circuits. It is also shown in Section V that this sensitivity can be reduced by changing the orientation of the transistors. Section VI presents conclusions.

II. THEORETICAL BASICS

To create a MOSFET stress model for use in circuit simulations, an analysis has to be done of the impacts of stress on transistors. The analysis in this paper uses the wafer coordinate system shown in Fig. 1 where it is assumed that the silicon wafer is (100) type. For (111) wafer types, the analysis gets somewhat more complicated but the overall procedure remains the same [6].

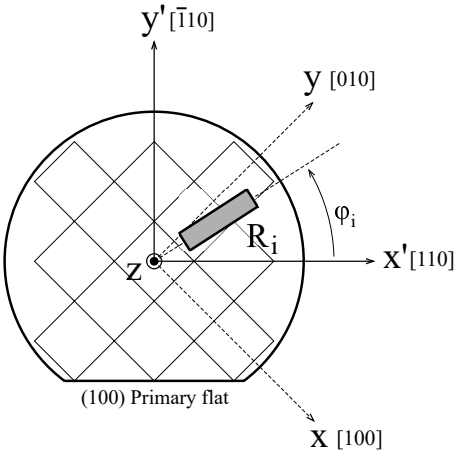


Fig. 1: Wafer coordinate system. The rectangular grid in the wafer represents the silicon crystal lattice orientation.

The change in the MOSFET drain current due to the mechanical stress is described by the following relation:

$$I_D = I_{D0} + \Delta I_D = I_{D0}\left(1 + \frac{\Delta I_D}{I_{D0}}\right) \tag{1a}$$

where I_{D0} is the unstressed MOSFET current in the linear region:

$$I_{D0} = K\left(U_{GS} - U_{GS0} - \frac{U_{DS}}{2}\right)U_{DS} \tag{1b}$$

or in the saturation region:

$$I_{D0} = \frac{K}{2}(U_{GS} - U_{GS0})^2 \tag{1c}$$

where K is the current coefficient of the MOSFET:

$$K = \mu C'_{OX} \frac{W}{L} \tag{1d}$$

where μ is the carrier mobility, C'_{OX} is the thin-oxide capacitance per unit area and W and L are the transistor width and length. It can be seen in (1) that the change of the MOSFET drain current ΔI_D because of the stress can be due to the carrier mobility μ variation, transistor dimensions W/L variation and/or threshold voltage U_{GS0} variation. Extensive research has been done on the magnitudes of these effects on the MOSFET current and the overall consensus in literature is that the impact of the carrier mobility variation vastly outweighs all other variations [7] in the strong inversion. Stress-dependent threshold shifts significantly alter the stress responses in moderate and weak inversion [1]. Regardless of the operating region, the proposed MOSFET stress model can still be used to simulate approximate stressed states in circuits and for stress-optimized circuit design.

So by neglecting all other variations, the change in the MOSFET current due to the stress can be directly linked with the change in the carrier mobility:

$$\frac{\Delta I_D}{I_{D0}} \approxeq \frac{\Delta \mu}{\mu}. \tag{2}$$

The resistance of a material is defined by:

$$R(\mu) = \frac{1}{\mu} \frac{L}{qntW} \tag{3}$$

where q is the elementary charge, n is the charge density and the t, W and L are the height, width and length of the material respectively. So the change of the material resistance (in this case, the MOSFET channel resistance) due to the change in the carrier mobility is:

$$\frac{\Delta R}{R} = \frac{R(\mu) - R(\mu_0)}{R(\mu_0)} = -\frac{\mu - \mu_0}{\mu} = -\frac{\Delta \mu}{\mu}. \tag{4}$$

By combining (1a), (2) and (4), the final equation for modeling the transistor behavior with respect to stress is extracted:

$$I_D = I_{D0} \left(1 - \frac{\Delta R}{R}\right). \tag{5}$$

The change of the MOSFET channel resistance $\Delta R/R$ due to stress is described by (6) [7]:

$$\begin{aligned} \frac{\Delta R}{R} = &\left[\frac{\pi_S + \pi_{44}}{2}\sigma'_{11} + \frac{\pi_S - \pi_{44}}{2}\sigma'_{22}\right]\cos^2\varphi \\ &+ \left[\frac{\pi_S - \pi_{44}}{2}\sigma'_{11} + \frac{\pi_S + \pi_{44}}{2}\sigma'_{22}\right]\sin^2\varphi \\ &+ \pi_D \sigma'_{12} \sin 2\varphi + f(T) \end{aligned} \tag{6a}$$

where function $f(T)$ models the temperature dependence of the resistance:

$$f(T) = \alpha_1 T + \alpha_2 T^2 + \alpha_3 T^3 \cdots . \tag{6b}$$

The coefficients $\pi_S = \pi_{11} + \pi_{12}$, $\pi_D = \pi_{11} - \pi_{12}$, π_{11}, π_{12} and π_{44} are the piezoresistive coefficients. They are the characteristics of the material in which integrated components are made. Equation (6a) is standard in literature and describes the change of the resistance of a material as a function of normal stresses σ'_{11} and σ'_{22} and shear stress σ'_{12}. The component σ'_{11} is the stress in the x' (Fig. 1) direction acting on a plane with a normal in the x' direction. Likewise, σ'_{22} is the stress in the y' direction acting on a plane with a normal in the y' direction. The value σ'_{12} is the stress in the y' direction acting on a plane with a normal in the x' direction. The angle φ is the direction of the current through the resistance shown in Fig. 1 (counterclockwise rotation has positive angles). Some standard values for lightly doped silicon are shown in Table I. For the greatest match in behaviour under stress between the modeled and manufactured MOSFETs, piezoresistive coefficients should be measured for the silicon technology the circuit designer is using.

TABLE I: Typical piezoresistive coefficient values for lightly doped silicon TPa^{-1} [1].

	n-type silicon	p-type silicon
π_{11}	-1022	66
π_{12}	534	-11
π_{44}	-136	1381

III. MODEL CALIBRATION

Although the approximate circuit behaviour can be simulated with the stress MOSFET models using the coefficients in Table I, to get more accurate results, a piezoresistive coefficient extraction for the used silicon technology should be done. Since (6) depends on π_S, π_D, and π_{44}, these values have to be measured. To get accurate results, these coefficients should be temperature compensated. In [8] authors use simple methods for calibration but π_S and π_D are not temperature compensated so the temperature during measurement has to be precisely controlled. To mitigate this issue, we use the more complex off-axis rosette MOSFET sensor for calibration [9].

The off-axis rosette is constructed by creating a coordinate system $x'' - y''$ which is rotated by θ (considered as a positive angle as drawn in Fig. 2) in the $x' - y'$ plane. Then placing 3 transistors in three directions: first in the x'' direction, second in the y'' and third in the direction that is rotated 45° counterclockwise from the direction x''. The off-axis rosette design is shown in Fig. 2.

Next, to analyze the current through the off-axis rosette transistors using (5), a coordinate system shift from $x' - y'$ to $x'' - y''$ is done. The stress vectors σ_{jk} are transformed using the matrix rotation operation in (7) [9]:

$$\begin{bmatrix} \sigma''_{11} \\ \sigma''_{22} \\ \sigma''_{12} \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin^2\theta & -2\sin\theta\cos\theta \\ \sin^2\theta & \cos^2\theta & 2\sin\theta\cos\theta \\ \sin\theta\cos\theta & -\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{22} \\ \sigma'_{12} \end{bmatrix} \tag{7a}$$

and written with matrix symbols:

$$\boldsymbol{\sigma}'' = \boldsymbol{R}\boldsymbol{\sigma}'. \tag{7b}$$

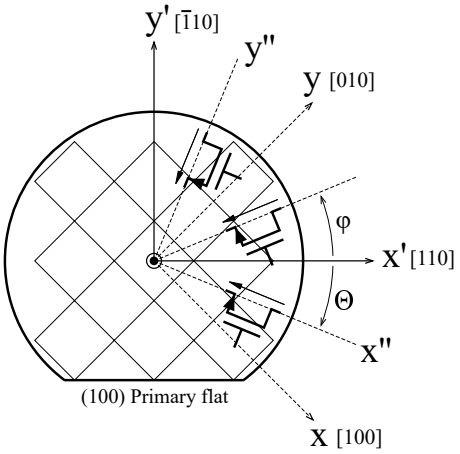


Fig. 2: Off-axis rosette wafer coordinate system.

By substituting σ'_{jk} with σ''_{jk} in (6) and calculating the results for the angle φ of the three off-axis transistors: $\varphi_1 = -\theta$, $\varphi_2 = -\theta + 45°$ and $\varphi_3 = -\theta + 90°$, the $\Delta R_i/R_i$ expressions can be derived for the transistors. If the piezoresistive coefficients are extracted from the calculated $\Delta R_i/R_i$ and if the uniaxial stress is applied in the y'' direction $\sigma''_{22} = \sigma$ ($\sigma''_{11} = \sigma''_{12} = 0$), the following expression can be derived:

$$\begin{bmatrix} \Delta R_1/R_1 \\ \Delta R_2/R_2 \\ \Delta R_3/R_3 \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{44} \end{bmatrix} \cdot \sigma \tag{8}$$

where $\boldsymbol{C}$ is the coefficient matrix which will be explained later. For the calibration, only the stress in one direction is needed. Calibrating the model by generating the stresses σ''_{11} and σ''_{22} should yield the same piezoresistive coefficients. Combining (2), (4) and (8) we get:

$$\begin{bmatrix} \Delta I_{D1}/I_{D0} \\ \Delta I_{D2}/I_{D0} \\ \Delta I_{D3}/I_{D0} \end{bmatrix} = -\boldsymbol{C} \begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{44} \end{bmatrix} \cdot \sigma = \boldsymbol{k} \cdot \sigma \tag{9}$$

where the matrix $\boldsymbol{k}$ represents the slopes of the $\Delta I_{Di}/I_{D0}$ vs. σ graphs:

$$\boldsymbol{k} = -\boldsymbol{C} \cdot \boldsymbol{\pi}. \tag{10}$$

By rearranging (10), the piezoresistive coefficients can be calculated from the normalized MOSFET current change vs. stress graph slopes:

$$\begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{44} \\ \pi_D \end{bmatrix} = -\boldsymbol{D} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} \tag{11}$$

where the index i in k_i denotes the respective MOSFET for which k is measured. $\boldsymbol{D} = \boldsymbol{C}^{-1}$ with the added fourth row. This row is constructed by subtracting the first and second rows of $\boldsymbol{D}$ (π_D = π_{11} - π_{12}). The generalized form of $\boldsymbol{D}$ is shown in (12a):

$$\boldsymbol{D} = \begin{bmatrix} -\frac{2\cos^2\theta - 1}{4A} & \frac{2\cos^2\theta - 1}{2A} & \frac{4A - 2\cos^2\theta + 1}{4A} \\ \frac{4A + 2\cos^2\theta - 1}{4A} & -\frac{2\cos^2\theta - 1}{2A} & \frac{2\cos^2\theta - 1}{4A} \\ -\frac{\cos 2\theta - \sin 2\theta}{\cos 2\theta} & -\frac{2\sin 2\theta}{\cos 2\theta} & \frac{\cos 2\theta + \sin 2\theta}{\cos 2\theta} \\ -\frac{2\cos^2\theta - 1 + 2A}{2A} & \frac{2\cos^2\theta - 1}{A} & -\frac{2\cos^2\theta - 1 - 2A}{2A} \end{bmatrix} \tag{12a}$$

where

$$A = \cos\theta \sin\theta. \tag{12b}$$

The final procedure is then constructed for the stress MOSFET model calibration:

1) Design an off-axis rosette with an angle of $\theta = <0°, 45°>$ (excluding $0°$ and $45°$).
2) Cut a long silicon stripe in the y'' direction with the rosette in the middle and apply a controlled stress using the four-point bending stress test [10].
3) Calculate the $\boldsymbol{D}$ matrix for the chosen angle θ.
4) Measure the $\Delta I_{Di}/I_{D0}$ for a range of stress values σ and extract the graph slopes k_i using the least squares method for the three transistors.
5) Calculate the piezoresistive coefficients with the measured k_i values using (11).

Equation (13) shows the matrix $\boldsymbol{D}$ if the angle θ is chosen to be 22.5°:

$$\boldsymbol{D}|_{\theta=22.5°} \begin{bmatrix} -0.5 & 1 & 0.5 \\ 1.5 & -1 & 0.5 \\ 0 & -2 & 2 \\ -2 & 2 & 0 \end{bmatrix}. \tag{13}$$

By analyzing the generalized form of the matrix $\boldsymbol{D}$, it can be observed that the third and fourth rows always add up to 0 regardless of the angle θ. If it is assumed that the temperature impacts all k_i values in an identical way (the temperature characteristics of transistors are expected not to depend on their direction), then the π_{44} and π_D measurements should be temperature compensated. Consequently, π_{11} and π_{12} are not temperature compensated. Since π_S is the sum of π_{11} and π_{12}, their temperature dependence will also be summed. Therefore, π_S is not temperature compensated. Alternatively, this coefficient can be estimated from $\pi_D = 3\pi_S$ based upon well-known theoretical results for electrons in silicon ($\pi_{11} = -2\pi_{12}$) [8]. Since, π_D is temperature compensated, π_S will be as well.

Using this method the piezoresistive coefficients have to be extracted for both nMOS and pMOS transistors as well as for all other used silicon devices (e.g. if resistors are used, then an off-axis resistor rosette has to be constructed and $\Delta R_i/R_i$ measurements have to be taken). For more precise measurements, the frequency method, the laser interferometer technique and the quasi-static method may be used [11].

IV. STRESS MOSFET MODEL

As stated before in (5), the current through a MOSFET transistor under mechanical stress is described by the following expression:

$$I_D = I_{D0} - I_{D0}\frac{\Delta R}{R} \tag{14}$$

where $\Delta R/R$ is the transistor channel resistance change due to stress and is described by (6). The problem with modeling the impact of stress on the transistor current is that the channel resistance is impacted by the carrier mobility variation but, in most cases (including this one), the only transistor parameters that can be changed are the channel dimensions W and L. So a stress-dependent transistor model has to be created around the existing stress-invariant MOS transistors. Fig. 3 shows a model that mimics (14). It works in the following way: a black box is created like the one in Fig. 3. A Verilog-A model is written to create an ideal current source which measures the current through the transistor I_{D0} and generates a current equal to $I_{D0}\cdot\Delta R/R$ and is parallel to the transistor. This current source calculates $\Delta R/R$ and is controlled by the voltages $U_{\sigma 11'}$, $U_{\sigma 22'}$ and $U_{\sigma 12'}$, which correspond to the values of the stresses σ'_{11}, σ'_{22} and σ'_{12} (1 V is equivalent to 1 Pa). A user defined value of φ (in degrees) determines the direction of the transistor relative to the x' direction. By analyzing this model, the following equations can be extracted:

$$I_X + I_{D0}\frac{\Delta R}{R} = I_{D0} \tag{15a}$$

$$I_Y + I_{D0}\frac{\Delta R}{R} = I_{D0}. \tag{15b}$$

So from the outside of the black box the following currents are measured:

$$I_X = I_Y = I_{D0} - I_{D0}\frac{\Delta R}{R}. \tag{16}$$

I_X and I_Y are identical to (14) which means that the black box will act like a transistor under mechanical stress and its current will change accordingly. This black box can be made as a new component which has φ and the transistor dimensions W and L set as user-editable variables. In this way, stress simulations can easily be incorporated into existing circuits by replacing the stress-invariant MOSFET models with the stress-dependent models.

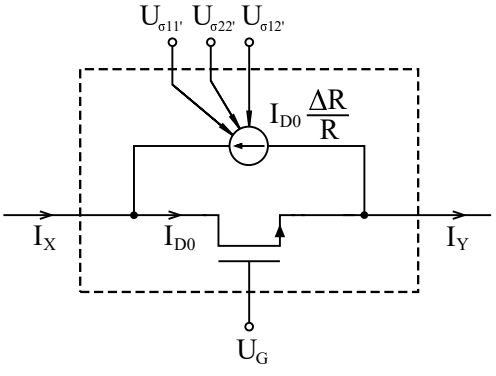
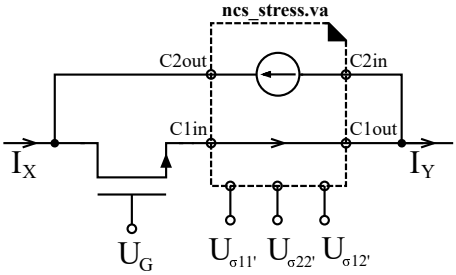
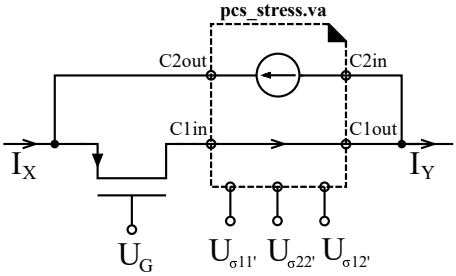


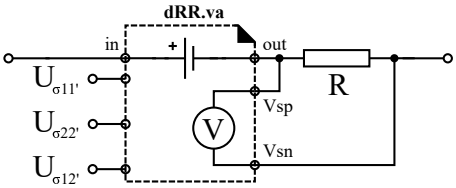
Fig. 3: Stress-dependent MOSFET model which can be implemented using Verilog-A.



(a) nMOS transistor model.



(b) pMOS transistor model.

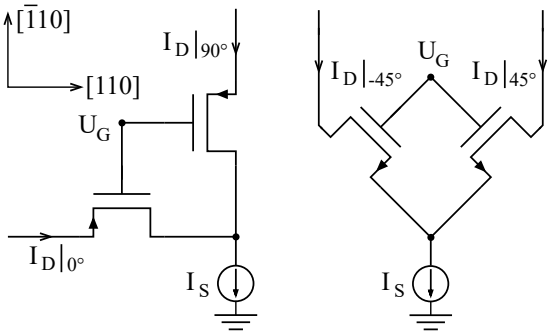


(c) Resistor model.

Fig. 4: Verilog-A implementation of the stress-dependent transistor and resistor models.

The biggest advantage of using this model is that all nonlinearities, temperature and corner variations of the transistor model are kept in I_{D0}. The transistor stress-dependent model can be simulated for temperature and corner variations the same way one would do for a standard transistor model. This model can be used for both pMOS and nMOS transistors.

Fig. 4a shows how the stress-dependent nMOSFET model is implemented. A current controlled current source written in Verilog-A (code in Appendix A) measures the current through the ports $C1in - C1out$, multiplies it by $\Delta R/R$ and forces that current through the ports $C2in - C2out$. $\Delta R/R$ is calculated using the input control voltages $U_{\sigma 11'}$, $U_{\sigma 22'}$ and $U_{\sigma 12'}$ and the CDF parameter phi (φ). This parameter is set up to allow the user to set a custom angle value for each stress-dependent transistor. The same model can be made with a pMOS transistor but the piezoresistive coefficients in the Verilog-A code have to substituted for those of the p-type silicon. Using these models, a comprehensive stress analysis can be made on CMOS circuits. By using global control voltages for σ_{jk}, the stress magnitude and direction can be dynamically altered. A chip designer can then pinpoint the critical circuit parts that are sensitive to stress.



(a) Normal stress pMOS sensor.

(b) Shear stress nMOS sensor.

Fig. 5: Differential pair MOSFET stress sensors.

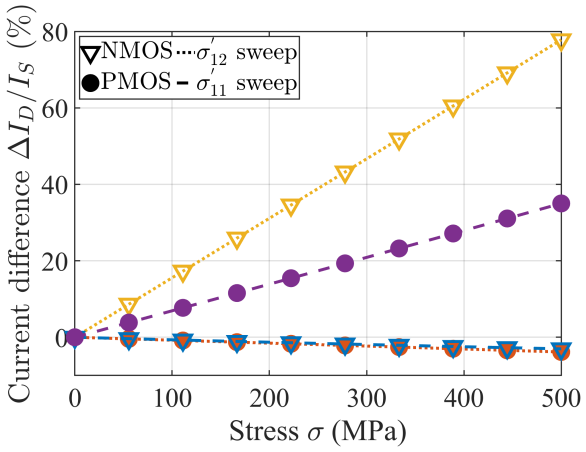


Fig. 6: Normalized current differences of MOSFET differential pairs in Fig. 5 in response to a stress sweep. Dotted and dashed lines represent the σ'_{12} and σ'_{11} stress sweeps respectively. The triangles and circles represent nMOS and pMOS sensors respectively.

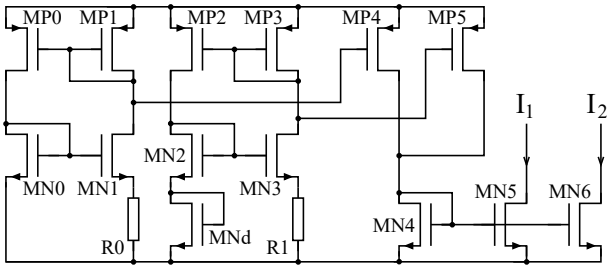


Fig. 7: CMOS reference current source.

To make this model user-friendly in Cadence Virtuoso, the adjustments which are described in Appendix B should be performed. This will enable the circuit designer to change the length, width and direction of each stress MOSFET instance separately in an identical way it would be done for a regular transistor circuit model.

The methodology of creating a stress model shown in this chapter can also be applied on resistors. In that case, the model in Fig. 4c is used. The Verilog-A block $dRR.va$ is similar to $ncs_stress.va$ but with a different analog block:

```
analog begin
      V(in,out) <+ ΔR/R *V(Vsp,Vsn);
end
```

where $\Delta R/R$ is shown in (6). In short, this model mimics the added voltage drop on a resistor due to the change of the resistance ΔR caused by the stress.

V. USE CASES AND EXAMPLES

The stress MOSFET model can be used to design and calibrate new stress sensors. The following examples use the 180-nm CMOS technology node and the piezoresistive coefficients from literature shown in Table I. Fig. 5 shows the normal stress pMOS sensor (Fig. 5a) and a shear nMOS sensor (Fig. 5b). In literature, it is accepted that the normal stress sensors are less sensitive to shear and should be made from pMOSFETs due to their increased sensitivity to normal stress [7]. Likewise, the shear sensors are less sensitive to normal stress and should be made from nMOSFETs due to their increased sensitivity to shear. To test this fact, a schematic is made where normal stress and shear sensors are made from both pMOS and nMOS transistors. Then the normal stress σ'_{11} and shear σ'_{12} are swept separately and the sensor currents are plotted in Fig. 6. It can be clearly seen that the slope (sensitivity) of the pMOS normal stress sensor is larger then that of the nMOS normal stress sensor. The opposite is true for the shear stress sensors.

The models presented in this work can also be used to analyze the impact of stress on more complicated circuits. In that case, the parts of the circuit that are most sensitive to stress can be identified. The example in Fig. 7 shows a schematic of a standard temperature-invariant CMOS reference current source. It is designed so that the output reference currents are: I_1 = 1 µA and I_2 = 4 µA. The simulation results in Fig. 8a show the change of the output reference currents when all elements are at an angle of φ = 0°. The output currents are sensitive to normal stresses σ'_{11} and σ'_{22} (but are less sensitive to shear σ'_{12}). In the worst case, the absolute output current sensitivity is 16.4 %/GPa. Fig. 8b shows the change of the output reference currents when the transistors MP4 and MP5 are both rotated by φ = 25°. With this modification, the worst case absolute output current sensitivity is 6.0 %/GPa. By changing the angle of just two transistors, the reference current source has been improved.

VI. CONCLUSION

This paper presents a stress-dependent MOSFET and resistor models that can be easily fitted around existing models. These models can be used by circuit designers to simulate the effects of stress caused by packaging on their circuits. A detailed procedure for calibrating the models is presented so that the simulations best represent the behaviour under stress of the used silicon technologies. For approximate simulation results, previously measured silicon piezoresistive coefficients can also be used without the need for model calibration. Simulations show that the

models can be used for stress sensor design/optimization, to decrease the reference source sensitivity to stress and to design active subcircuits that minimize the offsets caused by stress.

APPENDIX A
CURRENT CONTROLLED CURRENT SOURCE VERILOG-A CODE *ncs_stress.va* FROM FIG. 4A

```
`include "constants.vams"
`include "disciplines.vams"
module ncs_stress(C2in,C2out,C1in,C1out
    ↪ ,O11,O22,O12);
       electrical C2in, C2out, C1in,
          ↪ C1out;
       electrical O11, O22, O12;
       parameter real PI44 = -136E-12;
       parameter real PIS = -488E-12;
       parameter real PID = -1556E-12;
       parameter real phi = 0; //CDF
       localparam real radphi = `M_PI
          ↪ /180*phi;
       analog begin
         I(C2in,C2out)<+(((PIS+PI44)/2*V
            ↪ (O11) + (PIS-PI44)/2*V(
            ↪ O22))*$pow($cos(radphi)
            ↪ ,2) + ((PIS-PI44)/2*V(O11
            ↪ ) + (PIS+PI44)/2*V(O22))*
            ↪ $pow($sin(radphi),2) +
            ↪ PID*V(O12)*$sin(2*radphi)
            ↪ )*I(C1in,C1out);
       end
endmodule
```

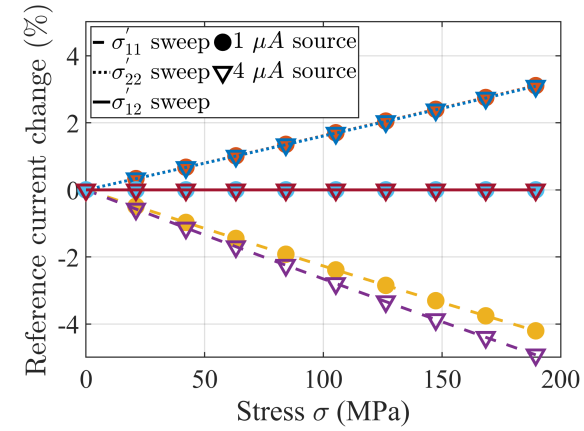
Code 1: Current controlled current source Verilog-A code *ncs_stress.va* from Fig. 4a.

APPENDIX B
CADENCE VIRTUOSO MODEL SETUP

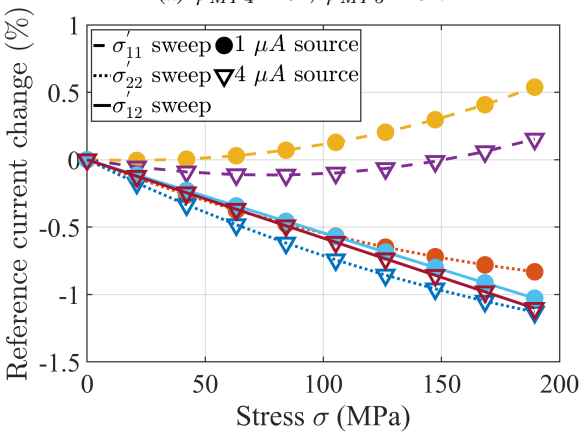
Select the *ncs_stress.va* cell in Fig. 4a and click *Q*. Go to *CDF parameter of View* → *veriloga* and in the *phi* input prompt write *pPar("Angle")*. Do the same for the transistor width and length and write *pPar("Width")* and *pPar("Length")* in the respective input prompts. In the Virtuoso terminal go to *tools* → *CDF* → *edit*. Select *CDF Layer* as *base* and select the cell in Fig. 4a. Add "Width", "Length" and "Angle" rows and set them as "string" (Parse as CEL: yes, Parse as Number: yes).

ACKNOWLEDGMENT

This work is supported by ams-OSRAM AG, Premstaetten, Austria.



(a) $\varphi_{MP4} = 0°$, $\varphi_{MP5} = 0°$.



(b) $\varphi_{MP4} = 25°$, $\varphi_{MP5} = 25°$.

Fig. 8: Reference current source (Fig. 7) output current drift in response to stress. Dashed, dotted and whole lines represent the σ'_{11}, σ'_{22} and σ'_{12} stress sweeps respectively. Circles and triangles represent the I_1 (1 μA) and I_2 (4 μA) currents respectively.

REFERENCES

[1] R. C. Jaeger and J. C. Suhling, "First- and second-order piezoresistive coefficients of CMOS FETs from strong into weak inversion," *IEEE Sensors Journal*, vol. 19, no. 23, pp. 11 309–11 317, 2019.

[2] S. Schlipf, C. Sander, A. Clausner, J. Paul, S. Capecchi, L. Wambera, K. Meier, and E. Zschech, "IC package related stress effects on the characteristics of ring oscillator circuits," in *2021 22nd EuroSimE*, 2021, pp. 1–5.

[3] Q. Nguyen, M. Rahim, J. Roberts, J. Suhling, and R. Jaeger, "Characterization of die stresses in plastic packages subjected to moisture and thermal exposures," *ASME 2015*, 07 2015.

[4] D. Wang and P. K. Chan, "An electrical model for nanometer CMOS device stress effect in design and simulation of analog reference circuits," *IEEE Transactions on VLSI Systems*, vol. 26, no. 5, pp. 958–968, 2018.

[5] Z. Y. Heidrun Alius, Horst Rempp and T. Gneiting, "A new MOSFET model for the simulation of circuits under mechanical stress," *MOS-AK Workshop presentation*, 2014.

[6] J. Suhling and R. Jaeger, "Silicon piezoresistive stress sensors and their application in electronic packaging," *IEEE Sensors Journal*, vol. 1, no. 1, pp. 14–30, 2001.

[7] R. Jaeger, J. Suhling, R. Ramani, A. Bradley, and J. Xu, "CMOS stress sensors on [100] silicon," *IEEE JSSC*, vol. 35, no. 1, pp. 85–95, 2000.

[8] S. Hussain, R. C. Jaeger, and J. C. Suhling, "Current dependence of the piezoresistive coefficients of CMOS FETs on (100) silicon." IEEE Computer Society, 11 2014, pp. 74–77.

[9] R. Jaeger, J. Suhling, M. Carey, and R. Johnson, "A piezoresistive sensor chip for measurement of stress in electronic packaging," in *Proceedings of IEEE 43rd ECTC '93*, 1993, pp. 686–692.

[10] R. Beaty, R. Jaeger, J. Suhling, R. Johnson, and R. Butler, "Evaluation of piezoresistive coefficient variation in silicon stress sensors using a four-point bending test fixture," *IEEE CHMT*, vol. 15, no. 5, pp. 904–914, 1992.

[11] J. Fialka and P. Beneš, "Comparison of methods for the measurement of piezoelectric coefficients," *IEEE Transactions on I&M*, vol. 62, no. 5, pp. 1047–1057, 2013.

Low-Power Frequency-Locked Loop Circuit with Static Frequency Offset Cancellation

Leo Gočan*, Andro Žamboki*, Niko Bako†, Josip Mikulić†, Gregor Schatzberger†, Tomislav Marković* and Adrijan Barić*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
† ams-OSRAM AG, Premstaetten, Austria
leo.gocan@fer.hr

***Abstract*—A phase-locked loop (PLL) is an important and commonly used electronic circuit in various electronic systems. Its main drawback is the use of an RC low-pass filter which takes up the majority of the PLL area on the chip. The RC low-pass filter is necessary to ensure the PLL stability. To mitigate this issue, a frequency-locked loop (FLL) is used because the stability of an FLL system depends on the Miller capacitance inside of the operational amplifier, which drastically reduces the capacitor size and thus the chip area. This paper presents an improved design of a fully integrated FLL. It is based on a single frequency-to-voltage converter (FVC) which uses a single capacitor and a single charging current for the frequency-to-voltage conversion of both the input and output frequencies. The use of one FVC reduces the static frequency offset caused by the mismatch between the FVCs. The circuit is implemented in a 180-nm CMOS process. The measurements show that the new FLL design has increased precision and accuracy and similar chip area compared to the previous design. It has higher power consumption, increased delay time, overshoot and settling time, but they are comparable to those of the previous design.**

***Keywords*—frequency-locked loop, frequency-to-voltage converter**

I. Introduction

A frequency-locked loop (FLL) is a circuit similar to a phase-locked loop (PLL) [1] as both circuits generate an output signal which tracks an input reference signal. The output signal of the PLL is synchronised with the input signal in phase, and therefore in frequency, whereas the output signal of the FLL is synchronised with the input signal only in frequency. The main advantage of the FLL over the PLL is a smaller die area needed to create a fully integrated circuit due to the absence of the area-consuming RC low-pass filter needed in the PLL for stabilisation.

A low-power FLL circuit with a short settling time has been reported in [2]. It is based on an FLL architecture from [3] and on a frequency-to-voltage converter (FVC) from [4] which was optimised for short settling time. Several advances in the FLL architecture based on the works [3] and [4] have been reported in [5]–[8].

This paper presents an improved design of the FLL presented in [2] in order to remove the static frequency offset caused by the mismatch between the two FVCs used in the design. Instead of the two FVCs, the FLL in this work uses a single FVC with an upgraded architecture. This results in greater accuracy and precision of the output frequency of the FLL.

This paper is organized in six sections. Section II explains the origin and the elimination of the static frequency offset caused by the mismatch between the FVCs. Section III describes the architecture and the operation of the FLL and the upgraded FVC and its control circuit. In Section IV, the measurement results of the FLL are presented. In Section V, the measurement and simulation results of the FLL from this work are compared to the measurement and simulation results of the FLL from [2]. The conclusion is given in Section VI.

II. Static Frequency Offset

A. FLL Architecture without Offset Cancellation

The measurements of the FLL design from [2] show that it exhibits a static frequency offset. The main causes of the offset are the mismatch between the two FVCs used in the design and the offset of the operational amplifier (opamp), with the former being the dominant one. In this work, the static frequency offset due to the mismatch of the two FVC circuits used in the FLL is removed.

The simplified schematic of the connection of the FVCs and the opamp from [2] is shown in Fig. 1. The two main causes of the offset are the mismatch between the charging currents I_c and $I_\mathrm{c}{}'$ and the mismatch between the capacitors C_1 and $C_1{}'$ of the FVCs because both the charging current I_c and the value of C_1 are directly related to the output voltage of the FVC through the equation [2]

$$V_{\mathrm{out,FVC}} = \frac{I_\mathrm{c}}{C_1} \cdot \frac{1}{2f_\mathrm{in}} \tag{1}$$

where f_in is the input frequency of the FVC. The capacitors C_2 and $C_2{}'$ serve only as charge storage elements, and therefore their potential mismatch does not result in any offset.

If there is a mismatch between the FVCs, in order for their output voltages to be equal, they require different input frequencies according to (1). The output voltages of the FVCs have to be equal because the opamp makes them so, thus generating a control voltage different than the one needed to generate the correct FLL output frequency. This creates the frequency offset.

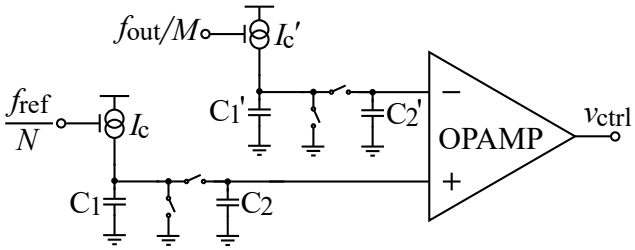


Fig. 1: The simplified schematic of the connection of the FVCs and the opamp from [2].

B. FLL Architecture with Offset Cancellation

The simplified schematic of the connection of the FVC and the opamp in this work is shown in Fig. 2. Since there is only one FVC, and since it uses a single charging current I_{c} and a single capacitor C_1, the possibility of the offset caused by the mismatch between the FVCs is removed. The capacitors $C_{2\mathrm{x}}$ and $C_{2\mathrm{y}}$ serve only as charge storage elements, and therefore their potential mismatch does not result in any offset, as it can be seen in (1).

III. Architecture and Operation

The block diagram of the proposed FLL with the static frequency offset cancellation is shown in Fig. 3. It is composed of two frequency dividers: one dividing by M and one by N, an FVC, an opamp and a voltage-controlled oscillator (VCO). The frequency dividers in this work divide by $M = 20$ and $N = 2$. The feedback loop is composed of the VCO, the frequency divider by M, the FVC and the opamp.

The operation of the FLL is as follows. The reference frequency f_{ref} is divided by N and converted into the voltage v_+ which is connected to the non-inverting input

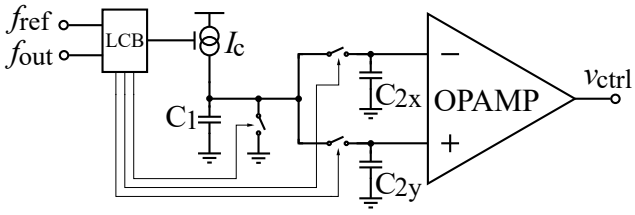


Fig. 2: The simplified schematic of the connection of the FVC and the opamp in this work.

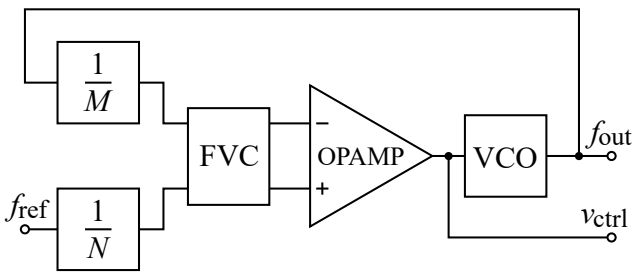


Fig. 3: The block diagram of the FLL with the static frequency offset cancellation ($M = 20$, $N = 2$).

of the opamp. The FLL output frequency f_{out}, which is also the VCO output frequency, is divided by M and converted into the voltage v_- which is connected to the inverting input of the opamp. The opamp tries to equalise the voltages at its inputs by changing its output voltage which is the control voltage v_{ctrl} of the VCO. The VCO output frequency changes until the opamp input voltages equalise. At that moment, the FLL is considered locked to the reference frequency.

The output frequency f_{out} of the FLL is defined by [2]

$$f_{\mathrm{out}} = \frac{M}{N} \cdot f_{\mathrm{ref}}. \tag{2}$$

A. Frequency-to-Voltage Converter (FVC)

The schematic of the new FVC architecture used in the offset-cancellation FLL is shown in Fig. 4. It is composed of three capacitors, a current source, several transistors that act as switches, two transmission gates with the corresponding dummy transmission gates, and a logic controller block (LCB). The dummy transmission gates are intended to compensate for the charge injection of the transmission gate. The transistors in each of them have half the width of the corresponding transistors in the transmission gate.

Unlike the former FVC architecture from [2], the new one has two distinct time intervals of conversion which alternate. During the entire first interval, the capacitor $C_{2\mathrm{y}}$ is isolated from the rest of the circuit. When the signal $F3$ goes low, the transistor MP3 turns on and the frequency f_{ref} is converted into voltage on the capacitor C_1 by means of the charging current I_{c}. When the signal $F3$ goes high, the transistor MP3 turns off, and C_1 stops charging. At the same time, $\Phi 1x$ goes high for a period of time τ_1 which turns on the transmission gate, thus connecting C_1 and $C_{2\mathrm{x}}$. The charges are redistributed between C_1 and $C_{2\mathrm{x}}$. The charge stored on the capacitor $C_{2\mathrm{x}}$ generates the voltage $v_{\mathrm{out},1}$. When the charge redistribution is finished, $\Phi 1x$ goes low, the transmission gate turns off, and $C_{2\mathrm{x}}$

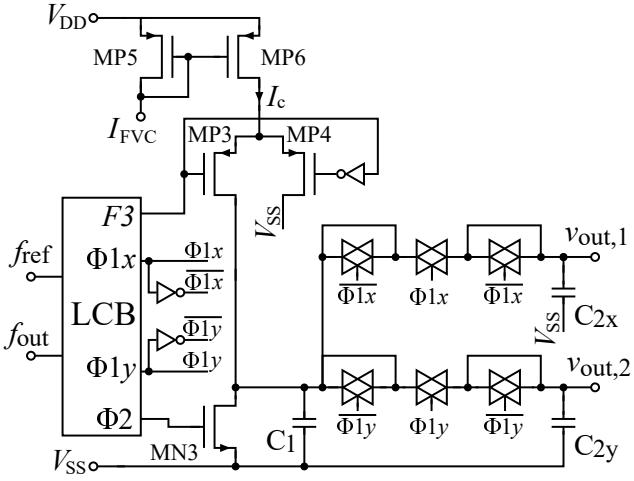


Fig. 4: The schematic of the frequency-to-voltage converter (FVC) architecture.

is isolated from the rest of the circuit. When $\Phi 1x$ goes low, $\Phi 2$ goes high for a period of time τ_2 which turns on MN3 and C_1 discharges. During the second interval, the capacitor C_{2x} is isolated from the rest of the circuit, the frequency f_{out} is converted into voltage on the capacitor C_1, and the charges are redistributed between the capacitors C_1 and C_{2y}. The charge stored on the capacitor C_{2y} generates the voltage $v_{out,2}$. Over several cycles, charges accumulate on the capacitors C_{2x} and C_{2y} and generate the voltages $v_{out,1}$ and $v_{out,2}$, respectively, which are proportional to the frequencies f_{ref} and f_{out}, respectively.

The time intervals of conversion do not have fixed values; they track the frequencies f_{ref} and f_{out} in such a way that one interval contains $N/2$ periods of the signal f_{ref} and the other interval contains $N/2$ periods of the signal f_{out}.

B. Logic Controller Block (LCB)

The new FVC architecture requires a new LCB architecture which includes an addition of an LCB controller circuit. The schematics of the new LCB architecture and of the LCB controller are shown in Figs. 5 and 6, respectively.

The LCB generates four control signals: $\Phi 1x$, $\Phi 1y$, $\Phi 2$ and $F3$, which control the elements in the FVC (Fig. 4). The signals $\Phi 1x$, $\Phi 1y$, $\Phi 2$ are pulses. The signals $\Phi 1x$, $\overline{\Phi 1x}$, $\Phi 1y$ and $\overline{\Phi 1y}$ control the transmission gates and the dummy transmission gates, the signal $\Phi 2$ controls the transistor MN3, and the signal $F3$ controls the transistors MP3 and MP4. The duration of the pulse $\Phi 1x$ must be long enough for the charge redistribution between C_1 and C_{2x} to occur. The same is applicable to the pulse $\Phi 1y$ and the capacitors C_1 and C_{2y}. The duration of the pulse $\Phi 2$ must be long enough to allow for a full discharge of C_1. The pulses $\Phi 1x$ and $\Phi 1y$ must not overlap with $\Phi 2$.

The operation of the LCB controller is as follows. When the output of the divider by N ($N = 2$ in this work) in Fig. 6 changes from 0 to 1, the SR latch SR2 is reset ($Q_2 \rightarrow 0$), and the SR latch SR1 is set ($Q_1 \rightarrow 1$) by means of a pulse synchronised with the signal f_{ref}. The pulse is generated by the circuits I0–3. With the signal Q_1 being set to 1, the NAND gate I8 generates a signal with the frequency f_{ref}, whereas with the signal Q_2 being set

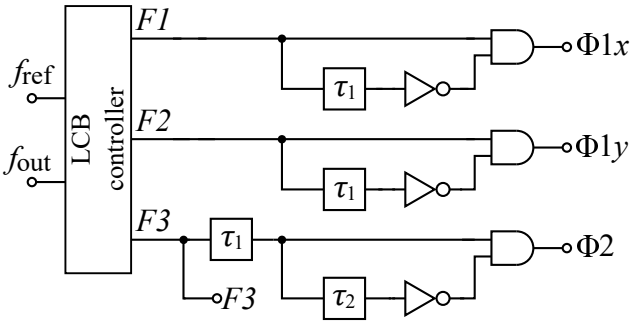


Fig. 5: The schematic of the logic controller block (LCB) architecture.

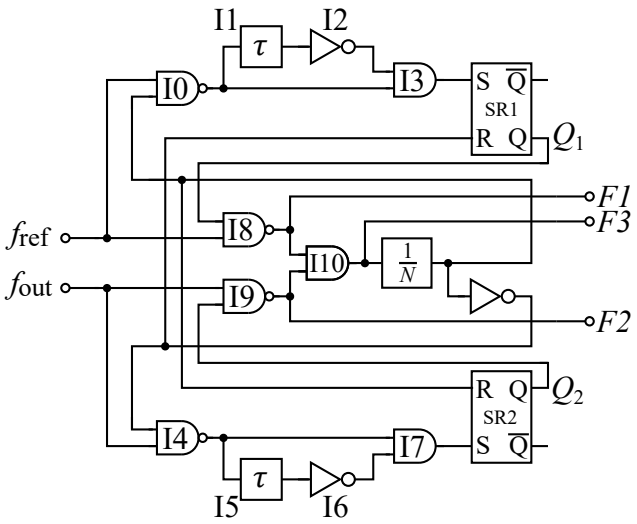


Fig. 6: The schematic of the controller for the logic controller block (LCB).

to 0, the NAND gate I9 output is held at 1. Since one of the inputs of the AND gate I10 is 1 and the other receives the signal with the frequency f_{ref}, the input of the divider by N receives the same signal. After $N/2$ periods of f_{ref}, the output of the divider by N changes from 1 to 0 and the equivalent process begins for the signal f_{out}.

The signal $F1$ contains $N/2$ periods of f_{ref}, the signal $F2$ contains $N/2$ periods of f_{out}, and the signal $F3$ is the result of the logic operation AND between $F1$ and $F2$. The signals $F1$ and $F2$ produce the signals $\Phi 1x$ and $\Phi 1y$, respectively, whereas the signal $F3$ produces the signal $\Phi 2$ and enables the charging current of the FVC.

IV. Measurement Results

The output frequency versus the reference frequency of 20 FLL chips is shown in Fig. 7. For the range of the reference frequencies from 1.5 MHz to 4 MHz, the characteristics can be considered affine, and therefore they can be approximated with lines. The slope of the lines represents the frequency multiplication factor of the FLL M/N from (2). In this work, the theoretical frequency

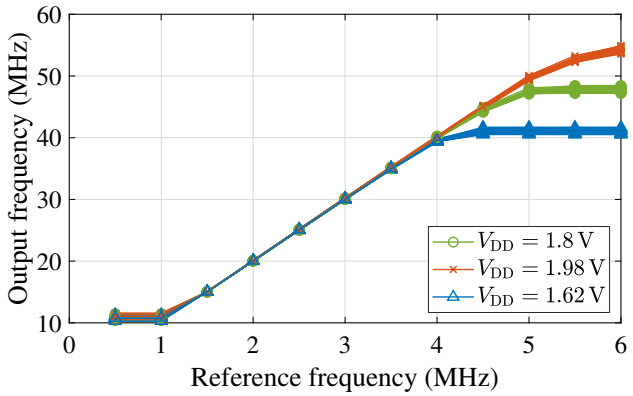


Fig. 7: The output frequency versus the input frequency of 20 FLL chips ($V_{DD} = 1.8\,V \pm 10\%$, $T = 25\,°C$).

multiplication factor is 10. For the nominal supply voltage and room temperature, the minimum slope is 9.95 and the maximum is 10.08, which is an improvement over the FLL architecture from [2] which had the minimum slope of 9.33 and the maximum of 9.94. The remaining deviation from the theoretical slope value can be reduced by eliminating the offset of the opamp. The output frequency levelling for the reference frequencies higher than 4 MHz is caused by the VCO because its output frequency levels off for sufficiently high control voltage v_{ctrl}, i.e., for sufficiently high reference frequency.

The distribution of the output frequency values of 20 FLL chips for the reference frequency of 2 MHz is shown in Fig. 8. Both the mean value of 20.04 MHz and the standard deviation of 32.79 kHz represent an improvement over the respective values of 19.22 MHz and 200.80 kHz for the FLL architecture from [2].

The period jitter and the cycle-to-cycle jitter are calculated using the definitions from JEDEC standard [9]. The period jitter is defined as the deviation in cycle time of a signal with respect to the ideal period over a number of randomly selected cycles [9]. The cycle-to-cycle jitter is defined as the variation in cycle time of a signal between adjacent cycles, over a random sample of adjacent cycle pairs [9].

The distribution of the values of the period jitter of one FLL chip is shown in Fig. 9. The period jitter is calculated as the difference between the individual period durations and the mean period duration. The peak-to-peak jitter value is calculated using the expression

$$\text{Peak-to-peak value} = \pm 3.090 \cdot \sigma \tag{3}$$

where the value of 3.090 is taken from the Gaussian probability density function table and σ is the standard deviation. With the sample size of 1000, 999 samples will, on average, fall within $\pm 3.090\sigma$ from the mean value. Both the standard deviation of 0.38 ns and the peak-to-peak jitter value of ± 1.18 ns represent an improvement over the respective values of 0.62 ns and ± 1.92 ns for the FLL architecture from [2].

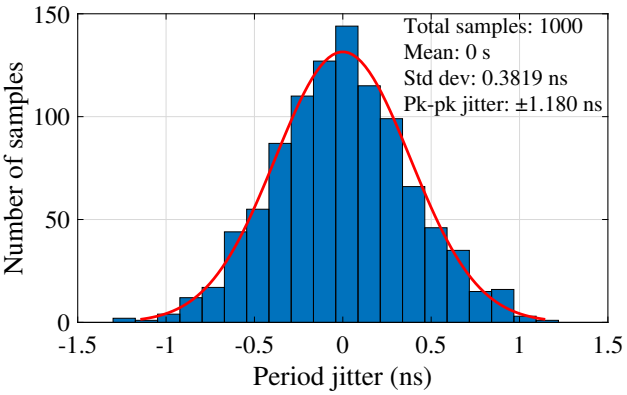


Fig. 9: The distribution of the values of the period jitter of one FLL chip ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $T = 25\,^{\circ}\mathrm{C}$, $f_{\mathrm{ref}} = 2\,\mathrm{MHz}$).

The distribution of the values of the cycle-to-cycle jitter of one FLL chip is shown in Fig. 10. The cycle-to-cycle jitter is calculated as the difference between the durations of every two adjacent periods. Both the standard deviation of 0.65 ns and the peak jitter value of 2.46 ns represent an improvement over the respective values of 1.07 ns and 3.89 ns for the FLL architecture from [2].

The reduction of the period jitter and the cycle-to-cycle jitter is a result of the use of a single charging current source and a single capacitor C_1 in the FVC (see Fig. 4). They are the critical analogue elements, and their mismatch increases the jitter, but by removing the possibility of the mismatch, the jitter is reduced.

V. Comparison with Previous Work

The comparison of the typical values of the parameters of both FLL designs is shown in Table I. The mean value and the standard deviation of the output frequency are improved in the new design, but the power consumption and the 0.1-% settling time are degraded. The new design has a marginally increased area which is still significantly

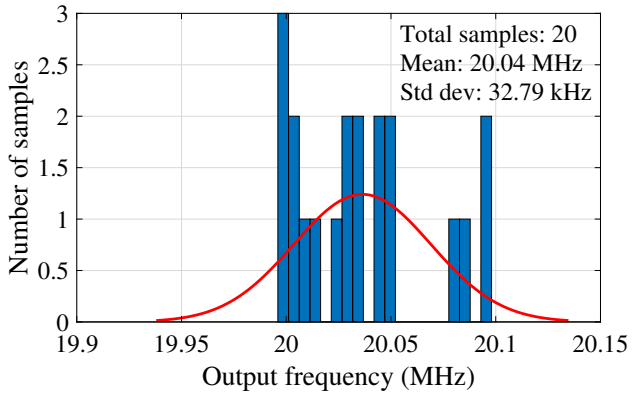


Fig. 8: The distribution of the output frequency values of 20 FLL chips for the reference frequency of 2 MHz ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $T = 25\,^{\circ}\mathrm{C}$).

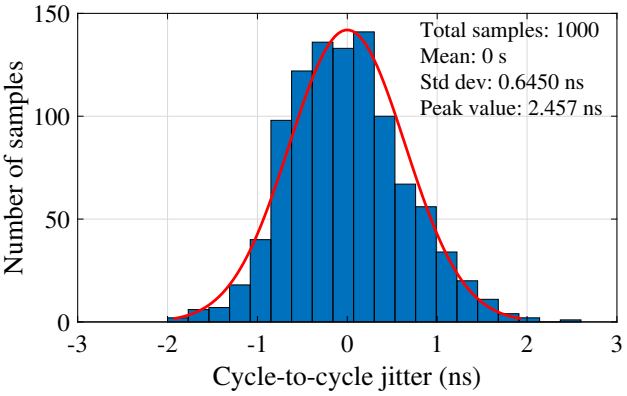


Fig. 10: The distribution of the values of the cycle-to-cycle jitter of one FLL chip ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $T = 25\,^{\circ}\mathrm{C}$, $f_{\mathrm{ref}} = 2\,\mathrm{MHz}$).

TABLE I: The measured typical values of the parameters of both FLL designs ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $T = 25\,^{\circ}\mathrm{C}$, $f_{\mathrm{ref}} = 2\,\mathrm{MHz}$). The sample size for the old design is 10 chips and for the new one is 20 chips.

Parameter	Unit	Old FLL design [2]	New FLL design
f_{out} mean	MHz	19.22	20.04
f_{out} std. dev.	kHz	200.80	32.79
I_{DD}	µA	91.85	111.13
Power consumption	µW	165.33	200.03
0.1-% settling time	µs	13.36	20.54
Area	$\mu\mathrm{m}^2$	40 998	41 535

smaller than the area of a PLL with the same output frequency.

The extreme values of the simulated transient parameters of both FLL designs across technology corners are shown in Table II. The simulation was conducted from a steady state by changing the reference frequency f_{ref} from $2\,\mathrm{MHz}$ to $2.1\,\mathrm{MHz}$. The degradation of the transient parameters in the new FLL design is due to the use of a single FVC in which the charge redistribution step for each FVC output voltage must wait for the previous charge redistribution step to finish, and therefore the opamp input voltages are generated alternately. Conversely, the FVCs in the FLL design from [2] allowed for an independent and simultaneous generation of the opamp input voltages. This resulted in an overall faster circuit compared to the FLL design in this work.

The comparison of the output frequencies of both FLL designs for different temperatures is shown in Table III. The output frequency of the old FLL design is temperature dependent, whereas the output frequency of the new FLL design shows no significant temperature dependence. The temperature coefficient of the output frequency is $4839\,\mathrm{Hz}/^{\circ}\mathrm{C}$ for the old FLL design and $-91.9\,\mathrm{Hz}/^{\circ}\mathrm{C}$ for the new one.

The current consumption of both FLL designs for different temperatures is shown in Table IV. The current consumption of the new FLL design is larger than the current consumption of the old FLL design. The temperature coefficient of the supply current is $38\,\mathrm{nA}/^{\circ}\mathrm{C}$ for the old FLL design and $147\,\mathrm{nA}/^{\circ}\mathrm{C}$ for the new one.

TABLE II: The extreme values of the simulated transient parameters of both FLL designs across technology corners ($f_{\mathrm{ref}}: 2\,\mathrm{MHz} \rightarrow 2.1\,\mathrm{MHz}$).

Parameter		Old FLL design [2]	New FLL design
Delay	Min	1.2	5.0
(µs)	Max	10.7	18.7
0.1-% settling time	Min	10.7	20.5
(µs)	Max	58.3	77.7
Overshoot (% of steady state)	Max	1	1.19

TABLE III: The measured output frequencies of both FLL designs for different temperatures ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $f_{\mathrm{ref}} = 2\,\mathrm{MHz}$). The sample size for the old design is 10 chips and for the new one is 20 chips.

	Output frequency (MHz)	
Temperature (°C)	Old FLL design [2]	New FLL design
−40	18.84 ± 0.27	20.06 ± 0.04
25	19.22 ± 0.20	20.04 ± 0.03
75	19.39 ± 0.17	20.05 ± 0.03

TABLE IV: The measured current consumption of both FLL designs for different temperatures ($V_{\mathrm{DD}} = 1.8\,\mathrm{V}$, $f_{\mathrm{ref}} = 2\,\mathrm{MHz}$). The sample size for the old design is 10 chips and for the new one is 20 chips.

	Supply current (µA)	
Temperature (°C)	Old FLL design [2]	New FLL design
−40	89.34 ± 1.36	99.97 ± 1.07
25	91.85 ± 1.09	111.13 ± 1.65
75	93.71 ± 1.05	116.76 ± 0.98

VI. Conclusion

This paper presents an improved design of a frequency-locked loop (FLL) circuit from [2] which suffered from a static frequency offset caused by the mismatch between the frequency-to-voltage converters (FVC). The FLL design in this work uses a single FVC with an upgraded architecture, thus eliminating the aforementioned static frequency offset. The circuit is implemented in a 180-nm CMOS process. The measurement results show that the output frequency of the FLL design from this work is more accurate and more precise than that of the design from [2]. The measurements also show that the new FLL design has an increased power consumption compared to the old one. The simulation results show that the delay, the 0.1-% settling time and the overshoot are increased for the new FLL design. The area of the new FLL design is marginally increased.

Acknowledgment

This work is supported by ams-OSRAM AG, Premstaetten, Austria.

References

[1] B. Razavi, *Design of CMOS Phase-Locked Loops: From Circuit Level to Architecture Level*. Cambridge University Press, 2020.

[2] A. Žamboki, L. Gočan, N. Bako, J. Mikulić, G. Schatzberger, and A. Barić, "Low-power frequency-locked loop circuit with short settling time," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 174–179.

[3] A. Djemouai, M. Sawan, and M. Slamani, "New frequency-locked loop based on CMOS frequency-to-voltage converter: design and implementation," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 5, pp. 441–449, 2001.

[4] ——, "High performance integrated CMOS frequency-to-voltage converter," in *Proceedings of the Tenth International Conference on Microelectronics (Cat. No.98EX186)*, 1998, pp. 63–66.

[5] H. T. Bui and Y. Savaria, "Design of a high-speed differential frequency-to-voltage converter and its application in a 5-GHz frequency-locked loop," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 3, pp. 766–774, 2008.

[6] ——, "High-speed differential frequency-to-voltage converter," in *The 3rd International IEEE-NEWCAS Conference, 2005.*, 2005, pp. 373–376.

[7] X. Ding, J. Wu, and C. Chen, "A fast settling frequency-to-voltage converter with separate voltage tracing technique for FLLs," in *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2021, pp. 700–703.

[8] B. Jamali and A. Babakhani, "A 0.2-2.6GHz instantaneous frequency-to-voltage converter in 90nm CMOS," in *2016 IEEE Radio and Wireless Symposium (RWS)*, 2016, pp. 45–47.

[9] JEDEC, "Definition of skew specifications for standard logic devices," September 2003, document number: JESD65B.

Low-Power CMOS Frequency Comparator

Martin Kováč, Miroslav Potočný, Daniel Arbet, Róbert Ondica, Richard Ravasz and Viera Stopjaková
Department of IC Design and Test,
Institute of Electronics and Photonics,
Slovak University of Technology, Bratislava, Slovakia
e-mail: martin_kovac@stuba.sk

***Abstract*—This paper is focused on the design and analysis of a fully on-chip frequency comparator (FC) implemented in 65 nm CMOS technology. The proposed FC employs a digital pre-processing stage followed by a rail-to-rail input and output (RRIO) voltage comparator, with low power consumption in units of μW. Evaluation accuracy is maintained for input frequencies up to 1 MHz. The FC is designed for the supply voltage of 1.2 V. The FC accepts signals with independent duty cycles and delays, without any correlation between the signals, making it an universal building block of more complex integrated systems. Presented simulation results show robustness of the proposed topology over process, temperature and supply voltage variations (PVT). The presented FC was used in complex ultra-low power System on-Chip (SoC).**

I. Introduction and Background

Today's electronic systems implemented on a single SoC require advanced control mechanisms (changing modes, adaptive switching, etc.) in order to achieve the required performance from the functional point of view, while also improving the energy efficiency of the whole system. By utilizing this approach, effective control of complex mixed-signal integrated circuits (ICs) can be leveraged to achieve better power efficiency of the system as a whole, ultimately leading to prolonged battery life in applications such as Internet-of-Things (IoT). To implement such an adaptive control approach sensing of selected parameters, which can affect the efficiency of whole system, is required. Using the sensed parameter as an input, a control mechanism can be introduced into the system. This controller can then be used to adjust some system parameters, so that energy consumption is minimized while maintaining the performance at a level required for the function of the device.

In many applications the frequency of a signal is used to determine the appropriate operating mode of a system. To reduce the time required for the change of the operating mode, thus minimizing any unwanted effects associated with the mode switch, a reliable method of detecting the signal frequency is required. The above mentioned FC is a simple way to accomplish this task. The FC compares an input signal frequency to a reference, and evaluates if it is higher or lower than the reference signal frequency. Various implementations of the FC are possible. These can be based on purely digital or mixed-signal circuits. One digital-only implementation, described in [1], uses an edge detector, a re-triggerable monostable multivibrator and a D-type Flip-flop (DFF). Here, the multivibrator timing interval is set to a duration derived from the reference frequency. This results in the multivibrator being re-triggered before timing out if the input frequency is higher than the reference. Another approach, described in [2], is based on two separate counters, one for the input and reference signal each. Similar methods are presented in [3], [4]. Here the two signals are also driving two counters [3], or simple DFFs. The output of these is instead of direct evaluation connected to a low-pass filter (LPF). An analog comparator is then used to compare the LPFs output signals.

To the extend of authors knowledge, the number of research papers dedicated to this topic is rather limited. This presents an opportunity to introduce novel topologies suited for on-chip implementation and focusing on low power consumption, while maintaining compatibility with standard CMOS technology nodes. For these reasons, the paper presents design and analysis of a fully on-chip FC, implemented in 65 nm CMOS technology. The paper is structured as follows: Section II describes design of the FC as well as its major circuit blocks. Section III presents achieved results of the FC circuit simulation. Conclusions are drawn in Section IV.

II. Proposed FC Topology

A. Frequency comparator

A simplified schematic diagram of the proposed FC is shown in Fig. 1. It is mainly based on the topology presented in [4], with small modifications made necessary by the requirements of fully on-chip implementation. Input frequencies (F_1 and F_2) are fed to the DFFs trough edge detectors (EDs). The EDs are used because of the master-slave architecture of DFFs in the particular technology node used. The outputs of the DFFs are compared in a NAND gate, which generates a reset signal for the DFFs. In the case when $F_1 < F_2$ the output of the first DFF (Q_F1) will exhibit a smaller pulse density than Q_F2. Since the DFF output is integrated by a LPF, the mean value of signal $CTRL_F1$ will be smaller then mean value of signal $CTRL_F2$. If we increase the frequency of $F1$, the voltage in the node $CTRL_F1$ will also be increase. When the frequency of F_1 is increased above F_2, the comparator will change the state of its output and indicate that $F_1 > F_2$. The designed comparator can evaluate the state of the inputs correctly even if the input and reference frequencies are significantly lower than cut-off frequency of the LPF. Details of how this is achieved are given in the following section. To ensure the reliability of the proposed FC, the initialization circuit block containing

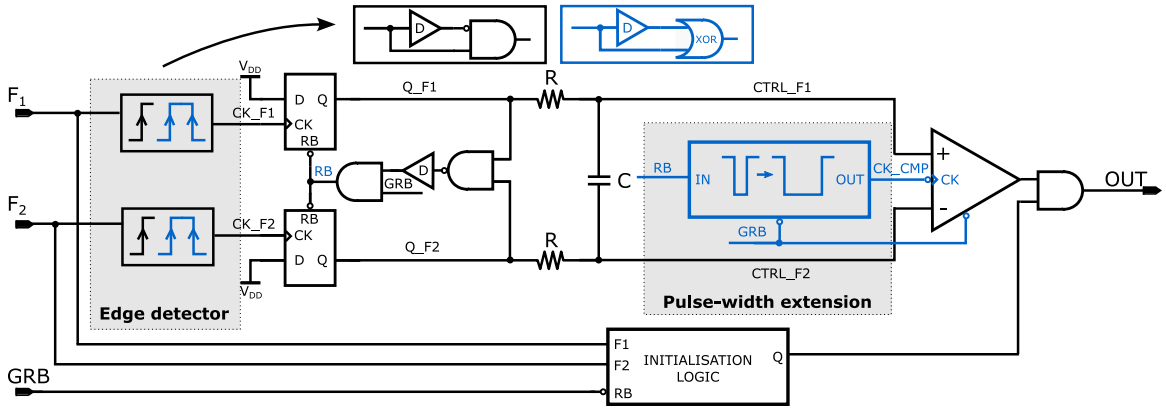


Fig. 1. Schematic diagram of the proposed FC circuit.

the phase's alignment detection logic was added. This sets a known state at the output of the FC after an initialization reset (GRB).

B. Analog comparator with hysteresis

The most important block of the proposed topology is an analog comparator with hysteresis, which is used to compare signals in the nodes $CTRL_F1$ and $CTRL_F2$. To obtain the widest possible input frequency range, and the capability to evaluate frequencies significantly lower than the cut-off frequency of the LPF, the analog comparator needs to have RRIO capability combined with internal hysteresis. Propagation delay and power consumption of the analog comparator have to be taken into consideration, to achieve low-power consumption of the FC as a whole. All these requirements need to be taken into account in the design phase of the analog comparator. The schematic diagram of the analog comparator used in the proposed FC is shown in Fig. 2.

In order to obtain a wide voltage range of operation, a rail-to-rail input stage was used in the designed analog comparator [5]–[7]. Therefore the NMOS input differential pair M1-M2 is connected in parallel with the PMOS input pair M3-M4. Transistors MNX and MPX were added to effectively equalize the transconductance of the combined input stage over the whole common-mode input range. Transistors M5-M6 represent the tail current sources, while diode-connected transistors M7-M10 represent the load of the NMOS and PMOS input differential pairs. The outputs of this input stage, the node $ST1_P$ and $ST1_N$, are connected to a second stage. In the second stage, labeled *Hysteresis stage* in in Fig. 2, voltage hysteresis of the analog comparator can be adjusted. Transistors M13-M14 form the differential input pair of the second (*hysteresis*) stage. The load of the second stage differential pair is a negative resistance, implemented by the use of transistors M16-M17. The voltage hysteresis can be set by the sizing of transistor M16-M19, which form a positive feedback network. To increase the output voltage swing of the analog comparator, a level translation block was employed. This block utilizes the commonly used cross-coupled voltage level shifter topology. Diode-connected transistors M26-M27 are used to limit the pull-up strength of M28-M29 and to increase the pull-down speed [8]. To limit the short-circuit current in the level shifter, transistors M30-M31 were added. A similar topology, with different input transistor arrangement, and its detailed description are presented in [9]. The last block of the designed analog comparator is a JK latch circuit. The latch is used as the evaluation stage. The JK latch uses the commonly used NAND gate architecture. Layout of the designed analog comparator, as well as of the entire FC, is shown in Fig. 3. The dimension of the designed block is 28 μm x 65 μm and thus occupy chip area of 1820 μm^2. Most of the

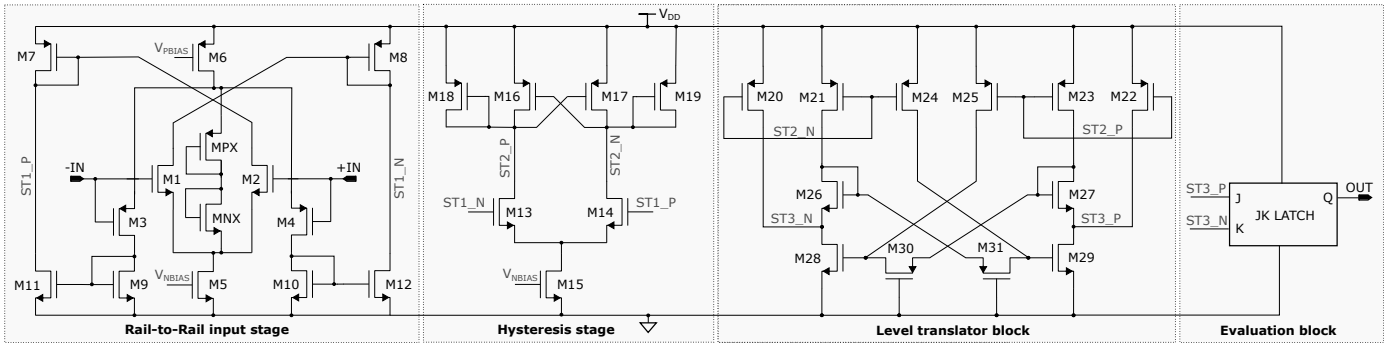


Fig. 2. Schematic diagram of the designed comparator.

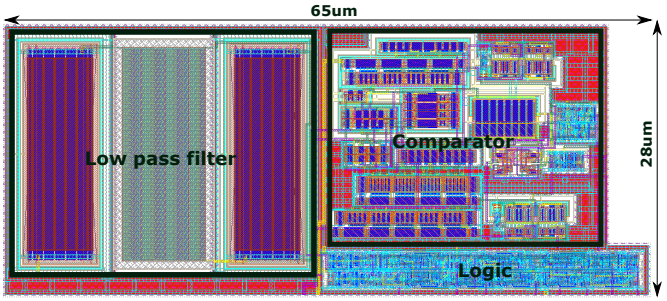


Fig. 3. The layout of proposed FC (with designed analog comparator) realized in standard UMC 65 nm CMOS process.

area is occupied by the LPF and analog comparator.

It is important to note, that instead of the analog comparator used in the proposed FC architecture, a dynamic clocked comparator (DCC) can be employed. This approach could improve some important features and decrease the overall power consumption. Such realization as well as some other improvements are discussed in the next section (Section II-C) and results are presented in Section III.

C. Possible improvements of proposed FC

Since the proposed topology of the FC is not purely digital, a disadvantage can be seen in the total power consumption. It can be especially noticeable at low input frequencies, where the static bias current of the analog comparator is the dominant source of power consumption. To further improve the power consumption of the proposed FC, a digital comparator needs to be employed. The improved FC architecture, utilizing a DCC, is highlighted in blue in Fig. 1. We used the comparator presented in [10] and shown in Fig. 4 to provide a comparison to the analog comparator. The clock signal for the DCC is generated by a pulse extension block, which uses the reset signal of the DFFs as its input. In order to increase detection speed, a XOR-based edge detector sensitive to both rising and falling edges can be used in the FC (improved FC architecture is depicted in blue in Fig. 1). This comes at the expense of increased power consumption and produces valid results only if the duty cycle of both the input and reference signals is 50%. For this reason, to ensure reliability of detection under all circumstances, using a edge detector sensitive to only one edge type is highly recommended.

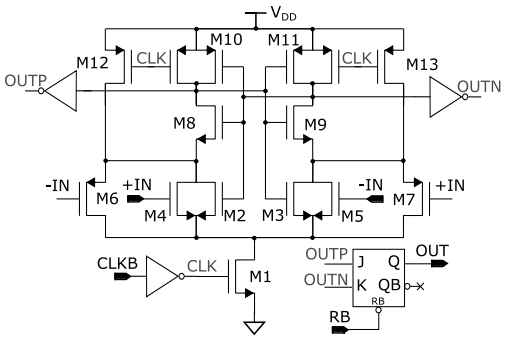


Fig. 4. Schematic of a dynamic clocked comparator presented in [10].

III. Achieved results

The results obtained by pre-layout simulation of the proposed FC and comparator are presented in this section. Fabrication process fluctuations were taken into account, and FC was simulated using Corner Analysis in the temperature range from $-20°$C to $85°$C. In these simulations F_1 is used as the input frequency ranging from $0.5F_{REF}$ to $1.5F_{REF}$ with the duty cycle of 50%. F_2 is used as the reference frequency F_{REF} with the same 50% duty cycle as F_1. The input frequency was swept with a selected slope (ramp) and the compared frequencies were in phase at the starting point of ramp, i.e. $F_1(t = 0) = 0.5F_{REF}$.

Waveforms of important circuit nodes, representing the operation of the proposed FC, are shown in Fig. 5. As can be observed, when the input frequency is higher than the reference frequency, the output goes high and vice-versa. Results in Fig. 5 have been achieved under typical condition with a ramp slope of 1000 MHz/s and the reference frequency value of 500 kHz.

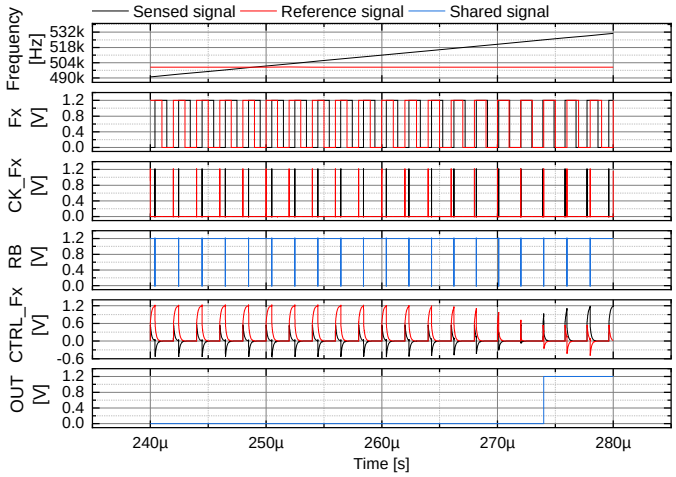


Fig. 5. Basic operation of the proposed FC, V_{DD} = 1.2 V, T = 23°C, R = 100 kΩ, C = 500 fF, F_{REF} = 500 kHz, ramp slope = 1000 MHz/s.

The investigation of frequency detection threshold dependence on LPF cut-off frequency is shown in Fig. 6. In this analysis, we consider a fixed value of capacitor (C = 500 fF in Fig. 1), while the resistor value was swept from 1 kΩ to 100 MΩ. Results were obtained for different values of the reference frequency, with the slope of the input frequency ramp changing accordingly to the reference frequency value (as described in Fig. 6). It can be observed, that the frequency thresholds are slightly above and below the reference frequency (about $\pm$ 10%) which is a result of the asynchronous nature of the signals. For low values of the LPF cut-off frequency (approaching value of F_{REF}), the frequency thresholds are shifted more apart. Therefore, the cut-off frequency of the LPF should be adjusted one decade

higher than the input and reference frequencies which are compared in FC.

Fig. 7 shows the detection delay time for different F_{REF}. It is obvious that detection delay time depends on the cut-off frequency of LPF used in the FC. It can be also observed that the detection delay time is constant for cut-off frequencies one decade higher than F_{REF}. This parameter also deviates from 25 μs to 56 μs depending on the reference frequency and ramp slopes. However, for LPF cut-off frequencies below F_{REF}, delay can be higher than 100 μs. Therefore, the LPF needs to be adjusted to higher frequencies (10x higher than F_{REF}) to improve the detection delay time of the FC circuit.

Detection delay time as well as frequency thresholds depend on the change rate of the input frequency (F_1) and phase alignment/shift at the near cross-point of the processed frequencies. In order to investigate this dependency, a ramp of the input frequency was swept with aligned 50% duty cycled signals. Simulation results of this analysis are shown in Fig. 8. Simulations were performed with the reference frequency $F_2 = F_{REF} = 250\ kHz$, while the input frequency (F_1) was swept from $0.5F_{REF}$ to $1.5F_{REF}$. A jagged behaviour without smooth trend of detection delay time dependency on the ramp slope of F_1 can be observed. Similar result was also obtained for frequency thresholds as a consequence of asynchronous nature of processed signals. On the other hand, if the slope increases, the signal's alignment (characterized by exchanging of the leading and lagging signal) is more likely to happen sooner. Therefore, the trend of detection delay time will have a declined tendency (see Fig. 5). This might be caused by evaluation mechanisms used in the proposed FC.

Another important parameter is the power consumption of the designed FC. The dependency of power consumption on the LPF cut-off frequency is shown in Fig. 9. One can conclude that the power consumption is increased for higher LPF cut-off frequencies. This is caused by the filter capacitor fully discharging/charging during each cycle. Otherwise, for

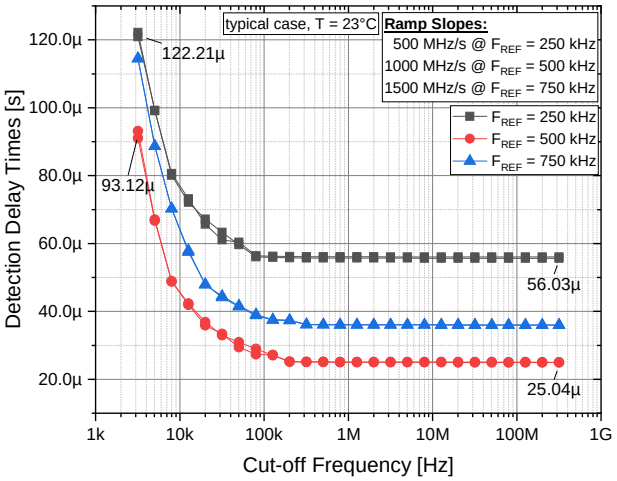


Fig. 7. Detection delay time, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 1 kΩ - 100 MΩ.

lower LPF cut-off frequencies, the power consumption below 1 μW can be achieved. Naturally, if $F_1 < F_2$, power consumption will be lower than in the case when $F_1 > F_2$. The power consumption will increase for higher frequencies (F_1 and F_2) being compared. Thus, for $F_2 = 750\ kHz$, the power consumption will be higher than 1.5 μW, which is still acceptable result.

Fig. 10 shows the power consumption of FC over the process corners and temperature. These results were obtained for a fixed value of LPF resistors (R = 100 kΩ, corresponding to $\approx$ 3 MHz LPF cut-off frequency). The lowest power consumption can be observed in SS corner (below 1μW), while FF corner represents the worst case scenario (2.83 μW) in terms of power consumption. In the typical corner, the power consumption of 1.05 μW was reported (for the input frequency of 250 kHz).

The main parameters of the proposed FC are summarized

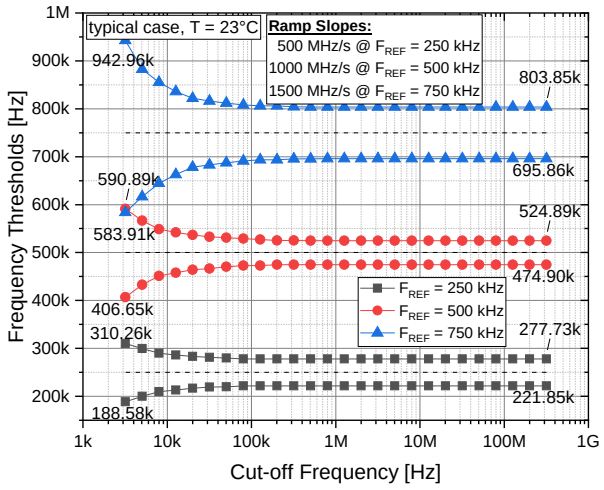


Fig. 6. Frequency detection thresholds, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 1 kΩ - 100 MΩ.

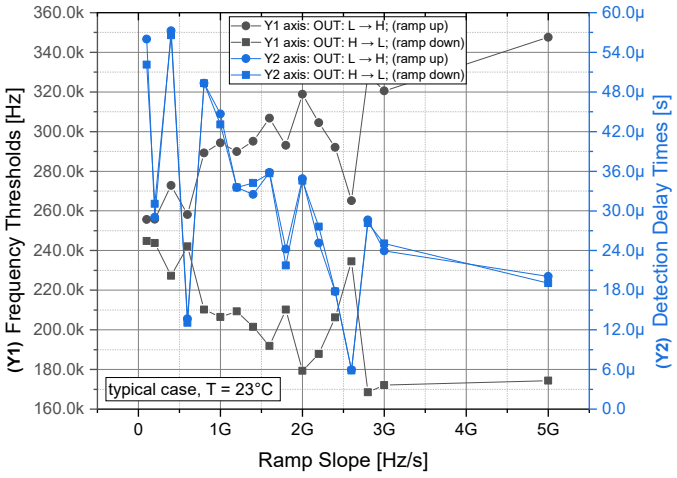


Fig. 8. Frequency thresholds and detection delay in relation to on ramp slope, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 100 kΩ, F_{REF} = 250 kHz.

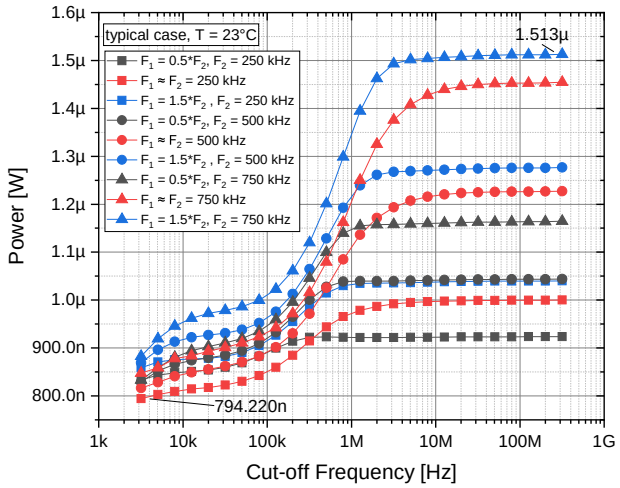


Fig. 9. Dependence of power consumption on the cut-off frequency, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 1 kΩ - 100 MΩ.

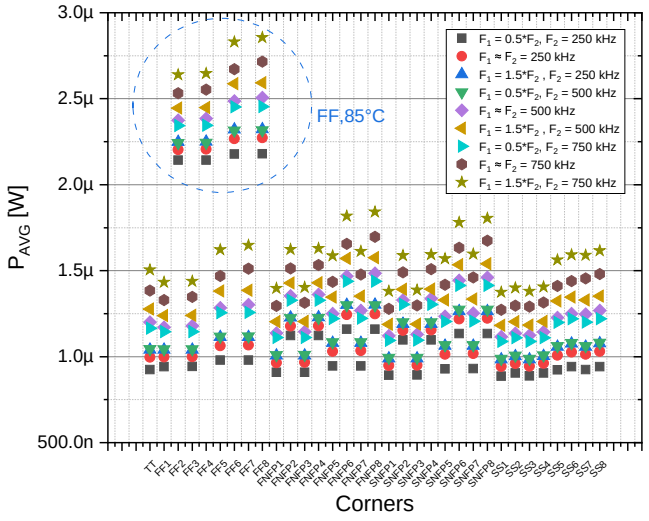


Fig. 10. Power consumption of the F circuit C over process corners, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 100 kΩ.

in Table I. These parameters were obtained using the reference frequency of 250 kHz. The proposed topology of FC is not sensitive to variations of process parameters, with the exception of increased mismatch in Monte Carlo analysis while the input frequency has a very low slope. This is achieved by the use of a differential topology, use of digital pre-processing circuits and the entire structure design approach. Therefore, only typical values of frequency thresholds are introduced in Table I. An exception is the power consumption, which deviates over process corners, and in the worst case, it will reach the value of 2.83 μW.

The performed observation and analysis presented above shows the efficiency of proposed FC with the designed analog comparator. However, the total power consumption is mainly caused by the static power consumption of the analog comparator. Therefore, the improved FC architecture utilizing

TABLE I
MAIN PARAMETERS OF THE PROPOSED FC

Parameter	Min	Typ	Max	Unit
Supply Voltage	-	1.2	-	V
Temperature Range	-20	27	85	°C
Total Power	0.94	1.05	2.83	μW
Frequency threshold L to H	-	277.7	-	kHz
Frequency threshold H to L	-	221.8	-	kHz
Detection Delay Time	-	56	-	μs
Area	-	1820	-	μm^2

a DCC has been proposed. This approach could outperform the previous design in terms of power consumption (for the case when comparing low frequency signals).

Fig. 11 shows signals in selected circuit nodes of the improved architecture of FC (version with DCC). The time interval around the output change is also zoomed in. This gives better understanding of the FC operation. The improved FC architecture is similar to the version with an analog comparator, except for the evaluation phase, which is triggered by the *RB* and *CK_CMP* signals. Also the frequency of pulses is doubled, due to the use of XOR-based edge detectors.

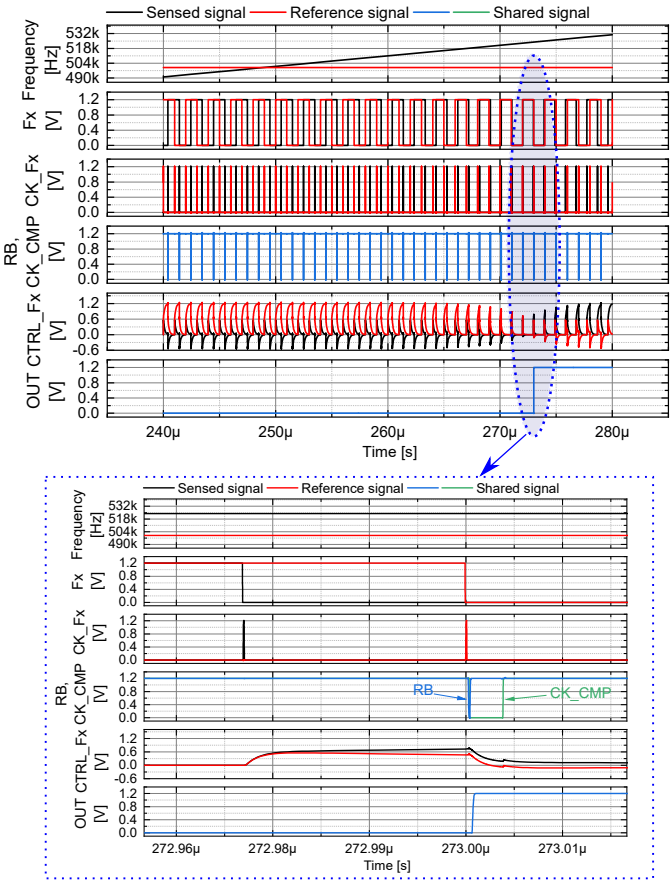


Fig. 11. Important signals of improved FC, V_{DD} = 1.2 V, T = 23°C, R = 100 kΩ, C = 500 fF, F_{REF} = 500 kHz, ramp slope = 1000 MHz/s.

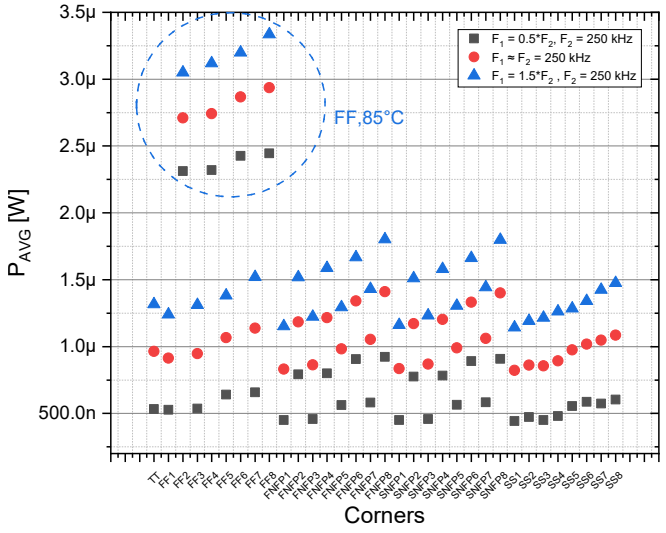


Fig. 12. Power consumption of improved FC, V_{DD} = 1.2 V, T = 23°C, C = 500 fF, R = 100 kΩ.

Fig. 12 shows the power consumption of the improved FC over the process corners and temperature. It can be observed that the improved architecture (that uses the DCC) has the power consumption of 0.53 μW in the typical case. This represents an improvement of about 50 %. On the other hand, the worst case power consumption is higher (higher than 3 μW in the FF corner) for the improved version of the FC. This is caused by the XOR-based edge detectors. As they are sensitive to both rising and falling edges, the switching frequency of the whole system is doubled, therefore the power consumption is increased in all process corners. While the XOR-based edge detectors improve speed of the FC, higher power consumption and strict 50% duty cycle of the input signals are major drawbacks and should be considered. In other words, the one-edge detector should be likely preferred for general purpose of FC and low-power applications.

IV. Conclusion

A new frequency comparator, suitable for full on-chip implementation, was designed in 65 nm CMOS technology. The proposed topology uses a RRIO analog comparator with hysteresis, which can be also replaced by a pure DCC solution to improve the overall power consumption. The proposed topology of the FC is comparatively insensitive to corner-based PVT process variations. Since the power consumption is below 2 μW (in the typical case for processed frequencies under 1 MHz), the proposed FC architecture represent an ultra low-power solution for complex mixed-signal SoC.

Acknowledgment

This work was supported in part by the grant VEGA 1/0731/20 and the Slovak Research and Development Agency under grant APVV-19-0392. This work has also received funding from the Electronic Components and Systems for European Leadership Joint Undertaking under grant agreement No 876868. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Slovakia, Netherlands, Spain, Italy.

References

[1] D. A. Pezzutti, "Digital Frequency Comparator," U.S. Patent US3 753 130A, 8 14, 1973. [Online]. Available: https://patentimages.storage.googleapis.com/8f/8c/74/5a44d92ceb468d/US3753130.pdf

[2] A. E. Tsuguhiro Okada, "Digital frequency comparator circuit," U.S. Patent US3 987 365A, 10 19, 1976. [Online]. Available: https://patentimages.storage.googleapis.com/be/54/18/eb832db6514064/US3987365.pdf

[3] G. M. M. M. I. K. Marius-Gheorghe Neag, Mici McCullagh, "Frequency Comparator and Early-Late Detector," U.S. Patent US20 160 191 035A1, 6 30, 2016. [Online]. Available: https://patentimages.storage.googleapis.com/9f/f6/2f/4a840eebb2c9ba/US20160191035A1.pdf

[4] "Frequency comparator - 74HCT74." [Online]. Available: https://pe2bz.philpem.me.uk/Power/-%20Pulse/P-106-FreqCompare/cir10.htm

[5] D. Arbet, G. Nagy, G. Gyepes, and V. Stopjaková, "Design of rail-to-rail operational amplifier with offset cancelation in 90nm technology," in *2012 International Conference on Applied Electronics*, 2012, pp. 17–20.

[6] B. Razavi, *Design of Analog CMOS Integrated Circuits*, 1st ed. McGraw-Hill, 2001.

[7] R. Hogervorst, J. Tero, R. Eschauzier, and J. Huijsing, "A compact power-efficient 3 v cmos rail-to-rail input/output operational amplifier for vlsi cell libraries," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 12, pp. 1505–1513, 1994.

[8] H. Shao and C.-Y. Tsui, "A robust, input voltage adaptive and low energy consumption level converter for sub-threshold logic," in *ESSCIRC 2007 - 33rd European Solid-State Circuits Conference*, 2007, pp. 312–315.

[9] S. Ali, S. Tanner, and P. A. Farine, "A robust, low power, high speed voltage level shifter with built-in short circuit current reduction," in *2011 20th European Conference on Circuit Theory and Design (ECCTD)*, 2011, pp. 142–145.

[10] M. A. Al-Qadasi, A. Alshehri, A. Alturki, A. S. Almansouri, K. N. Salama, H. Fariborzi, and T. Al-Attar, "Rail-to-rail complementary input strongarm comparator for low-power applications," *IET Circuits, Devices & Systems*, vol. 14, no. 6, pp. 898–900, 2020. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-cds.2019.0361

Measurement System for Characterization of a Resistor Array in 180-nm CMOS Technology

Ivana Franković*, Franjo Mikić*, Josip Mikulić†, Niko Bako†, Gregor Schatzberger†, Adrijan Barić*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
† ams-OSRAM AG, Premstaetten, Austria
ivana.frankovic@fer.hr

***Abstract*—Integrated circuits are very sensitive to variations of temperature. All components used in IC design (transistors, resistors, capacitors and other components) are characterized with respect to the change of temperature by defining the temperature coefficients, usually of the first and second order. The temperature coefficient is usually specified as a fixed number for a given type of a resistor, i.e. it is assumed that the temperature coefficient does not depend on the dimension *L* (length) and *W* (width) of the resistor, which is usually not the case. This paper describes the system for characterization of the temperature coefficients of resistors having various lengths, widths and number of fingers, as well as different resistor types. The system is based on an array of resistors consisting of 20 rows and 10 columns, which is sufficient to characterize 200 resistors. Each resistor cell can be selected separately, i.e. only one cell is active during the measurement. The measurement is based on the 4-wire method, i.e. on forcing the current and measuring the voltage on the resistor. Digital circuits and analogue switches are built into the integrated circuit to enable the selection of each resistor cell.**

***Keywords*—resistor array, temperature coefficient, 4-wire sensing, Kelvin sensing**

I. INTRODUCTION

Nowadays, precise and temperature compensated operation is required in many applications. Examples are different types of oscillators, current/voltage references and others [1]–[4]. One of the critical components required for realising these systems are various types of resistors. To achieve required system performance, detailed temperature characterization of the used resistors is necessary.

Thewes et al. [5] presented a model which shows a relation between polysilicon resistors mismatch and their grain size in 0.65-µm CMOS technology. Their design consists of a resistor array arranged in rows and columns where each resistor can be selected for 4-wire measurement. A similar test structure is used in this paper. Different structures which use resistor arrays with each resistor connected directly to the pads without a switching matrix are described in [6] and [7]. Temperature characterization of polysilicon resistors to determine coefficients of CMC-R2 compact model is shown in [8].

The main goal of this paper is to describe a system which can be used to characterize resistor temperature coefficient dependence on width and length of polysilicon and diffusion resistors. These resistors are arranged in an array which consists of 20 rows and 10 columns where each cell can be selected for the 4-wire measurement. The influence of possible sources of errors on the resistor temperature coefficient characterization is taken into consideration.

This paper is organized as follows. Section II describes the architecture of a resistor array. Simulations of resistance measurements and their results are shown and discussed in Section III. Finally, a conclusion is given in Section IV.

II. RESISTOR ARRAY ARCHITECTURE

There are multiple different resistor types available in the used 180-nm CMOS technology (poly, diffusion and well resistors) which differ in their sheet resistance and temperature coefficients (T_{Ci}). The T_{Ci} of these resistors depend on their width (*W*) and length (*L*) [9] which is not listed in the device specifications where they are given as constants. The temperature dependence of a resistance can then be modelled as:

$$R(T) = R_0 \cdot (1 + T_{C1} \cdot \Delta T + T_{C2} \cdot \Delta T^2), \qquad (1)$$

where R_0 is the resistance at the nominal temperature T_0, T_{C1} and T_{C2} are the first and second-order temperature coefficients and ΔT is the difference between current temperature T and T_0.

Based on the specified first-order T_C in this technology, two resistor types are chosen for the characterization of the T_C dependence on *W* and *L*: *p+* diffusion resistor RDIFFP and *p+* poly resistor RPOLY1PC as they have the lowest positive and the highest negative T_C, respectively. Their first-order temperature coefficients are shown in Table I.

The system for T_C characterization is designed as an array of resistor cells where each cell is defined by the following parameters: number of serially connected unit resistors it is comprised of and dimensions of a single unit resistor (*W* and *L*). These parameters also define a naming convention for cells used in this paper: RP_FX_WY_LZ for RPOLY1PC and RD_FX_WY_LZ for RDIFFP, where *X* is the number of identical units (fingers), *Y* is the width *W* and *Z* is the length *L*. The resistor cells are made with

TABLE I: First-order T_C of selected resistors

Resistor type	T_{C1} [ppm/°C]
RPOLY1PC	-238
RDIFFP	1383

1, 5 or 9 unit resistors where each one of them can have four different values of widths (0.5 µm, 1 µm, 2 µm, 4 µm) and lengths (5 µm, 20 µm, 50 µm, 100 µm). Table II shows the maximum and minimum resistances for each resistor type using previously mentioned combinations.

Fig. 1 shows an example of a layout of two resistor cells. Each cell has a guard ring which is used to ensure a stable potential of a local substrate. The guard rings are connected to the external high-level voltage V_{NWELL} for diffusion resistors and to the reference voltage V_{SS} for poly resistors. Each resistor cell contains two dummy unit resistors connected to the guard ring and positioned above and below other unit resistors to achieve better matching [10]. Two columns of contacts at both ends of each unit resistor are used to reduce the resistance of a connection to the metal layer.

The measurement of each cell is performed using the 4-wire method as high accuracy is desired. Fig. 2 shows the implementation of this method where both current-forcing and voltage-sensing connections can be disconnected from the resistor cell using controllable analogue switches. Each switch is implemented as shown in Fig. 3. The dimensions of pMOS MP1 are W = 4 µm, L = 0.5 µm and of nMOS MN1 are W = 2 µm, L = 0.5 µm.

The resistor cells are organized as an array with selectable rows and columns. Only a single resistor cell can be measured at a time by selecting its corresponding row and column, i.e. when its inputs $RSEL$ and $CSEL$ are high level. Additionally, two high-level inputs FSW and SSW have to be provided to separately enable current-forcing and voltage-sensing connections.

The measured resistance is then by the definition:

$$R_M = \frac{V_V}{I_B}, \tag{2}$$

where V_V is the voltage measured by the voltmeter and I_B is the current forced by the current source. The finite input resistance R_V of the voltmeter and non-zero resistance of switches in the voltage-sensing loop introduce an error in the measurement of an actual resistance of a resistor cell. This effect can be seen when (2) is rewritten as:

$$R_M = \frac{R_{real} \cdot R_V}{R_{real} + R_V + 2R_S}, \tag{3}$$

where R_{real} is the actual resistance of a resistor cell, R_V is the input resistance of the voltmeter and R_S is the resistance of each analogue switch in on-state. It is visible that the error can be reduced by increasing input resistance of the voltmeter or reducing on-state resistance of switches. Also, a lower error is obtained when measuring lower resistance values. Usually, the on-state resistance of the switches can be neglected if R_{real} is known to be of the relatively higher value. In that case, the actual resistance of the resistor cell can be simply calculated as:

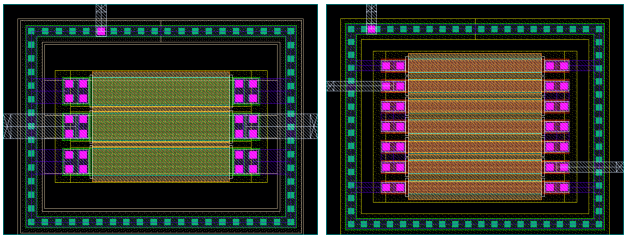
$$R_{real} = \frac{R_V \cdot R_M}{R_V - R_M}. \tag{4}$$

The encircled part of the schematic in Fig. 2 represents a basic building block of the resistor array and its layout is shown in Fig. 4. The switches and the resistor cell beneath them are placed in the right part of the block while the bus which connects them to the other blocks is placed in the left part.

The resistor array with 10 columns and 20 rows, totalling in 200 blocks is then built by placing these blocks

TABLE II: Minimum and maximum resistance in the array

Resistor type	Minimum	Maximum
RPOLY1PC	441.80 Ω	619.2 kΩ
RDIFFP	187.10 Ω	219.6 kΩ



(a) RD_F1_W1_L5 (b) RP_F5_W0.5_L5

Fig. 1: Example of a layout for both resistor types: (a) diffusion resistor with one unit and dummy units above and below the resistor, (b) poly resistor with 5 units and two dummy units.

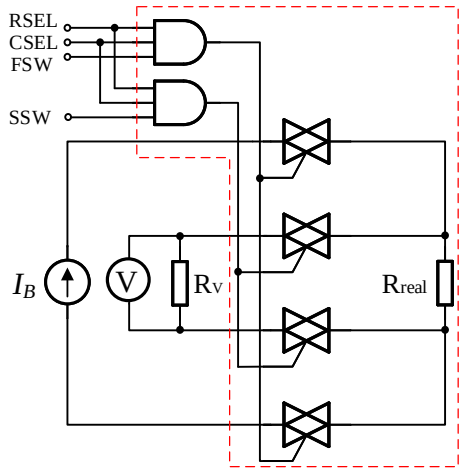
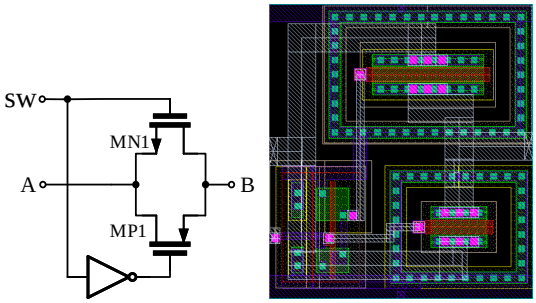


Fig. 2: Schematic of the 4-wire method for measuring resistance. R_{real} represents a measured resistor.



(a) (b)

Fig. 3: Schematic and layout of an analogue switch.

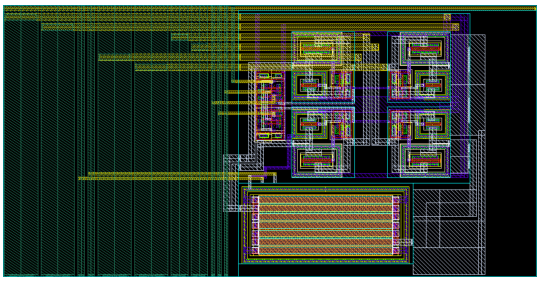


Fig. 4: Layout of a block with switches and resistor.

next to each other as shown in Fig. 5. There are 64 different blocks used, most of them being repeated 3 or 4 times to fill out the array. As previously explained, each block can be selected for measurement by setting corresponding $CSEL$ and $RSEL$ signals to high-level. Then, after also enabling switches in current-forcing and voltage-sensing loops with FSW and SSW signals, the test current flowing through FP and FN can be set and the voltage across VSP and VSN can be measured. According to the measured voltage and known test current, resistance of the selected block can be determined.

The signals for selecting particular row $RSEL_1$ to $RSEL_{20}$ and column $CSEL_1$ to $CSEL_{10}$ are generated with two separate decoders. The row and column decoders are formed using a 2-to-4 decoder with output enabling signal CS which is shown in Fig. 6(a). The column decoder shown in Fig. 6(b) is built by combining five 2-to-4 decoders which results in 16 outputs, out of which only 10 are used. Then, by combining two column decoders and an additional digital logic at the inputs, a row decoder is formed as shown in Fig. 6(c). Again, only 20 outputs of available 32 are used. The required four inputs of the column decoder and five inputs of the row decoder are externally provided through pins of the chip, the same as the shared input CS which enables outputs of decoders. Fig. 7 shows the layout of decoders on the left side and the first nine blocks on the right side. Table III lists external pins of the chip.

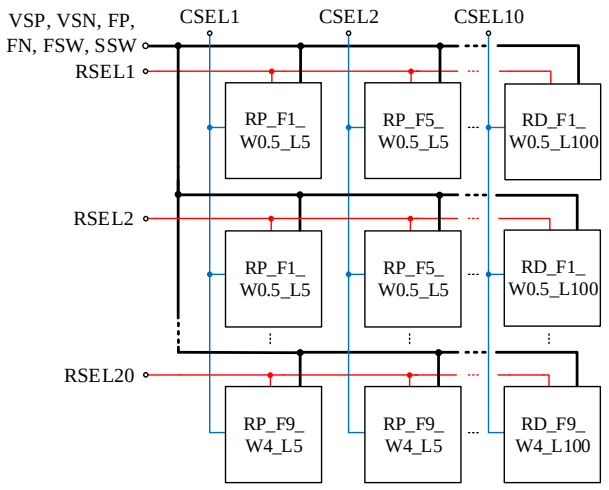


Fig. 5: Schematic of the resistor array.

TABLE III: Pins

Pin name	Description
V_{DD}	power supply
V_{SS}	reference voltage (gnd)
V_{NWELL}	n-well polarization voltage
FP	force pin (current in)
FN	force pin (current out)
VSP	voltage sense plus
VSN	voltage sense minus
SSW	sense switch select
FSW	force switch select
CS	chip select
IN_C1 to IN_C4	column select 1 to 4
IN_R1 to IN_R5	row select 1 to 5

III. SIMULATIONS AND RESULTS

As previously explained, only a single block can be measured at a time. To measure the resistance of all 200 blocks, the input signals for decoders are generated so that each 10-ms a block in the following column is measured, row by row. The simulation of this measurement for the blocks in the first row is shown in Fig. 8. This simulation is performed at the nominal conditions (T = 27 °C, V_{DD} = 1.8 V) and with parasitic components extracted from the layout of the whole design. The input resistance of the voltmeter R_V is set to 1 MΩ and the test current to 1 µA. The same approach is used in all other simulations.

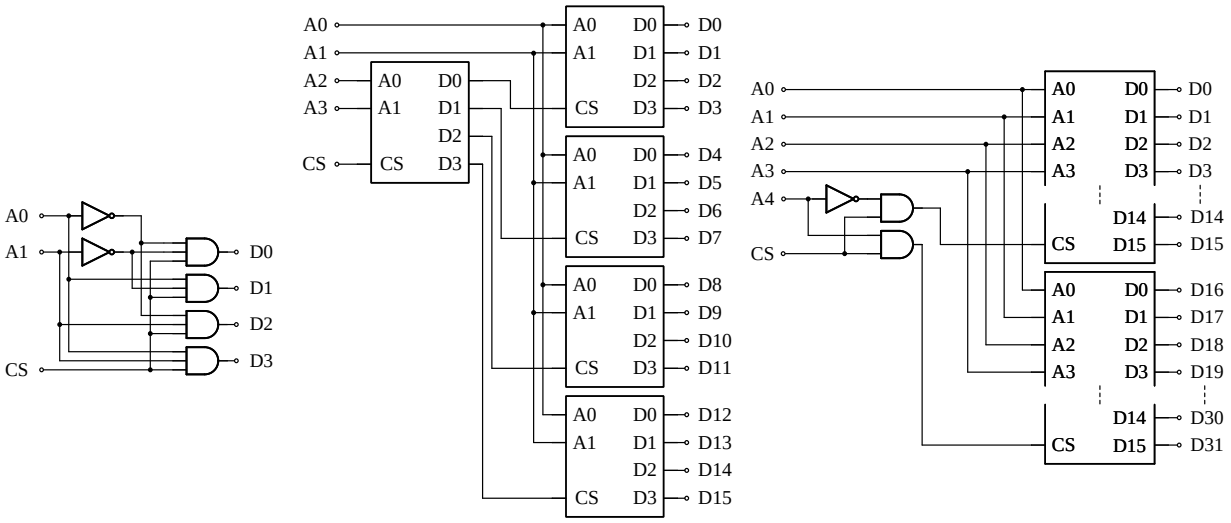
The change in temperature and the power supply voltage V_{DD} affects the value of the measured resistance. The simulations at temperatures -40 °C and 125 °C at the nominal V_{DD} are performed to compare these results with those acquired at nominal conditions. Similarly, the impact of the supply voltage variation on the measured resistance is evaluated for $\pm 10\%$ change of the nominal V_{DD} at the nominal temperature. Only a subset of the available blocks is used for simulations to show general trends and these results are presented in Table IV. As the absolute change of resistance depends on the nominal resistance, a relative change of the measured resistance is used for comparison:

$$\Delta R_r(T, V_{DD}) = \frac{R_M(T, V_{DD}) - R_M(T_0, V_{DD0})}{R_M(T_0, V_{DD0})}, \quad (5)$$

where V_{DD0} is the nominal supply voltage.

The measured resistance changes with respect to the temperature due to the temperature coefficient of resistors, as previously shown in (1). The measured resistance value of RPOLY1PC resistors decreases as the temperature rises due to their negative T_C. The opposite is true for RDIFFP resistors, for which the measured resistance increases as the temperature rises because of their positive T_C. The relative change of the measured resistance is larger for diffusion resistors as they have a larger absolute T_C value than poly resistors.

Simulations of the supply voltage variations show that the measured resistance increases for both types of resistors as the voltage increases. The cause of this error is the change of the on-state resistance of analogue switches



(a) 2-to-4 decoder (b) column decoder (c) row decoder

Fig. 6: Schematics of decoders.

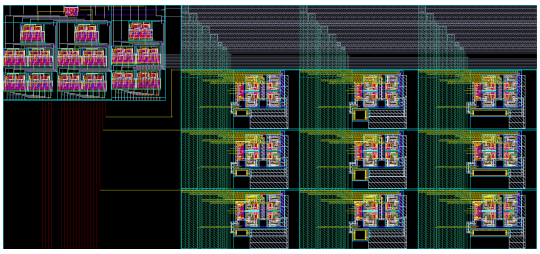


Fig. 7: Layout of a resistor array (decoders and first nine blocks).

which impacts the measured resistance. When the supply voltage is lower, the overdrive voltage $|V_{GS}-V_{TH}|$ is also lower which increases the channel resistance of transistors. This increase of the on-state resistance of switches reduces the measured resistance of the block as can be seen from (3). Similarly, the supply voltage increase reduces the on-state resistance of the switches which results in higher measured resistance.

Monte Carlo simulations with 50 samples at the nominal conditions are performed for three different blocks of both resistor types as shown in Table V. For both types of resistors the same triplets are chosen with dimensions (W and L) and a number of unit resistors such that they cover a wide range of resistances. These simulations are also performed for the same resistors as in the chosen blocks but without the measurement system. This would show the impact of the measurement system on the measured resistance. For easier comparison, the input resistance of the voltmeter is removed from the measured resistance as in (4) to obtain the actual resistance of the block. This value is then compared to the resistance measured without the impact of the measurement system. The results show that the standard deviation increases for higher resistance of a block, which is expected but is not a good indicator of the measured resistance dispersion. A better indicator is the coefficient of variation which shows dispersion relative to the value of the measured resistance. It shows that for the both cases (with and without the impact of the measurement system), a relative dispersion is higher for blocks which use unit resistors that have one small dimension. This occurs because the same deviation of one dimension results in its relative change to be higher if that dimension is small and this consequently results in higher relative resistance change. Impact of the measurement system can be seen in the mean values of the measured resistances. Measurement system decreases the measured resistance due to the nonzero on-state resistance of the switches as previously described. Resistors show similar dispersions for both cases.

Existence of leakage currents that flow through unselected blocks can change the value of the measured resistance. The main contributors to the leakage are the currents that flow from the substrate to the sources/drains of transistors inside the switches and similarly from the substrate connections to the terminals of diffusion resistors. The leakage currents obtained by the simulations are shown in Table VI. These simulations are performed by using two voltage sources where one of them with voltage values of 0.2 V or 1.1 V is connected to the external pin FP while the other one with voltage values of 0 V or 0.9 V is connected to FN. In this way, a constant 200 mV voltage drop across the resistor is achieved with

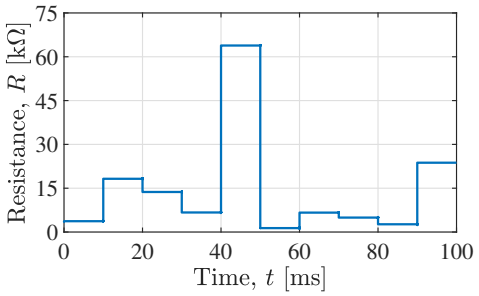


Fig. 8: Simulation of the first row measurement.

TABLE IV: Various temperature and voltage simulation results

Resistor	V_{DD} = 1.8 V T = 27 °C	V_{DD} = 1.8 V T = -40 °C		V_{DD} = 1.8 V T = 125 °C		V_{DD} = 1.62 V T = 27 °C		V_{DD} = 1.98 V T = 27 °C	
	Res (Ω)	Res (Ω)	ΔR_r (%)	Res (Ω)	ΔR_r (%)	Res (Ω)	ΔR_r (%)	Res (Ω)	ΔR_r (%)
RP_F1_W1_L5	1803.40	1848.70	2.51	1767.10	-2.01	1799.12	-0.24	1806.30	0.16
RP_F5_W0.5_L20	64.96 k	65.99 k	1.60	64.38 k	-0.90	64.79 k	-0.26	65.07 k	0.17
RP_F9_W0.5_L100	380.70 k	384.25 k	0.93	378.15 k	-0.67	380.31 k	-0.10	380.93 k	0.06
RD_F1_W1_L5	715.20	664.43	-7.10	802.90	12.26	713.29	-0.27	716.55	0.19
RD_F5_W0.5_L20	24.21 k	22.24 k	-8.13	27.53 k	13.70	24.16 k	-0.22	24.26 k	0.20
RD_F9_W0.5_L100	179.53 k	166.45 k	-7.29	200.25 k	11.54	179.18 k	-0.20	179.74 k	0.12

TABLE V: Monte Carlo simulation results (T = 27 °C, V_{DD} = 1.8 V)

Resistor	With the measurement system			Without the measurement system		
	Mean, μ (Ω)	Std dev, σ (Ω)	Coefficient of variation, σ/μ (%)	Mean, μ (Ω)	Std dev, σ (Ω)	Coefficient of variation, σ/μ (%)
RP_F1_W0.5_L5	3.74 k	270.60	7.24	3.77 k	269.40	7.16
RP_F5_W2_L20	16.65 k	604.10	3.63	16.75 k	601.50	3.59
RP_F9_W0.5_L100	617.40 k	45.65 k	7.39	621.50 k	46.03 k	7.41
RD_F1_W0.5_L5	1.35 k	104.50	7.72	1.36 k	104.60	7.69
RD_F5_W2_L20	6.91 k	257.16	3.72	6.95 k	255.40	3.67
RD_F9_W0.5_L100	220.10 k	15.32 k	6.96	221.30 k	15.51 k	7.01

TABLE VI: Leakage currents simulation results (V_{DD} = 1.8 V)

Temperature (°C)	27	27	125	125	27	27	125	125
Corner	nominal	nominal	nominal	nominal	wp	wp	wp	wp
V_{FP} (V)	0.2	1.1	0.2	1.1	0.2	1.1	0.2	1.1
V_{FN} (V)	0	0.9	0	0.9	0	0.9	0	0.9
I_{FP} (A)	-667.97 p	33.46 p	25.38 n	-623.31 p	-509.16 p	33.028 p	84.01 n	-1.09 n
I_{FN} (A)	-1.84 n	-135.08 p	-45.23 n	-841.21 p	-3.78 n	-135.68 p	-105.47 n	-1.33 n

the possibility of setting two different common-mode voltages of the resistor terminals. Then, while none of the blocks is selected, the current through each pin is measured with the currents being positive if they flow from the outside into the chip. Simulations are performed for all 8 possible combinations of temperature (27 °C, 125 °C), process corner (nominal, worst power *wp*) and input voltages (0 V and 0.2 V, 0.9 V and 1.1 V) while the supply voltage is set to the nominal value. The worst power corner and temperature of 125 °C are chosen in addition to the nominal values as the leakage currents increase in these conditions. The results show that in addition to the worst power corner and temperature increase, input voltages set to 0 V and 0.2 V also result in higher leakage currents, with the worst case being for all these conditions combined. In that case, the leakage currents reach 10% of the previously used test current of 1 µA. Therefore, the input voltages should be set to approximately half of the supply voltage for the actual measurement to reduce the error.

IV. Conclusion

This paper shows the architecture of a resistor array with 200 blocks arranged in 20 rows and 10 columns which is used for characterization of temperature coefficients of two resistor types. The resistance is measured using the 4-wire method and error sources in these measurements are explained. Temperature simulations show the dependence of the measured resistance on the temperature coefficient, while the voltage simulations show the impact of the analogue switch on-state resistance. Monte Carlo simulations demonstrate that the relative resistance dispersion changes with the respect to the unit resistor dimensions. Leakage currents can be minimised by properly setting the input voltages as shown in simulations.

ACKNOWLEDGMENT

This work is supported by ams-OSRAM AG, Premstaetten, Austria.

References

[1] H. Çetinkaya, A. Zeki, A. Girgin, and T. C. Karalar, "Composite Resistor Technique for Process and Temperature Compensations of Low Power Ring Oscillators," in *2019 IEEE 10th Latin American Symposium on Circuits & Systems (LASCAS)*, 2019, pp. 29–32.

[2] H. Chun and T. Lehmann, "CMOS Current Reference Generator Using Integrated Resistors," in *2010 International Conference on Electronics and Information Engineering*, vol. 1, 2010, pp. V1–290–V1–294.

[3] T.-C. Lu, H.-W. Zan, and M.-D. Ker, "Temperature Coefficient of Poly-Silicon TFT and Its Application on Voltage Reference Circuit With Temperature Compensation in LTPS Process," *IEEE Transactions on Electron Devices*, vol. 55, no. 10, pp. 2583–2589, 2008.

[4] C. C. Sheng and M. T. Hua, "The research of temperature compensation technology of high-temperature pressure sensor," in *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, vol. 5, 2011, pp. 2267–2270.

[5] R. Thewes, R. Brederlow, C. Dahl, U. Kollmer, C. Linnenbank, B. Holzapfl, J. Becker, J. Kissing, S. Kessel, and W. Weber, "Explanation and Quantitative Model for the Matching Behaviour of Poly-Silicon Resistors," in *International Electron Devices Meeting 1998. Technical Digest (Cat. No.98CH36217)*, 1998, pp. 771–774.

[6] H. Tuinhout, G. Hoogzaad, M. Vertregt, R. Roovers, and C. Erdmann, "Design and Characterization of a High-Precision Resistor Ladder Test Structure," *IEEE Transactions on Semiconductor Manufacturing*, vol. 16, no. 2, pp. 187–193, 2003.

[7] W. Tian, P. Steinmann, E. Beach, I. Khan, and P. Madhani, "Mismatch Characterization of a High Precision Resistor Array Test Structure," in *2008 IEEE International Conference on Microelectronic Test Structures*, 2008, pp. 11–16.

[8] B. Landgraf, A. Vujasin, and B. Ankele, "Modelling of Silicided and Blocked Poly-Si Resistors in 90 nm CMOS with the CMC-R2 Model," in *Proceedings of the 18th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2011*, 2011, pp. 99–102.

[9] H.-M. Chuang, K.-B. Thei, S.-F. Tsai, and W.-C. Liu, "Temperature-Dependent Characteristics of Polysilicon and Diffused Resistors," *IEEE Transactions on Electron Devices*, vol. 50, no. 5, pp. 1413–1415, 2003.

[10] R. J. Baker, *CMOS Circuit Design, Layout, and Simulation*, 4th ed. Hoboken, NJ: Wiley, 2019.

Performance Analysis of 1-MHz Voltage-Controlled Ring Oscillator Designed in 180-nm CMOS Technology for Phase-Locked Loop

Alma Traživuk, Adrijan Barić
University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia
Tel: +385 (0)1 6129547, Fax: +385 (0)1 6129653, Email: alma.trazivuk@fer.hr

***Abstract*—This paper presents a voltage-controlled ring oscillator designed and implemented in 180-nm CMOS technology to be used in a phase-locked loop. The performance of the circuit as a function of the temperature and supply voltage is measured. The measurements of the voltage-controlled oscillator are compared with the postlayout simulations.**

***Keywords*—voltage-controlled oscillator, phase-locked loop**

I. Introduction

The key component of every phase-locked loop (PLL) is a voltage-controlled oscillator (VCO) [1]. Application of VCOs today includes clock generation in high-speed microprocessors and carrier synthesis in communication systems [2]. Most of these applications require different oscillation frequencies at different times i.e. that the oscillators are tunable in terms of output frequency. For example, a microprocessor runs with a high-speed clock for computationally-intensive tasks but for less demanding tasks it needs a low-speed clock [1]. For VCOs, as their name implies, the controlling input signal is the voltage.

The 1-MHz voltage-controlled ring oscillator in this paper is designed and processed in 180-nm CMOS technology. The VCO is built of an odd number of delay cells that are forming a ring structure as shown in Fig. 3. Fig. 1 shows the schematic of one delay cell. The cell is powered down by $pdvco$ signal. When $pdvco$ is low, the VCO begins to oscillate. The signal $pdvco$ goes through two inverters $INV1$ and $INV2$ before it arrives to the $M6$ gate and therefore has a delay comparing with $pdvco$ signal on $M7$ gate [3]. This ensures the start of the oscillations. The inverters $INV3$ and $INV4$ are added to outputs of each delay cell so the each cell is loaded with the same parasitic capacitance. To ensure proper operation of the VCO, bulks of pMOS transistors are connected to supply voltage whereas bulks of nMOS transistors are connected to ground. Each of 21 cells is driven by the current coming from the Voltage-to-current converter (V2I), shown in Fig. 2. Namely, the output frequency of ring oscillators depends on semiconductor parameters, respectively on the supply voltage [4]. So, these oscillators require improved supply rejection. This is ensured with V2I converter, which converts control voltage (V_{ctrl}) into the current supplying the cells. The amount of that current will depend on V_{ctrl}, considering that V2I will give the highest current for V_{ctrl} = 0 V and for every greater V_{ctrl} the current will be lower. So, the voltage that controls the operation of the VCO is V_{ctrl}. The VCO circuit is designed and presented in [3]. This paper gives the measurements and comparison with simulations.

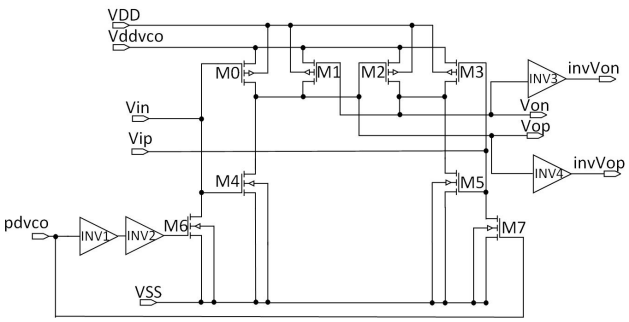


Fig. 1. Schematic of a single delay cell.

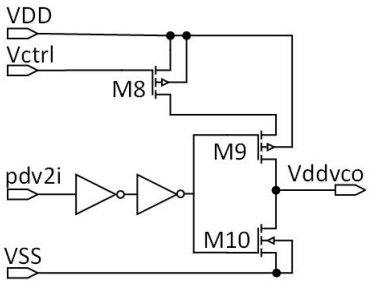


Fig. 2. Schematic of V2I converter.

This paper is organized in five sections. Section II presents an overview of the supply voltage dependence. Temperature dependence results are presented in Section III. Section IV compares measured results of the manufactured chips with simulation results. Finally, Section V brings the conclusion.

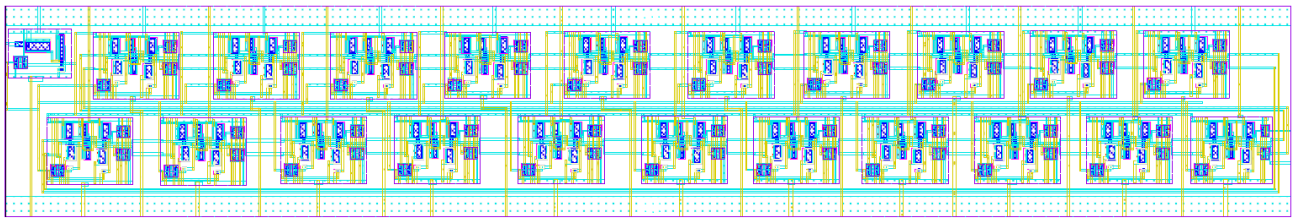


Fig. 3. Layout of the voltage-controlled ring oscillator consisting of 21 cells and V2I converter (top left).

II. Supply Voltage Dependence

The measurement setup is based on National instruments PXIe-1095 platform. Three PXI-4139 source measure unit (SMU) modules are used for the measurements. The first module provides the supply voltage for the VCO, the second module provides 3.3 V for the 10 pads and the third module provides the control voltage V_{ctrl} which is a variable. A fourth module is DMM (digital multimeter) and it is connected to a Pt100 sensor used to measure the temperature in the test chamber. The output frequency of the VCO is measured with the frequency counter Keysight 53220A. The measurement process is automated in LabVIEW so that V_{ctrl} is swept from 0 V to 1 V in 50 mV steps for three different supply voltage values: nominal (1.8 V), decreased by 10% (1.62 V) and increased by 10% (1.98 V). The frequency characteristics measured for 10 chip samples with the nominal supply voltage at room temperature are presented in Fig.4. Table I contains standard deviation σ and values of output frequency matching the characteristics in Fig. 4.

TABLE I
AVERAGE FREQUENCY AND σ FOR V_{DD} = 1.8 V AND $\overline{T}$ = 22.4 °C FOR 10 SAMPLES.

Vctrl [V]	f [MHz]	σ [kHz]
0	2.69	19.3
0.1	2.62	19.0
0.2	2.53	18.7
0.3	2.43	18.4
0.4	2.29	18.1
0.5	2.12	18.0
0.6	1.87	17.9
0.7	1.58	16.7
0.8	1.28	15.1
0.9	0.99	13.3
1	0.71	10.8

The characteristics show that, for the nominal supply voltage and room temperature, the VCO achieves the design goal because it provides output frequency of approximately 1 MHz for the control voltage V_{ctrl} around 0.9 V.

An important parameter when designing a VCO for a PLL is the frequency derivation by control voltage called VCO gain or K_{VCO}. It is desirable to minimize the variation of K_{VCO} across the tuning range of the VCO for better settling behavior of a PLL [5]. Fig. 5 plots

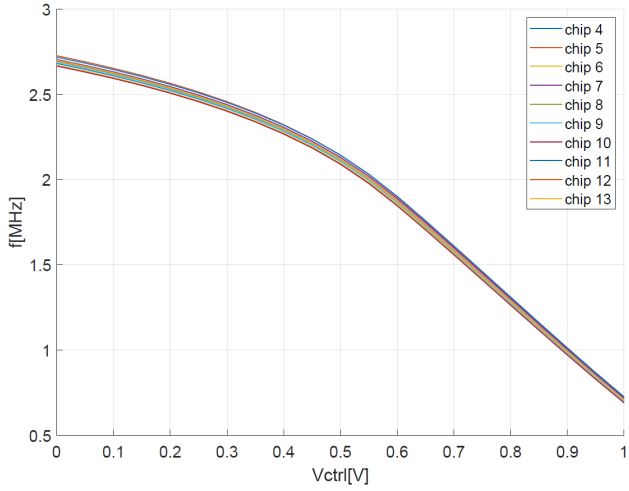


Fig. 4. Measured frequency vs. V_{ctrl} for V_{DD} = 1.8 V and room temperature ($\overline{T}$ = 22.4 °C).

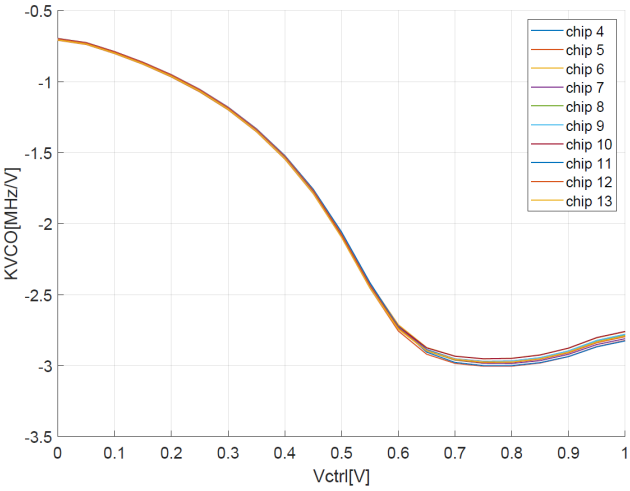


Fig. 5. Measured VCO gain vs. V_{ctrl} for V_{DD} = 1.8 V and room temperature ($\overline{T}$ = 22.4 °C).

the K_{VCO} versus V_{ctrl} for the nominal supply voltage of 1.8 V and room temperature. The slope of the curves for all 10 chip samples indicate that the characteristics are approximately constant around the value of V_{ctrl} from 0.6 V to 1 V required for the desired frequency of 1 MHz.

Table II shows the results of the VCO output frequency and standard deviation σ for the supply voltage decreased by 10% from the nominal value to 1.62 V. The change

of supply voltage to lower values causes slightly lower output frequency of the VCO, so the targeted 1 MHz output is achieved already for V_{ctrl} = 0.7 V, while the standard deviation is approximately the same as for the nominal supply voltage.

TABLE II
AVERAGE FREQUENCY AND σ FOR V_{DD} = 1.62 V AND $\overline{T}$ = 22.4 °C FOR 10 SAMPLES.

Vctrl [V]	f [MHz]	σ [kHz]
0	2.34	17.8
0.1	2.26	17.4
0.2	2.17	17.3
0.3	2.05	16.8
0.4	1.89	16.7
0.5	1.66	16.6
0.6	1.38	15.7
0.7	1.08	13.9
0.8	0.79	11.9
0.9	0.52	9.5
1	0.28	6.6

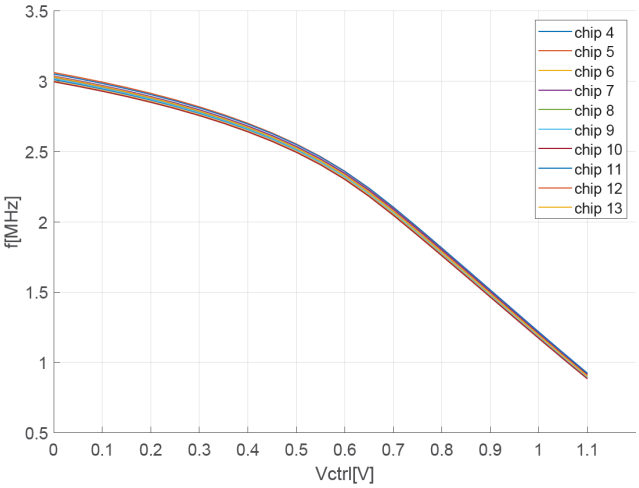


Fig. 6. Measured frequency vs. V_{ctrl} for V_{DD} = 1.98 V and room temperature.

Similarly, the increased supply voltage by 10% causes the VCO to give the desired output for slightly higher control voltage of 1 V with σ = 14.4 kHz. This is shown in Fig.6. The measurements for the supply voltage range of +/-10% from the nominal value show that there is a tendency of variation of the VCO output frequency with the variation of supply voltage. This is understandable due to the fact that ring oscillators typically provide decreased frequencies if the supply voltage is reduced [6].

Before the manufacturing of the chip, simulations with extracted parasitic components are conducted in Cadence Virtuoso tool. The results of these simulations for the output frequency of the VCO are different from the measurements. Table III shows the comparison of the results of simulations and measurements for the VCO, when supply voltage is altered +/-10% and for the control voltage of 1 V.

TABLE III
COMPARISON OF SIMULATIONS AND MEASUREMENT RESULTS FOR V_{ctrl} = 1 V.

	frequency	
VDD	simulations	measurements
1.62 V	0.38 MHz	0.28 MHz
1.8 V	1.18 MHz	0.71 MHz
1.98 V	2.18 MHz	1.19 MHz

The output frequency of the processed and manufactured VCO is lower than the frequency in simulations. This is expected because of the parasitic resistances and capacities in the layout. This effect is also desirable because the required frequency is achieved for lower V_{ctrl}. There are two sources of possible errors: V2I and ring oscillator. The ring oscillator operates in a wide voltage range and the extraction of the capacities is probably not as good as we expected. The influence of the V2I converter is probably smaller because it regulates only the DC current flowing into the ring oscillator.

III. TEMPERATURE DEPENDENCE

The manufactured chips are supposed to be characterized in a temperature range from -40 °C to 75 °C. Due to the imperfections of the temperature chamber, reached temperatures that were closest to these were approximately -32 °C and 73 °C. Fig. 7 presents the results of measurements for this temperature range and nominal supply voltage. 10 chip samples were measured for each temperature.

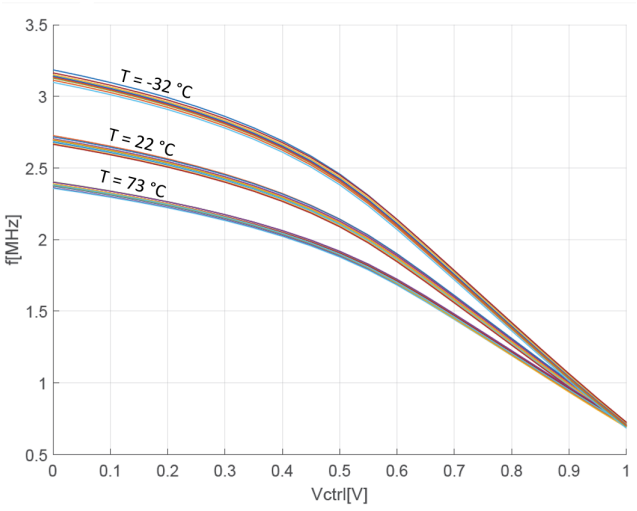


Fig. 7. Measured frequency vs. control voltage for T = -32 °C, 22 °C, 73 °C and V_{DD} = 1.8 V.

At higher temperatures, the output frequency of the VCO is lower. The divergence between the characteristic decreases as the control voltage increases. The output frequency of 1 MHz is achieved for the value of V_{ctrl} located in the area were the characteristics are only

slightly moved apart. The result is satisfying because the targeted output on temperature extremes is achieved for approximately the same V_{ctrl} as for the room temperature. K_{VCO} for the same temperature range is shown in Fig.8. At temperature extremes, K_{VCO} varies +/-17% from the nominal value achieved on room temperature.

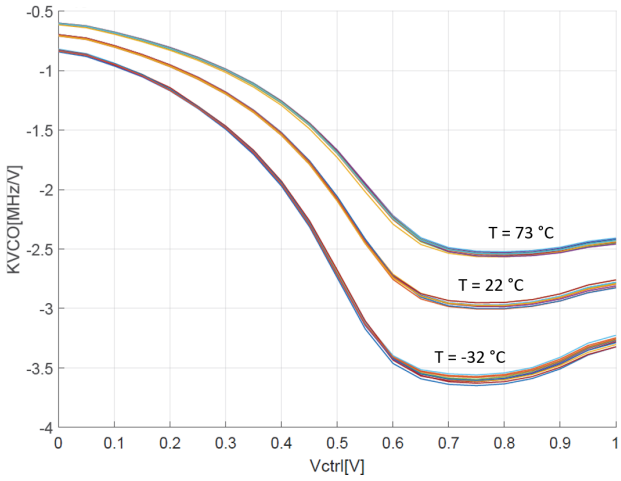


Fig. 8. Measured K_{VCO} vs. control voltage for T = -32 °C, 22 °C, 73 °C and V_{DD} = 1.8 V.

For the supply voltage altered +/-10% across the temperature range of -32 °C to 73 °C, the output frequency characteristics show good results because V_{ctrl} necessary for targeted VCO output is similar as for the room temperature. This is shown in Fig. 9 and Fig. 10. During the measurements it is established that for the minimum supply VCO operates near it's limit of V_{ctrl} and stops functioning at lowest output frequency of 0.5 MHz. The measurement in LabVIEW is then adapted so that V_{ctrl} is swept from 0 to 0.9 V. This is visible in Fig. 9, as the characteristics for the temperature extremes are slightly shorter than nominal.

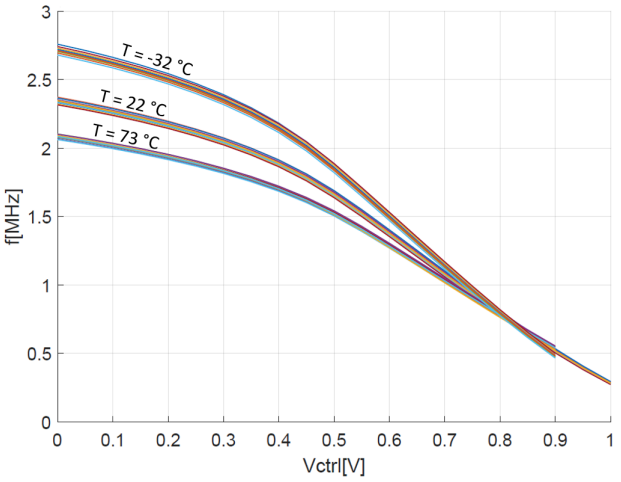


Fig. 9. Measured frequency vs. control voltage for T = -32 °C, 22 °C, 73 °C and V_{DD} = 1.62 V.

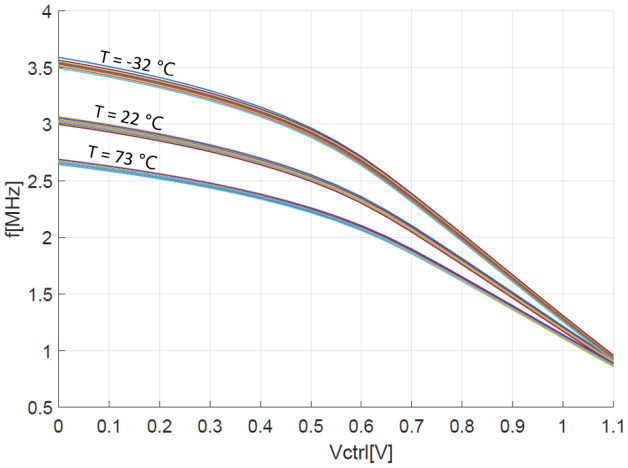


Fig. 10. Measured frequency vs. control voltage for T = -32 °C, 22 °C, 73 °C and V_{DD} = 1.98 V.

IV. Comparison of Processed Chip and Postlayout Simulations

Before the manufacturing the chip is simulated in Cadence Virtuoso tool. Postlayout simulations are conducted for the nominal corner (T = 27 °C and V_{DD} = 1.8 V) with extracted RC parasitic components included. The results of postlayout simulations showed that the VCO oscillates at lower frequencies when extracted parasitic components are taken into account. Table IV shows the comparison of the VCO output between the postlayout simulation results and measurement results of the processed and manufactured chip. Targeted output frequency of the VCO in postlayout simulations is gained for V_{ctrl} slightly above 1 V, more accurate 1.05 V. For this control voltage, the output frequency in simulations was 1.18 MHz.

TABLE IV
Comparison of Simulations and Measurement Results, V_{DD} = 1.8 V, T = 22 °C.

	frequency	
Vctrl	postlayout	manufactured
0 V	5.44 MHz	2.68 MHz
100 mV	5.31 MHz	2.62 MHz
200 mV	5.14 MHz	2.53 MHz
300 mV	4.93 MHz	2.43 MHz
400 mV	4.68 MHz	2.29 MHz
500 mV	4.34 MHz	2.12 MHz
600 mV	3.88 MHz	1.87 MHz
700 mV	3.29 MHz	1.58 MHz
800 mV	2.67 MHz	1.28 MHz
900 mV	2.06 MHz	0.991 MHz
1 V	1.46 MHz	0.707 MHz
1.05 V	1.18 MHz	0.574 MHz

However, when manufactured chip was measured, smaller V_{ctrl} was required to gain the targeted frequency of 1 MHz. Precisely, the output frequency of 0.99 MHz is obtained for the control voltage of 0.9 V. So for the same

control voltages, the measured output frequencies of the VCO are smaller than those gained with simulations, as it is shown in Table IV. The reason of this difference lies in the fact that processed and manufactured chip has some parasitic resistances and capacities that are not included in Cadence Virtuoso postlayout simulations, even though these simulations give results with RC parasitic components of the layout included.

V. Conclusion

The performance of the 1-MHz voltage-controlled ring oscillator manufactured in 180-nm CMOS technology is evaluated. The measurement results of the manufactured VCO show that there is a small output frequency variation with the change of work conditions, temperature and supply voltage. The supply voltage decreased by 10% results in a smaller output frequency which means that a smaller control voltage is required for the desired output frequency. Similarly, increased supply voltage results in a higher output frequency and V_{ctrl}. Regarding the temperature dependence, chip samples are tested in a temperature range from -32 °C to 73 °C for the nominal supply voltage and +/-10%. The results show that the targeted output frequency of 1 MHz is achieved for approximately the same V_{ctrl} value as for room temperature, for all three different supply voltages. The measurement results of the manufactured VCO are compared with postlayout simulation results. The comparison shows that on manufactured chip, the required frequency is achieved for smaller V_{ctrl} than in simulations. The difference is caused by parasitic resistances and capacities in the layout.

Acknowledgement

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures FASCON (IP-2019-04-8959).

References

[1] B. Razavi, *Design of CMOS Phase-Locked Loops: From Circuit Level to Architecture Level, 1st edition*. Cambridge University Press, 2020.

[2] T. Miyazaki, M. Hashimoto, and H. Onodera, "A performance comparison of PLLs for clock generation using ring oscillator VCO and LC oscillator in a digital CMOS process," in *ASP-DAC 2004: Asia and South Pacific Design Automation Conference 2004 (IEEE Cat. No.04EX753)*, 2004, pp. 545–546.

[3] A. Traživuk and A. Barić, "1-MHz voltage-controlled ring oscillator designed in 180-nm CMOS technology for implementation in phase-locked loop," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 180–184.

[4] E. J. Pankratz and E. Sanchez-Sinencio, "Multiloop high-power-supply-rejection quadrature ring oscillator," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 9, pp. 2033–2048, 2012.

[5] B. Razavi, *Design of Analog CMOS Integrated Circuits, 1st edition*. McGraw-Hill, 2001.

[6] I. Skeledžija, J. Mikulić, and A. Barić, "200-MHz and 400-MHz self-biased temperature-compensated ring oscillators in 180-nm CMOS technology," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2021, pp. 113–118.

Analysis of energy consumption for SPI and I2C communications in ultra-low power embedded systems

Josip Zidar*, Ivan Aleksi*, Tomislav Matić*
* Faculty of Electrical Engineering, Computer Science and Information Technology, Osijek
jzidar@ferit.hr; ialeksi@ferit.hr; tmatic1@ferit.hr

***Abstract*—Developing ultra-low-power devices requires low-power components, algorithms, and communication protocols. For environmental monitoring along the supply chain, products may travel a long way from the distributor to the customer. In addition, extreme climatic circumstances, such as temperatures below 0 °C, affect battery capacity. In order to monitor package's environment, it is crucial to ensure that all embedded system parts use as little energy as possible. This paper addresses how much power the SPI and I2C communication protocols use. These protocols are widely used in low-power embedded devices for sensor communication. Based on its results, this research article presented recommendations for developing ultra-low power embedded devices that use the SPI/I2C protocol.**

Keywords—I2C, SPI, low power, embedded system, supply chain

I. INTRODUCTION

Inter Integrated Circuit (I2C) and Serial Peripheral Interface (SPI) communication protocols are typically used in modern embedded systems to communicate with on-board peripherals at low or medium data transfer rates. Article [1] discusses the different parameters that must be considered while picking a protocol, including features, advantages, disadvantages, the number of devices, data transfer rate, and power consumption.

I2C is a serial communication protocol invented by Philips labs that allows numerous devices to be linked to a single bus, with one device acting as the controller and the others as targets. I2C is commonly used for low-speed devices, like microcontrollers, sensors, actuators, real-time clocks, monitoring devices, power management devices, memories, and so on, to communicate with each other. I2C allows for the connection of up to 128 devices to a single bus, and it is able to support communication speeds of up to 3.4 Mbps. The I2C standard's most recent modifications are in [2].

SPI is a synchronous serial communication protocol originally developed by Motorola that permits several devices to be linked to a single bus, with one device acting as the controller and the others functioning as targets. On an embedded system, SPI is often used for communication between high-speed devices, including control devices, cameras, display controllers, communications, and memory banks. Supported are communication speeds of up to 60 Mbps. The number of target devices on a bus can range from one to several hundred, but is typically restricted to a dozen (e.g., four, eight, or sixteen). Certain specialised hardware implementations may be able to accommodate numerous target devices, each with its own chip-select pin. Consequently, the number of available input/output (I/O) pins on a microcontroller represents significant constraints. It is possible to utilise additional hardware, like an I/O expander or a multiplexer, to overcome the constraint on the number of available I/O pins [3].

I2C is commonly used for connecting low-speed devices, whereas SPI is used for connecting high-speed devices. I2C is simpler, more straightforward, and slower than SPI, which is more complicated, sophisticated, and faster. Although both protocols coexist and are utilised for inter-chip communication, designers of ultra-low power embedded systems should consider the energy consumption of the implemented design.

I2C communication protocol's main advantages are:

- communication requires only two wires, and the microcontroller requires fewer pins
- multi-target functionality enables communication with multiple target devices which are connected to the same bus
- easy to implement and does not require any special hardware or software
- supports data transfer rates of up to 3.4 Mbps
- supports multi-controller systems and is adaptable to a broad range of applications
- flow control and error handling, which results in a more reliable protocol.

I2C communication protocol's main disadvantages are:

- half-duplex communication, in which only one direction of communication is possible at once
- restricted communication range, not suitable for long-distance use
- in a multi-controller system, there is a possibility of address collision between devices
- a maximum of 128 devices are allowed
- comparatively slower than other communication technologies such as SPI or RS-232
- limited data transfer capabilities may not be suitable for applications requiring high speeds.

SPI communication protocol's main advantages are:

- full-duplex communication permits the simultaneous transfer of data in both directions
- well-defined interface and intuitive signalling
- It enables data transfer rates ranging from a few kilobits per second to 60 Mbps, making it suited for high-speed applications
- addressing-free, making it an efficient protocol for communicating with a single peripheral device
- supports numerous target devices, allowing a single controller to communicate with multiple targets.

SPI communication protocol's main disadvantages are:

- restricted communications up to a few metres in range
- chip select lines determine the maximum number of devices connected to the bus
- each extra target device adds four new wires to the circuit board layout
- needs additional circuitry or software to detect and repair faults since it lacks a built-in error detection mechanism.

The current state-of-the-art in low-power I2C and SPI communication focuses on reducing power consumption while preserving or enhancing communication speed and reliability. This is accomplished by combining hardware and software optimisation with low-power FPGA (Field-Programmable Gate Array) components.

Authors in [4] accurately simulated a reduced power model for low-power systems using peripherals. The method can be applied to various instruction sets and peripherals. Power supply fluctuations and number-rounding inaccuracies generate energy estimating issues, according to the research. This work introduces a new power measuring method using high-precision power measurement hardware and low-level software control to accurately evaluate power consumption.

In [5], the authors proposed a low-power synchronous data line protocol design that attempts to reduce energy consumption by utilising only two data lines, which is beneficial for PCB design. The proposed system consumed 12 mW, while the standard 3-Wire SPI design consumed 19 mW of power, which is a significant improvement. It also enables simplex and half-duplex data transfer at the same rate. The design was validated using FPGA as the controller and the chip as the target.

The authors of [6] provide an exhaustive evaluation of current power efficiency techniques and introduce new ones focused on lowering the quantity of data delivered and processed by devices. It aims to improve the power efficiency of the I2C communication protocol by reducing leakage current. The ıTurbo approach formulates optimal pull-up resistors. With increasing SDA and SCL pull-up resistance values, energy usage is significantly reduced as the current decreases.

Gated and non-gated low-power SPI switch interfaces have been compared by researchers in [7]. The authors intend to reduce the SPI switch interface's power consumption while preserving its performance. The paper describes the design and implementation of the FPGA XP2 low-power SPI switch interface. The results demonstrate that the proposed low-power SPI switch interface consumes much less power than existing methods while maintaining high performance.

The authors of [8] compare the power consumption of various serial interfaces, such as Universal asynchronous receiver-transmitter (UART), SPI, I2C, and others, in order to identify the most power-efficient approach. The report compares the results of a comprehensive analysis of the power consumption of several serial interfaces under various scenarios. The results suggest that the I2C interface is the most power-efficient alternative for low-data-rate applications, but the SPI interface is better suited for high-data-rate applications. Also, the authors present various optimisation strategies to improve the power efficiency of serial interfaces.

The paper [9] describes the implementation of both the I2C and SPI protocols on an FPGA. The authors evaluate these two FPGA-implemented protocols' performance, energy consumption, and design complexity. While implementing these protocols on an FPGA, the authors also discuss the design tradeoffs and difficulties involved. The results indicate that the SPI protocol has superior performance and lower power consumption per Mbps than the I2C protocol, although the I2C protocol has a simpler design.

The framework implementation in [10] involves the implementation of the Flex-SPI protocol into a software architecture, which provides a basis for communication between various devices. The Flex-SPI protocol is characterised by a comprehensive study of its performance, reliability, and power consumption. The energy consumption analysis of the Flex-SPI protocol involves measuring the power utilised by the communication link in various data transfer scenarios. This information is then compared to the I2C standard's energy consumption.

In [11], the energy consumption of SPI and I2C with various pull-up resistors was compared. They determined that a pull-up resistor higher than 47 kΩ would require a reduction in parasitic bus capacitance to function properly. SPI transfer was finished approximately 200 milliseconds faster than predicted, but I2C took 50 ms longer (with accounted NACK and CONTINUE commands on the bus). The influence of bus capacitance on the communication protocol causes the timing disparity.

In [12], the authors suggest a design for a digital circuit that can convert signals between the SPI and I2C communication protocols, presenting a novel approach to SPI and I2C. The control module employs state machines to regulate the data transmission between the two interfaces and guarantees that the communication adheres to the protocol specifications of each interface. The simulation results demonstrate that the conversion circuit can convert signals between the two protocols while preserving the data's integrity.

This article compares the power consumption of I2C and SPI communications on the same chip (eg. register write, register read, ...). This operation can be carried out using a microcontroller connected to a chip supporting the I2C and SPI communication protocols.

The rest of the paper is organised as follows. Section II defines the experimental setup used to perform measurements. Results and discussions are presented in Section III, while Section IV concludes the paper and gives future remarks.

II. SETUP

Two hardware configurations are compared:

1) microcontroller connected to I2C devices (Fig. 1)
2) microcontroller connected to SPI devices (Fig. 2).

I2C contains two pull-up resistors, the resistance of which influences the energy consumption results. Energy consumption of pull-up resistors decreases with decreasing resistance values adjusted to the level of reliable signal integrity for the I2C connection with the wire length chosen.

I2C bus impedance and signal rise time decrease with lower pull-up resistor values but it might affect signal quality causing voltage drops and bus noise. Lowering the pull-up resistor increases data transmission rate, but it must be done carefully to avoid damaging the bus and devices. Reading bus devices datasheets and specs makes it easier to choose a pull-up resistors value.

Only one device is connected to the SPI or I2C bus when measurements are performed.

Table I gives a list of structural characteristics in order to properly specify the setup. It is important to note that 47 kΩ pull-up resistors are the most energy-efficient option for I2C connections as per [11]. Well estimated pull-up resistors may be a deciding factor in terms of power usage throughout the hardware design phase.

The chip select (CS) pin of the controller is set to push-pull GPIO mode. Typically, pull-up resistors on chip select lines are advised for systems where chip select pins of the controller may enter an undefined state. Added pull-up resistor would increase power consumption, which is undesirable for this research.

TABLE I: The characteristics of the measuring setup.

Characheristic	*Value*
MCU model	STM32L010K4T6
I2C/SPI device	LPS22HBTR
Device type	MEMS pressure sensor
MCU operating frequency [KHz]	1048
Operating voltage [V]	3.3
I2C pull − up resistor values [kΩ]	47
Used wire lenght [cm]	15
SPI clock polarity ($CPOL$)	High
SPI clock phase ($CPHA$)	2 Edge
Power supply	GW Instek PSB-1400L
Measuring multimeter	Keysight 34465A

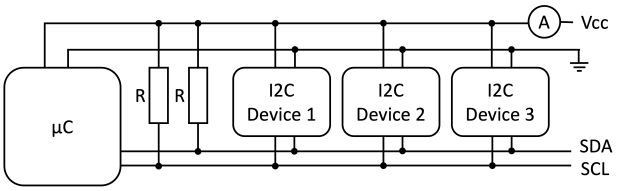


Fig. 1: I2C configuration diagram.

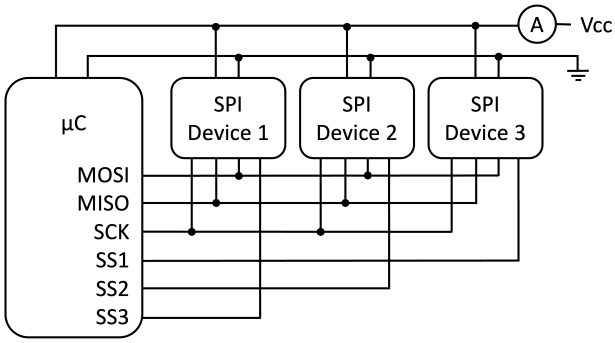


Fig. 2: SPI configuration diagram.

III. EXPERIMENTAL RESULTS

The power consumption of I2C and SPI interfaces depends on several factors, including clock frequency, voltage level, and implementation details. I2C often consumes less power than SPI due to its slower clock speed and more efficient data transport protocols. However, the actual power consumption of I2C versus SPI can vary considerably based on the application and other variables, such as power control tools, which can have a substantial impact. However, it is recommended to consult the utilised devices' datasheets to ascertain each interface's actual power consumption.

Current measurements were performed with Keysight 34465A 6 ½ digit multimeter (listed in Tab. I) that provides high level of accuracy, speed and resolution. The microcontroller model and clock frequency are shown in Tab. I, as well as the sensor used as the target device.

According to the measured values in Tab. II and Tab. III, the distinction between open drain output and push-pull output is apparent. When the I2C bus idles, both SDA and SCL signals are in the logic high. In that period, the lowest energy consumption is achieved. This is due to the I2C bus consuming energy on every logic low signal, which happens when sending or receiving data.

Because of different data transfer approaches, I2C uses KHz while SPI uses KBit/s as a measure of transfer speed. I2C utilizes synchronous communication, in which a controller handles the timing of data delivery. With I2C, the controller transmits a clock signal to the target device(s), which transmit or receive data. According to the standard, since the frequency of the clock signal directly correlates to the transfer speed, the KHz measure is used. SPI is also a synchronous serial communication protocol,

TABLE II: Current consumption during SPI communication with the device's registers.

SPI		
Communication speed [KBit/s]	*Read current* [μA]	*Write current* [μA]
idle	458,92	458,92
4,093	269,57	606,45
8,187	270,07	686,12
16,375	271,44	867,54
32,75	273,59	1.124,30
65,5	274,02	1.231,91
131	275,26	1.001,59

TABLE III: Current consumption during I2C communication with the device's registers.

I2C		
Communication speed [KHz]	*Read current* [μA]	*Write current* [μA]
idle	270,05	270,05
4	523,49	539,17
8	505,96	519,05
16	475,77	486,67
32	441,18	456,16
65	403,80	417,72
131	376,97	404,45

but it is designed to operate at much higher speed rates compared to I2C. The duplex property of the SPI protocol enables data to be sent simultaneously. Read and write operations do not require equal clock signal cycles. Hence the transfer rate is measured in KBit/s.

According to SPI communication measurements, results have shown that reading data consumes less energy than writing data. For a read operation, the controller sends a data request, and the target device sends back the requested data by shifting it out on the MISO (Master In Slave Out) line. For the write operation, the controller sends both the address of the register and the data which will be written into the register. Since the target device has to write received data into the register, power consumption for the write operation rises. As a result, a read operation typically consumes less power than a write operation.

IV. CONCLUSION

When designing ultra-low power embedded systems, it is recommended that all on-board peripherals use minimum energy. One of the aspects that affect energy consumption is the communication protocol used. This paper analyses the energy consumption of SPI and I2C communication protocols with speeds in the range of 100 KHz.

Based on the results, a conclusion can be made for three scenarios which can occur when utilizing SPI or I2C.

- Communication occurs sometimes, and the communication line spends in the idle mode most of the time. In this case I2C is more suitable when it comes to energy consumption.
- SPI is used to continuously retrieve data (i.e. measurements from sensor) without frequent register writes. In this case SPI is more suitable when it comes to energy consumption.
- Communication is balanced, a number of register reads and register writes are equally represented. In this case I2C is more suitable when it comes to energy consumption.

In applications with multiple I2C devices, results may vary due to the calculated resistance of pull-up resistors. It is also worth noting that number of read and write operations can also impact energy consumption when using SPI communication protocol, which isn't the case for I2C communication protocol.

Due to SPI not having a formal standard, and possible variations and options of SPI configuration, the results of this study are not uniform across the board. More experimental cases should be evaluated for modes which are not assessed in this research. Options such as SPI clock polarity (CPOL), clock phase (CPHA), and other types of SPI configurations, such as daisy chain or expander configuration should be considered.

Future work will include research for developing new ultra-low power communication protocol and developing methods for improving energy consumption in embedded systems.

ACKNOWLEDGMENT

This work was funded by the European Union through the European Regional Development Fund, under project "Smart sticker for measuring and monitoring storage and transportation conditions of products" KK.01.1.1.04.0116.

REFERENCES

[1] F. Leens, "An introduction to i2c and spi protocols," *IEEE Instrumentation & Measurement Magazine*, vol. 12, no. 1, pp. 8–13, 2009.

[2] *I2C-bus specification Rev 7*, NXP Semiconductors.

[3] *SPI Block Guide*, Motorola, Inc. [Online]. Available: https://www.nxp.com/files-static/microcontrollers/doc/ref_manual/S12SPIV4.pdf

[4] G. Berthou, K. Marquet, T. Risset, and G. Salagnac, "Accurate power consumption evaluation for peripherals in ultra low-power embedded systems," in *2020 Global Internet of Things Summit (GIoTS)*, 2020, pp. 1–6.

[5] S. L. Chen, T.-K. Chi, M.-C. Tuan, C.-A. Chen, L.-H. Wang, W.-Y. Chiang, M.-Y. Lin, and P. A. Abu, "A novel low-power synchronous preamble data line chip design for oscillator control interface," *Electronics*, vol. 9, p. 1509, 09 2020.

[6] Y. Song, S. Li, C. Zhang, S. Li, and L. Lu, "Rethinking power efficiency for next-generation processor-free sensing devices," *Sensors*, vol. 22, no. 8, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/8/3074

[7] S. Kaur and N. Sharma, "Designing and implementation of low power spi-switch interface on fpga," *International Journal of Engineering Research and*, vol. V4, 05 2015.

[8] K. Mikhaylov and J. Tervonen, "Evaluation of power efficiency for digital serial interfaces of microcontrollers," in *2012 5th International Conference on New Technologies, Mobility and Security (NTMS)*, 2012, pp. 1–5.

[9] A. Oudjida, M. L. Berrandjia, R. Tiar, L. Ahmed, and K. Tahraoui, "Fpga implementation of i2c & spi protocols: A comparative study," 12 2009.

[10] P. Visconti, G. Giannotta, R. Brama, P. Primiceri, R. De fazio, and A. Malvasi, "Framework implementation, firmware development and characterization of flex-spi communication protocol: Energy consumption analysis and comparison with i2c standard," *International Journal on Smart Sensing and Intelligent Systems*, vol. Vol.10, pp. pp. 754 – 792, 12 2017.

[11] T. Solheim and M. Grannæs, "A comparison of serial interfaces on energy critical systems," in *2015 Nordic Circuits and Systems Conference (NORCAS): NORCHIP & International Symposium on System-on-Chip (SoC)*, 2015, pp. 1–4.

[12] D. Trivedi, A. Khade, K. Jain, and R. Jadhav, "Spi to i2c protocol conversion using verilog," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018, pp. 1–4.

Genetic Algorithm Based Optimization of Circular Planar Coil Geometry With Homogeneous Magnetic Field Distribution

Domagoj Bilandžija*, Davor Vinko* and Luka Filipović*
* Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Department of Communications, Osijek, Republic of Croatia
dbilandzija@ferit.hr

***Abstract*—A prominent problem in planar resonant wireless power transfer (WPT) systems is instability of efficiency and load power due to position misalignment of the receiver (RX) coil relative to postion of the transmitter (TX) coil. In order to keep stable efficiency and load power even in such misaligned condition it is necessary to design transmitter coil which generates homogeneous magnetic field in charging plane. In this paper, a geometry of planar transmitter coil which consist of multiple concentric circles is optimized by utilizing genetic algorithm. Optimization goal is achieving as large as possible homogeneous region, i.e. portion of charging plane which is characterized by homogeneous magnetic field. Magnetic field evaluations are done by FEMM software. Proposed coil achieves continuous homogeneous region at 44.4% surface of charging plane. Optimized coil and its homogenous region is compared to the state-of-the-art.**

***Keywords*—coil, genetic algorithm, optimization, resonant wireless power transfer**

I. INTRODUCTION

In recent decades, interest in wireless power transfer (WPT) is growing rapidly. This is due to an increase in usage of rechargeable battery powered devices: consumer electronics, [1], electric vehicles (EVs), [2], and implanted medical devices, [3]. In comparison with conventional wired charging, WPT reduces electrical shock hazard as a result of not involving wired connection between source of energy and battery. Besides, using WPT eliminates development of mechanical damage of power connectors in conventional and manual pluging on a daily basis. Instead of doing surgical operation in order to replace batteries of implanted medical devices, more convenient WPT is applied to reduce health risks.

Depending upon time variant physical fields which are the mediators for energy transfer, there is capacitive (electric fields) and inductive (magnetic fields) WPT. Working principle of a two coil inductive WPT system is that changing magnetic field of transmitter (TX) coil induces voltage at receiver (RX) coil. Efficient power transfer is limited to about 1 cm range, [4]. Groundbreaking paper, [5], introduced resonant inductive wireless power transfer (RIWPT) system. Simplified model of such system consists of two LC oscillators which operate at the same resonant frequency. General properties of efficient RIWPT system are: high quality factors of TX and RX coils and low value of magnetic coupling coefficient k. In RIWPT system, k is typically less than 0.15, [4], [6].

Generally, coupling coefficient value is dependent on the distance, d, between TX coil and RX coil. Small coupling coefficient value in RIWPT system implies larger transfer distance and eventually, a transfer of energy towards multiple RX coils that are significantly smaller compared to TX coil. Although efficient power transfer is feasible even for small coupling coefficient value (larger transfer distance), conventional RIWPT system (TX coil and RX coil are situated in parallel planes) requires precise or coaxial alignment of RX coil position relative to TX coil position. Situation in which centers of coupled TX coil and RX coil are horizontally displaced is referred to as lateral misalignment, [7]. However, accurate positioning of coils is not so easy task in real life scenarios. For example, precise alignment of coil integrated in EV, coaxial alignment of RX coil of implanted medical device under the tissue and so on.

Since lateral misalignment causes nonstable efficiency and load power in conventional RIWPT, authors propose two approaches to keep them more stable. First, tracking resonant frequency shift and tuning system to new resonant frequency which occurs in phenomenon called frequency splitting, [4], [8], [9]. Adapting to new resonant frequency is equivalent to procedure of impedance matching. Second, generating homogeneous magnetic field distribution is proposed to prevent frequency splitting and impedance mismatching in lateral misalignment condition. In other words, various TX coil designs are proposed to generate homogeneous magnetic field distribution at the charging plane of RIWPT.

Proposed 2D TX coil designs would be systemized as follows: hybrid, spiral and specific structure. Hybrid coils are composed of outer ring coil and inner spiral coil connected in series, [10], [11]. Spiral coils are charcterized by specific and nonuniform spacing between adjacent turns, [12]–[17]. Coils with specific design are: coil formed of concentric circles, [18], Taichi coil, [19], and coil with intertwined turns, [20]. Proposed 3D coils are the ones that occupy certain volume which keeps them distinct from proposed 2D coils. Hybrid, [21], funnel type, [22] and rectangular coils, [23]–[25], are proposed as 3D coil

solutions.

Following section analyzes nonhomogeneity of magnetic fields in charging plane above conventional TX coils. Third section deals with the optimization of a coil made up of concentric circles. Genetic algorithm was used as an optimization tool. At the end, the conclusion is derived.

II. MAGNETIC FIELD DISTRIBUTION ANALYSIS

RX coils in planar RIWPT are placed within charging plane which is lightly shaded in Fig. 1. Surface of charging plane corresponds to surface enclosed by TX coil outline. Hence, charging plane radius, r, is equal to TX coil radius, r_{out}. Charging plane (RX coils) and TX coil are vertically separated by distance d. Magnetic field intensity distribution at charging plane is a matter of interest.

According to well known Faraday-Lenz law, the voltage induced in a coil is proportional to the time varying magnetic flux, ϕ. Equation which describes magnetic flux which passes through some area A is following:

$$\phi = \vec{B} \cdot \vec{A} = \mu \cdot |\vec{H}| \cdot |\vec{A}| \cdot \cos(\alpha) = \mu \cdot |\vec{H_y}| \cdot |\vec{A}|, \quad (1)$$

where, $\vec{B}$ is the magnetic flux density, $\vec{A}$ is the vector normal to an area A. Magnitude of vector $\vec{A}$ is equal to surface of observed area A through which magnetic flux is passing. The magnetic field intensity, $\vec{H}$, and $\vec{B}$ are mutually proportional and the constant of proportionality is the permeability, μ. Finally, α is the angle between $\vec{A}$ and $\vec{H}$. From (1) and according to the Fig. 1, it is obvious that the only component of magnetic field intensity that is perpendicular to the charging plane, $\vec{H_y}$, contributes to magnetic flux. Thus, solely vertical component of the magnetic field intensity is considered in the rest of this paper.

Conventional planar ring coil and concentric circles coil are shown in the Fig. 2. Both coil designs are created in FEMM (Finite Element Method Magnetics) software and simulations are carried out to determine magnetic

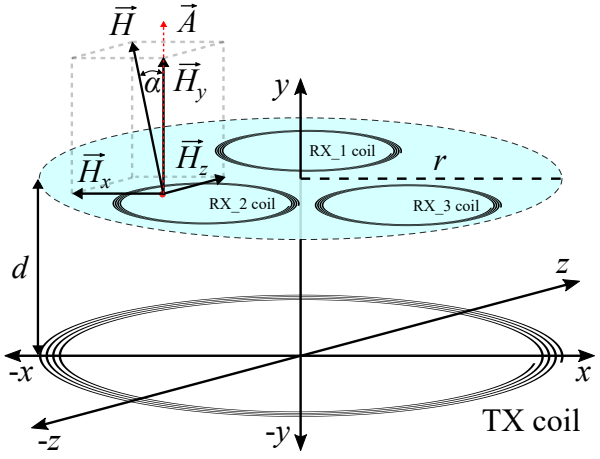
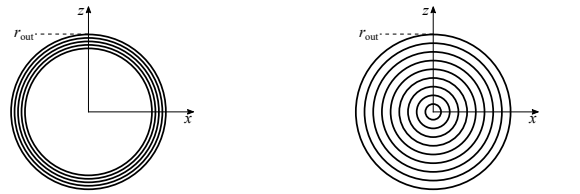


Fig. 1: Sketch of planar RIWPT



(a) Ring coil (b) Concentric circles coil

Fig. 2: Conventional coils

TABLE I: Simulated coils parameters

Parameter	**Description**	**Ring coil**	**Concentric circles coil**
N, [mm]	number of turns	5	9
r_{out}, [mm]	outer radius of coil	45	45
th, [mm]	trace height of turn	0.035	0.035
tw, [mm]	trace width of turn	0.5	1
s, [mm]	adjacent turns spacing	0.5	4

field intensity in charging plane that is 10 mm above (d = 10 mm). Turns of coils are made of copper and surrounding environment is air ($\mu = \mu_0 \cdot \mu_{air} \approx \mu_0$). Parameters of both simulated coils are shown in the Table I. In FEMM, both coils are defined as axisymmetric problem since doing a 360° rotation of turns cross sections around axis forms complete 3D coil models. For the reason that problem is axisymmetric, examining magnetic field intensity in charging plane is simplified to charging line that is eventually a radius of circular charging plane, r, depicted as dashed line in the Fig. 1.

Normalized magnetic field intensity in charging line above ring coil and concentric circles coil is shown in Fig. 3 and Fig. 4, respectively. Magnetic field intensity distribution above ring coil is concave while convex distribution of magnetic field intensity is generated above concentric circles coil. Change of RX coil(s) position(s) in such

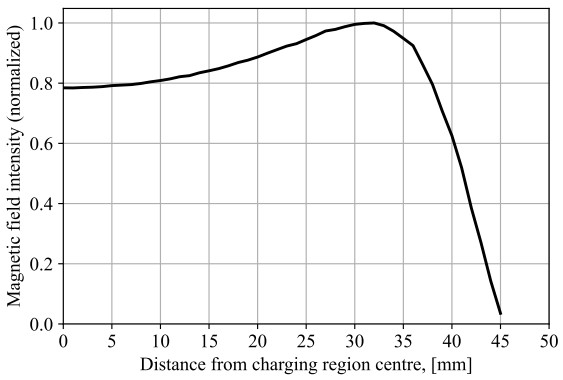


Fig. 3: Magnetic field intensity distribution (vertical component) across charging line of ring shaped coil

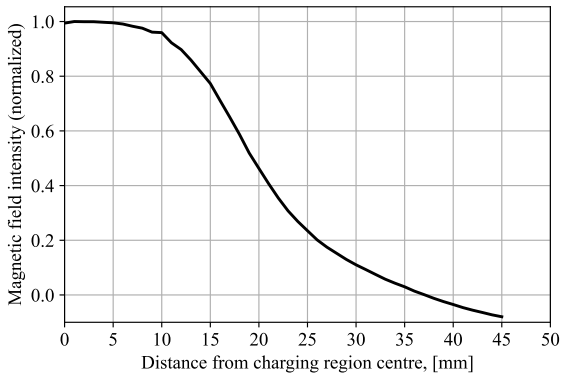


Fig. 4: Magnetic field intensity distribution (vertical component) across charging line of concetric circles sheped coil

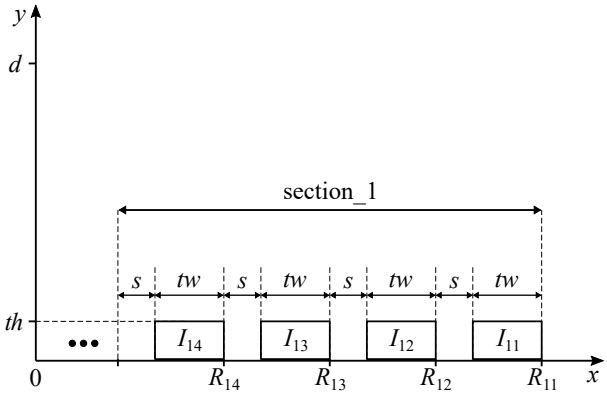


Fig. 5: Simplified representation of optimized TX coil

nonhomogeneous magnetic field intensity distributions at charging plane causes change in magnetic coupling factor, k, which should be stable to keep efficiency and load power at desired level. Additionally, when fundamental resonant frequency is applied, maximum power transfer (along with 50% efficiency) is achieved at specific value of magnetic coupling factor which is called as critical coupling factor, k_c, [4], [8].

Flattening of magnetic field distribution by superposition of magnetic fields of combined outer ring coil (concave distribution) and inner concentric circles coil (convex distribution) is an innovative idea that came up in previous scientific papers, [10], [21], [22]. However, majority of proposed solutions are based on approximate position adjustment of coil turns and observing magnetic field distribution rather than executing mathematical optimization. In this paper, GA is used as a tool to find optimal design of concentric circles coil with predefined radius of outermost coil circle, r_{out}.

III. GENETIC ALGORITHM BASED OPTIMIZATION OF TX COIL

TX coil axisymmetric model is shown in the Fig. 5. Coil consists of concentric circular turns that are separated in nine equal sections. Each section consists of four turns. First section (section_1) is outermost and radii of corresponding turns are denoted as R_{11}, R_{12}, R_{13}, R_{14}, (see Fig. 5). First number in the subscript is the ordinal number of section and second number in the subscript is the ordinal number of turn in corresponding section. Following eight sections are not shown in Fig. 5, but they use the same logic to distinct remaining turns of coil. According to Fig. 5, between each turn of the coil is space denoted by $s = 0.5$ mm. Coil is intended to be made in PCB technology, so trace (turn) height, *th*, is set to be 0.035 mm and trace (turn) width, *tw*, is set to be 0.75 mm. Also, currents passing through turns of section_1 are labeled as I_{11}, I_{12}, I_{13}, I_{14}. Labeling currents of each section adheres to the same rule as labeling radii of turns of coil sections.

According to previously described model, coil is formed out of 36 turns and adjacent turns are equally spaced by 0.5 mm. However, during optimization, current of each turn could be either 0 or 1 (binary digit). If the turn current equals to zero, then that turn is actually omitted. On the other hand, if the turn current is equal to one, then corresponding turn is the part of the optimized coil. Thus, all turns that conduct current equal to one belong to coil and all turns which current is equal to zero are not taken into account and left floating.

Optimization is performed in Python-FEMM cosimulation. FEMM is used to calculate magnetic field intensity in charging line. FEMM do not have its own optimization module so it was required to work in cosimulation with other software which is compatible with FEMM. The module pyFEMM is a Python package that allows for the operation of FEMM via a library of Python functions, [26]. Thus, pyFEMM module provides creating model of coil externally from the Python thus allowing to define variables which describe coil and which are optimized (decision variables). Python library pygmo was used for optimization.

There are nearly 20 heuristic global optimization algorithms that are currently available to use in pygmo library. Genetic algorithm in pygmo library is called Simple Genetic Algorithm (SGA), and is specified as stochastic, single-objective, unconstrained and integer programming, [27]. SGA is suitable for solving optimization problem with regard to suggested coil model because there is one (single) objective of generating as large as possible homogeneous region within the charging plane. The mathematical notation of the optimization problem is as follows:

$$f(\mathbf{x}) \rightarrow min., \;\; \mathbf{x} \in \mathbb{Z} \tag{2}$$

subject to constraints:

$$\begin{aligned} c_e(\mathbf{x}) &= 0, \\ c_i(\mathbf{x}) &\leq 0. \end{aligned} \tag{3}$$

$\mathbf{x}$ is a vector of decision variables or chromosome that involves n integers. Function that is to be minimized, (2), should be specified to express objective of optimization.

In this optimization problem, equality and inequality constraints are not present, (3). For that reason, optimization problem is said to be unconstrained. Lower and upper bound of $\mathbf{x}$ are as follows:

$$0 \leq \mathbf{x} \leq 1, \tag{4}$$

that is in accordance with allowed binary value of turn current. Fitness function, (5), is used to evaluate how good the solution represented by a chromosome is, [28].

$$f(\mathbf{x}) = -(r_{hom.}/r), \tag{5}$$

where $r_{\text{hom.}}$ is a fraction of radius r that is characterized by homogeneous magnetic field intensity. Magnetic field intensity is calculated at 45 points that are placed on charging line. Distance between two adjacent points is 1 mm. Magnetic field intensity at certain point on charging line is homogeneous if the following inequallity is satisfied:

$$|\vec{H}_y| \geq 0.95 \cdot |\vec{H}_{ymax}|. \tag{6}$$

Before executing optimization, it is necessary to set SGA parameters. SGA is an algorithm from larger group of related algorithms called evolutionary algorithms (EAs). SGA imitate biological evolution in highly simplified way. Main idea of this iterative process is to keep good solutions through generations along with randomly exploring solution space to avoid stucking in local optimum. To mimic biological evolution, genetic operators used in SGA are: selection, crossover and mutation. Basic SGA parameters applied for coil optimization are listed in Table II.

SGA finds optimal solution as a coil that consists of 12 turns. Nine of turns are concentrated at the edge of coil and each adjacent turns are uniformly separated by *s*. That nine turns includes each turn of first two sections and outermost turn of third section. Remaining turns that form optimal coil are closer to coil center, turn with radius R_{51}, turn with radius R_{82} and turn with radius R_{83}. Therefore, optimal coil design is a specific combination of uniformly and nonuniformly separated turns.

Magnetic field intensity distribution in charging line that are 10 mm above optimaly designed coil is shown in the Fig. 6. Continuous homogeneous magnetic field intensity is generated in 66.67% of charging line length.

In comparison to conventional coils and theirs magnetic field intensity distribution, homognenous magnetic field

TABLE II: SGA parameters

SGA parameters	Description	Type	Value
gen	number of generations	integer	50
pop	population size	integer	50
selection	[27]	tournament	3
crossover	[27]	exponential	n/a
cr	crossover probability	real number, cr ∈ [0,1]	0.9
mutation	[27]	polynomial	n/a
m	mutation probability	real number, m ∈ [0,1]	0.07

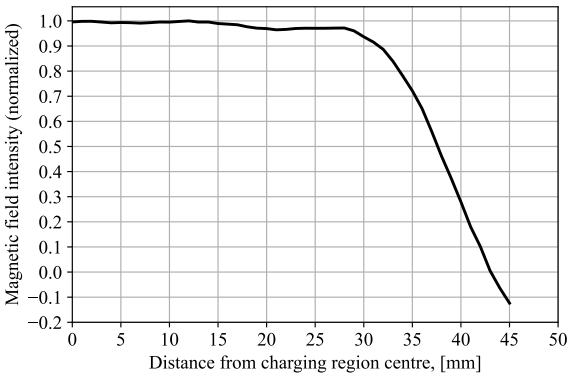


Fig. 6: Magnetic field intensity distribution (vertical component) accros charging line of optimal coil design

TABLE III: Comparison with other 2D TX coil designs

Paper	$\mathbf{H_y/H_{ymax}}$	Transfer distance, [mm]	Homogeneous region, [%]
[12]	≥ 0.8	1	∼36.0
[15]	≥ 0.8	50	∼51.8
[16]	≥ 0.904	100	∼42.3
[17]	≥ 0.7	15	∼64.0
[20]	≥ 0.928	18	∼49.0
This paper	≥ 0.95	10	∼44.4

intenisity distribution is achieved. However, in comparison with other solutions which address the same problem, this solution achieves nearly large homogeneous region, Table III.

Future plans for increasing homogneous region of optimized concentric circles coil include: parallel connection of turns which implies different current distribution than binary (but positive integer) and using both layers of PCB instead of one that is used in this paper (coil becomes 3D).

IV. Conclusion

Homogeneous magnetic field intensity in charging plane ensures stable efficiency and load power in RIWPT. In order to achieve as large as possible homogeneous region, concentric circles TX coil is optimized using SGA of pygmo library along with FEMM. Proposed coil generates homogeneous region that occupies 44.4% surface of charging plane. Such performance is comparable with other proposed coils intended for generating homogeneous region. Future goal is to enlarge homogeneous region by introducing different current distribution and utilizing PCB in terms of printing optimal coil design on both PCB sides.

References

[1] E. Waffenschmidt and T. Staring, "Limitation of inductive power transfer for consumer applications," in *2009 13th European Conference on Power Electronics and Applications*, 2009, pp. 1–10.

[2] A. Mahesh, B. Chokkalingam, and L. Mihet-Popa, "Inductive wireless power transfer charging for electric vehicles-a review," *IEEE Access*, vol. 9, pp. 137 667–137 713, 2021.

[3] M. Kiani, "Wireless power transfer and management for medical applications: Wireless power," *IEEE Solid-State Circuits Magazine*, vol. 14, no. 3, pp. 41–52, 2022.

[4] A. P. Sample, D. T. Meyer, and J. R. Smith, "Analysis, experimental results, and range adaptation of magnetically coupled resonators for wireless power transfer," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 2, pp. 544–554, 2011.
[5] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, Jul. 2007. [Online]. Available: https://doi.org/10.1126/science.1143254
[6] T. Imura and Y. Hori, "Maximizing air gap and efficiency of magnetic resonant coupling for wireless power transfer using equivalent circuit and Neumann formula," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4746–4752, 2011.
[7] K. Fotopoulou and B. W. Flynn, "Wireless power transfer in loosely coupled links: Coil misalignment model," *IEEE Transactions on Magnetics*, vol. 47, no. 2, pp. 416–430, 2011.
[8] R. Huang, B. Zhang, D. Qiu, and Y. Zhang, "Frequency splitting phenomena of magnetic resonant coupling wireless power transfer," *IEEE Transactions on Magnetics*, vol. 50, no. 11, pp. 1–4, 2014.
[9] X. Liu, X. Yuan, C. Xia, and X. Wu, "Analysis and utilization of the frequency splitting phenomenon in wireless power transfer systems," *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 3840–3851, 2021.
[10] J. Kim, H.-C. Son, and Y.-J. Park, "Multi-loop coil supporting uniform mutual inductances for free-positioning WPT," *Electron. Lett.*, vol. 49, no. 6, pp. 417–419, Mar. 2013. [Online]. Available: https://doi.org/10.1049/el.2013.0135
[11] T. H. Kim, G. H. Yun, W. Y. Lee, and J. G. Yook, "Asymmetric coil structures for highly efficient wireless power transfer systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 7, pp. 3443–3451, 2018.
[12] J. J. Casanova, Z. N. Low, J. Lin, and R. Tseng, "Transmitting coil achieving uniform magnetic field distribution for planar wireless power transfer system," in *2009 IEEE Radio and Wireless Symposium*, 2009, pp. 530–533.
[13] T. D. Yeo, D. H. Kim, S. C. Chae, S. T. Khang, and J. W. Yu, "Design of free-positioning wireless power charging system for AAA rechargeable battery," in *2016 46th European Microwave Conference (EuMC)*, 2016, pp. 759–762.
[14] L. Shen, W. Tang, H. Xiang, and W. Zhuang, "Uniform magnetic field of the planar coil with new winding structure for displacement-insensitive WPT," in *2014 IEEE International Conference on Communiction Problem-solving*, 2014, pp. 394–396.
[15] S. Wang, Z. Hu, C. Rong, C. Lu, J. Chen, and M. Liu, "Planar multiple-antiparallel square transmitter for position-insensitive wireless power transfer," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 2, pp. 188–192, 2018.
[16] Y. Zhang, L. Wang, Y. Guo, and Y. Zhang, "Optimisation of planar rectangular coil achieving uniform magnetic field distribution for EV wireless charging based on genetic algorithm," *IET Power Electronics*, vol. 12, no. 10, pp. 2706–2712, Aug. 2019. [Online]. Available: https://doi.org/10.1049/iet-pel.2018.6202
[17] Q. Xu, Q. Hu, H. Wang, Z. H. Mao, and M. Sun, "Optimal design of planar spiral coil for uniform magnetic field to wirelessly power position-free targets," *IEEE Transactions on Magnetics*, vol. 57, no. 2, pp. 1–9, 2021.
[18] Y. Yang, X. Kuang, P. Yang, Y. Jing, X. Su, and Y. Cheng, "Parallel connected transmitting coil for achieving uniform magnetic field distribution in WPT," in *2015 IEEE 16th International Conference on Communication Technology (ICCT)*, 2015, pp. 529–532.
[19] Y. Li, J. Zhao, Q. Yang, L. Liu, J. Ma, and X. Zhang, "A novel coil with high misalignment tolerance for wireless power transfer," *IEEE Transactions on Magnetics*, vol. 55, no. 6, pp. 1–4, 2019.
[20] J. Li, J. Sun, R. Qin, and D. Costinett, "Transmitter coil design for multi-load wireless power transfer systems," in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2020, pp. 1032–1038.
[21] X. Liu and S. Y. Hui, "Optimal design of a hybrid winding structure for planar contactless battery charging platform," *IEEE Transactions on Power Electronics*, vol. 23, no. 1, pp. 455–463, 2008.
[22] J. Huang, T. Hong, M. Bojarski, F. de León, and D. Czarkowski, "Design algorithm of a uniform magnetic field transmitter intended for the wireless charging of electric vehicles," in *2014 IEEE International Electric Vehicle Conference (IEVC)*, 2014, pp. 1–6.
[23] W.-S. Lee, H. Lim Lee, K.-S. Oh, and J.-W. Yu, "Uniform magnetic field distribution of a spatially structured resonant coil for wireless power transfer," *Appl. Phys. Lett.*, vol. 100, no. 21, p. 214105, May 2012. [Online]. Available: https://doi.org/10.1063/1.4719585
[24] D. Vinko, D. Bilandžija, and V. Mandrić Radivojević, "Optimization of a two-layer 3D coil structure with uniform magnetic field," *Wireless Power Transfer*, vol. 2021, p. 6303628, 2021. [Online]. Available: https://doi.org/10.1155/2021/6303628
[25] D. Bilandžija, D. Vinko, and M. Barukčić, "Genetic-algorithm-based optimization of a 3d transmitting coil design with a homogeneous magnetic field distribution in a WPT system," *Energies*, vol. 15, no. 4, 2022.
[26] D. Meeker, "Finite Element Method Magnetics: pyFEMM-User's Manual," Mar. 2018.
[27] *pygmo 2.19.0 documentation*.
[28] A. P. Engelbrecht, *Computational Intelligence An Introduction*, 2nd ed. John Wiley and Sons, Ltd, 2007.

Analysis of the Effect of Differential-Mode and Common-Mode Impedance Matching on the Common-Mode Rejection Ratio of a Differential Electro-Optical Voltage Probe

Hrvoje Štimac*, Adrijan Barić*
*University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia
Tel: +385 (0)1 6129547, Fax: +385 (0)1 6129653, Email: hrvoje.stimac@fer.hr

***Abstract*—A differential electro-optical voltage probe is characterized. The two main parts of the probe circuit are the attenuator circuit and the laser diode. The probe circuit is characterized by measuring S-parameters. The probe circuit has a stable common-mode rejection ratio (CMRR) up to 4 GHz, with a quick drop at higher frequencies. The probe circuit performance is simulated by combining the individual attenuator and laser characteristics connected in cascade. The differential-mode and common-mode signal propagation is analyzed in terms of voltage and impedance matching. The impedance matching between the attenuator and the laser, in combination with their individual characteristics, determines the CMRR frequency characteristic of the probe circuit. By suppressing the common-mode signal propagation, relative to the differential-mode signal, an electro-optical system with a CMRR higher than the nominal CMRR of the laser diode can be designed.**

***Index Terms*—CMRR, common-mode, differential-mode, electro-optical probe, impedance matching.**

I. Introduction

Characteristics of a broadband differential electro-optical voltage probe are measured [1]. The probe circuit consists of two main parts, the attenuator circuit and the laser diode (Fig. 1). The attenuator is used to lower the input signal level in order not to overdrive the laser. The laser diode converts the measured radio frequency (RF) signal into an optical signal, which is transmitted to the photodetector. The laser is biased using a dedicated bias circuit. The probe circuit design is discussed in [2].

The characteristics of the differential electro-optical voltage probe are measured. The differential-mode component of the input signal presents the wanted signal, while the common-mode component presents the unwanted signal. The probe has a stable common-mode rejection ratio (CMRR) up to 4 GHz, which drops quickly at higher frequencies. The voltage attenuation ratio of the probe circuit is 125:1 (42 dB). The maximum input differential voltage amplitude is 25 V. The nominal differential-mode input impedance is 1092 Ω, and the nominal common-mode input impedance is 275 Ω.

Research typically focuses on overall external probe characteristics, with a desired CMRR of 30–40 dB up to 5 GHz [3], [4]. In contrast, the motivation for this paper is to understand the internal signal propagation through the probe circuit building blocks and its impact on the CMRR characteristics.

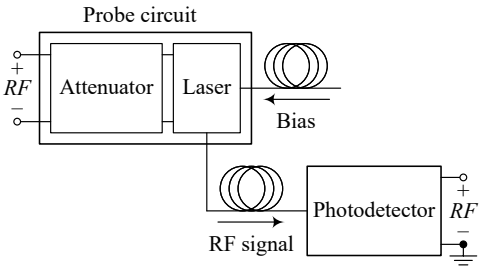


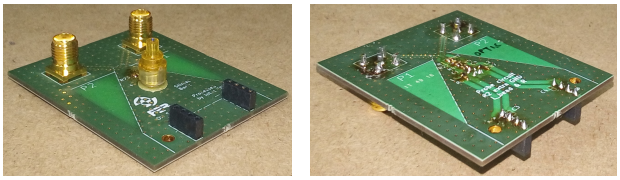
Fig. 1. Differential electro-optical voltage probe schematic.

In this paper, the differential-mode and common-mode signal propagation through the probe circuit is analyzed. The two main parts of the probe circuit, the attenuator and the laser, are analyzed separately. The performance of the probe circuit is simulated by combining the characteristics of these two building blocks. Based on the differential-mode and common-mode voltages and impedance matching between the attenuator and the laser, the CMRR characteristics of the probe are evaluated.

The probe characterization and simulation setup are presented in Section II. The signal propagation through the probe circuit is analyzed in Section III. The effect of the attenuator and laser characteristics on the CMRR of the probe circuit is discussed in Section IV. The conclusion is given in Section V.

II. Characterization and Simulation Setup

Three-port S-parameters of the probe circuit characterization structure shown in Fig. 2 are measured. The mea-



(a) Top side. (b) Bottom side.

Fig. 2. Probe circuit characterization structure.

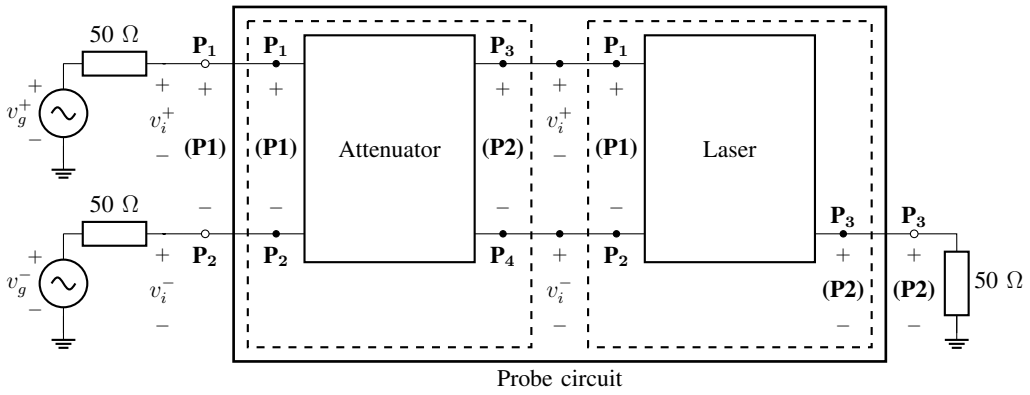


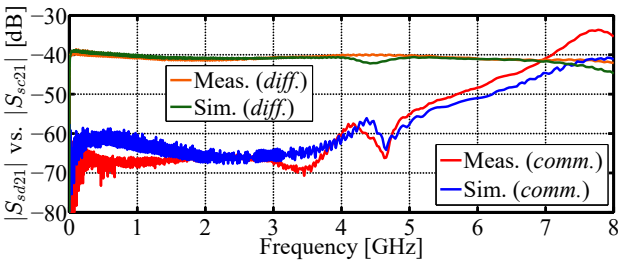
Fig. 3. Probe circuit represented as a cascade connection of the attenuator circuit and the laser.

surements are performed using a two-port vector network analyzer (VNA) [5], in the frequency range from 1 MHz to 8 GHz. The VNA output power is set to 0 dBm. The three-port standard single-ended S-parameters are de-embedded and converted to mixed-mode S-parameters [6], [7]. The CMRR of the probe circuit is calculated [8].

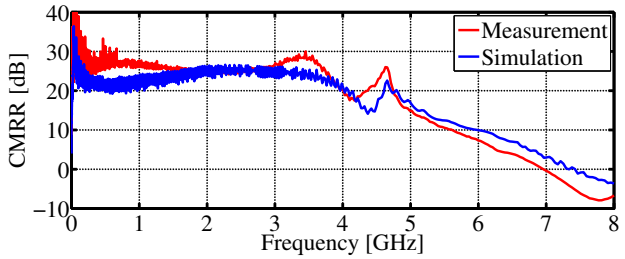
The schematic of the probe circuit characterization structure is shown in Fig. 3. In the probe circuit, the attenuator and the laser are connected in cascade. The goal is to simulate the performance of the probe circuit by combining the characteristics of the attenuator and the laser, that are characterized separately. Electromagnetic (EM) simulations of the attenuator structure, performed using a commercially available EM solver based on the finite element method (FEM) [9], are used in the probe circuit simulation. The mixed-mode S-parameters of the attenuator are calculated from the four-port standard single-ended S-parameter EM simulation results [6], [7]. The biased laser diode is measured separately. The VNA output power is set to –7 dBm, making sure that the laser operates in the linear region. Three-port S-parameter measurement results of the laser are used in the probe circuit simulation.

The probe circuit measurement and simulation results are compared in Fig. 4. The simulation results of the differential-mode transmission coefficient S_{sd21} show good matching to the measurement results. The simulation results match the general trend of the common-mode transmission coefficient S_{sc21} measurements. Taking into account the matching of the differential-mode and the common-mode signal transmission, the simulation results match the general trend of the common-mode rejection ratio measurements. The difference between the simulation and the measurement results varies between 3–5 dB. The simulated CMRR of the probe circuit remains above 20 dB up to 4.1 GHz. The fast drop in the measured CMRR at frequencies above 4.5 GHz is replicated by the probe circuit simulation results.

The differences between the simulation and measurement results are expected, because of the differences between the laser samples used, the components used in the attenuator circuit, the limitations of the attenuator EM simulations, as well as the coupling between the components of the probe circuit structure, which is not present when characterizing the parts individually.



(a) Differential-mode and common-mode signal transmission.



(b) Common-mode rejection ratio.

Fig. 4. Comparison of the measurement and the simulation of probe circuit. The magnitude of the differential-mode transmission coefficient S_{sd21}, the common-mode transmission coefficient S_{sc21}, and the common-mode rejection ratio (CMRR) are compared.

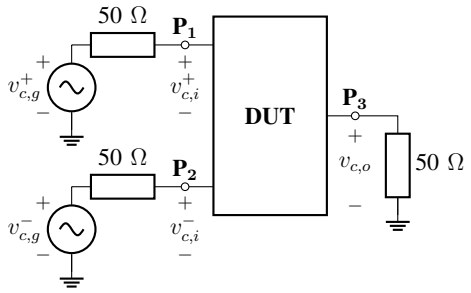
III. Signal Propagation

A. Probe Circuit Input Voltage

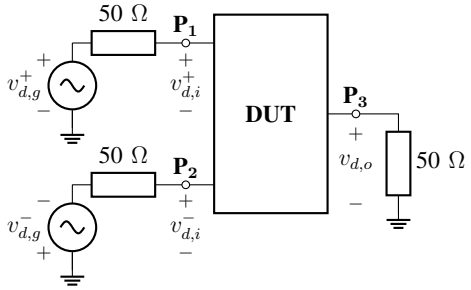
The probe circuit characterization setup is represented using the voltage sources v_g^+ and v_g^- with a 50 Ω internal impedance, and the 50 Ω termination load, shown in Fig. 5. In-phase generators $v_{c,g}^+$ and $v_{c,g}^-$ are used for the common-mode simulation. Counter-phase generators $v_{d,g}^+$ and $v_{d,g}^-$ are used for the differential-mode simulation. The generator voltages are set as follows:

$$v_{c,g}^+ = v_{c,g}^- = v_{d,g}^+ = v_{d,g}^- = 1 \text{ V}. \quad (1)$$

The common-mode voltage at the generator $v_{c,g}$ and the differential-mode voltage at the generator $v_{d,g}$ in the

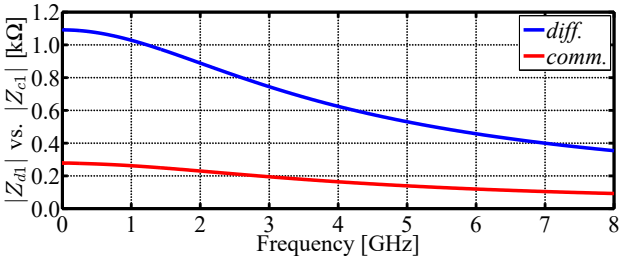


(a) Common-mode characterization setup.

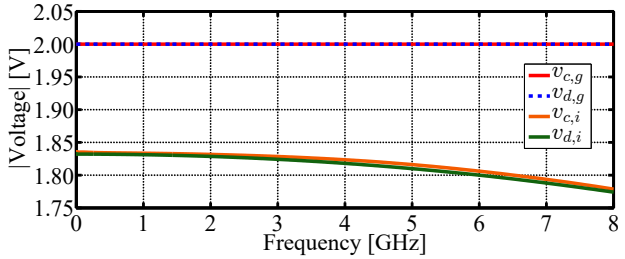


(b) Differential-mode characterization setup.

Fig. 5. Common-mode and differential-mode probe circuit characterization setup schematic.



(a) Differential-mode and common-mode input impedance.



(b) Generator and input voltage.

Fig. 6. Probe circuit simulation results. The magnitude of the differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} are compared (top). The magnitude of the common-mode generator voltage $v_{c,g}$, the differential-mode generator voltage $v_{d,g}$, the common-mode input voltage $v_{c,i}$, and the differential-mode input voltage $v_{d,i}$ are compared (bottom).

probe circuit simulation setup are defined as follows:

$$v_{c,g} = v^+_{c,g} + v^-_{c,g}, \tag{2}$$

$$v_{d,g} = v^+_{d,g} - v^-_{d,g}. \tag{3}$$

The common-mode voltage at the input balanced logical port (P1) of the probe circuit $v_{c,i}$ and the differential-mode voltage at the input of the probe circuit $v_{d,i}$ in the probe circuit simulation setup are defined as follows:

$$v_{c,i} = v^+_{c,i} + v^-_{c,i}, \tag{4}$$

$$v_{d,i} = v^+_{d,i} - v^-_{d,i}. \tag{5}$$

The differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} of the simulated probe circuit at the balanced logical port (P1) are calculated [10]. The differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} of the simulated probe circuit are shown in Fig. 6(a). Both the differential-mode input impedance and the common-mode input impedance decrease with frequency, from the nominal value of 1092 Ω and 275 Ω, respectively. The impedances decrease to approximately 1/3 of the nominal value at 8 GHz, while the relative ratio between the two impedances remains around 4:1 in the entire frequency range. The drop in impedance at higher frequencies is a result of the attenuator circuit component and printed circuit board (PCB) parasitics. The input impedance of the probe circuit matches the input impedance of the attenuator structure very similarly, given that the input impedance of the laser, that is connected to the output of the attenuator, is matched to the 100 Ω differential-mode characteristic impedance.

The common-mode and differential-mode voltages at the generator and at the input of the simulated probe circuit are compared in Fig. 6(b). Given that the voltages of the individual generators are set to 1 V, both the differential-mode voltage at the generator $v_{d,g}$ and the common-mode voltage at the generator $v_{c,g}$ is fixed to 2 V. Both the differential-mode input voltage $v_{d,i}$ and the common-mode input voltage $v_{c,i}$ are close to the generator voltage of 2 V, given that both the differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} of the probe circuit are significantly higher than the generator differential-mode internal impedance of 100 Ω, and the generator common-mode internal impedance of 25 Ω. Both input voltages slightly decrease with frequency given that both the differential-mode and the common-mode input impedance decrease with frequency, while the internal impedance of the generator remains constant. Despite the common-mode input impedance being lower, the common-mode input voltage is slightly higher than the differential-mode input voltage, because the common-mode input impedance is higher relative to the 25 Ω generator common-mode internal impedance, than the differential-mode input impedance is relative to the 100 Ω generator differential-mode internal impedance.

B. Attenuator Signal Transmission

The mixed-mode S-parameter simulation results of the attenuator used in the probe circuit are shown in Fig. 7. Both the differential-to-differential transmission coefficient S_{dd21} and the common-to-common transmission coefficient S_{cc21} are very stable over the entire frequency range. The common-mode transmission coefficient is slightly higher than the differential-mode transmission

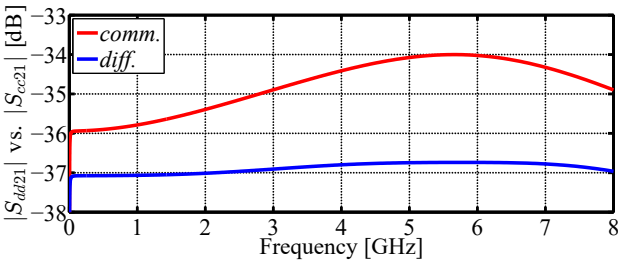
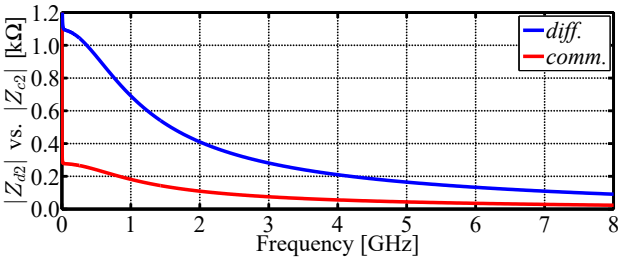


Fig. 7. EM simulations of the attenuator used in the probe circuit simulation. The magnitude of the differential-to-differential transmission coefficient S_{dd21} and the common-to-common transmission coefficient S_{cc21} are compared.

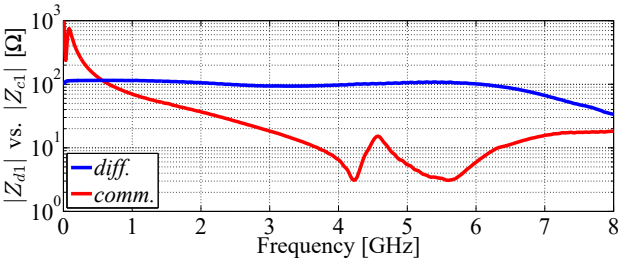
coefficient, meaning that the common-mode signal is less attenuated than the differential-mode signal. The mode conversion in the attenuator circuit EM simulations is limited. The mode conversion transmission coefficients S_{dc21} and S_{cd21} extracted from the simulation are significantly lower than it would be expected in practice. For this reason, the mode conversion in the attenuator circuit is not taken into account in this analysis.

C. Impedance Matching Between the Attenuator and the Laser

The differential-mode output impedance Z_{d2} and the common-mode output impedance Z_{c2} of the attenuator circuit are shown in Fig. 8(a). Both the differential-mode output impedance and the common-mode output impedance decrease with frequency, from the nominal value of 1092 Ω and 275 Ω, respectively. The impedances decrease to about 1/12 of the nominal value at 8 GHz,



(a) Attenuator differential-mode and common-mode output impedance.



(b) Laser differential-mode and common-mode input impedance.

Fig. 8. Comparison of the EM simulations of the attenuator and the measurements of the laser used in the probe circuit simulation. The magnitude of the attenuator differential-mode output impedance Z_{d2} and the common-mode output impedance Z_{c2} are compared (top). The magnitude of the laser differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} are compared (bottom).

while the relative ratio between the two impedances remains around 4:1 in the entire frequency range. The output impedance of the attenuator circuit is lower than the input impedance in most of the frequency range.

The magnitude of the differential-mode input impedance Z_{d1} and the common-mode input impedance Z_{c1} of the measured laser diode is shown in Fig. 8(b). The differential-mode input impedance of the laser is matched to the 100 Ω differential-mode internal impedance of the generator. At frequencies above 6.1 GHz, the differential-mode input impedance drops off a bit. The common-mode input impedance of the laser is very high at lower frequencies and drops gradually, with the minimum around 4.2 GHz, followed by a resonance and an antiresonance. The common-mode input impedance increases for frequencies above 5.6 GHz, but remains lower than the differential-mode input impedance.

D. Laser Input Voltage

The voltage components at the input balanced logical port (P1) of the laser, for the differential-mode and common-mode characterization of the laser by itself are shown in Fig. 9(a). The voltage components at the input of the laser extracted from the probe circuit simulation (Fig. 3), where the laser is connected in cascade with the attenuator circuit, for the differential-mode and common-mode characterization are shown in Fig. 9(b). The common-mode input voltage component $v_{c,i,c}$ and the differential-mode input voltage component $v_{c,i,d}$ for the common-mode characterization are defined as follows:

$$v_{c,i,c} = v_{c,i}^{+} + v_{c,i}^{-}, \tag{6}$$

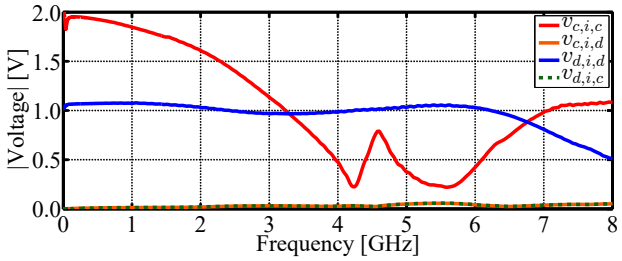
$$v_{c,i,d} = v_{c,i}^{+} - v_{c,i}^{-}. \tag{7}$$

The differential-mode input voltage component $v_{d,i,d}$ and the common-mode input voltage component $v_{d,i,c}$ for the differential-mode characterization are defined are defined as follows:

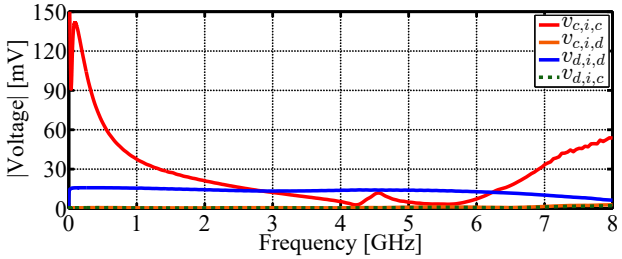
$$v_{d,i,d} = v_{d,i}^{+} - v_{d,i}^{-}, \tag{8}$$

$$v_{d,i,c} = v_{d,i}^{+} + v_{d,i}^{-}. \tag{9}$$

When the laser is characterized by itself, the differential-mode input voltage $v_{d,i,d}$ follows the trend of the differential-mode input impedance Z_{d1}, while the common-mode input voltage $v_{c,i,c}$ follows the trend of the common-mode input impedance Z_{c1} of the laser shown in Fig. 8(b). For the differential-mode characterization, the differential-mode signal component $v_{d,i,d}$ is around 1 V, because the differential-mode input impedance of the laser is matched to the 100 Ω generator differential-mode internal impedance. At frequencies above 6.1 GHz, the differential-mode voltage component drops with the differential-mode input impedance. For the common-mode characterization, the common-mode voltage component $v_{c,i,c}$ at low frequencies is around 2 V, because of the much higher common-mode input impedance of the laser compared to the 25 Ω generator common-mode internal



(a) Input voltage of the laser characterized by itself.



(b) Input voltage of the laser used in the cascade connection with the attenuator in the probe circuit simulation.

Fig. 9. Comparison of the voltage at the input of the laser diode extracted from the measurements of the laser by itself (top), and extracted from the probe circuit simulation using the cascade connection of the attenuator and the laser (bottom). The magnitude of the common-mode voltage component $v_{c,i,c}$ and the differential-mode voltage component $v_{c,i,d}$ in the common-mode characterization, and the differential-mode voltage component $v_{d,i,d}$ and the common-mode voltage component $v_{d,i,c}$ in the differential-mode characterization are compared.

impedance. At higher frequencies the common-mode voltage component decreases as a result of the changes in the common-mode input impedance. Due to the asymmetry of the laser, a differential-mode voltage component $v_{c,i,d}$ is present in the common-mode characterization, and a common-mode voltage component $v_{d,i,c}$ is present in the differential-mode characterization. These components slowly increase with frequency.

In the probe circuit, the laser is connected in cascade with the attenuator circuit (Fig. 3). The differential-mode and the common-mode voltage at the input of the laser are dependent on the attenuation of the differential-mode and the common-mode signal in the attenuator shown in Fig. 7, and the impedance matching between the output impedance of the attenuator and the input impedance of the laser shown in Fig. 8. More common-mode signal passes through the attenuator than the differential-mode signal. The differential-mode voltage $v_{d,i}$ and the common-mode voltage $v_{c,i}$ at the input of the probe circuit, shown in Fig. 6(b), are very similar and decrease very slightly with frequency, so this does not have a significant impact on the laser input voltage.

The common-mode output impedance Z_{c2} of the attenuator drops gradually with frequency, while the common-mode input impedance Z_{c1} of the laser drops quickly up to 4.2 GHz, but starts increasing again after 5.6 GHz. For the common-mode characterization the common-mode input voltage component $v_{c,i,c}$ roughly follows the trend of the common-mode input impedance of the laser. However, compared to the nominal case when the laser is characterized by itself, the common-mode input voltage $v_{c,i,c}$ is higher relative to the differential-mode input voltage $v_{d,i,d}$ in the lower frequency range up to 1.8 GHz and in the high frequency range above 5.5 GHz, while it is lower in the frequency range between 1.8 GHz and 4.2 GHz, as shown in Fig. 10.

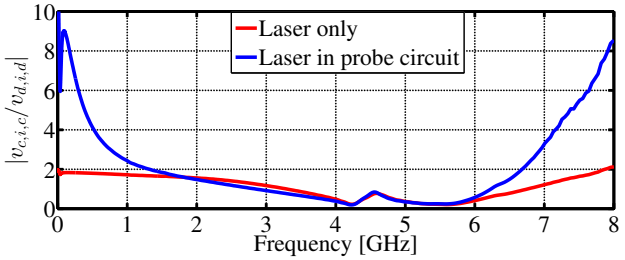


Fig. 10. Comparison of the voltage at the input of the laser extracted from the measurements of the laser by itself, and extracted from the probe circuit simulation using the cascade connection of the attenuator and the laser. The magnitude of the ratio of the common-mode voltage component in the common-mode characterization $v_{c,i,c}$ and the differential-mode voltage component in the differential-mode characterization $v_{d,i,d}$ is compared.

The differential-mode output impedance Z_{d2} of the attenuator drops gradually with frequency, while the differential-mode input impedance Z_{d1} of the laser remains almost constant up to 6.1 GHz. This means that more differential-mode voltage $v_{d,i,d}$ is applied at the input of the laser at lower frequencies, compared to the nominal case when the laser is characterized by itself.

E. Laser Signal Transmission

The mixed-mode S-parameter measurements of the laser used in the probe circuit simulation are shown in Fig. 11. The differential-mode transmission coefficient S_{sd21} is relatively stable over the entire frequency range. At frequencies above 6.6 GHz, there is a slight drop in the differential-mode signal level. The common-mode transmission coefficient S_{sc21} increases gradually with frequency over the measurement frequency range. The common-mode rejection ratio of the laser drops gradually with frequency, as shown in Fig. 12. It is above 30 dB up to 600 MHz and above 20 dB up to 3.1 GHz. It should be noted that the sudden drop in the CMRR above 4.5 GHz that is observed in the probe circuit characteristic is not

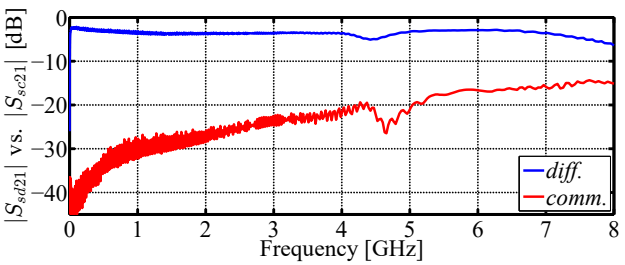


Fig. 11. Measurements of the laser used in the probe circuit simulation. The magnitude of the differential-mode transmission coefficient S_{sd21} and the common-mode transmission coefficient S_{sc21} are compared.

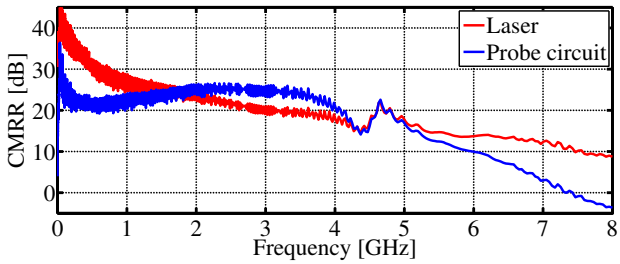


Fig. 12. Comparison of the common-mode rejection ratio (CMRR) of the probe circuit simulation and the measurement of the laser used in the probe circuit simulation.

present in the CMRR characteristic of the laser, nor the differential-mode and the common-mode characteristic. The same is true for the attenuator mixed-mode S-parameter characteristics shown in Fig. 7.

IV. Discussion

The high frequency drop in the CMRR that is observed in the probe circuit measurement and simulation results, shown in Fig. 4, is a result of the interaction between the attenuator circuit and the laser. Taking into account the differential-mode and common-mode signal propagation through the attenuator circuit, due to the impedance matching between the attenuator circuit and the laser, the common-mode voltage at the input of the laser increases significantly in the frequency range up to 1.8 GHz and above 4.2 GHz, relative to the differential-mode voltage, as shown in Fig. 10. Because of the higher common-mode signal relative to the differential-mode signal at the input of the laser, compared to the nominal case when the laser is characterized by itself, the CMRR of the probe circuit is degraded. The probe circuit CMRR is decreased in the frequency range up to 1.7 GHz and above 4.9 GHz. In turn, in the frequency range between 1.9 GHz and 4.3 GHz, where the common-mode signal is more attenuated than the differential-mode signal relative to the nominal case, the simulated probe circuit has a higher CMRR than the laser characterized by itself, as shown in Fig. 12.

The impedance matching between the attenuator circuit and the laser has a crucial effect on the CMRR characteristics of the probe circuit. In particular, the higher common-mode laser input voltage at lower frequencies, relative to the differential-mode input voltage, when compared to the characterization of the laser by itself, results in a relatively stable CMRR up to around 4 GHz. The typically higher CMRR of the laser at lower frequencies is flattened out by the higher common-mode voltage at the input of the laser, that is connected in cascade with the attenuator. Using the same mechanism, the increasing common-mode input voltage at high frequencies, relative to the differential-mode input voltage, when compared to the characterization of the laser by itself, accelerates the rate at which the CMRR drops at frequencies above 4.5 GHz, resulting in a steep slope.

A higher CMRR of the probe circuit could potentially be achieved by reducing the parasitic capacitance in the attenuator circuit and the laser diode layout, controlling the ground paths, and employing more effective common-mode signal filtering techniques.

V. Conclusion

An important mechanism of the common-mode signal propagation through the probe circuit is demonstrated. It is shown that not only the suppression of the common-mode signal and the mode conversion in each part of the probe circuit is important, but also the transmission of the differential-mode and the common-mode signal between the attenuator and the laser. While the CMRR of the probe circuit is primarily dependent on the CMRR of the laser used for the electro-optical conversion of the signal, the CMRR of the system can be increased above the CMRR of the laser, by optimizing the attenuator circuit design. The CMRR of the system can be improved by controlling the output impedance of the attenuator and the input impedance of the laser, where the portion of the common-mode signal that is transmitted to the laser should be minimized, while achieving a high transmission of the differential-mode signal.

Acknowledgment

This work was supported in part by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures (IP-2019-04-8959).

References

[1] H. Štimac, "Design and characterization of multiport electro-optical circuits for broadband voltage measurements in an electromagnetically polluted environment," Ph.D. dissertation, University of Zagreb Faculty of Electrical Engineering and Computing, 2020.

[2] H. Štimac, R. Gillon, and A. Barić, "Common-mode rejection ratio characterisation of a broadband electro-optical differential esd voltage probe," *Electronics Letters*, vol. 55, no. 19, pp. 1047–1049, Sep 2019.

[3] G. Colman, J. Bauwelinck, R. Gillon, A. Wieers, and J. Vandewege, "Wideband measurement system for on-chip esd waveform characterisation," *Electronics Letters*, vol. 48, no. 3, pp. 150–152, February 2012.

[4] Chong Ding and D. Pommerenke, "Laser optical in-circuit measurement system for immunity applications," in *2006 IEEE International Symposium on Electromagnetic Compatibility, 2006. EMC 2006.*, vol. 3, Aug 2006, pp. 569–574.

[5] *R&S ZVB Vector Network Analyzer Specifications*, Rohde & Schwarz GmbH & Co. KG, June 2011. [Online]. Available: https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_common_library/dl_brochures_and_datasheets/pdf_1/ZVB_dat_sw_en.pdf

[6] *Application Note: Three and Four Port S-parameter Measurements*, Anritsu, May 2002. [Online]. Available: http://ytdp.ee.wits.ac.za/11410-00279B.pdf

[7] K. Jung, L. A. Hayden, O. D. Crisalle, W. R. Eisenstadt, R. M. Fox, P. Hanaway, R. L. Campbell, C. McCuen, and M. Lewis, "A new characterization and calibration method for 3-db-coupled on-wafer measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 5, pp. 1193–1200, May 2008.

[8] W. Eisenstadt, *Microwave differential circuit design using mixed-mode S-parameters*. Boston: Artech House, 2006.

[9] *Keysight EEsof EDA EMPro*, Keysight Technologies, Inc., Sep 2017. [Online]. Available: https://www.keysight.com/us/en/assets/7018-02343/brochures/5990-4819.pdf

[10] C. T. Carrasco, C. J. Sieiro, J. M. Lopez-Villegas, N. Vidal, R. R. Gonzalez-Echevarría, and M. E. Roca, "Mixed-mode impedance and reflection coefficient of two-port devices," *Progress In Electromagnetics Research*, vol. 130, pp. 411–428, 2012.

Measurement Setup for Characterizing Immunity of Integrated Circuits to Pulsed Electric Fields using the IC Stripline

Can Cetin, Marko Magerl and Christian Stockreiter
ams-OSRAM AG, Premstätten, Austria
can.cetin@ams-osram.com

***Abstract* - A measurement setup is presented for characterizing the immunity of an ambient light sensor (ALS) integrated circuit (IC) to pulsed electric fields for consumer electronics products. The charge transferred to the photodiode during the rising and falling edges of the pulses is integrated and processed by the sensor, and it generates a shift in the ALS reading mean value. The pulsed electric fields up to 50 kV/m are applied to the IC using an open-terminated IC stripline with a septum height of 1.5 mm above the IC that is driven by voltage pulses with positive and negative amplitudes up to 80 V. The circuit for generating the voltage pulses synchronized with the on-chip photodiode integration phase is presented, and the achieved rising and falling edge characteristics are measured. The ALS reading shift is measured as a function of the voltage pulse amplitude, DUT orientation, level of illumination, and integrator gain. The relationship between the ALS reading shift and the electric field amplitude is established as a figure-of-merit for the immunity of the IC to pulsed electric fields.**

Keywords - pulsed electric field immunity; IC stripline; electromagnetic compatibility; ambient light sensors; charge injection

I. INTRODUCTION

The scope of electromagnetic compatibility testing continuously increases with the number of integrated circuits and the complexity of consumer electronics products. A subset of EMC tests are the radiated immunity measurements [1]. The radiated immunity of complex systems, such as electrical/electronic subassemblies (ESA) at the vehicle level [2], can be validated by testing each component separately [3].

At the level of individual ICs, such as integrated voltage-controlled oscillator (VCO) circuits [4], the radiated immunity is characterized according to the IEC standard 62132-8 [1] using the IC stripline, a component that enables applying electric and magnetic fields in the radio-frequency (RF) band to the device-under-test (DUT) through capacitive and inductive coupling [5].

Various types of open and closed IC striplines are designed and modelled in literature in order to achieve good load matching, i.e. having the voltage standing wave ratio (VSWR) below 1.25 in the frequency range up to 4 GHz [4], [5], [6]. The influence of the IC package on the EM field distribution inside the stripline is also considered when designing the IC stripline [7], [8].

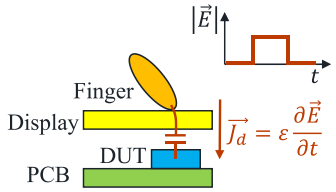


Figure 1: A consumer electronics application example motivating the pulsed electric field (*E*-field) testing requirement for ICs.

Ambient light sensors (ALS) are ICs that use a photodiode (PD) for generating a photo-current in the presence of visible light. The photo-current is converted into a digital value proportional to the illumination intensity using an on-chip current integrator and an analogue-to-digital converter (ADC).

Figure 1 presents an ALS application in a consumer electronics product. The finger touching the display is capacitively coupled to the DUT inside the product. When the charged finger presses onto the display, a pulsed electric field E is generated. During the rising and falling edges of the pulse, a non-zero displacement current J_d flows into the photodiode area of the sensor, equivalently to a photo-current, and it generates an erroneous ALS reading. In addition to this example, pulsed *E*-fields may also originate from the display operation itself.

In this paper, an experimental setup for the characterization of the immunity of ICs to pulsed electric fields (*E*-fields) using the closed IC stripline is presented. The IC stripline is chosen because it enables defining the precise distance between the DUT and the metal plate that applies the *E*-field pulses. In contrast to the 50 Ω termination required by the IEC standard [1], the IC stripline is open-terminated in order to reduce the DC current requirements and the magnetic field component within the IC stripline cavity.

The paper is organized as follows. Section II presents the pulsed *E*-field immunity measurement setup, including the voltage pulse generator circuit, the pulse edge characteristics, and the setup noise characteristics. Section III presents the measurement results of an industrial ambient light sensor over several illumination levels, device orientations, and device settings. The results are discussed in Section IV, and Section V concludes the paper.

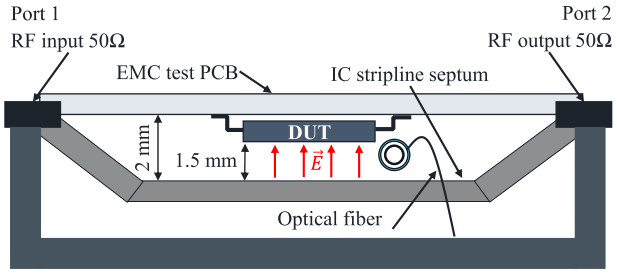


Figure 2: Cross-section of the closed IC stripline with a sidewall optical fiber input for controlling the illumination level within the cavity.

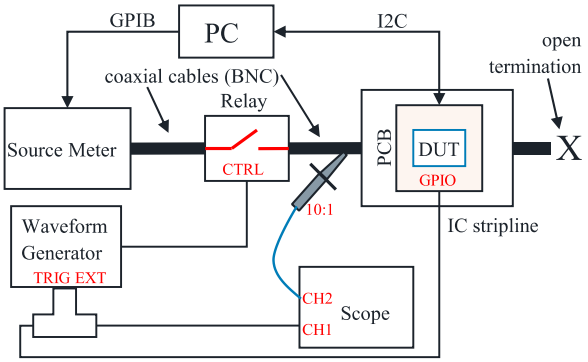


Figure 3: Block diagram of the pulsed *E*-field immunity measurement setup. Port 2 of the IC stripline is terminated in open circuit.

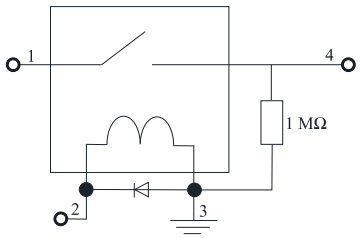


Figure 4: The schematic of the relay circuit. The Coto 9012-05-10 relay [9] connects the input pin 1 to the output pin 4 based on the control voltage applied across pin 2 and pin 3. The freewheeling diode between pin 2 and pin 3 protects the circuit against negative control voltage transients. The output (pin 4) has a weak pull-down resistor of 1 MΩ that defines the output voltage in the open circuit condition.

II. Measurement Setup

A. IC stripline for pulsed E-field immunity measurements

Figure 2 presents the cross-section of the closed IC stripline that generates pulsed *E*-fields from the septum that is 1.5 mm away from the DUT, which is soldered on the bottom layer of the 105 mm by 105 mm EMC test PCB on FR4 substrate. In this way, the distance of 1.5 mm between the DUT and the septum is precisely defined. Four DUT orientations relative to the septum are achieved by rotating the EMC test PCB. The setup is placed in a dark ambient, and the DUT is illuminated using an LED light source connected to the sidewall input by optical fiber. The illumination level is controlled by the LED current I_{LED}.

The *E*-field is uniform in the region between the septum and the ground plane, and the field strength is defined by the ratio of the voltage applied to the septum and the distance from the septum to the ground plane below the DUT [7]. Since the DUT size is comparable to the septum

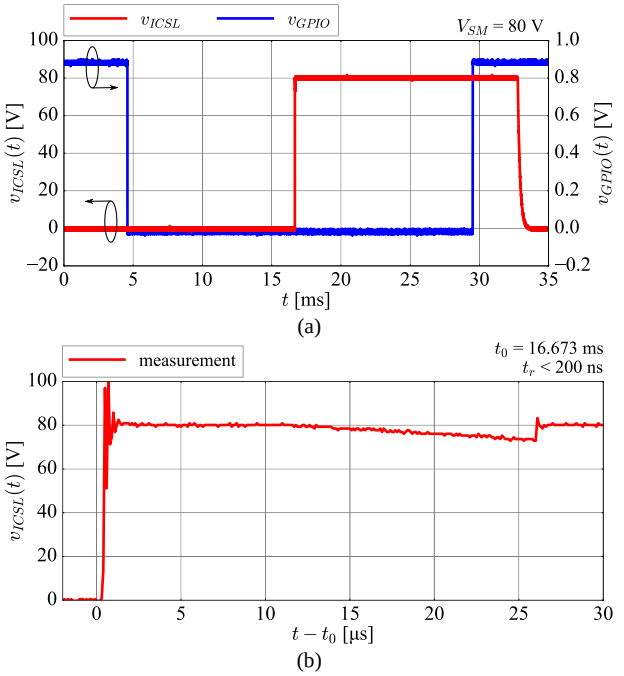


Figure 5: Rising voltage pulse edge: (a) timing relationship between the GPIO signal from the DUT (blue) and the signal applied to the IC stripline input (red), (b) zoom-in detail of the rising edge with the rise-time t_r under 200 ns. The additional voltage pulse at 27 µs is attributed to reflections due to the mechanical characteristics of the relay.

distance, the *E*-field is calculated using the 1.5 mm distance between the septum and the DUT surface, i.e. the active area of the IC that includes the ground plane. Therefore, a pulsed *E*-field amplitude of 50 kV/m is generated by applying a pulsed voltage amplitude of 75 V into the RF input of the IC stripline (Port 1 in Figure 2). Since most arbitrary waveform generators cannot generate voltages above 20 V, the voltage pulses are generated using the measurement setup presented in Figure 3.

The source meter (SM) that can generate DC voltages up to 100 V is connected to the IC stripline input by coaxial cables through the relay circuit shown in Figure 4. When the relay is closed, the IC stripline input voltage (pin 4 in Figure 4) is equal to the SM output voltage V_{SM} (pin 1 in Figure 4). When the relay is open, the 1 MΩ resistor pulls the IC stripline input voltage to 0 V. In this way, the relay circuit enables generating rising and falling voltage pulse edges with large amplitudes at the IC stripline input (port 2 in Figure 2) using low-voltage control signals applied to the relay control input (pin 2 in Figure 4).

If the IC stripline output (port 2 in Figure 2) would be terminated in 50 Ω according to [1], the maximum current flowing through the relay at SM output set to 80 V would exceed 1.5 A, which is above the maximum relay carry current of 0.5 A [9], thus destroying the relay. The IC stripline output port is therefore terminated in open circuit.

In order to synchronize the voltage pulse edges with the on-chip ALS integration, the relay control input is driven by a waveform generator that is externally triggered by the DUT signal (GPIO in Figure 3), which indicates that ALS integration is active. Figure 5 presents the rising voltage pulse edge measured using an oscilloscope, with channel 1 monitoring the GPIO trigger signal and channel 2 probing

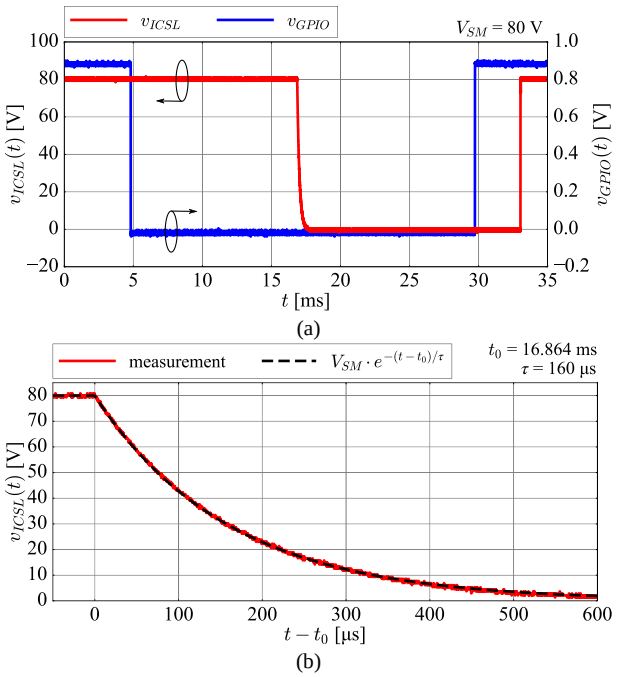


Figure 6: Falling voltage pulse edge: (a) timing relationship between the GPIO signal from the DUT (blue) and the signal applied to the IC stripline input (red), (b) zoom-in detail of the falling edge with the first order exponential shape defined by the time constant τ of 160 µs, resulting in a fall-time t_f of 352 µs.

the relay circuit output using a 10:1 probe. The rise time t_r is below 200 ns, with additional reflections due to the relay mechanical characteristics observed after 27 µs. Figure 6 presents the measured falling voltage pulse edge. The time constant τ of 160 µs is defined by the 1 MΩ pull-down resistance and the equivalent capacitance of 160 pF seen at pin 4 in Figure 4 when the relay is in the open position. Since the DUT responds to the total integrated displacement current, its output is not sensitive to the exact current profile. It was therefore decided to keep the 1 MΩ pull-down resistor regardless of the slow fall-time.

B. *Measurement setup noise floor*

The noise floor of the measurement setup is divided into two components: the ALS circuit noise, and the LED light source noise. The overall noise level is characterized by measuring a long time-series of ALS readings (expressed in "least significant bits", LSB) without any pulsed *E*-field applied. Figure 7a presents a time-series of 1000 samples measured at the LED light source current I_{LED} of 1 mA. The data are sorted into a histogram shown in Figure 7b with 20 bins. A normal distribution probability density function (PDF) is fitted onto the histogram, and the standard deviation σ of 25.4 LSB is extracted.

Table 1 presents the overview of the normal distribution parameters, i.e. mean value μ and standard deviation σ, obtained in this way over several illumination levels, two perpendicular device orientations on the IC stripline, and two device analogue gain (AGAIN) settings. The mean values under 90 degree orientation are higher than in the 0 degree case for the same LED current because the optical fiber illuminates the ALS photodiode under a different incident angle. The standard deviation of 8 LSB observed at 0 mA is attributed to the ALS circuit noise. The additional increase of the standard deviation with the light

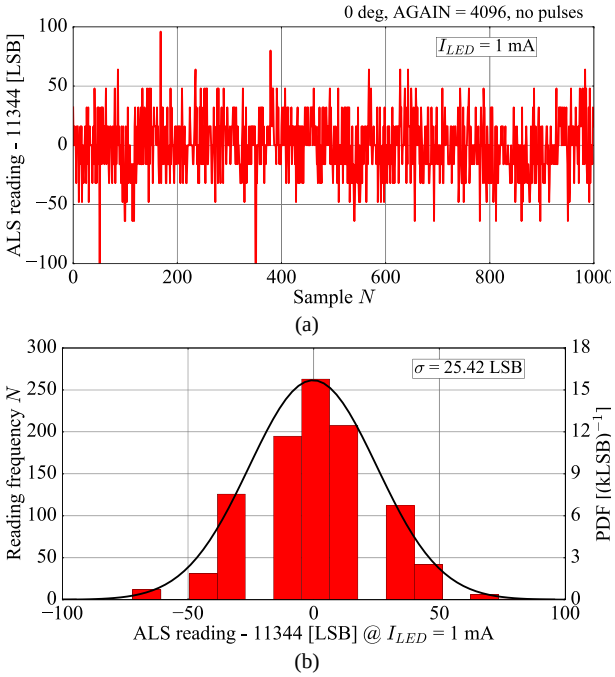


Figure 7: (a) Consecutive ALS reading samples under the illumination corresponding to LED current of 1 mA in the 0 degree orientation, (b) histogram of the sample reading frequency sorted into 20 bins with the fitted probability density function (PDF). The extracted standard deviation σ is 25.42 LSB.

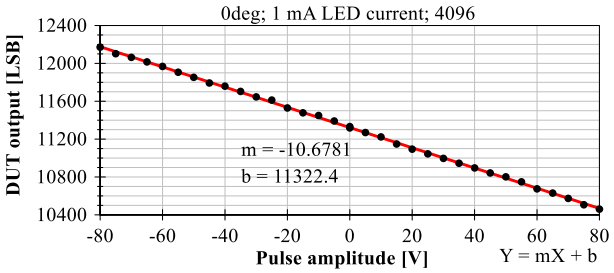
TABLE 1: MEAN VALUE μ AND STANDARD DEVIATION σ OF A 1000 SAMPLE LONG TIME-SERIES UNDER SEVERAL ILLUMINATION LEVELS, TWO PERPENDICULAR DEVICE ORIENTATIONS, AND TWO DEVICE ANALOGUE GAIN (AGAIN) VALUES.

AGAIN	I_{LED} [mA]	orientation [deg]	μ [LSB]	σ [LSB]
4096	0.0	0	0	8.16
		90	0	8.08
	0.2	0	355	11.15
		90	640	18.42
	0.5	0	4116	22.27
		90	7812	32.26
	1.0	0	11344	25.42
		90	21287	43.50
512	1.0	0	5886	10.55

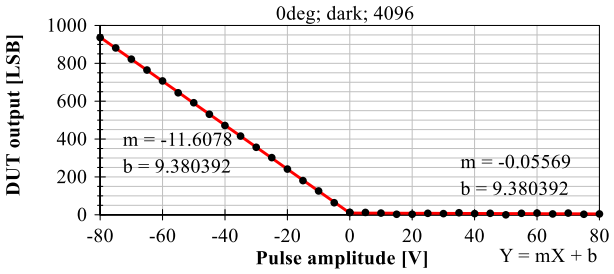
level is attributed to the LED light source noise, and it is within 35 LSB in the full range of illumination levels.

When the AGAIN setting is reduced from 4096 to 512, the ALS reading under the same illumination level decreases both in mean value and in standard deviation.

The noise floor characteristics of the measurement setup with the particular DUT presented in Table 1 are used as the baseline for determining the required averaging factor for the pulsed *E*-field immunity measurements as a function of the voltage pulse amplitude. Reducing the averaging factor from 1000 samples down to 10 samples significantly reduces the measurement time per pulse amplitude, while the achieved standard deviation remains within the same range as in Table 1. No significant increase in the standard deviation is observed when the voltage pulses are applied to the IC stripline input.



(a)



(b)

Figure 8: ALS reading (DUT output) measured as a function of the voltage pulse amplitude (black dots) and the interpolated function (red lines), taken in two illumination conditions: (a) LED current I_{LED} of 1 mA, (b) dark condition. The slope *m* and offset *b* are extracted separately for each piecewise linear part of the measured DUT output. The DUT analogue gain is set to 4096.

III. Experimental Results

The measurement setup introduced and characterized in Section II enables applying pulsed *E*-fields with both positive and negative pulse edges, which are synchronized with the on-chip ALS integration. The ALS reading in LSB is measured as a function of the voltage pulse amplitude in the range from -80 V (i.e. falling edge from 80 V to 0 V shown in Figure 6) to +80 V (i.e. rising edge from 0 V to 80 V shown in Figure 5) with a voltage step-size of 5 V.

Figure 8 presents the measurement results at two illumination levels. In Figure 8a, the LED current I_{LED} is set to 1 mA. The ALS reading at 0 V represents the undisturbed DUT output with the *E*-field pulse amplitude set to zero, which is equivalent to having no *E*-field pulses applied. The observed undisturbed ALS reading of 11322 LSB in Figure 8 is consistent with the value of 11344 LSB observed in the noise floor measurements in the same illumination condition given in Figure 7.

The ALS reading is observed to have a linear change with the voltage pulse amplitude, therefore the measurement data are interpolated according to the linear equation $Y = mX + b$, where X is the pulse amplitude (x-axis in Figure 8) and Y is the ALS reading (y-axis in Figure 8). The slope m of -10.7 LSB/V is the main pulsed *E*-field immunity characteristic of the DUT, and the offset b defines the corresponding DUT illumination level. The maximum observed absolute ALS error is 850 LSB.

The data in Figure 8a show that the difference in the rise time t_r and fall time t_f of the voltage pulses presented in Section II-A does not affect the slope m in the negative pulse amplitudes compared to the positive amplitudes.

In Figure 8b, the DUT is in the dark ambient, with the undisturbed ALS reading of 9 LSB. A linear relationship between the DUT output and the pulse amplitude is observed for negative pulse amplitudes, however at positive pulse amplitudes the ALS reading is clipped into the noise floor of 9 LSB, because the DUT does not support negative ALS values. The slope m and offset b are extracted separately for each piecewise linear part of the measured DUT output, and the maximum observed ALS output error is 925 LSB. The extracted slope m of -11.6 LSB/V at negative pulse amplitudes is in a similar range as in the illuminated case in Figure 8a.

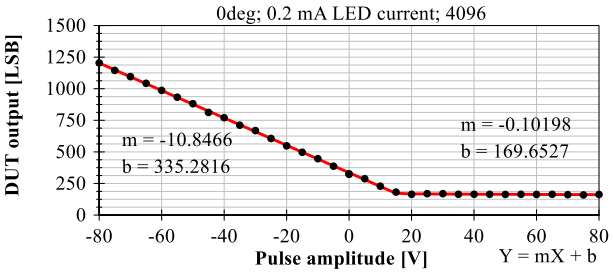


Figure 9: ALS reading (DUT output) measured as a function of the voltage pulse amplitude (black dots) and the interpolated function (red line), taken at LED current I_{LED} of 0.2 mA. The level at which the ALS reading is clipped increases with the illumination level. The DUT analogue gain is set to 4096.

TABLE 2: OVERVIEW OF THE INTERPOLATED SLOPE *m* AND OFFSET *b* UNDER SEVERAL ILLUMINATION LEVELS, TWO PERPENDICULAR DEVICE ORIENTATIONS, AND TWO DEVICE ANALOGUE GAIN (AGAIN) VALUES.

AGAIN	I_{LED} [mA]	orientation [deg]	*m* [LSB/V]	*b* [LSB]
4096	0.0	0	-11.61	9.4
		90	-11.61	8.4
	0.2	0	-10.85	335.3
		90	-10.98	627.4
	0.5	0	-11.87	4199.1
		90	-11.03	7803.7
	1.0	0	-10.68	11322.4
		90	-10.96	21232.7
512	1.0	0	-3.04	3126.3
		90	-3.04	5877.2

Figure 9 presents the measurement data with the LED current set to 0.2 mA, i.e. a medium illumination level. These measurement data show that the level at which the ALS reading is clipped is not always equal to 0 LSB, but rather it increases with the ambient illumination level. This behaviour is attributed to the ALS circuit operation.

Table I presents the extracted slope m and offset b over the following measurement conditions in the non-clipped part of the measurement data: two perpendicular device orientations, several LED current (I_{LED}) values, and two analogue gain (AGAIN) settings. All other DUT configuration settings are kept constant.

The slope m is approximately equal to -11 LSB/V across a wide range of illumination levels with AGAIN set to 4096. Reducing the AGAIN from 4096 to 512 under the same illumination level reduces the slope m down to -3 LSB/V, i.e. by a similar factor as AGAIN.

IV. Discussion

The presented measurement setup for characterizing the pulsed *E*-field immunity of ICs uses relay switching to enable generating voltage pulse amplitudes higher than what can be achieved by most arbitrary waveform generators. The drawback of the presented methodology is that the rising and falling edges have slow and widely different rise- and fall-times. The presented measurement data show that the extracted slopes *m* are not highly sensitive to the exact rise- and fall-times of the voltage pulses. This is explained by the implementation of the ALS function, which is based on integration of the displacement current flowing through the photodiode, rather than the exact profile of the current over time. The setup is therefore considered appropriate for measuring this family of ambient light sensors.

The mean values μ in Table 1 of the long time-series measurements taken without *E*-field pulses are consistent with the extracted offset values *b* in Table 2, which correspond to the ALS output values at 0 V voltage pulse amplitude, i.e. the undisturbed DUT output. This demonstrates the stability of the measurement setup and the repeatability of the measurements.

The measured noise floor of the measurement setup is within 44 LSB in the ambient illumination corresponding to the undisturbed DUT output of 21230 LSB, or 0.2%. This noise level is considered to be sufficient for extracting stable and accurate values of the slope *m*.

The unit of the slope *m* in this paper (LSB/V) is specific to a particular distance between the DUT and the IC stripline septum. A more general unit of measure would be expressed as LSB per kV/m of the *E*-field amplitude. The presented value of *m* equal to -11 LSB/V at 1.5 mm septum distance is equivalent to -16.5 LSB/(kV/m). In practice, the septum distance can be considered constant for a particular lab setup, therefore using LSB/V as the unit is appropriate during the characterization phase. The conversion to the more general unit of LSB/(kV/m) can be done as a final step in the pulsed *E*-field immunity report.

The negative sign of the measured slope *m* presented in Table 2, as well as the ALS error clipping effects observed in Figure 8b and Figure 9, indicate that the device is more sensitive to falling edges of the voltage pulse than the rising edges depending on the ambient illumination level. The data also indicate that the slope *m* reduces proportionally to the device analogue gain setting. This information about the IC characteristics are a useful input for defining the EMC mitigation strategies on the level of the consumer electronics product during the design-in phase of product development.

V. Conclusion

A measurement setup for characterizing the immunity of integrated circuits to pulsed electric fields based on the open-terminated IC stripline is presented. The circuit for generating voltage pulses with amplitudes up to 80 V is presented, with rise-times below 200 ns and fall times of 352 µs. The voltage pulse rising and falling edges are synchronized with the on-chip timing. The noise floor of the measurement setup is characterized, and it is divided into the on-chip analogue noise and the light source noise components. The relationship between the ALS reading error and the voltage pulse amplitude is measured over several illumination conditions, device orientations, and device analogue gain settings.

Acknowledgment

This work was supported by ams OSRAM.

References

[1] Integrated Circuits - Measurement of Electromagnetic Immunity - Part 8: Measurement of radiated immunity - IC stripline method, Int. Electrotechn. Commission, Geneva, Switzerland, IEC Standard 62132-8, 2012.

[2] Proposal for a new 06 series of amendments to UN Regulation No. 10 (Electromagnetic compatibility), Retreived: February 1, 2023, from https://unece.org/DAM/trans/doc/2019/wp29/ECE-TRANS-WP29-2019-020e.pdf.

[3] N. F. Bedjiah, M. Klingler, M. Kadi and R. Rossi, "Methodology to Validate the Radiated Immunity of Very Complex Systems by a Succession of Simple Component Radiated Immunity Tests at System Level," 2022 IEEE International Symposium on Electromagnetic Compatibility & Signal/Power Integrity (EMCSI), Spokane, WA, USA, 2022, pp. 561-564.

[4] J. Hwang, Y. Han, H. Park, W. Nah and S. Kim, "Radiated electromagnetic immunity analysis of VCO using IC stripline method," 2015 10th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo), Edinburgh, UK, 2015, pp. 147-151.

[5] H. Kang, W. Jung, K. Kim, H. H. Park and S. Kim, "Equivalent circuit model of the IC-Stripline coupling to IC package," 2014 IEEE 18th Workshop on Signal and Power Integrity (SPI), Ghent, Belgium, 2014, pp. 1-4.

[6] T. Mandic, R. Gillon, B. Nauwelaers and A. Baric, "Design and modelling of IC-Stripline having improved VSWR performance," 2011 8th Workshop on Electromagnetic Compatibility of Integrated Circuits (EMC Compo), Dubrovnik, Croatia, 2011, pp. 82-87.

[7] F. Fiori and M. Perotti, "On the use of the IC stripline to evaluate the susceptibility to EMI of small integrated circuits," 2016 International Symposium on Electromagnetic Compatibility - EMC EUROPE, Wroclaw, Poland, 2016, pp. 306-309.

[8] M. Koohestani, M. Ramdani and R. Perdriau, "Impact of Mode Propagation on Radiated Immunity Characterization in Commonly Used TEM Cells," 2019 12th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo), Hangzhou, China, 2019, pp. 5-7.

[9] 9011, 9012 & 9117 miniature SIP Relays - Coto Technology, Retreived: January 20, 2023, from https://www.cotorelay.com/wp-content/uploads/2014/09/9011_series_reed_relay_datasheet.pdf.

Characterization of Ground Bounce and Conducted Emissions of Integrated Isolated DC-DC Converter

Andrea Križanić*, Raul Blečić **, Andro Broznić** and Adrijan Barić*
*University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
** Rimac Technology d.o.o., Sveta Nedjelja, Croatia
andrea.krizanic@fer.hr, raul.blecic@rimac-technology.com, andro.broznic@rimac-technology.com, adrijan.baric@fer.hr

***Abstract* - Isolation is used in automotive industry for safety reasons to galvanically isolate different domains. Isolated DC-DC converters are used to provide power across the isolation boundary. Integrated isolated power converters provide simpler, smaller, and more robust solutions compared to discrete solutions such as push-pull converters. Integrated isolated DC-DC converters operate at switching frequencies in the range of tens of MHz which makes them potential sources of electromagnetic emissions. Ground bounce is the main cause of the electromagnetic emissions. A 5V-to-5V integrated isolated DC-DC converter based on ISOW774x with interlayer stitching capacitance is developed on a printed circuit board. The ground bounce is characterized by time-domain measurements. The conducted emissions of the converter are measured according to CISPR25 standard. EMI filtering design for the converter is discussed.**

***Keywords* – CISPR25, electromagnetic compatibility, electromagnetic interference, switched-mode power converter**

I. INTRODUCTION

Isolation technology protects parts of a system from high voltage surges and can also shield devices from signal interference. The most important role isolation has is safety. In the automotive industry isolation plays a vital role as electric vehicles are now running on 400 V or 800 V battery packs and it is necessary to isolate parts of the system that operate at a different voltage [1].

Isolated DC-DC converters are used to provide the necessary isolation and lower voltage levels. Integrated solutions for isolated power converters are smaller, simpler, and cheaper than converters with discrete components, such as push-pull converters. The downside of integrated isolated converters is they are operating at higher switching frequencies, in range of tens of MHz, which makes them potential sources of electromagnetic (EM) emissions of these frequencies and their harmonics.

For every electronic device, the allowed maximum levels of EM emissions are defined by specific electromagnetic compatibility (EMC) standards. For the automotive industry, this is Comite International Special des Perturbations Radioelectriques 25 (CISPR25) standard [2]. For measuring the conducted emissions, this standard defines the necessary test setup and limits for the frequency range from 150 kHz to 108 MHz in specific frequency bands [3]. In this paper, class 5 limits (the most stringent in CISPR25) are used.

Chapter II of this paper describes the designed printed circuit board and the used device. Chapter III presents the results of time-domain measurements, while chapter IV shows conducted emissions measurement results. Chapter V presents the conclusion.

II. PRINTED CIRCUIT BOARD

A. *ISOW774x*

The main component of the developed printed circuit board is the ISOW774x quad-channel digital isolator with integrated DC-DC converter. It has two possible configurations for the output voltage (3.3V or 5V) depending on whether the VSEL pin is shorted to GND2 or to V_{ISOOUT} pin. In this paper, the 5V output configuration is used. The maximum output power of the device is 550 mW, with up to 46% efficiency at maximum load. The integrated isolated converter operates at 25 MHz switching frequency. Information about the output voltage is sent to the primary side through an isolated feedback channel and the switching duty cycle is modified accordingly. Beside the feedback channel, there are four isolated digital channels. Fig. 1 shows simplified functional diagram of ISOW774x [4].

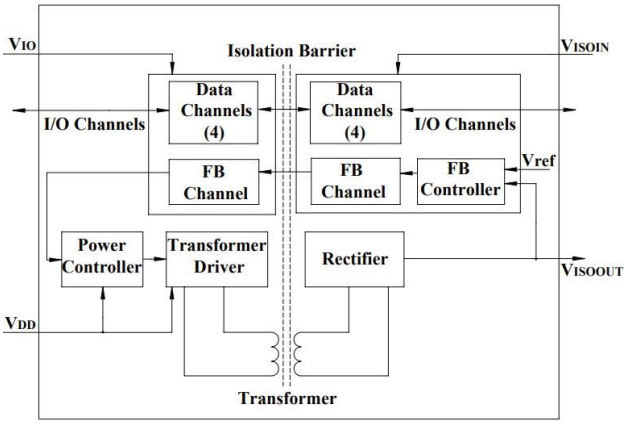


Figure 1. ISOW774x functional block diagram [4].

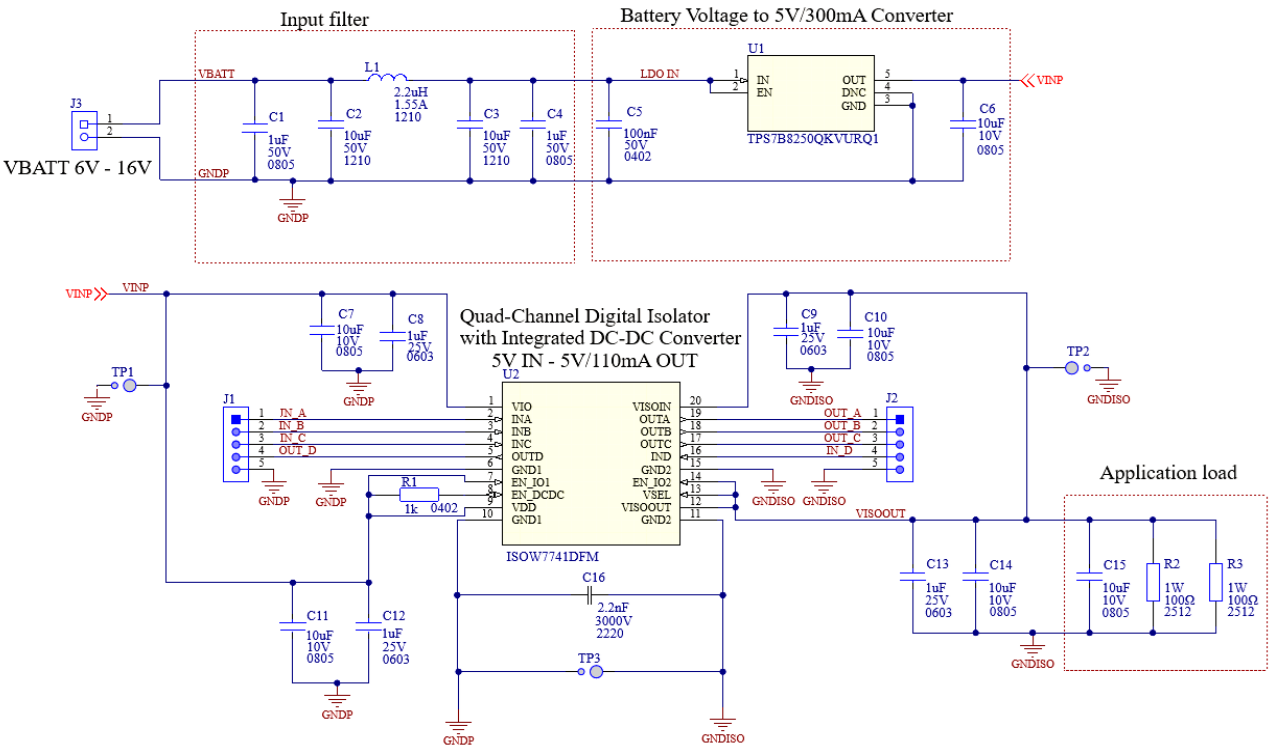


Figure 2. Printed circuit board schematic.

B. *Schematic and layout*

The printed circuit board (PCB) is designed in software tool Altium Designer [5]. Fig. 2 shows the schematic with important sections of the design marked, such as the passive input filter and LDO used for converting battery voltage to 5 V necessary for ISOW774x. The layout in 3D view is presented in Fig. 3. The board contains four metal layers, with copper thickness of 35 µm, and dielectric thickness between layers 1-2 and 3-4 of 100 µm. The dimensions of the board are 100 mm x 50 mm x 1 mm. Two ground planes GNDP and GNDISO are separated by 8 mm at the top and bottom layer, but overlap in two inner layers under the device creating a stitching capacitance. This technique is used to improve the EMC performance of the device [6].

III. TIME-DOMAIN MEASUREMENTS

A. *Setup*

Fig. 4 shows the measurement setup and the used instruments are marked as follows:

1. DC voltage supply
2. Oscilloscope with floating ground
3. Function generator
4. Programmable DC electronic load

The DC voltage supply is set to 7 V and it is connected to the VBATT connector on the board. The oscilloscope channels are connected to the three test points on the board and to the channel A output. The function generator is connected to the channel A input, and it is used as a source of the square wave signal with 1 MHz frequency and 5-V peak-to-peak amplitude. Instead of the application load marked in Fig. 2, a programmable DC electronic load is used to adjust the output load.

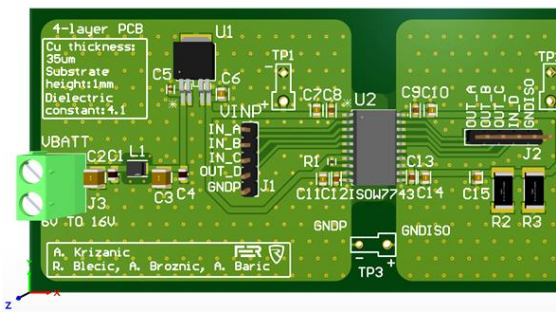


Figure 3. Printed circuit board 3D layout.

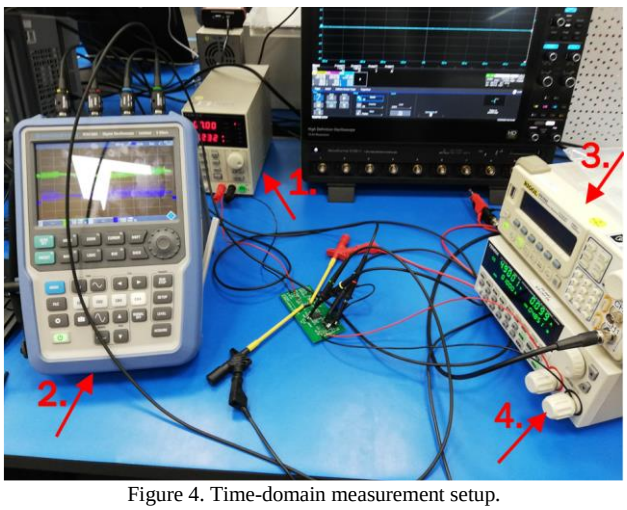


Figure 4. Time-domain measurement setup.

B. *Results*

The first step of measurement is checking the converter output voltage to confirm that the device is working properly. Table I shows the measured output voltage, along with the values of load resistance, input current, and voltage. These values are used to calculate the efficiency at a given load. The result is approximately 42% efficiency, which is a little lower than 46% given in the datasheet [4].

Fig. 5 and 6 show the AC output of the device without load and with 100 mA load. The change of the duty cycle depending on the voltage V_{ISOOUT} is visible from the on- and off-part of the cycle marked in the figures. With a higher current, the output capacitors are discharging much faster, so the off-time of the device is shorter and the duty cycle rises to approximately 0.5 when the load is applied. During the on-time, switching noise is visible as higher amplitude oscillations.

When a square wave signal is connected to the channel A it interferes with the converter output at every rising and falling edge, as it is shown in Fig. 7.

Fig. 8 shows the voltage between GNDP and GNDISO which is different from zero and confirms that these ground planes are isolated. During the on-phase of the converter the ground bounce is visible. This oscillation of the ground voltage is one of the main causes of electromagnetic emissions.

TABLE I.
MEASURED VALUES FOR CALCULATING EFFICIENCY

Parameter	I_{IN}	V_{DD}	V_{ISOOUT}	R_{LOAD}
Value	223 mA	4.97 V	4.86 V	50.9 Ω

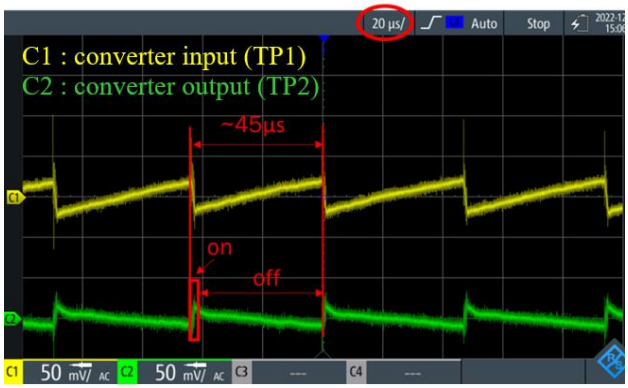


Figure 5. AC input and output without load.

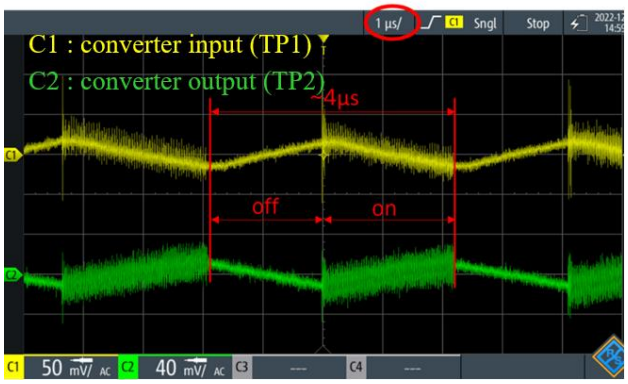


Figure 6. AC input and output with 100 mA load.

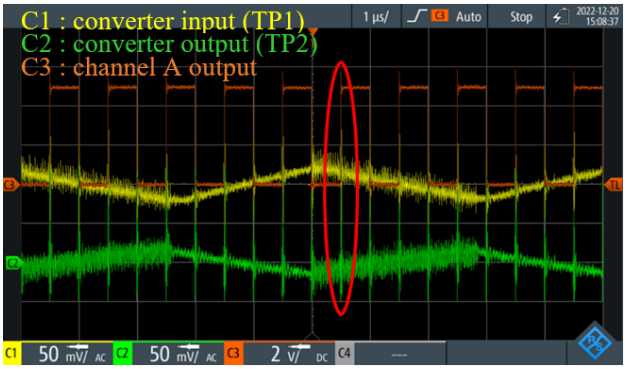


Figure 7. AC output with 100 mA load and channel A on.

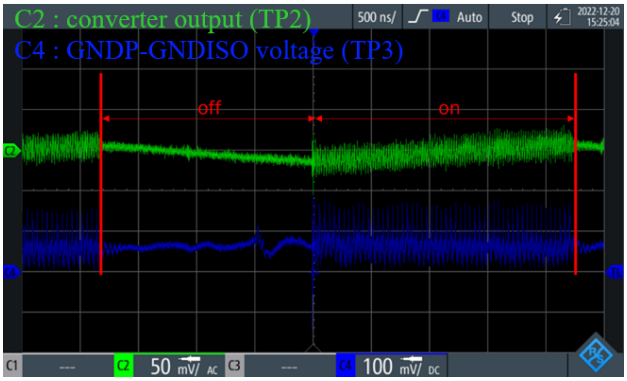


Figure 8. AC output and voltage between isolated ground planes with 100 mA load.

IV. CONDUCTED EMISSIONS MEASUREMENTS

A. *Setup*

To measure the conducted emissions an EMI-shielded box is used for testing in the lab. Inside the box, there is the device under test (DUT) placed on insulating material and connected to two Line Impedance Stabilization Networks (LISNs), as shown in Fig. 9. For these measurements a 5-µH LISN is specified by CISPR25 standard to "correspond with the approximate inductance of an automotive wiring harness" [3]. The power supply and EMI receiver are outside the shielded box and connected to LISNs through specific ports on the box. The frequency range for measurements is set from 150 kHz to 120 MHz, with a 9 kHz resolution bandwidth.

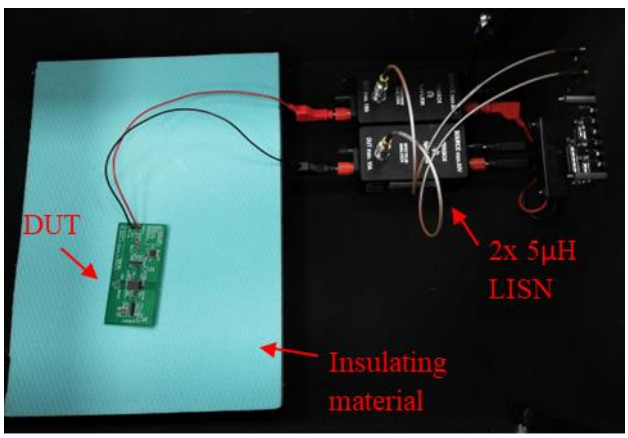
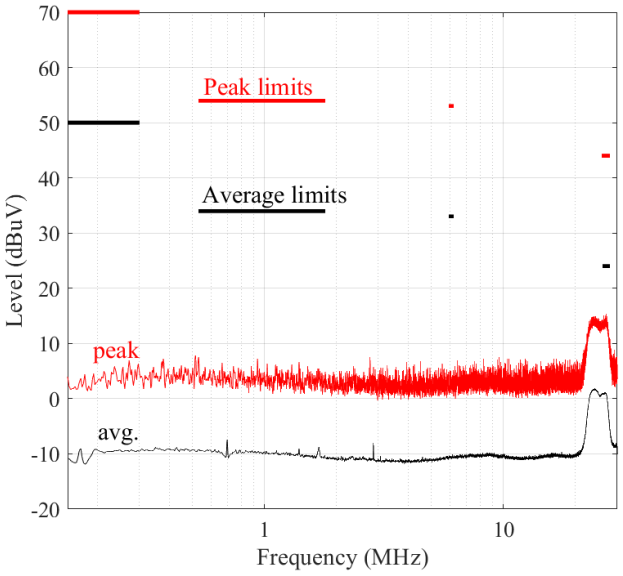


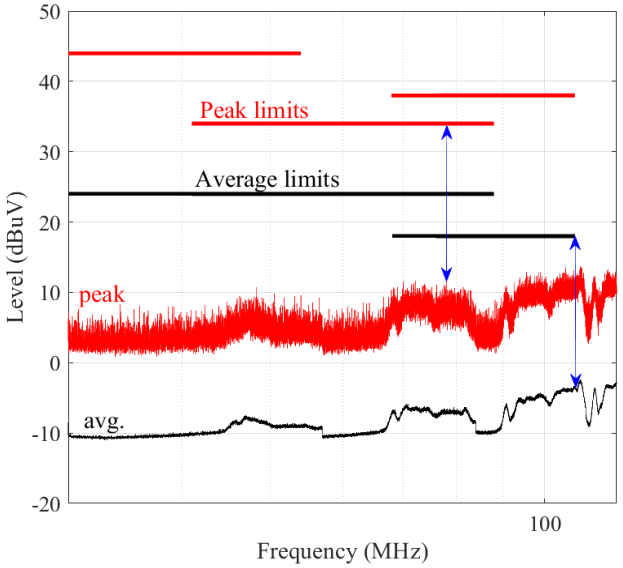
Figure 9. Conducted emissions measurement setup.

B. *Results*

The results are plotted in Matlab along with the limit lines of CISPR25 standard. Fig. 10 shows conducted emission measurements for a PCB without load and with an input filter. In Fig. 11 the values measured with a 51-Ω load are presented in two different configurations, with and without *C1*, *C2*, *C4* and *L1* components of the input filter. It is visible that both peak and average values are lower than the limits in both figures. There is also a clear difference in the amplitude of emissions with and without the load mounted on the board. Higher emissions with the load appear because of higher current and shorter off-time of the device. This is also shown in Table II where minimum margins for tests with load are much lower than without the load. The values of minimum margins are also marked on Fig. 10 and 11 with blue arrows.

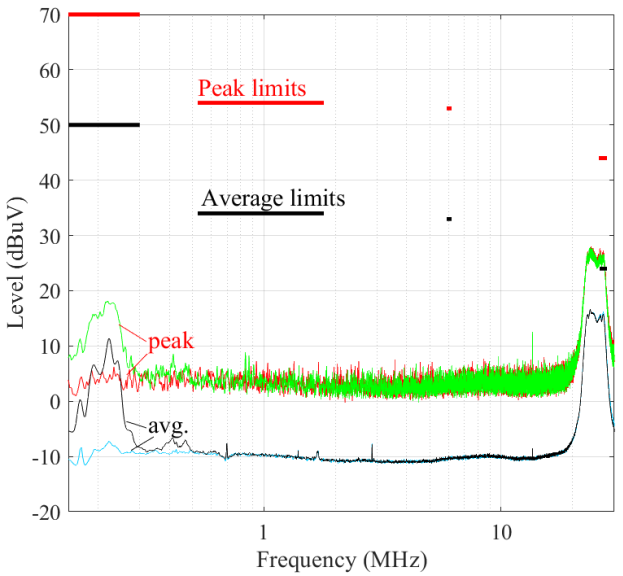


a) 150 kHz to 30 MHz range

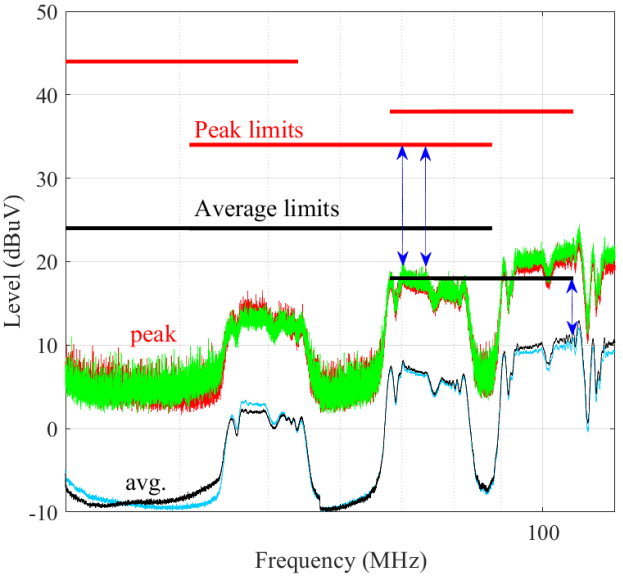


b) 30 MHz to 120 MHz range

Figure 10. CISPR25 conducted emmissions measurement without load (red line: peak values, black line: average values).



a) 150 kHz to 30 MHz range



b) 30 MHz to 120 MHz range

Figure 11. CISPR25 conducted emissions measurement with 51-Ω load (red line: peak values with filter, green line: peak values without filter, cyan line: average values with filter, black line: average values without filter).

To achieve the results under the limit lines, different techniques can be used. Mitigation of conducted EMI can be achieved at the source or along the propagation path [7]. The ISOW774x design includes advanced schemes to reduce the emissions at the system level, some of which are bigger on-chip decoupling capacitors and reduced common currents across the isolation barier [4]. From the presented results, it is visible that instead of high peak amplitude at the switching frequency of 25 MHz there is a wider spectrum of frequencies with lower amplitudes. This is called a spread spectrum technique and it is used for reducing EM emissions in many applications, including DC-DC converters [8].

For mitigating emissions along the propagation path external and internal filters can be used [7]. In this paper, an example of external filter is the passive input filter used for attenuating differential mode noise.

TABLE II.
MINIMUM MARGINGS FOR MEASURED RESULTS

Test	Minimum margin AV detector, dB	Frequency of minimum margin AV detector, MHz	Minimum margin PK detector, dB	Frequency of minimum margin PK detector, MHz
Without load, with filter	21.7	108	22.4	78
With load, with filter	7.2	108	14.5	74.4
With load, without filter	6.3	108	13.9	70.3

It dominantly works in a frequency range below 30 MHz, and its role is visible in Fig. 11 a) where there is a rise in amplitude without the filter. Another filter used is an interlayer stitching capacitance achieved with specific layout design. The overlaping ground planes under the device in the second and third layer create a common-mode capacitive filer [6].

V. CONCLUSION

A 5V-to-5V DC-DC converter based on ISOW774x developed on a printed circuit board is presented. Time-domain measurements confirm the functionality of the device and show the change of the duty cycle with varying load, interference from the digital channel and the ground bounce. The results of conducted emission measurements are below the CISPR25 class 5 limit lines. This is a result of the chip-level improvements in ISOW774x, input filter, and interlayer stitching capacitance on the PCB. The next step in the research is designing a 5V-to-5V isolated DC-DC converter in 4 different configurations on one PCB. It will include the one presented in this paper and also one designed with discrete components, both made with and without interlayer stitching capacitance. Those different configurations will then be tested and compared.

ACKNOWLEDGMENT

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures (IP-2019-04-8959).

REFERENCES

[1] TI (2022), Why isolation matters for your power electronics Accessed on: March 2023. [Online], Available: https://news.ti.com/blog/2022/08/30/why-isolation-matters-for-your-power-electronics.

[2] “Vehicles, boats and internal combustion engines–Radio disturbance characteristics – Limits and methods of measurement for the protection of on-board receivers.” CISPR 25:2016, fourth edition (or EN 55025:2017). CISPR: Geneva, Switzerland, October 2017.

[3] Timothy Hegarty, “An Engineer’s Guide to Low EMI in DC/DC Regulators”, Accessed on : December 2022. [Online], Available: https://www.ti.com/lit/eb/slyy208/slyy208.pdf.

[4] “ISOW774x datasheet”, Quad-Channel Digital Isolator with Integrated Low-Emissions, Low-Noise DC-DC Converter, 2022, [Online], Available: https://www.ti.com/lit/ds/symlink/isow7741.pdf.

[5] Altium Designer [PCB design software], Available: https://www.altium.com/altium-designer.

[6] Robert Martinez, Zhemin (Jimmy) Zhan, “Fundamentals of EMI Requirements for an Isolated DC/DC Converter”, Texas Instruments white paper, Accessed: December 2022., [Online], Available: https://www.ti.com/lit/wp/slyy202/slyy202.pdf.

[7] K. Mainali and R. Oruganti, "Conducted EMI Mitigation Techniques for Switch-Mode Power Converters: A Survey," in IEEE Transactions on Power Electronics, vol. 25, no. 9, pp. 2344-2356, Sept. 2010, doi: 10.1109/TPEL.2010.2047734.

[8] F. Pareschi, R. Rovatti and G. Setti, "EMI Reduction via Spread Spectrum in DC/DC Converters: State of the Art, Optimization, and Tradeoffs," in IEEE Access, vol. 3, pp. 2857-2874, 2015, doi: 10.1109/ACCESS.2015.2512383.

Characterization of Class-E Resonant Converter Operating in MHz-Range

I. Krois and A. Barić
University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
adrijan.baric@fer.hr

***Abstract* - A class-E resonant converter operating at 2 MHz is designed and characterized. The values of the inductive and capacitive components used in the designed converter are calculated using a set of mathematical equations. Time-domain measurements are performed to analyze the steady-state operation of the designed converter and to verify an impact of the converter parameters on the zero-voltage-switching behavior. An impact of switching frequency on output voltage and impact of duty cycle on zero-voltage-switching effect is analyzed. Finally, conducted electromagnetic emissions of the converter are measured at the nominal operating point.**

Index Terms* - *DC-DC converter, electromagnetic emissions, switching power supply, zero-voltage switching

I. INTRODUCTION

Reduction of physical dimensions of switching power converters is the major reason for the on-going attempts to push the operating frequency of switching power converters to higher values [1]. The conventional topologies of switching power converters suffer from hard-switching effects that cause a rather large switching loss. Furthermore, due to the rich spectrum of their switching voltage and current waveforms, fast switching converters conduct noise currents back to the power line at the harmonics of the switching frequency and cause conducted and radiated electromagnetic emissions [2].

The zero-voltage switching (ZVS) and zero-voltage-derivative switching (ZVDS) effects that occur in the class-E resonant converters significantly reduce switching losses ad generated EM emissions caused by the switching behavior of the converter [3].

The methodology used to calculate inductive and capacitive components used to shape the time-domain voltage and current waveforms of a transistor and a diode for a specified operating condition of a resonant converter is described in detail in [4], while a unified design theory for various topologies of the class-E resonant converter is proposed in [5]. An impact of inductors having different inductances, case sizes, saturation currents and quality factors on the operation of a resonant converter is simulated in SPICE and presented in [6].

In this paper, a non-isolated class-E resonant DC-DC converter designed to operate at 2 MHz is presented. The values of the inductive and capacitive components used in the converter are calculated using the analytical method described in [4]. The time-domain measurements and estimation of conducted EM emissions are performed to characterize the operation of the designed converter.

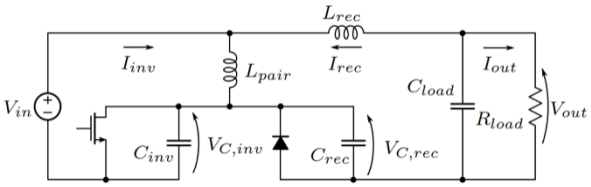


Fig. 1. Schematic of a non-isolated class-E resonant converter [7].

Section II presents the designed class-E resonant converter. The time-domain measurement results and conducted EM emissions are presented in Section III. Conclusion is given in Section IV.

II. DESIGNED CLASS-E RESONANT DC-DC CONVERTER

A schematic representation of a non-isolated class-E resonant converter analyzed in this paper is shown in Fig. 1 [7]. The converter is designed to operate at 12-V input voltage, 5-V output voltage, 1-A output current (R_{load} = 5 Ω) and switching frequency of 2 MHz. A 100-V N-channel MOSFET from Vishay [8] is used in the inverter side of the converter, while a PMEG6030ELPX Schottky diode with a current rating of 3 A is used as a free-wheeling Schottky diode in the rectifier side. A 5-Ω SMD load resistor is used to define the nominal output current of 1 A.

The design of the converter, i.e. the calculation of the inductive and capacitive components (L_{pair}, L_{rec}, C_{inv} and C_{rec}) is performed using the state-of-the art methodology presented in [4]. Denormalized values of the converter that are obtained are the following: $L_{pair} = L_{rec}$ = 1.68 µH, C_{inv} = 3.82 nF, C_{rec} = 0.86 nF.

A converter is designed on a dual-sided printed circuit board (PCB) which enables its insertion into the 62-pin connector to the 100 mm x 100 mm motherboard that can be used for far-field estimation in a (G)TEM cell. The values of the inductive and capacitive components that are soldered to the PCB, i.e. used in the assembled resonant converter, are the following: $L_{pair} = L_{rec}$ = 1.5 µH, C_{inv} =

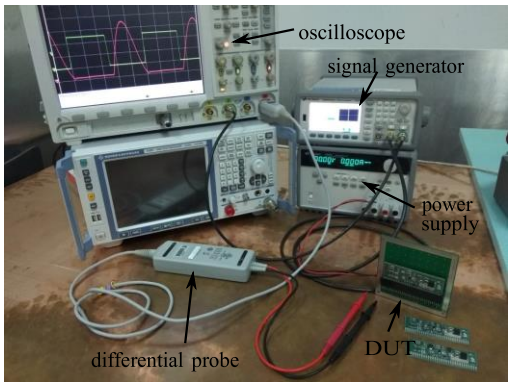


Fig. 2. Photo of the measurement setup for time-domain measurements.

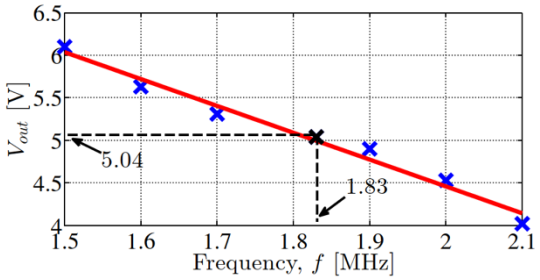


Fig. 3. Measured output voltage of the designed converter at different switching frequencies.

3.90 nF, C_{rec} = 0.82 nF.

III. Measurement Results

A. Time-domain

The measurement setup used for time-domain characterization of the designed resonant converter is shown in Fig. 2. A signal generator is used to generate a 2-MHz pulse signal for the operation of the transistor switch. The driver circuit is supplied by the 12-V DC voltage that is separated from the input voltage of the designed converter to enable the control of the input voltage.

The voltage across C_{inv}, i.e. the drain-source voltage of the transistor, and the voltage across C_{rec}, i.e. the Schottky diode, are measured using a differential voltage probe having a bandwidth of 100 MHz. Time-domain waveforms of the currents through L_{pair} and L_{rec} are measured using the high-sensitivity clamp-on current probe attached to the pieces of wire placed in series to the 1.5-µH inductors. A misalignment between the time delays of voltage and current probe is corrected using the deskew fixture [9]. The time-domain voltage and current waveforms are recorded using a 4-channel oscilloscope.

An impact of the switching frequency of the designed converter on the measured output voltage is shown in Fig. 3. The output voltage of the converter at 2-MHz switching frequency is measured to be 4.53 V, while 5.04 V at the output of the converter is achieved at switching frequency of 1.83 MHz. This discrepancy between measured and calculated data can be attributed to the large impact of the component parameters, due to the tolerance and parasitics,

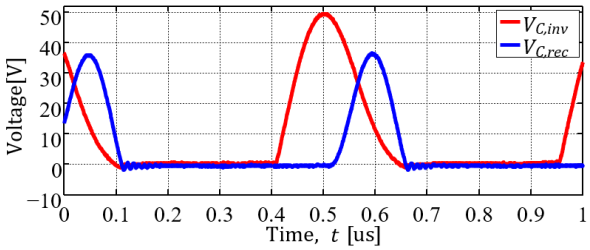


Fig. 4. Time-domain waveforms of the voltages across transistor $V_{C,inv}$ and diode $V_{C,rec}$ (D = 54%).

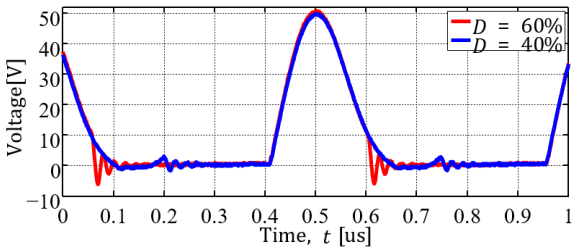


Fig. 5. Voltage across the transistor $V_{C,inv}$: impact of the duty cycle D on ZVS effect.

on the operation of the resonant converter.

The power efficiency η of the converter is estimated at 1.83 MHz and the output voltage of 5.04 V as follows:

$$\eta = \frac{V_{OUT}^2/R_L}{V_{IN} \cdot I_{IN}} = \frac{5.04^2/5.09}{11.89 \cdot 0.988} = 0.425 = 42.5\%.$$

The input and output voltage of the converter are measured using the digital multimeter, while the input current is measured by the DC voltage source applied to the input of the converter. The load resistance R_{load} of 5.09 Ω is measured using the precise four-point measurement method. This relatively low value of converter efficiency is due to losses in the ferrite cores of the used coils. These are coils that are commercially available and have a more favorable ratio of operating frequency, rated current and core volume. Increasing the efficiency of the converter should go in the direction of using RF air coils while increasing the frequency of the converter.

The time-domain waveforms of the voltages across the transistor switch ($V_{C,inv}$ in Fig. 1) and across the diode ($V_{C,rec}$ in Fig. 1) are shown in Fig. 4. The converter is set to reduce the 12-V input voltage to 5-V output voltage, but the peak values of 50 V across the transistor and almost 40 V across the diode can be observed, which puts a stringent limitation on capacitor selection in terms of breakdown voltage.

A slow increase from zero and slow decrease to zero value of the voltages across transistor and a diode at their turn-on and turn-off events indicate that ZVS operation of the designed converter is achieved. The waveforms shown in Fig. 4 correspond to the duty cycle of 54% at which the smoothest waveforms and lowest oscillations are obtained. An impact of the duty cycle on the time-domain waveform of the voltage across the transistor can be observed in Fig. 5. It can be seen that both increase or decrease of the duty ratio from the nominal value causes a significant increase

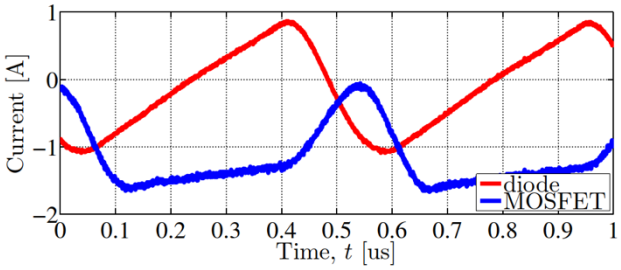


Fig. 6. Time-domain waveforms of the currents flowing through transistor and diode.

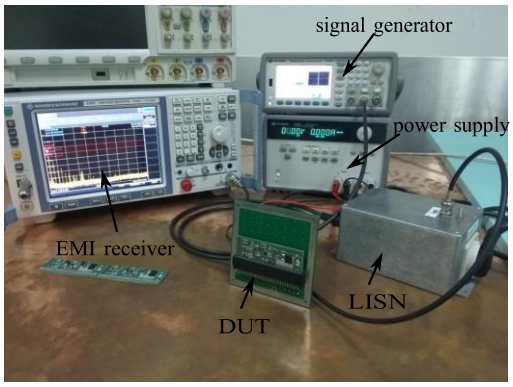


Fig. 7. Photo of the measurement setup used for estimation of conducted EM emissions.

of the oscillations observed in time-domain waveforms.

Finally, the time-domain waveforms of the currents flowing through the transistor and the diode are measured using the clamp-on current probe and shown in Fig. 6.

B. Conducted EM emissions

The measurement setup used for estimation of conducted EM emissions generated by the designed converter is shown in Fig. 7. The conducted emissions are measured between 150 kHz and 30 MHz using the EMI receiver and an average detector, 9-kHz RBW filter, logarithmic frequency sweep with 10 sweep counts and 8001 points per one frequency sweep.

The conducted EM emissions are measured for the input voltage of 12 V, switching frequency of 1.83 MHz and duty cycle of 54%, which produced the output voltage of 5.04 V. The red curve in Fig. 8 shows the noise generated by the measurement setup when the converter is switched off. The blue curve shows the measured conducted EMI voltage. The spectrum consists of the fundamental frequency at 1.83 MHz and odd harmonics. The limit lines for conducted emissions according to the CISPR standard are shown in black.

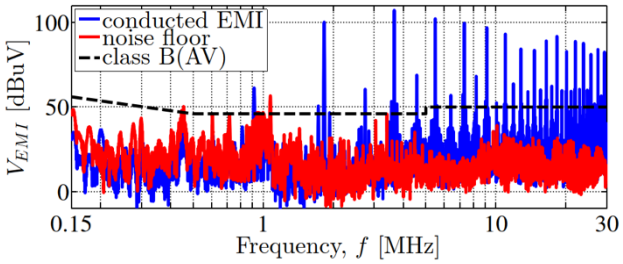


Fig. 8. Conducted EMI voltage of the designed resonant converter.

IV. CONCLUSION

A class-E resonant DC-DC converter that reduces the 12-V input voltage to the 5-V output voltage is designed and characterized. The designed converter operates at 1.83 MHz, while the duty cycle of 54% is selected to obtain the smoothest transitions between the on-and off-state. It is observed that the shift of the duty cycle from the nominal value causes a significant increase of the oscillations at the voltage waveforms. The conducted EM emissions of the converter are measured at the nominal operating point.

ACKNOWLEDGEMENT

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures (IP-2019-04-8959).

REFERENCE

[1] R. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*, 2nd ed. Kluwer Academic Publishers, 2001.

[2] H. Ott, *Electromagnetic Compatibility Engineering*, 1st ed. Wiley, 2009.

[3] M. Kazimierczuk and D. Czarkowski, *Resonant Power Converters*, 2nd ed. Wiley-IEEE Press, 2011.

[4] Bertoni, Nicola and Frattini, Giovanni and Massolini, Roberto G. and Pareschi, Fabio and Rovatti, Riccardo and Setti, Gianluca, "An Analytical Approach for the Design of Class-E Resonant DC-DC Converters," *IEEE Transactions on Power Electronics*, vol. 31, no. 11, pp. 7701–7713, 2016.

[5] F. Pareschi, N. Bertoni, M. Mangia, R. Rovatti, and G. Setti, "A Unified Design Theory for Class-E Resonant DC-DC Converter Topologies," *IEEE Access*, vol. 7, pp. 83 825–83 838, 2019.

[6] J. Bačmaga, R. Blečić, F. Pareschi, R. Rovatti, G. Setti, and A. Barić, "Frequency-Domain Characterization of Power Inductors for Class-E Resonant Converters," in *IEEE 23rd Workshop on Signal and Power Integrity (SPI)*, 2019, pp. 1–4.

[7] N. Bertoni, G. Frattini, P. Albertini, F. Pareschi, R. Rovatti, and G. Setti, "A First Implementation of a Semi-Analytically Designed Class-E Resonant DC-DC Converter," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015, pp. 221–224.

[8] *N-Channel 100 V (D-S) MOSFET – Si2392ADS*, Vishay, 2014, available at https://www.vishay.com/.

[9] *R&S RT-ZF20 Power Deskew Fixture*, Rohde&Schwarz, 2019, available at https://www.rohde-schwarz.com/.

Design and Characterization of the Output Network of a High-Current Buck Converter

Krešimir Šolaja*, Raul Blečić†, Andro Broznić†, Adrijan Barić*,
* University of Zagreb, Faculty of Electrical Engineering and Computing/Department of Electronics, Microelectronics, Computer and Intelligent Systems, Zagreb, Croatia
† Rimac Technology d.o.o./Embedded Hardware, Sveta Nedelja, Croatia
kresimir.solaja@fer.hr, raul.blecic@rimac-technology.com, andro.broznic@rimac-technology.com, adrijan.baric@fer.hr

***Abstract*—A 10-A automotive buck converter is designed and realized on a four layer printed circuit board. The output network is designed in the electromagnetic solver to account for the trace and package parasitic elements. The output network of the converter is simulated in the SPICE solver. The converter is characterized in the time domain by the load transient measurements and in the frequency domain by the shunt measurement method. Three configurations of the output network are compared. The results in the frequency domain and the results of the SPICE simulations are correlated.**

Keywords—automotive, PDN, target impedance, VRM, voltage overshoot

I. INTRODUCTION

Nowadays there is an increasing need for high current power supply that can power up different types of microcontrolers, processors and other electronics [1]. Efficiency of about 90% and higher is the main reason why DC-DC converters are commonly used for this purpose [2]. Buck converter is a form of DC-DC converter used for reducing the applied DC input voltage level to desired output voltage level. Buck converter is a part of a power distribution network (PDN).

The objective of the PDN is primarily to provide clean and reliable power to the active devices on the system. The PDN consists of the chip-level power distribution with thin-oxide decoupling capacitors, the package-level planes and mid-frequency decoupling capacitors and board-level power distribution with planes, low-frequency ceramic and bulk decoupling capacitors, and the voltage regulator module (VRM) [3]. Design of the PDN is typically done in the frequency domain by analyzing the impedance profile. Accurate method for characterization of low-impedance PDN is two-port shunt method. Main advantage of this method is reduced effect of series connection [4].

The second chapter describes the target impedance and the output network of the buck converter. Designed printed circuit board (PCB) of the buck converter with three versions of output network are presented. In chapter III, the results of simulation and measurements of the PDN impedance profiles are correlated. The load transient measurements and efficiency of buck converter for three versions of the output network are shown. Chapter IV concludes the paper.

II. POWER DISTRIBUTION NETWORK

A. Target impedance, output capacitance and inductor

The target impedance sets the limit on the highest impedance of the power rail for desired voltage change due to the amount of current drawn trough the rail. The objective is to have the output impedance below the target impedance. The value can be determined by the equation

$$Z_{target} = \frac{\mathrm{d}V}{\mathrm{d}I} \tag{1}$$

where dV represents the desired voltage ripple on the power rail and dI represents the maximum current change. The dV value is set to 5% of output voltage which is equal to 250 mV, and dI is equal to 10 A which then leads to target impedance value of 25 mΩ.

A group of different-value capacitors with low equivalent series resistance (ESR) is used to ensure that the PDN impedance value is below the target impedance. The purpose of capacitors with lower values is to control the impedance at higher frequencies and purpose of capacitors with larger capacitance is to control the impedance value at lower frequencies [5]. In this paper, six different capacitance values are used for covering frequency range as wide as possible. The total nominal capacitance is 255.3 µF.

The value of the output inductor ensures that the buck converter is always in the continuous conduction mode of operation. The selected value of the inductor is 1.5 µH [6].

Important part of designing output network of buck converter is the behavior of output voltage during load transient. The chosen value for undershoot and overshoot is 5% of the nominal output voltage which is equal to 250 mV. The following equations

$$C_{out} > \frac{L \cdot \mathrm{d}I^2}{2 \cdot V_{over} \cdot V_{out}} \tag{2}$$

$$C_{out} > \frac{L \cdot \mathrm{d}I^2}{2 \cdot V_{under} \cdot D \cdot (V_{in} - V_{out})} \tag{3}$$

which are given in [7], where C_{out} is the output capacitance, L is the inductance of an output inductor, dI the maximum current change, V_{over} the overshoot value, V_{out} the output voltage, V_{under} the undershoot value, D duty cycle and V_{in} is the input voltage give minimum capacitance value for the desired overshoot and undershoot

value. When the data from Table I is used the result is that the output capacitance needs to be larger than 95.4 µF.

TABLE I: Design parameters of a buck converter.

Parameter	Value
V_{in}	13.5 V
V_{out}	5 V
dI	10 A
L	1.5 µH
D	0.37

B. Buck converter LM61495Q1

LM61495Q1 10 A is a high power density buck converter with wide range of input voltage (3 -36 V) and adjustable output voltage which can be setup with the combination of resistors in the voltage feedback loop. The converter can have two behaviors while lightly loaded, when output current is smaller then 1/10 of the maximum output current. One behavior, called auto mode operation, allows a seamless transition between normal current mode operation while heavily loaded and in highly-efficient light-load operation. The other behavior, called FPWM mode, maintains constant frequency even when unloaded. The mode can be selected on the pin *SYNC/MODE* pin [6]. The frequency is also adjustable, it can be set on the pin *RT* with a resistor or when connected to VCC for 400 kHz and connected to GND for 2.2 MHz. Spread spectrum is configurable using the *SPSP* pin. Spread spectrum eliminates peak emissions at specific frequencies by spreading these peaks across a wider range of frequencies than a part with fixed-frequency operation [8].

C. Two-port shunt method

Two-port shunt method is used for measuring low-impedance structures such as PDN. The main advantage of using this method is that by using the two-port impedance measurements the effect of series connection discontinuities can be greatly reduced, which is the largest error contributor of one-port measurements [4]. Figure 1 shows measuring setup. The impedance can be calculated by the equation

$$Z_{DUT} = \frac{Z_{VNA}}{2} \cdot \frac{S_{21}}{1 - S_{21}} \tag{4}$$

while S_{21} is measured scattering parameter with the vector network analyzer (VNA).

D. Description of PCB

Layout of the four layer PCB is designed following the evaluation module board [9]. Stackup is shown in Table II. Wide and thick output plane is used to reduce the parasitic inductance and for better thermal performance of the buck converter. Ground vias are added beneath the converter for better thermal conductivity. The jumpers on the board are added to enable or disable converter, to switch the frequency (400 kHz or 2.2 MHz), to enable or disable spread spectrum and to choose the mode of operation (auto or FPWM). The board is designed to have three different

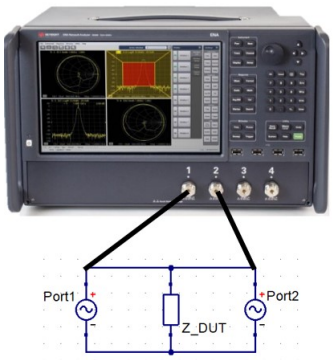


Fig. 1: Shunt method setup.

TABLE II: PCB layer stackup.

Name	Material	Type	Thickness
Top layer	Copper	Signal	70 µm
Dielectric 1	FR-4		100 µm
Inner layer 1	Copper	Ground plane	35 µm
Core	Core		1.2 mm
Inner layer 2	Copper	Signal	35 µm
Dielectric 2	FR-4		100 µm
Bottom layer 2	Copper	Signal	70 µm

versions of the output network, version A (Figure 2) that has all capacitors close to the IC, version B (Figure 3a) where all capacitors are placed close to the output and version C (Figure 3b) where the capacitors are again close to the output but some capacitors are removed leaving the total nominal capacitance of 179.5 µF.

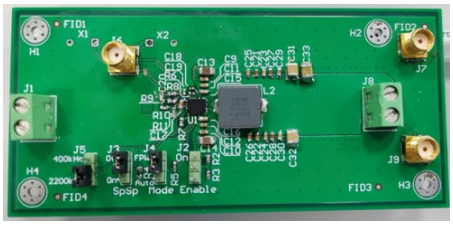
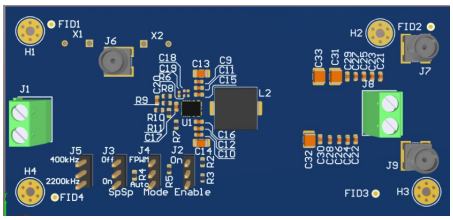


Fig. 2: Assembled printed circuit board - version A.

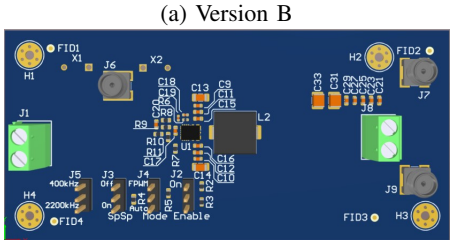
E. Simulation

Simulation process starts with SPICE simulation where the SPICE models of capacitors, available on the manufacturer's website [10], are added and scattering parameters are simulated in the shunt method configuration.

For a better estimate of the output network impedance, 3D electromagnetic simulations are performed using finite element method (FEM) solver. For this kind of simulations it is necessary to have the PCB layout and stackup with material parameters defined. The permittivity and dissipation factor of the core and prepreg materials are modeled by the frequency-depended Djordjevic-Sarkar model [11]. The mesh is adaptive at 100 MHz, while the sweep is interpolating with 100 points per decade (logarithmic scale) in the frequency range from 10 kHz to 100 MHz. Two circuit ports [12] are defined between the middle pin and the ground pin of the SMA connectors.



(a) Version B



(b) Version C

Fig. 3: Versions of output network shown in Altium Designer 3D viewer.

The differences between simulation types are shown in Figure 4. The main reason for such a difference between the simulations is that in SPICE simulation the parasitic elements of the PCB are not included.

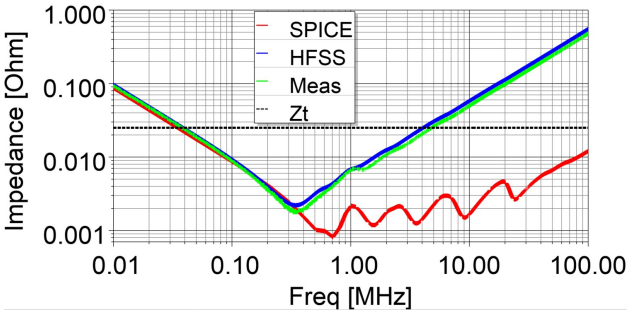


Fig. 4: Comparison of different simulation types and the measurement.

III. Results

A. Efficiency

The efficiency is defined as the ration of the output and input power commonly expressed in percentage. Table III shows the efficiency of three different cases of the output network.

TABLE III: Measured efficiency.

Case	V_{IN} [V]	I_{IN} [A]	V_{OUT} [V]	I_{OUT} [A]	Efficiency [%]
A	13.11	4.14	4.96	9.99	90.3
B	13.17	4.21	4.98	9.97	89.5
C	13.21	4.13	4.99	9.98	91.3

B. Comparison of simulation results and measurements in the frequency domain

The measurements are done using the shunt method at two SMA connectors that are placed on the PCB (J7 and J9), as shown in Figure 5. The PDN impedance is extracted from the measured scattering parameters. Figure 6 shows simulation and measurements results of output network impedance. It can be observed that cases A and B are matching at lower frequencies because they have the same capacitance, but they differ at high frequencies because of different parasitic inductance due to capacitors placement. The case C has lower capacitance and impedance characteristic is higher at lower frequencies. Table IV shows extracted values for capacitance at 10 kHz, equivalent series inductance (ESL) at 100 MHz and ESR extracted at the self-resonant frequency of each simulated and measured test case when buck converter in not powered.

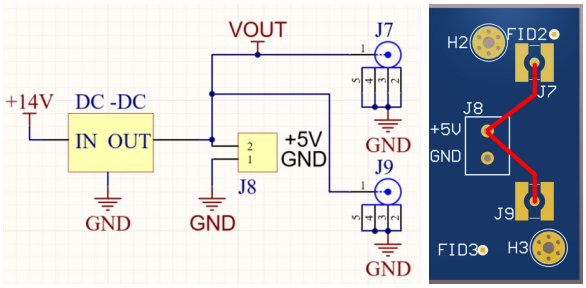
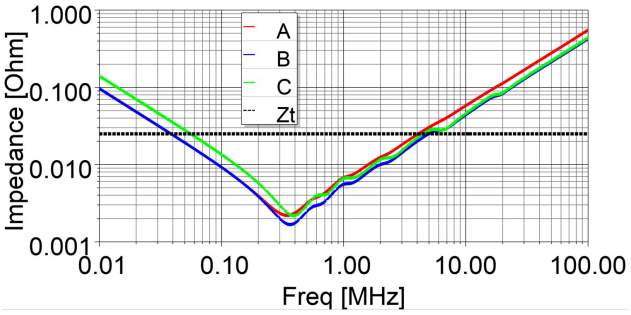
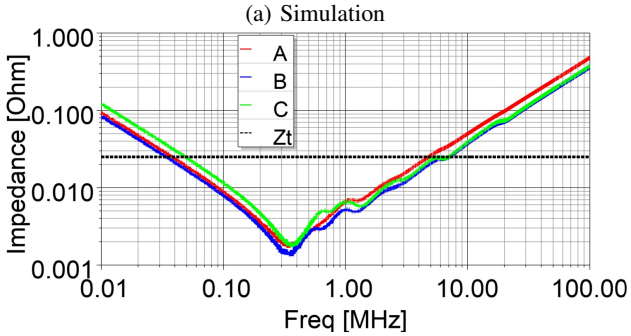


Fig. 5: Shunt method configuration: schematic (left), layout (right).



(a) Simulation



(b) Measurements

Fig. 6: Output impedance when the converter is off.

Figure 7 shows output impedance when the converter is turned on, off (without bias) and when the bias voltage is applied at the output connector. The red curve shows how the impedance characteristic is changed from blue when the bias voltage is applied, capacitance value is decreasing and the lower part of frequency characteristic is shifted to the right. At the high frequency there is no change

TABLE IV: Extracted values from simulation and measurements.

Case	A		B		C	
	Sim.	Meas.	Sim.	Meas.	Sim.	Meas.
C (10 kHz) [µF]	164	173	164	195	113	130
ESL (100 MHz) [pH]	879	765	675	570	700	592
ESR [mΩ]	1.8	2.2	1.3	1.7	1.8	2.2

due to the bias. Table V shows numerical values extracted from manufacturer's website, how capacitance is changed when the bias is applied [13]. The green curve on the graph shows the characteristic when the buck converter is turned on, output current is 5 A. The impact of the converter can be observed at lower frequencies. High frequency part depends on the total inductance of the output network (ESL, parasitic inductance of the plane, etc.). The characteristic with output current of 5 A does not differ from characteristic with output current 100 mA or 10 A.

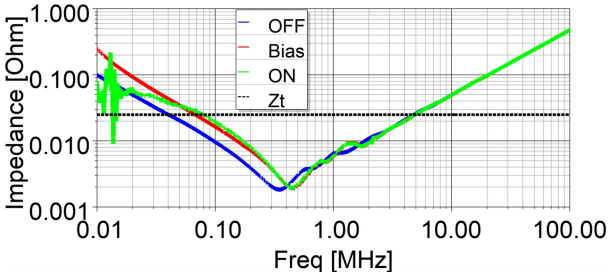


Fig. 7: Output impedance when the buck is turned on, off and with the bias voltage applied.

TABLE V: Capacitance change due to bias of 5 V

C [µF]	C bias [µF]	Qty.	∑ C [µF]	∑ C bias [µF]
100	46	1	100	46
47	26.8	2	94	53.58
22	5.5	2	44	11
4.7	2.8	3	14.1	8.32
1	0.73	3	3	2.19
0.1	0.096	2	0.2	0.192
			255.3	121.3

Figure 8 shows the impedance characteristic of three cases when the buck converter is turned on. As well as in the case when buck is turned off, cases A and B are different only at high frequencies due to different parasitic inductance, and the case C characteristic, with lower capacitance is shifted at lower frequencies.

C. Measurements in the transient domain

Transient response describes how a supply responds to a sudden change in load. That sudden change can cause large voltage transient undershoots and overshoots. Capacitors are used for maintaining stable voltage output, to provide short-term source of charge during the transient. Figure 9 shows the load transient response of the buck converter. Due to imperfection of the used active load device which has 3-A dip during transient from high to low level current, settings on the device is set from 10 A to 3 A to achieve the transient from 10 A to 0 A. All three cases have similar response, an overshoot of 300 mV. According to [14] current profiles have frequency content depending on duration and when that change is in microseconds it only depends on the VRM (Voltage Regulator Module) and not on the other parts of the PDN like capacitors, PCB planes, package and on-die capacitance.

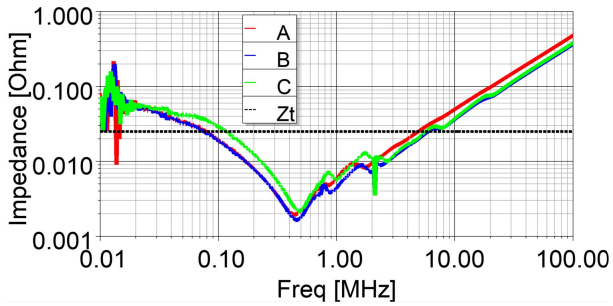


Fig. 8: Output impedance when the converter is on.

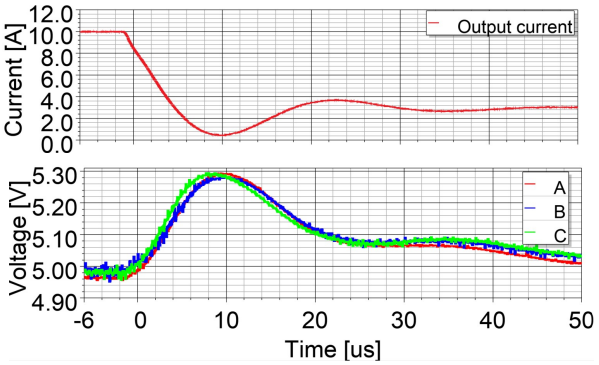


Fig. 9: Load transient.

IV. Conclusion

This paper shows design, simulation and measurements of the PDN. Simulations are done in the SPICE simulator and commercial EM solver. The most important factor in the design of the PDN is defining the target impedance. The 3D electromagnetic solver gives more realistic results than SPICE when compared to the measurements. The change of capacitance due to the applied bias voltage has to be considered during design. Load transient depends on the magnitude of the current change dI and on how fast that change is. If the change is in the range of microseconds then the voltage overshot is defined by the voltage regulator. Future work includes testing the device with faster load transients.

Acknowledgment

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures (IP-2019-04-8959).

REFERENCES

[1] R. W. Erickson and D. Maksimović, *Fundamentals of Power Electronics*. Springer, 2020.

[2] ROHM. (2016) Efficiency of buck converter. [Online]. Available: https://fscdn.rohm.com/en/products/databook/applinote/ic/power/switching_regulator/buck_converter_efficiency_app-e.pdf

[3] T. Instruments. (2022.) Sitara™ processor power distribution networks: Implementation and analysis. [Online]. Available: https://www.ti.com/lit/ml/sprac76f/sprac76f.pdf

[4] I. Novak, "Measuring milliohms and picohenrys in power distribution networks," *Design Con*, 2000.

[5] J. Arrigo, "Input and output capacitor selection," Texas Instruments, Tech. Rep., 2006.

[6] T. Instruments. (2021.) Lm62460, lm61480, and lm61495 pin-compatible 6-a/8-a/10-a buck converter optimized for power density and low emi. [Online]. Available: https://www.ti.com/lit/ds/symlink/lm61495.pdf

[7] R. Nowakowski, "Solving common buck converter design challenges," Texas Instruments, Tech. Rep., 2020.

[8] S. Jaffe, "The pros and cons of spread-spectrum implementation methods in buck regulators," 2021.

[9] T. Instruments. (2021) Lm61495rphevm evm user's guide. [Online]. Available: https://www.ti.com/lit/ug/snvu682a/snvu682a.pdf

[10] M. Manufacturing. (2023) Simsurfing. [Online]. Available: https://ds.murata.co.jp/simsurfing/mlcc.html

[11] A. R. Djordjevic, R. M. Biljie, V. D. Likar-Smiljanic, and T. K. Sarkar, "Wideband frequency-domain characterization of fr-4 and time-domain causality," *IEEE Transactions on Electromagnetic Compatibility*, vol. 43, no. 4, pp. 662–667, 2001.

[12] A. Inc. (2033.) Ansys hfss 3d components boundary, conditions, ports and mesh. [Online]. Available: https://www.ansys.com/training-center/course-catalog/electromagnetics/ansys-hfss-3d-components-boundary-conditions-ports-and-mesh

[13] I. Novak, K. B. Williams, J. R. Miller, G. Blando, and N. Shannon, "Dc and ac bias dependence of capacitors," *Design Con*, 2011.

[14] L. Smith, "Impedance and transient current fundamentals," *Design Con*, 2015.

Developing a Model of Buck Converter for EMI Filter Optimization by Circuit Simulations

J. Baća*, R. Blečić†, A. Broznić†, A. Barić*
* Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
† Rimac Technology d.o.o, Sveta Nedelja, Croatia
e-mail: juraj.baca@fer.hr, raul.blecic@rimac-technology.com, andro.broznic@rimac-technology.com, adrijan.baric@fer.hr

***Abstract*—Switching DC-DC converters for automotive applications typically require an input EMI filter to comply with the regulations. An automotive buck converter with the 5-A maximum output current is designed. To validate the performance, the buck converter is modelled in the circuit simulator. The model consists of the power stage and the voltage-mode control loop. Discrete components are modelled by real models with parasitic elements. The performance of the converter without the EMI filter is provided.**

Keywords—voltage-mode control loop, parasitic elements, conducted emissions

I. INTRODUCTION

Most of electronic devices are affected by electromagnetic interference (EMI) in the presence of electromagnetic waves caused by a source device or proximity device which is an operational disturbance. This can cause malfunction or unwanted operation of electronic circuits and components. EMI control is one of the more difficult challenges in switching DC-DC converter design, beyond functional issues, robustness, cost, thermal and space constraints. The most commonly used switching DC-DC converter is a buck converter, which down-converts a DC voltage to a lower DC voltage of the same polarity [1]. The main components of the power stage of synchronous buck converter are the high-side and low-side power MOSFET, with the output low pass filter which consists of the power inductor and output capacitor [2]. The synchronous buck converter uses high-side and low-side MOSFET as a switch that alternately connects and disconnects the input voltage to an inductor.

This paper is organized as follows. Section II describes the buck converter circuit. Selection of the main components and voltage-mode control loop design are presented in Section III. Simulations of the designed buck converter are shown and discussed in Section IV. Finally, a conclusion is given in Section V.

II. BUCK CONVERTER CIRCUIT

By using design requirements, such as the input and output voltage, operating frequency, the values of components in the output filter are determined, and power MOSFETs are selected. The selection of inductor above the critical value makes the buck converter operate in continuous-conduction mode and determines the output ripple current while the output capacitance directly affects the output voltage ripple and loop stability with the overshoot and undershoot during load transients [3]. After selection of the main components in the power stage, a voltage reference, an error amplifier and delay circuit must be added to the the power stage in order to regulate the output voltage. The complete buck converter circuit combined of these basic blocks is shown in Fig. 1.

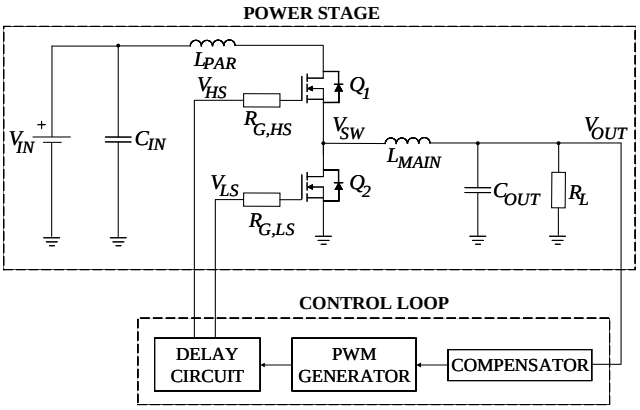


Fig. 1: Synchronous buck converter with voltage-mode control loop.

Conducted emissions are divided into differential-mode and common-mode interference [4]. Differential-mode interference is a noise signal which exists between line and return conductors. In contrast, common-mode noise is a signal which is in-phase on both the line and return conductors reffered to the safety ground. The focus is on conducted EMI in switching DC-DC converters generated by high current transients at the input and switching actions of the power MOSFETs.

III. BUCK CONVERTER DESIGN

A. Power MOSFET's model

When a MOSFET is used as a switch, its basic function is to control the drain current by the gate voltage. Fig. 2 shows the model of a power MOSFET. In order to model the power MOSFET, the voltage-contolled switch SW_1 with the on-state resistance R_{ds} is connected in series. Due to their structure, MOSFETs have parasitic capacitances and the switching performance of the device is determined by the time required to establish voltage across the parasitic capacitances [5]. In additon to the on-

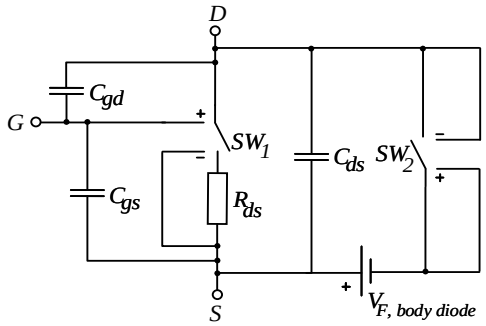


Fig. 2: Power MOSFET model.

state resistance, the gate-to-source capacitance C_{gs}, the gate-to-drain capacitance C_{gd} and the drain-to-source capacitance C_{ds} have been added to the model. Furthermore, due to p-n junction between the source/drain and the bulk, MOSFETs have an intrinsic body diode providing a path for inductive load current to bypass the MOSFET during the off-state. The body diode is modelled by using the voltage-controlled switch SW_2 and the 0.7-V reference voltage $V_{F,bodydiode}$. The MOSFET parameters selected from the automotive-compliant, 5 A, synchronous buck converter AP64502Q [6] are shown in Table I.

TABLE I: MOSFET parameters

Parameter	high-side MOSFET	low-side MOSFET
R_{ds}	45 mΩ	20 mΩ
C_{gs}	90 pF	200 pF
C_{ds}	350 pF	200 pF
C_{gd}	25 pF	20 pF

B. Inductor, input and output capacitor selection

Design parameters:

- Output voltage, V_{OUT} = 5 V
- Maximum load current, $I_{OUT_{max}}$ = 5 A
- Switching frequency, f_{sw} = 2.2 MHz
- Nominal input voltage, $V_{IN_{nom}}$ = 14 V
- Input voltage range, V_{IN} = 6 V to 28 V
- Duty cycle, D = 0.36
- Transient load step, I_{step} = 2.5 A
- Voltage overshoot V_{OS} = 1.03 · V_{OUT} = 5.15 V

Design requirements:

- Inductor peak-to-peak ripple current, $\Delta I_{L_{pp}}$ = 1 A
- Output peak-to-peak ripple voltage, $\Delta V_{OUT_{pp}}$ = 20 mV
- Allowed input peak-to-peak ripple voltage, $\Delta V_{IN_{pp}}$ = 100 mV

The required inductance is calculated from the desired peak-to-peak ripple current. The higher the inductor value, the higher is the output current because of the reduced ripple current. Commonly, the lower the inductor value, the smaller is the solution size. The following equation shows the needed inductance for the ripple current [3]:

$$L_{MAIN} > \frac{V_{OUT} \cdot (V_{IN_{nom}} - V_{OUT})}{\Delta I_{L_{pp}} \cdot f_{sw} \cdot V_{IN_{nom}}} \quad (1)$$

The inductance should be greater than 1.46 µH as calculated by (1). The inductor must always have a higher current rating than the maximum load current of 5 A due to current increase with decreasing inductance. A 1.5 µH Coilcraft power inductor from the XEL series is selected [7]. Parasitic elements of the inductor include the resistive element R_{DC} in series and the parallel capacitance $C_{P,MAIN}$ with the series resistance $R_{S,MAIN}$. To complete the model of the inductor, the parallel resistance $R_{P,MAIN}$ is added. Fig. 3 (a) shows the inductor model. The values of the parasitic frequency-dependent elements are calculated from [7], as shown in Table II.

TABLE II: Inductor parasitic elements (Coilcraft inductor)

Inductor	R_{DC}	$C_{P,MAIN}$	$R_{S,MAIN}$	$R_{P,MAIN}$
L_{MAIN}	4.5 mΩ	12 pF	30 Ω	1400 Ω

Due to the maximum output current $I_{OUT_{max}}$ of the DC-DC switching converter for a given application, the minimum value of the input capacitor is critical for stabilizing the input ripple voltage. Equation (2) estimates capacitance C_{IN} that will meet the ripple requirement [2]:

$$C_{IN} > \frac{D \cdot (1 - D) \cdot I_{OUT_{max}}}{f_{sw} \cdot \Delta V_{IN_{pp}}} \quad (2)$$

The input capacitance should be greater than 5.4 µF. With two 4.7 µF capacitors in parallel, the total effective capacitance is sufficient for stabilizing the ripple voltage.

Commonly the selection of the output capacitor is not driven by the desired ripple voltage, but by the load transient response. The output voltage fluctuates due to the time it takes the inductor to catch up with the increased or decreased output current. An empirical formula (3) is used to calculate the necessary output capacitance for the specified output overshoot or undershoot [8]:

$$C_{OUT} > \frac{I_{step} \cdot (\frac{1}{4 \cdot f_c} + \frac{1}{f_{sw}})}{2 \cdot V_{OS}} \quad (3)$$

A good starting value for the crossover frequency f_c is 1/20 to 1/10 of the switching frequency f_{sw}. For this application, f_c = 1/20 · f_{sw} = 110 kHz is selected. The minimum output capacitance C_{OUT} for the 3 % output voltage change with the crossover frequency f_c of 110 kHz is 22.7 µF. Connecting three 10 µF capacitors in parallel is enough to supply the current difference to maintain the output voltage within the specified range. Table III shows the chosen Murata MLCC capacitors parasitic elements from the GCM series.

The input and output capacitors are modelled with parasitic elements. The impedance-frequency characteristic is used for extracting equivalent series resistance (ESR) and the equivalent series inductance (ESL) of the output capacitors [9]. Real model of the capacitor with parasitic elements is shown in Fig. 3 (b).

TABLE III: Input and output capacitor parasitic elements (Murata MLCC capacitors)

Capacitor	ESL	ESR
4.7 µF (C_{IN} = 2×4.7 µF)	0.6 nH	4 mΩ
10 µF (C_{OUT} = 3×10 µF)	0.7 nH	3 mΩ

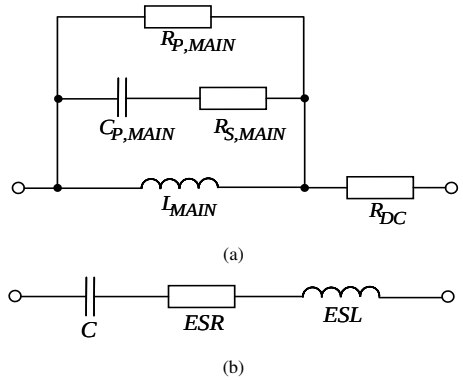


Fig. 3: Parasitic elements of the components. (a) Model of the inductor. (b) Model of the capacitor.

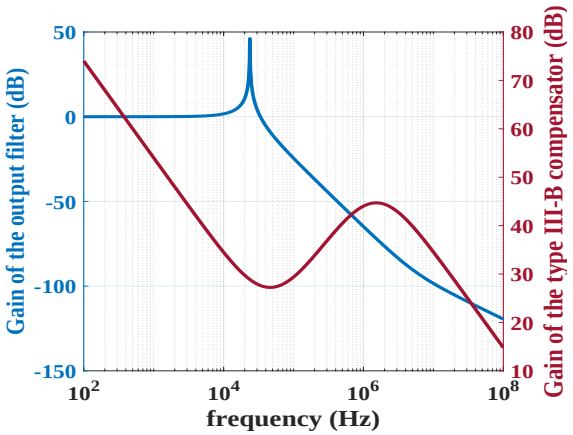


Fig. 4: Simulation of the amplitude-frequency chacteristic of the output filter and amplitude-frequency chacteristic of type III-B compensator for determined zeros and poles. The first zero of the compensator (f_{z1}) compensates the phase lag of the pole which is at the origin. The second zero (f_{z2}) is to compensate for one of the poles of the LC filter so that at f_c the slope of the bode plot of the loop is about -20 dB/dec. The second pole of the compensator (f_{p2}) and the zero of the ESR of the capacitor (f_{ESR}) cancel each other and the third pole (f_{p3}) is to provide more attenuation for frequencies above f_s/2.

C. Voltage-mode control loop design

In voltage-mode control the output voltage is returned through a feedback loop. The differential voltage, which is obtained to compare the output voltage with the reference voltage by an error amplifier, is compared with the triangular waves by a PWM generator [10]. As a result, the pulse width of the PWM signal is determined to control the output voltage.

In order to ensure a stable system, a slope of the loop gain should be about -20 dB with a phase margin greater than 45 ° for the overall stability. To make the analysis of the compensation network simpler, the ESL of the output capacitor and the series resistance R_{DC} with other parasitic elements of the output inductor are neglected. The transfer function of the output filter is a second order system with a double pole at a resonance frequency of the LC filter and a zero produced by the ESR of the output capacitor, as shown in Fig. 4. The double pole at 23.7 kHz causes the gain to fall with a slope of -40 dB/dec up to zero frequency of 5.3 MHz which compensates one of the poles.

As mentioned, having a stable closed loop buck converter with an appropriate performance, a properly designed compensator is required. The feedback is implemented with a type III-B compensation network [10] due to location of the desired zero crossover frequency f_c = 110 kHz, switching frequency f_{sw} = 2.2 MHz and charasteristics of the output filter. Using equations described in [10], the desired locations of the poles and zeros of the compensator are calculated. Fig. 4 shows the amplitude-frequency charasteristic of the type III-B compensator. Furthermore, the resistors and capacitors of the compensator (see Fig. 5) with the closest standard components values to the calculated results for achieving detemined poles and zeros are shown in Table IV.

TABLE IV: Values of parameters in compensation network

Parameter	Value
R_{F1}	23.7 kΩ
R_{F2}	3.83 kΩ
R_{F3}	768 Ω
C_{F3}	0.1 nF
R_{C1}	374 kΩ
C_{C1}	13 pF
C_{C2}	0.39 pF

An error amplifier added between the regulator output and the PWM generator holds the output voltage constant and controls the duty cycle of the PWM signal providing required gain and phase of the loop so desired phase margin and zero crossover frequency are attained. The error amplifier is implemented using an operational amplifier modelled by the input and output resistance with a voltage-controlled voltage source. In order to close the loop of the buck converter, the PWM generator is inserted so that a repeating rectangular output waveform in response to a control voltage is produced. As the control voltage increases, the duty cycle of the output increases as well. When the control voltage is greater than the voltage of the ramp signal, the output is continuously high. In contrast, the output of the comparator is continuously low if the control voltage is below the ramp voltage. The PWM generator is achieved by a sawtooth waveform with the peak voltage V_{OSC} = 5 V and the comparator modelled by the operational amplifier, as shown in Fig. 6.

When both the high-side MOSFET Q_1 and low-side MOSFET Q_2 turn on simultaneously, shoot-through happens which means the input voltage will be shorted to the ground. To prevent shoot-through, a delay circuit is required. After one MOSFET turns off, delay circuit will generate a short dead time t_d before another MOSFET

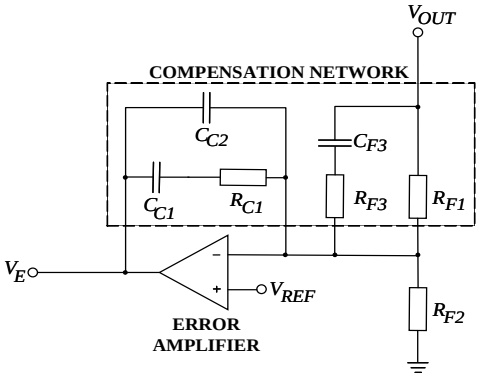


Fig. 5: Type III-B compensator.

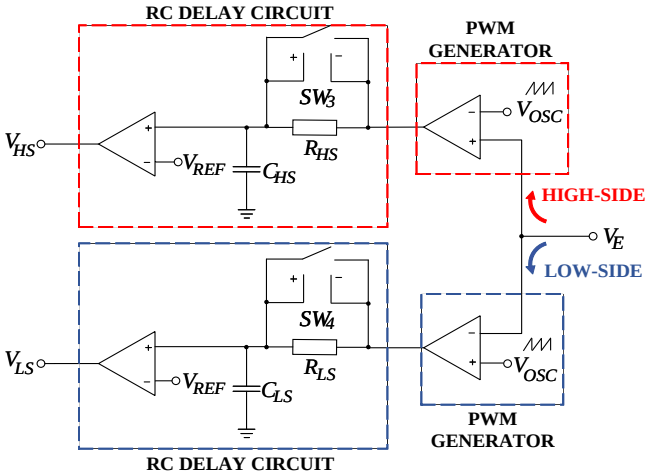


Fig. 6: RC delay circuit with PWM generator.

switches on. The inductor current will be conducted by the body diode of the low-side MOSFET to keep the current continuos during the dead time when both the high-side and low-side are off. Fig. 6 shows the delay circuit. The main component of the delay generator is an RC circuit in parallel with the Schottky diode modelled by the voltage controlled switches SW_3 and SW_4. At the output of the comparator, due to PWM signal's falling edge, the diode conducts and the circuit adds no delay due to very small RC constant. In contrast, during the falling edge of the PWM signal, the Schottky diode is bypassed and the RC circuit converts the PWM signal edge into the ramp signal. Inserting the voltage comparator between the delay circuit and the MOSFET gate resistance, the output ramp is flipped at V_{REF} = 3 V to ensure lagging of the delay circuit output rising edge compare to the PWM signal rising edge. The values of parameters in the RC circuit are shown in Table V.

TABLE V: RC contants for the high-side and low-side MOSFET

Parameter	HS MOSFET	LS MOSFET
R	47 Ω	47 Ω
C	127 pF	150 pF
t_d	5.2 ns	10.1 ns

Different RC constants for the high-side and low-side MOSFET are selected, essentially different dead times t_d (see Fig. 8), to have body diodes conducting about 5 ns.

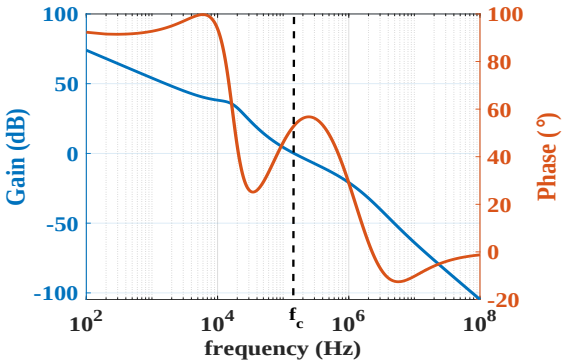


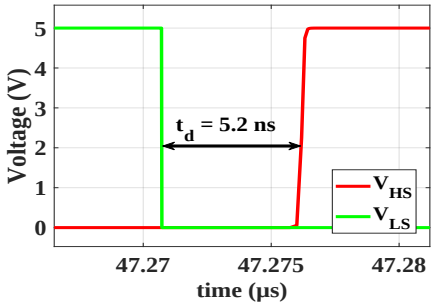
Fig. 7: Simulation of the Bode plot of the loop gain with the slope at crossover frequeny of -20 dB/dec and the phase margin of about 51 °.

Fig. 7 shows the Bode plot of the developed buck converter loop gain.

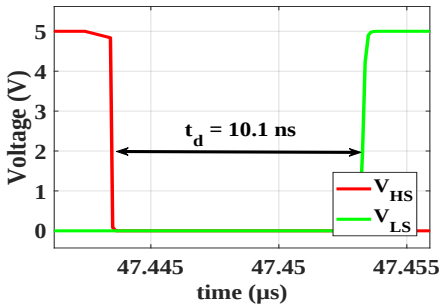
IV. SIMULATION ENVIRONMENT AND RESULTS

In order to verify the efficieny of the designed buck converter, the *Advanced Design System* (ADS) model is created. The simulation model, as shown on Fig. 1, has been operated under full load conditions, continuous current mode and closed loop control. The external 2-Ω gate resistors $R_{G,HS}$ and $R_{G,LS}$ are very important in limiting noise and ringing due to the parasitic capacitances and inductances, high voltage and current change that can cause unwanted performance. Small resistor values result in faster turn-on speed, but can cause the overshoot in the gate voltage. However, higher resistor values overdamp the oscillation and extend the switching time. Also, due to the parasitic elements on the printed circuit boards, the inductor L_{PAR} = 1 nH placed between the power supply and high-side MOSFET Q_1 generates high frequency ringing. For the proper performance of the buck converter, the PWM generator must apply the square wave voltage with the 5-V peak across the gate of MOSFETs with the delay time inserted between the high-side MOSFET Q_1 and low-side MOSFET Q_2 switch over, defined as dead time (see Fig. 8). The results of the model simulations are shown in Fig. 9.

Fig. 9 (a) shows that the output voltage V_{OUT} equals 5.03 V. The voltage spikes with 50 mV peak are caused by equivalent series resistance (ESR) of the output capacitors. As seen in Fig. 9 (b) while the high-side MOSFET is in on-state, the inductance current increases due to constant positive voltage across inductor. In contrast, negative voltage causes current decrease during the on-state of the low-side MOSFET. In addition, inductor current ripple of 1 A is as required. The switch node alternatly connects to the input voltage $V_{IN_{nom}}$ = 14 V or ground, and rings to about 9 V larger voltage than the input voltage, as shown in Fig. 9 (c).



(a)



(b)

Fig. 8: (a) Simulation of the dead time from low-side MOSFET off to high-side MOSFFET on. (b) Simulation of the dead time from high-side MOSFET off to low-side MOSFET on.

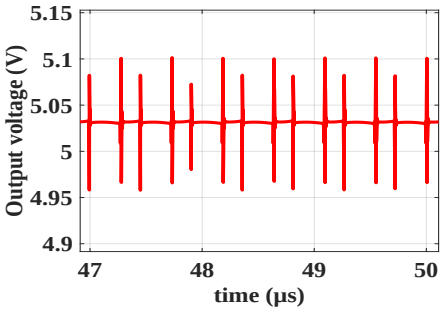
A. Load transient response

The performance of the feedback loop compensation and its stability can be tested by the transient response. The output of the under-compensated converter tends to oscillate, while the output of an over-compensated converter takes a long time to return to a steady state. The load transient response to a 2.5 A load step is shown in Fig. 10. The simulated voltage overshoot is about V_{OS} = 5.063 V.

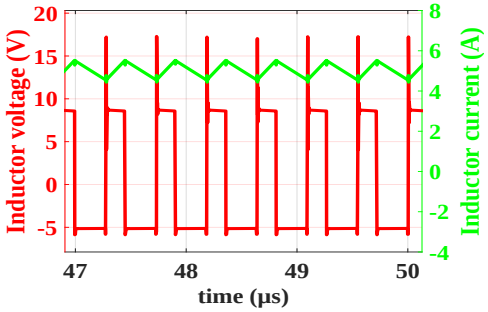
B. Conducted emission

The input port EMI noise comes from voltage ripple generated by the discontinuous current on the input capacitors. The fundamental frequency of the voltage ripple is the switching frequency f_{sw} = 2.2 MHz of the designed buck converter. Higher order harmonics of the fundamental frequency also exist in the noise spectrum. For consistent results, conducted EMI measurments are performed using a passive device called a LISN (Line Impedance Stabilization Network) which establishes a consistent source and measurement impedance allowing for repeatability of test results. The LISN is connected in series with the power supply to the DC-DC switching converter. The topology of 5-µH LISN is shown in Fig. 11.

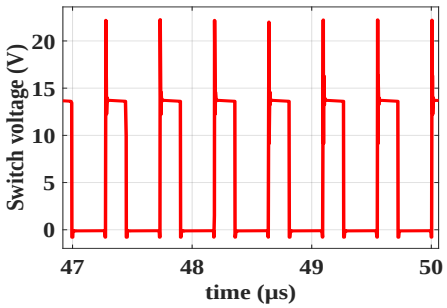
To measure or simulate conducted emissions, two seperate CISPR 25 LISN circuits are connected to the positive and negative power supply lines referenced to the ground plane. In additon, as seen in Fig. 12, the parasitic capacitance C_{PAR} = 100 pF is connected between the negative power supply line and reference ground plane.



(a)



(b)



(c)

Fig. 9: Simulated results. (a) Output voltage V_{OUT}. (b) Inductor current I_L and voltage V_L. (c) Switch voltage V_{SW}.

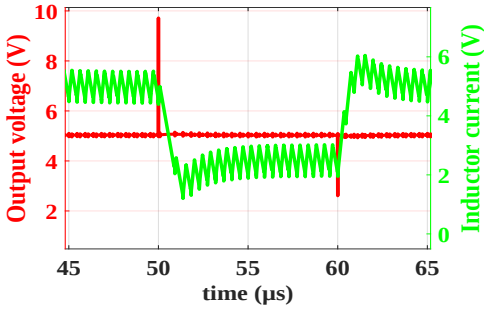
The most commonly used EMI measurement and conformance for automotive application standard is CISPR 25 which specifies limits of conducted EMI noise over a frequency range from 150 kHz to 108 MHz for peak (PK) and average (AVG) signal detectors. Fig. 13 shows the simulated differential-mode and common-mode conducted emissions described as

$$V_{DIFF} = V_{LINE,1} - V_{LINE,2}, \tag{4}$$

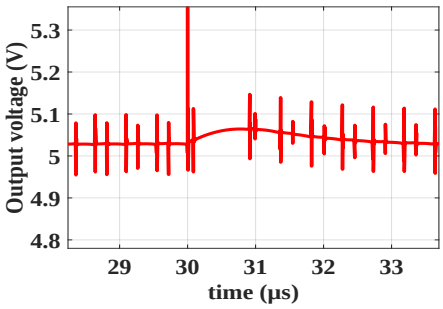
$$V_{COMM} = \frac{V_{LINE,1} + V_{LINE,2}}{2}, \tag{5}$$

where $V_{LINE,1}$ and $V_{LINE,2}$ are the noise voltages on the positive and negative power supply lines.

The most significant noise magnitude of 83.6 dBµV appers at the switching frequency f_{sw} = 2.2 MHz as expected. It is evident that the designed buck converter does not conform to CISPR 25 Class 5 standards. Therefore, in



(a)



(b)

Fig. 10: Simulated results. (a) Load transient response, I_L = 5 A to 2.5 A to 5 A. (b) Output voltage overshoot fits the calculated value.

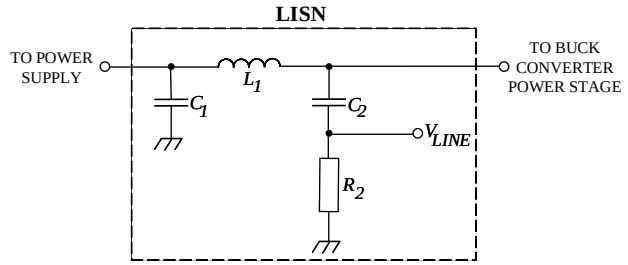


Fig. 11: LISN (Line Impedance Stabilization Network).

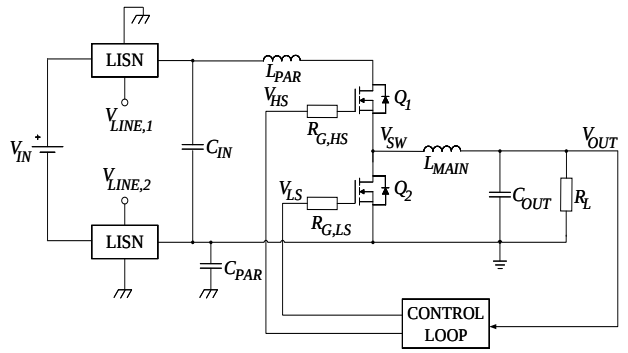


Fig. 12: Synchronous buck converter with voltage-mode control loop and LISN.

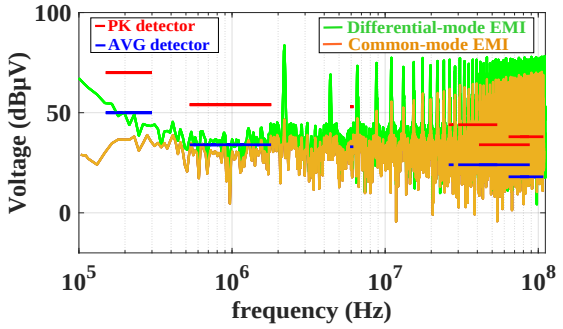


Fig. 13: Simulation of the differential-mode and common-mode conducted emissions.

order to satisfy the standard, the input EMI filter must be included in the design to reduce conducted emissions.

V. Conclusion

The model of a 5-A buck converter used for automotive applications is designed using Advanced Design System (ADS). The developed model consists of the power stage and voltage-mode control loop. The passive components are modelled by the model with parasitic elements using impedance-frequency characteristic. The buck converter performs as expected. Adding differential-mode EMI filter is required in order to reduce conducted emissions.

Acknowledgment

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures (IP-2019-04-8959).

References

[1] R. Nowakovski and N. Tang, “Efficiency of synchronous versus nonsynchronous buck converters,” Analog Applications Journal, 2009.

[2] M. Xie, “How to select input capacitors for a buck converter,” Analog Applications Journal, 2016.

[3] B. Hauke, “Basic calculation of a buck converter’s power stage,” Low Power DC/DC Applications, 2011.

[4] Hegarty, T., "An Engineer’s Guide to Low EMI in DC/DC Regulators," An overview of EMI management and mitigation techniques for DC/DC regulators, 2021.

[5] Barkhordarian, V. and Segundo E., "Power MOSFET Basics," 2021.

[6] "AP64502Q," Document number: DS42750, Diodes Incorporated 2020.

[7] "Coilcraft XEL Family molded power inductor," https://www.coilcraft.com/en-us/products/power/shielded-inductors/molded-inductor/xl/

[8] Pan L.,Su I., "How to Estimate Overshoot and Undershoot of Load Transients of a DC/DC Regulator," Low Power DC/DC Applications, 2017.

[9] "Design Support Software SimSurfing," https://www.murata.com/en-global/tool/simsurfing

[10] Rahimi, A. M., Parto, P. and Asadi, P., "Compensator Design Procedure for Buck Converter with Voltage-Mode Error-Amplifier," SupIRBuckT, 2016.

[11] Zhang, Y., amd Wang, J., "Understanding Dead-time Based On TPS51225/275/285," 2018.

[12] "Panasonic’s aluminum electrolytic capacitor," https://industrial.panasonic.com/ww/products/capacitors/aluminum-capacitors

Implementation of Voltage Regulation in a Spread-Spectrum-Clocked Buck Converter

Jurica Kundrata and Adrijan Barić
University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia

***Abstract*—Buck converters use spread-spectrum technique to reduce the electromagnetic emissions by distributing the electromagnetic interference in a frequency band around the switching frequency which consequently reduces the amplitude of the interference. Changing the switching frequency also changes the output voltage and causes an unwanted output voltage ripple. The output voltage ripple is then reduced by properly controlling the time delay used in the duty cycle controller of the buck converter. This paper presents a digital implementation of the duty cycle controller used in the buck converter which uses a phase-locked loop to generate the switching frequency. Changing the divider value of the phase-locked loop changes its output frequency, but this frequency change is not instantaneous and it is modelled as a second-order system response. The implementation of the duty cycle controller presented and analysed in this paper aims to compensate for this transient change of the output frequency, i.e. the switching frequency and keep the duty cycle of the buck converter constant. The paper focuses on the compensation algorithm used to generate a constant duty cycle, as well as on the challenges of its digital implementation.**

***Index Terms*—DC-DC converter, spread-spectrum, duty cycle compensation, on-time compensation, physical implementation**

I. Introduction

Recent years have seen an increase in the switching frequency of switch-mode power DC-DC converters, which has led to the emergence of fully-integrated converters. These converters offer various advantages over discrete component designs, including simpler use, reduced cost, and improved reliability [1]. Efficiency is a crucial characteristic of power converters, and the utilization of resonant soft-switching techniques can reduce switching losses and maintain high efficiency even at switching frequencies of tens of MHz [2], [3], [4].

One of the significant challenges in designing high-frequency DC-DC converters is the increase in electromagnetic emissions resulting from the increased switching frequency. The generated electromagnetic emissions must not exceed the specified limits in regulations [5]. Regulatory tests for electromagnetic compatibility (EMC) have shown that even low-power devices with frequencies exceeding 160 MHz have failed such tests [6] [7].

Integrated converters offer another advantage in the form of advanced management techniques that can be used at a minimal cost, primarily related to silicon area. One such technique is the spread-spectrum modulator, which reduces electromagnetic emissions by distributing its energy to the frequency band around it. The magnitudes of the spread-spectrum frequency components are reduced compared to the magnitude of the original signal, although the total energy of the signal during modulation is not altered.

The spread-spectrum technique has become increasingly popular as an efficient means of reducing electromagnetic emissions [8]. Various switching devices already take advantage of this technique including SATA controllers [9], audio amplifiers [10], [11] and power converters [8], [12]. However, despite being widely used in commercial products, it has been demonstrated in [13] that the spread-spectrum technique's parameters are not always optimized for specific applications.

The use of spread-spectrum techniques is typically a trade-off, while it improves electromagnetic compatibility (EMC), it may worsen other device characteristics such as efficiency, output voltage ripple, etc. Moreover, a spread-spectrum technique effective in reducing conducted EM emissions may not be equally effective in reducing radiated EM emissions, and vice versa. However, a spread-spectrum modulation technique called Dual Random Spread Spectrum (DRSS) was recently introduced which is effective in reducing both conducted and radiated electromagnetic interference (EMI) [14]. Commercially available converters typically use fixed spread-spectrum modulation, with user-configurable parameters being unavailable. The introduction of a configurable spread-spectrum modulator would enable a better understanding of these trade-offs and enable optimization of EM emission reduction without compromising device performance. Although, there are commercially available devices in which the output ripple introduced by the spread spectrum is greatly reduced by using the feed-forward technique [15].

Application of the spread-spectrum technique in DC-DC converters leads to e.g. a substantial output voltage ripple as the duty cycle of the output voltage changes with the changing switching frequency. In [16] the output voltage ripple, as well as several compensation techniques are analysed.

This paper presents an integrated spread-spectrum controller design which uses a first-order approach to compensate the transient change of the output oscillator frequency used in a buck converter thus reducing its output voltage

This work is supported by the Croatian Science Foundation (HRZZ) within the project Fast switching converters based on GaN devices and resonant architectures FASCON (IP-2019-04-8959).

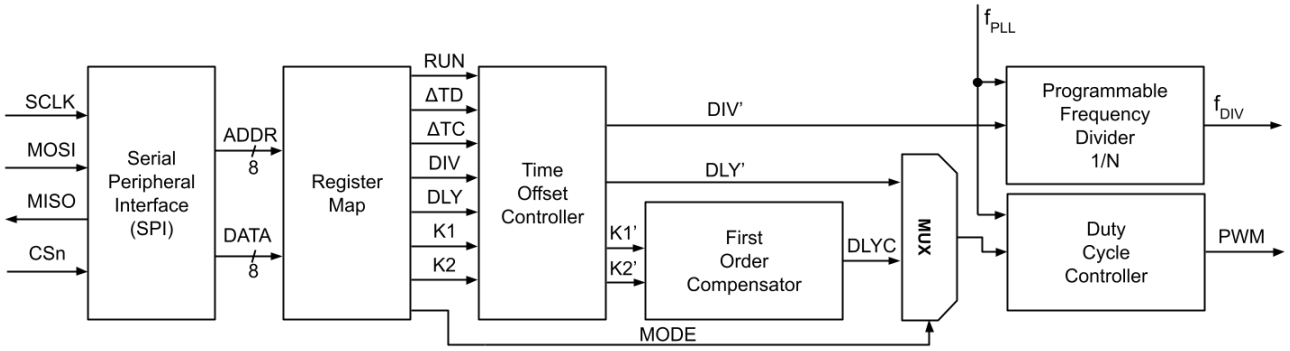


Fig. 1: The block diagram of the spread-spectrum controller including the first-order compensator.

ripple.

Section II presents the top level design of the spread-spectrum controller. Section III describes the design of the first-order compensator including its data flow. The spread spectrum controller is functionally verified by simulations in Section IV. Section V presents the physical implementation of the spread-spectrum controller followed by conclusions in Section VI.

II. Spread-spectrum controller

Fig. 1 shows the block diagram of the spread-spectrum controller including the first-order compensator. The spread-spectrum controller consists of the duty-cycle controller and the programmable frequency divider. The programmable frequency divider is connected in the feedback loop of an external Phase-Locked Loop (PLL) oscillator. The output frequency of the divider f_{DIV} represents the output frequency of the PLL oscillator f_{PLL} divided by a constant N and serves as the feedback signal in the PLL oscillator. The output frequency of the PLL oscillator is fed to the divider, but also to the duty-cycle controller which outputs a Pulse-Width Modulated (PWM) signal with a given value of the time-delay, i.e. the duty cycle. The output signal of the duty-cycle controller is used as the switching signal of the buck converter.

This spread-spectrum controller is designed to be used together with an external PLL oscillator and combining them with an output stage constitutes a complete buck converter controller.

The spread-spectrum controller is parametrized using the Serial Peripheral Interface (SPI). The communication interface itself consists of the clocking signal SCLK, chip select signal CSn and data-carrying signals MOSI and MISO. The spread-spectrum controller features an SPI slave which translates the SPI two-byte transfers to accesses to the register map. The individual SPI transfer consists of the address byte ADDR and the data byte DATA which are then used to configure the register map.

The parameters determined by the register map are fed to the time offset controller, first-order compensator, programmable frequency divider and duty cycle controller. The time offset controller uses timers which are loaded with the divider time offset ΔTD and compensator time offset ΔTC values. These values determine the time delays applied to the divider DIV, constant on-time delay DLY and compensator coefficients K1 and K2. The delayed parameters are respectively DIV', DLY', K1' and K2'. The time offset controller allows for fine tuning the start of the on-time compensation with respect to the PLL oscillators transient behaviour. The register map command RUN is used to start the time offset controller, while the command MODE is used to select the constant or compensator-generated DLYC on-time using the multiplexer shown in Fig. 1.

Table I shows the mapping of the register bank used in the spread-spectrum controller. The register map consists of 12 individual addresses where the first few addresses represent individual commands or values, while the rest of the addresses represent in pairs the high and low bytes of the parameters of the spread-spectrum controller. The address 0x00 represents the command RUN which enables/disables the operation of the spread-spectrum controller. The address 0x01 represents the mode selection command MODE which is used to select the operation mode of the spread-spectrum controller. There are two separate operation modes: the constant delay mode and the mode which uses the time delay values generated by the first-order compensator DLYC. The address 0x02 represents the divider value DIV used in the frequency divider of the PLL oscillator. The following addresses represent the lower and higher bytes of the constant time delay value DLY, first-order compensator constants K_1 and K_2 and accompanying time offset values for the divider ΔTD and first-order compensator ΔTC.

III. First-order compensator

This paper uses an exponential compensation technique of the on-time t_{ON} which uses a straightforward difference equation defined as

$$t_{ON}[n] = t_{ON}[n-1] + \frac{T_{CLK}}{\tau_C}(t_{ON,trgt} - t_{ON}[n-1]) \quad (1)$$

where $t_{ON}[n]$ and $t_{ON}[n-1]$ are the compensated on-time values at two consecutive calculation steps, T_{CLK} is the

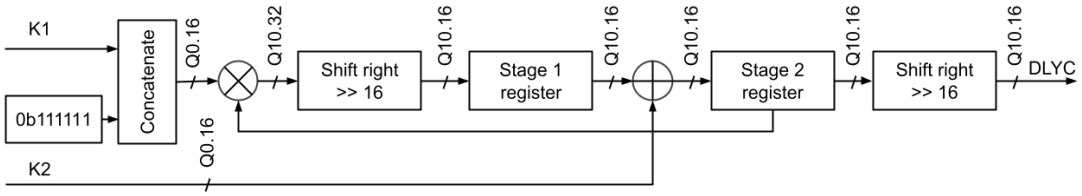


Fig. 2: The data flow of the first-order compensator.

time period of the controller clock, τ_C is the time constant of the exponential compensation technique and $t_{ON,trgt}$ is the targeted, final value of the compensated on-time value. The difference equation includes three arithmetic operation, i.e. subtraction, multiplication and addition. The difference equation can be reorganized in order to reduce the number of necessary operations and consequently simplify the digital implementation of the first-compensator. The simplification of the difference equation begins by expanding it as follows

$$t_{ON}[n] = t_{ON}[n-1] + \frac{T_{CLK}}{\tau_C} t_{ON,trgt} - \frac{T_{CLK}}{\tau_C} t_{ON}[n-1] \tag{2}$$

which leads to

$$t_{ON}[n] = t_{ON}[n-1]\left(1 - \frac{T_{CLK}}{\tau_C}\right) + \frac{T_{CLK}}{\tau_C} t_{ON,trgt} \tag{3}$$

and finally

$$t_{ON}[n] = K_1 \cdot t_{ON}[n-1] + K_2 \tag{4}$$

where

$$K_1 = 1 - \frac{T_{CLK}}{\tau_C} \text{ and } K_2 = \frac{T_{CLK}}{\tau_C} t_{ON,trgt} \tag{5}$$

where $t_{ON,trgt}$ is defined as

$$t_{ON,trgt} = \frac{D}{f_{SW}} \tag{6}$$

TABLE I: THE REGISTER MAPPING.

Address	Register Name	Register Description
0x00	RUN	Run command
0x01	MODE	Operation mode
0x02	DIV	Divider value
0x03	DLY_L	Constant delay, L byte
0x04	DLY_H	Constant delay, H byte
0x05	K1_L	Compensator constant K_1, L byte
0x06	K1_H	Compensator constant K_1, H byte
0x07	K2_L	Compensator constant K_2, L byte
0x08	K2_H	Compensator constant K_2, H byte
0x09	ΔTD_L	Divider time offset, L byte
0x0A	ΔTD_H	Divider time offset, H byte
0x0B	ΔTC_L	Compensator time offset, L byte
0x0C	ΔTC_H	Compensator time offset, H byte

where D is the targeted duty cycle of the buck converter output and f_{SW} is the switching frequency of the buck converter, i.e. the output frequency of the PLL oscillator f_{PLL}.

Fig. 2 shows the data flow of the first-order compensator. The data flow implements the sequence of the operations required to calculate the next time delay value as defined in relation (4). Furthermore, the two operations, i.e. the multiplication and addition are separated each in its own stage which constitutes a pipeline and facilitates the time closure of the digital design during synthesis.

The first-order compensator is designed to use a fixed-point fractional number format. The width of the output time delay value is determined by the size of the control word of the duty cycle controller, while the widths of the input constants K1 and K2 are chosen with respect to the value changes with the used time steps. Both constants K1 and K2 are less than 1 and consequently the Q0.16 number format is used. Additionally, constant K1 is very close to 1 as it will be shown in Section IV and the upper 6 bits of constant K1 are fixed. This is represented by the concatenation block in Fig. 2. Fixing the upper bits of the constant K1 also simplifies the synthesis of the following multiplication block.

The generated time delay value is internally kept in the Q10.16 format in order to preserve the accuracy of the calculation and to accommodate the small value changes due to very small ratio of the clock period and time constant, also discussed in Section IV. Consequently the output of the multiplication block is Q10.32 which is rounded to the Q10.16 format using the following shift right block. The value of the multiplication is kept in the Stage 1 register. The value of the Stage 1 register is then used in the addition with constant K2 which results in the internal value of the time delay and it is kept in the Stage 2 register. The output time delay value DLYC is finally generated by rounding the Stage 2 register value using the following shift right block.

IV. Digital simulation

The digital core which implements the spread-spectrum controller is functionally verified by digital simulation. The digital simulation is based on a testbench which communicates with the spread-spectrum controller via the SPI communication interface and sets the parameters of the first-order compensator for the testcase where the output PLL frequency starts at the mean value, then reduces to

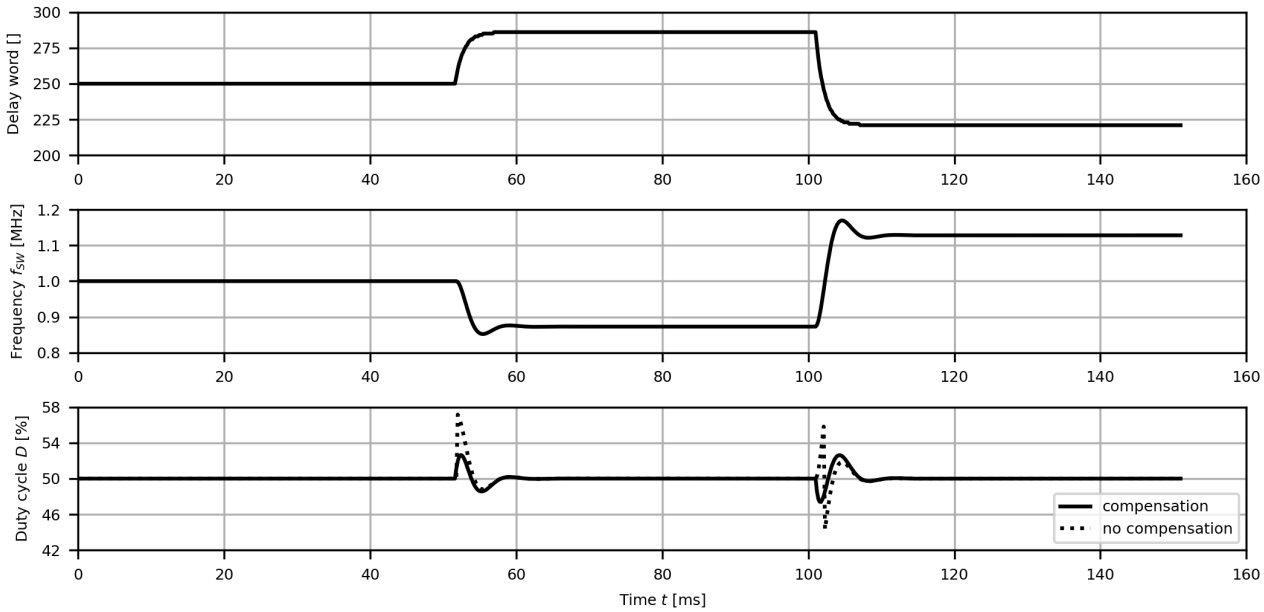


Fig. 3: The waveforms of the simulated compensator-generated delay word, switching frequency and resulting duty cycle.

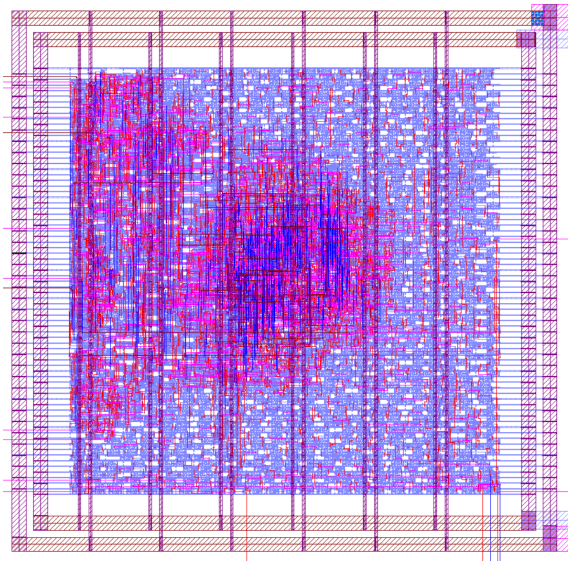


Fig. 4: The layout of the spread-spectrum controller.

the minimum value and finally increases to the maximum value.

Table II shows the parameters of the first-order compensator used in the testbench simulations. The testbench simulation uses three distinct PLL frequencies to illustrate the functioning of the first-order compensator. The simulation starts at the nominal frequency f_{PLL} = 1 MHz, in the next part of the simulation it drops to the minimum value f_{PLL} = 0.873 MHz and finally it steps up to the maximum value f_{PLL} = 1.128 MHz. The target delay times $t_{ON,trgt}$ shown in Table II are calculated using (6). The delay words DLY are the control words used in the digital spread spectrum controller to represent the given delay time $t_{ON,trgt}$ and they are calculated using

$$DLY = \lfloor \frac{t_{ON,trgt}}{\Delta t} \rfloor \tag{7}$$

where the resolution of the duty cycle controller is Δt = 2 ns which is defined by the used delay standard cell.

The parameters of the first-order compensator in Table II are calculated using (5) where the clock frequency, i.e. its period is T_{CLK} = 1 ns and the time constant τ_C = 1.25 ms. Furthermore, these parameters are presented in their fixed-point fractional number formats as used in the first-order compensator.

Fig. 3 shows the waveforms of the simulated compensator-generated delay word, switching frequency and resulting duty cycle. The upper waveform represents the delay words generated by the first-order compensator. The middle waveform shows the switching frequency generated by an example PLL oscillator which features the characteristic second-order transient change. The lower waveform shows the resulting duty cycle of the output PWM signal for the cases without (command MODE = 0) and with (command MODE = 1) first-order compensation. The resulting duty cycle is calculated using

TABLE II: THE COMPENSATOR PARAMETERS.

Frequency f_{SW} [MHz]	**0.873**	**1.0**	**1.128**
Delay time $t_{ON,trgt}$ [ns]	572	500	444
Delay word DLY []	286	250	222
Parameter $K1$ []	0.9992	0.9992	0.9992
Parameter $K2$ []	0.2288	0.2000	0.1776
Parameter $K1$ [Q0.16]	0xFFCB	0xFFCB	0xFFCB
Parameter $K2$ [Q0.16]	0x3A92	0x3333	0x2D77

TABLE III: THE AREA UTILIZATION OF THE SPREAD-SPECTRUM CONTROLLER.

Type	Instances	Area [μm^2]	Area [%]
Sequential	290	15950.323	22.4
Inverter	94	623.437	0.9
Buffer	1025	38222.823	53.8
Logic	1036	16290.579	22.9
Total	2445	71087.162	100.0

TABLE IV: THE AREA UTILIZATION OF THE FIRST-ORDER COMPENSATOR.

Type	Instances	Area [μm^2]	Area [%]
Sequential	52	2638.630	17.2
Inverter	81	533.434	3.5
Buffer	3	26.342	0.2
Logic	734	12102.138	79.1
Total	870	15300.544	100.0

$$D = DLY \cdot \Delta t \cdot f_{SW}. \tag{8}$$

The comparison of the output duty cycle in the cases with and without the first-order compensation in Fig. 3 shows that the duty cycle error is approx. $\pm 6\%$ when no compensation is used and it is approx. $\pm 2.5\%$ when the first-order compensation is used.

V. Physical Implementation

The spread-spectrum controller with the SPI communication interface and first-order compensator as shown in Fig. 1 is physically implemented in the 180-nm TSMC technology. Fig. 4 shows the layout of the physical implementation of the spread-spectrum controller. The size of the digital core including the power ring is approx. 400 by 400 μm^2. Most of the area in the digital core is taken up by the duty cycle controller and the delay cells it uses. The purple-coloured area in the layout corresponds to the area taken by the delay cells, while the red-coloured area corresponds to the rest of the logic of the digital core.

Table III shows the area utilization of the physical implementation of the spread-spectrum controller. Table shows that the digital core has a large number of buffer cells which take up more than half of the total cell instances and area. These buffer cells are mostly delay cells used by the duty cycle controller as observed earlier in Fig. 4. Rest of the cells used by the digital core are approximately evenly distributed between the sequential and logic standard cells and they implement the rest of the digital blocks shown in Fig. 1.

Table IV shows the area utilization of the physical implementation of the first-order compensator alone. Table shows that the first-order compensator constitutes approx. a third of the standard cell instances and approx. a fifth of the utilized area of the whole spread-spectrum controller. The difference in these ratios is due to the fact that the delay cell used in the duty cycle controller has a significantly larger area than the rest of the standard cells. The number of the sequential standard cells correspond to the Stage 1 and Stage 2 registers which hold the intermediary values in Q10.16 format. The logic standard cells implement the arithmetic operations of multiplication and addition in each stage. The physical implementation of this digital core estimates a power consumption of 42 μW.

VI. Conclusions

Spread-spectrum technique is used in buck converters to reduce electromagnetic emissions. However, changing the switching frequency to implement this technique causes unwanted output voltage ripple, which can be reduced by controlling the time delay used in the duty cycle controller. The paper presents a digital implementation of the spread-spectrum controller, which includes a first-order compensator of the on-time used in the duty cycle controller. The implementation aims to compensate for the transient change of the output frequency and keep the duty cycle of the buck converter constant. The first-order compensator uses a difference equation to generate new on-time values which is implemented in a pipeline with reduced number of arithmetic operations with a fixed-point fractional number representation. The spread-spectrum controller design is functionally verified by simulations which show that the compensator can reduce the duty cycle error by more than half. The physical implementation results show that the addition of the first-order compensator to the spread-spectrum controller increases the area of the implemented core by a quarter, while providing the reduction of the output voltage ripple of the buck converter.

References

[1] Baoxing Chen, "Fully integrated isolated dc-dc converter using micro-transformers," 2008 TwentyThird Annual IEEE Applied Power Electronics Conference and Exposition, Austin, TX, 2008, pp. 335- 338.

[2] R. Erickson and D. Maksimovic, Fundamentals of Power Electronics, Springer, 2001.

[3] Marian K Kazimierczuk, Dariusz Czarkowski, Resonant power converters, Wiley, 2011.

[4] N. Bertoni, G. Frattini, R. G. Massolini, F. Pareschi, R. Rovatti and G. Setti, "An Analytical Approach for the Design of Class-E Resonant DC–DC Converters," in IEEE Transactions on Power Electronics, vol. 31, no. 11, pp. 7701-7713, Nov. 2016.

[5] C. Paul, Introduction to Electromagnetic Compatibility, Wiley, 2006.

[6] Z. Tianting, Z. Yue and C. Baoxing, "An isolated DC-DC converter with fully integrated magnetic core transformer," 2017 IEEE Custom Integrated Circuits Conference (CICC), Austin, TX, 2017, pp. 1-4.

[7] Analog Devices, Application note AN-0971, 2014.

[8] F. Pareschi, R. Rovatti and G. Setti, "EMI Reduction via Spread Spectrum in DC/DC Converters: State of the Art, Optimization, and Tradeoffs," in IEEE Access, vol. 3, pp. 2857-2874, 2015.

[9] F. Pareschi, G. Setti and R. Rovatti, "A 3-GHz Serial ATA Spread-Spectrum Clock Generator Employing a Chaotic PAM Modulation," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 57, no. 10, pp. 2577-2587, Oct. 2010.

[10] B. Deutschmann, G. Winkler and T. Karaca, "Emission reduction in Class D audio amplifiers by optimizing spread spectrum modulation," 2015 10th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo), Edinburgh, 2015, pp. 1-6.

[11] X. Ming, Z. Chen, Z. Zhou and B. Zhang, "An Advanced Spread Spectrum Architecture Using Pseudorandom Modulation to Improve EMI in Class D Amplifier," in IEEE Transactions on Power Electronics, vol. 26, no. 2, pp. 638-646, Feb. 2011.

[12] V. Subotskaya, K. Cherniak, K. Hoermaier and B. Deutschmann, "Emission reduction with spread spectrum clocking for switched capacitor buck converter," 2018 IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility (EMC/APEMC), Singapore, 2018, pp. 297-302

[13] F. Pareschi, G. Setti, R. Rovatti and G. Frattini, "Practical Optimization of EMI Reduction in Spread Spectrum Clock Generators With Application to Switching DC/DC Converters," in IEEE Transactions on Power Electronics, vol. 29, no. 9, pp. 4646-4657, Sept. 2014.

[14] P. Curtis and E. Lee, "EMI Reduction Technique, Dual Random Spread Spectrum (Application Note)," Texas Instruments, 2020.

[15] S. Jaffe, "The pros and cons of spread-spectrum implementation methods in buck regulators," Analog Design Journal, Texas Instruments, 2021.

[16] J. Kundrata, I. Skeledžija and A. Barić, "Voltage Regulation in an Integrated Controller of a Spread-Spectrum-Clocked Buck Converter," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 185-190, doi: 10.23919/MIPRO55190.2022.9803464.

Uninterruptible Power Supply System for Railway Infrastructure with Power Factor Correction and Immunity to Energy Strike

Ante Hećimović*, Leo Patrlj*, Viktor Šunde* and Željko Ban*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
ante.hecimovic@fer.hr, leo.patrlj@fer.hr, viktor.sunde@fer.hr, zeljko.ban@fer.hr

***Abstract*— This article deals with an uninterruptible power supply (UPS) for infrastructure equipment used in the railway system. The main feature of this UPS is the high power factor and resistance to overvoltages that occur in the railway network. Power factor correction (PFC) is done by adding a boost DC converter. The resistance to overvoltages is achieved by means of protective elements, i.e. varistors and inductors, as well as an additional overvoltage detection in the power supply network. Based on the output of the mentioned device, the part of the topology intended for power factor correction can be switched off.**

Keywords – Uninterrupted Power Supply (UPS), Power Factor Correction (PFC), overvoltage protection, railway power supply system

I. INTRODUCTION

The importance of rail transport is extremely significant, and rail transport based on electric traction is particularly important. The traction vehicles are mainly supplied by a 25 kV, 50 Hz catenary [1], while the signal, safety, and security systems are supplied by the same line through a transformer and a uninterrupted power supply (UPS) with direct and/or alternating voltage (230 V, 50 Hz; 3 x 400/230 V, 50 Hz; 3 x 380 V, 83,33 Hz; 12 V_{dc}; 24 V_{dc}; 110 V_{dc}) [2]. The UPS usually consists of a rectifier in the input stage, a DC / DC link consisting of capacitors to reduce voltage ripple, and a DC / DC or DC / AC converter at the output [6]. In addition to noise immunity, an important requirement for UPS systems used to supply infrastructure equipment is a built-in power factor correction (PFC) in the input stage, i.e., compliance with power quality regulations such as IEC 61000-3-2 and IEEE 519 [3], [4]. Its main task is to reduce total harmonic distortion (THD) and electromagnetic interference (EMI).

Although the voltages of the power supply grid are strictly standardized, there are sections of the grid where high-voltage disturbances occur when a train passes through, causing damage to low voltage equipment supplied from the same grid. Disturbances present in the power supply network of the railway system, as shown in Figure 1, caused by the passing thyristor locomotive, disrupt proper operation and may even destroy the grid-connected UPS -es [2][5].

All the costs of publishing of this paper are co-financed by the project KK.01.2.1.02.0193- "MARETON II" under the Competitiveness and Cohesion Operational Program from the European Regional Development Fund.

This article analyzes a UPS that has a high degree of immunity to interference. The effectiveness of protective components against overvoltage is considered. In addition, the UPS structure is upgraded with a circuit for power factor and THD correction [7], as well as with a sensor for overvoltage detection.

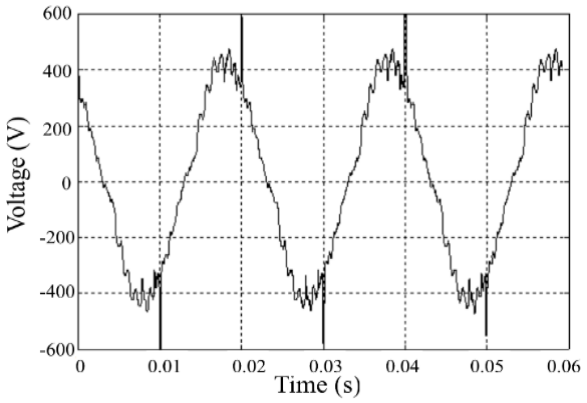


Figure 1: Input voltage with harmonics and overvoltage spikes

After the introduction, the 2nd chapter of the article discusses the structure of the UPS with varistors and inductors as protective elements. The chapter describes the model and parameters of the UPS, as well as its operation. In Chapter III, a simulation model was created, and simulation results of the output voltage caused by the input voltage with multiple higher harmonics and overvoltages are presented. The model of UPS with the voltage disturbance sensor and the PFC circuit is analyzed in Chapter IV. The simulation results show the behavior of the system and the improvement caused by the mentioned circuits. Concluding remarks are made in Chapter V, and then the list of references follows.

II. UPS STRUCTURE

The structure of the considered UPS is shown in Figure 2. The UPS is fed from the low-voltage side of the transformer, whose high-voltage side is connected to the power supply network of the railroad network (25 kV). This implies that all disturbances that occur on the high-voltage side of the transformer are reflected to the low-voltage side via an inductive link. These disturbances include higher harmonics, overvoltages and voltage spikes.

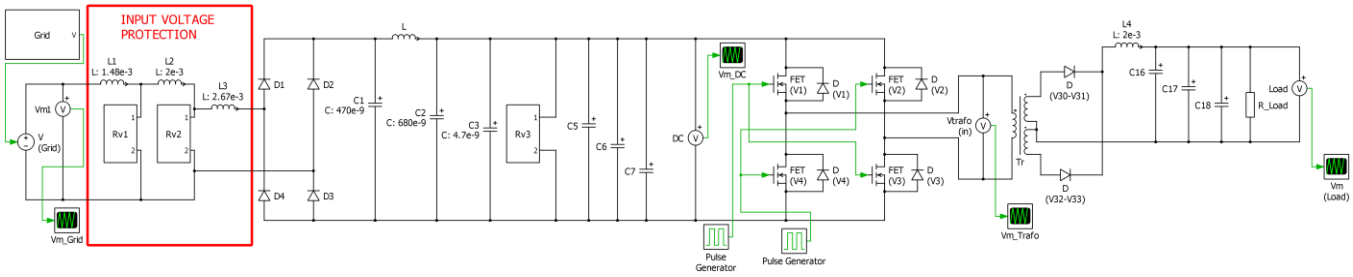


Figure 2: Model of UPS structure

A. *AC distribution and rectifier stage*

To reduce the harmful effects of overvoltages on the operation of UPS in the input part of the AC distribution, a protection consisting of inductors and varistors (L1, L2, Rv1, Rv2) is used. The varistors are from the manufacturer Epcos, types: VDR B40K320 and VDR B40K275. An example of such a varistor model and its response is shown in Figure 3. Several rectifier stages connected in parallel are connected to this input section of UPS (one such stage is shown in Figure 2). The rectifier stage was implemented with a single-phase diode bridge circuit. Between each rectifier stage and the input AC distribution there is an inductance which serves to reduce the charging current of the capacitors forming the DC link in case of an overvoltage. The inductance is L3=2.67 mH. The DC intermediate circuit consists of three capacitors connected in parallel, each with a capacitance of 2·470 μF. This part of the UPS also contains components that serve as EMI filter and another Epcos varistor, model S20K300 (Rv3). The EMI filter is implemented by connecting capacitors and inductor (C1, C2, C3, L), where the inductor L is essentially common mode inductor.

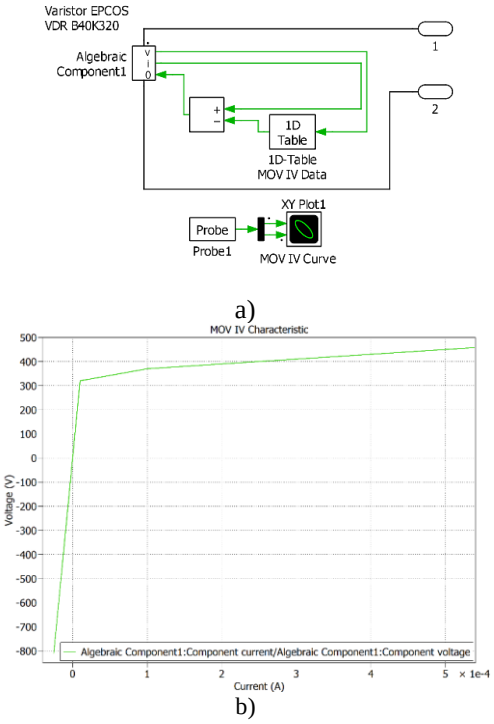


Figure 3: Varistor: a) model, b) *UI* characteristic

B. *Full-bridge DC/DC converter*

The full bridge DC /DC converter is connected after the DC link, (C5, C6, C7). The task of this stage of the UPS is to control the level of the output DC voltage and the galvanic isolation. On the primary side of the high-frequency transformer is a single-phase bridge circuit consisting of hybrid current-bidirectional switches. The switches are implemented by MOSFET and body diode. On the secondary (low voltage) side of the transformer, there is a single-phase diode rectifier with transformer midpoint connection. The additional low-pass filter is connected between the output of the rectifier and the load (battery). The switches on the primary side of the transformer are controlled to produce the three-level voltage.

III. UPS WITHOUT POWER FACTOR CORRECTION

A. *Model and model parameters*

The UPS model described in this chapter is based on a well-known commercial solution, (Figure 2), which will be extended in later stages of the article to include the input voltage disturbance sensor and PFC with the corresponding control algorithm. At this stage of the article, the model doesn't include any sensor and neither additional protection in this regard. The most relevant parameters of this model are listed in Table 1.

TABLE 1: PARAMETERS OF THE EXISTING MODEL

Macromodel / Function	**Element**	**Parameter**	**Value**
AC distribution and input protection circuit	L1	Inductance	1.48 mH
	L2	Inductance	2 mH
	Rv1	U-I char.	do 320 V
	Rv2	U-I char.	do 275 V
	Rv3	U-I char.	do 300 V
Rectifier	D1 - D4	Fwd. voltage	1 V
		R_{on}	10 mΩ
Input EMI filter	C1	Capacitance	470 nF
	C2	Capacitance	680 nF
	C3	Capacitance	4,7 nF
	L	Inductance	1.13 mH
Intermediate DC circuit	C5 - C7	Capacitance	2x470 μF
Bridge connection	FET	R_{on}	10 mΩ
	D	Fwd. voltage	0.7 V
		R_{on}	1 mΩ

The appearance of higher harmonics, overvoltages and other disturbances in the railway transport system is well

known. The power supply *V (Grid)* at the input UPS is directly connected to the block/macro model Grid, where the actual model of the grid is located. The macro model consists of a combination of base sine wave $230\sqrt{2}\,V, 50\,Hz$, superimposed by odd harmonics (up to and including the 17th harmonic), overvoltage of about 140 V, and low signal noise with random characteristics. By superimposing these disturbances on the network voltage with a sinusoidal waveform, the actual waveform of the voltage on the dedicated line of the double-track railway is simulated, Figure 4.

In order to preserve the proper operation of the existing UPS and protect the output voltage of the effects of the higher harmonics and overvoltage on the voltage waveform, two inductors (L1 and L2) and two varistors (Rv1 and Rv2) are used in the AC distribution and protection part. The main task of this protection is to try to eliminate (over)voltage spikes and, if possible, higher harmonics with significant amplitudes. In practice, other solutions such as surge arresters are also used, but recent research suggests that solutions from the field of power electronics may be helpful [8]. Between the AC distribution and protection on one side and the rectifier bridge on the other, inductor L3 is placed. Its function is to further reduce the voltage rise. The effect of these protection elements was tested in further simulations.

B. Simulation experiment

The objective of these simulation experiments is to confirm the effectiveness of the previously described protection elements with the input voltage shown in Figure 4. The disturbances present in the input voltage in the form of higher harmonics and overvoltages are active in the time span of 300 ms (between 600 - 900 ms). This form of input voltage occurs on the power line of one track of a double-track railway at the moment when the locomotive passes by the traction power substation. UPS operation with and without protection is simulated and the obtained results are compared. The waveforms to be considered are: (i) voltage stress of the diodes in the bridge circuit (reverse voltage), (ii) voltage of the DC link, (iii) voltage stress on the transistors in the bridge circuit, (iv) voltage at the transformer, and (v) load voltage.

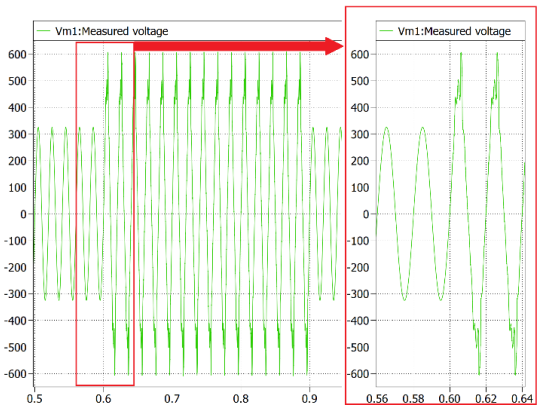


Figure 4: Grid voltage with harmonics and overvoltage spikes

A comparison of these waveforms with and without protection is shown in the following figures. The effect of the protection can be seen in the reverse voltage of the diode in the rectifier bridge circuit, Figure 5. With the input protection, the voltage load is significantly lower (510 V on average compared to 556 V without protection).

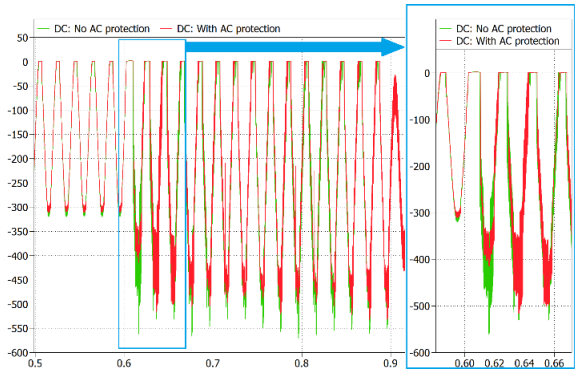


Figure 5: Reverse voltage of the diode in the rectifier bridge with and without the input protection

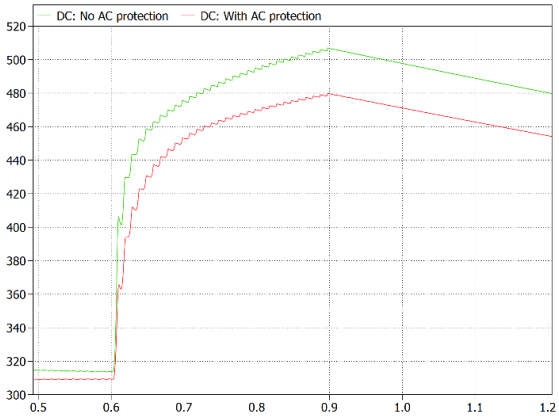


Figure 6: DC intermediate circuit voltage with and without the input protection

The results of the simulation of the DC link voltage with and without protection included are presented and shown in Figure 6. In the steady state, which occurs within the first 600 ms, a sinusoidal voltage appears at the input, after which higher harmonics, overvoltages and voltage spikes occur. The disturbances in the grid voltage cause significant increase of the DC link voltage. After 900 ms, the voltage of the DC link slowly decreases. When the protection is not used, the voltage of the DC link increases to 482 V, which is higher than the voltage of 461 V when the protection is used. After the locomotive passes TPSS, the influence of the higher harmonics slowly decreases, but the return to the stable voltage with the protection not used requires a longer time (1500 ms).

Similar to the previous considerations, the effect of voltage disturbances, but also the effect of protection, can be seen in the voltage waveforms of transistors with body diodes connected in a bridge circuit, Figure 7.

Figure 8 shows the voltage waveforms on the primary side (high voltage side) of the DC /DC converter transformer. When the protection is used in the presence of higher harmonics and voltage spikes in the input voltage, the change in the DC link voltage is minor (the maximum voltage here reaches 482 V).

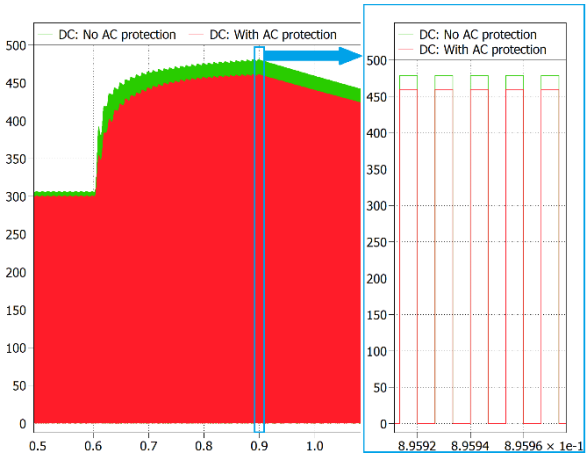


Figure 7: MOSFET voltage with and without the input protection

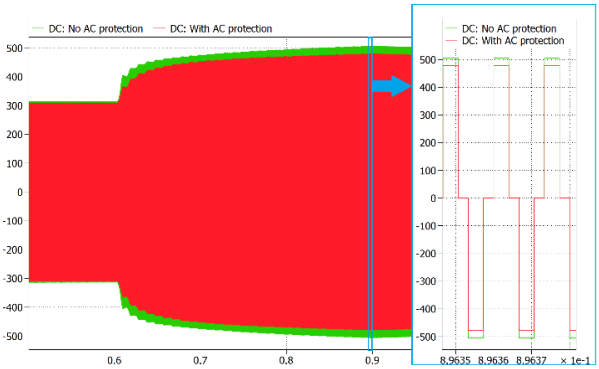


Figure 8: Transformer primary voltage with and without the input protection

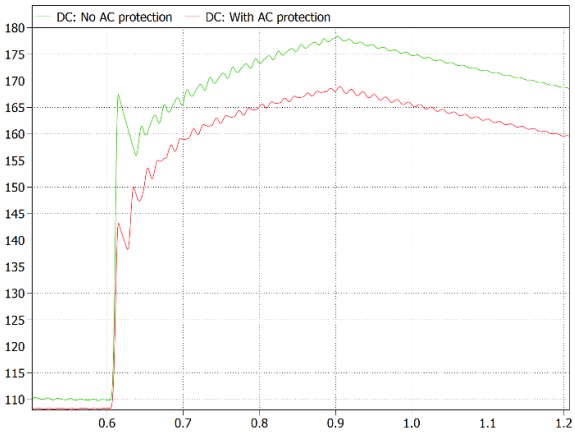


Figure 9: Load voltage with and without the input protection

Figure 9 shows voltage waveforms on the load side. The analysis of the simulation results presented earlier leads to the conclusion that the protection consisting of inductors and varistors in the input circuit of the UPS (described in chapter III, table 1) educes the voltage stress on the components of the UPS caused by overvoltages and higher harmonics occurring in the input voltage due to the passage of the thyristor locomotive. However, the question arises as to how effective this protection is when a PFC subsystem is included in the UPS in form of a boost DC /DC converter. Presence of overvoltage, even though they're reduced by the input voltage protection, can still damage semiconductor components, the transformer, the DC/DC link, or other circuitry or equipment connected to this converter, especially the boost converter used for PFC. For this reason, a sensor is introduced to detect such disturbances and shuts down part of the circuit if necessary.

IV. UPS WITH POWER FACTOR CORRECTION

A. *Model of PFC*

Power factor correction is achieved by introducing a DC/DC boost converter. The structure and arrangement of the boost converter in the overall system can be seen in Figure 11.

To control both current and voltage simultaneously, two control loops are used. The inner control loop is the current control loop and the outer one is the voltage control loop, as shown in Figure 10.

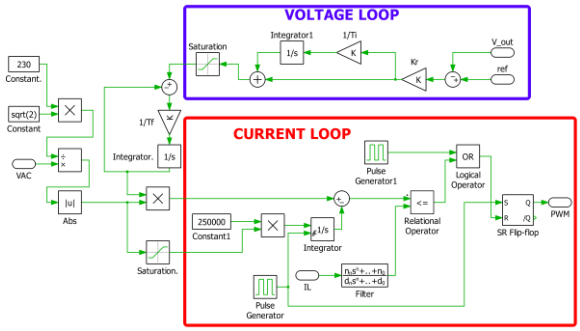


Figure 10: The boost DC/DC converter control schematic

The usual approach to control in PFC is cascade control with two PI controllers and PWM generator. However, in this case the occurrence of bifurcations is possible. To avoid this problem and achieve average current control (ACMC), peak current control (PCMC) with a low-pass filter in the feedback and an adaptive compensation ramp [10] is used, where the average current value is obtained by processing the feedback current signal with a low-pass filter. The time constant of the filter corresponds to switching frequency.

The requirement for obtaining the PFC is that the reference for the current through the inductor in the boost converter has the same waveform as the input voltage, i.e., a fully rectified sine wave. At the same time, the amplitude of this sine wave is determined by the output voltage of the boost converter. Thus, the reference for the inductor current is determined by multiplying the output of the voltage controller, a standard PI controller, by a rectified sine wave of amplitude 1 and the same frequency as the input voltage (100 Hz).

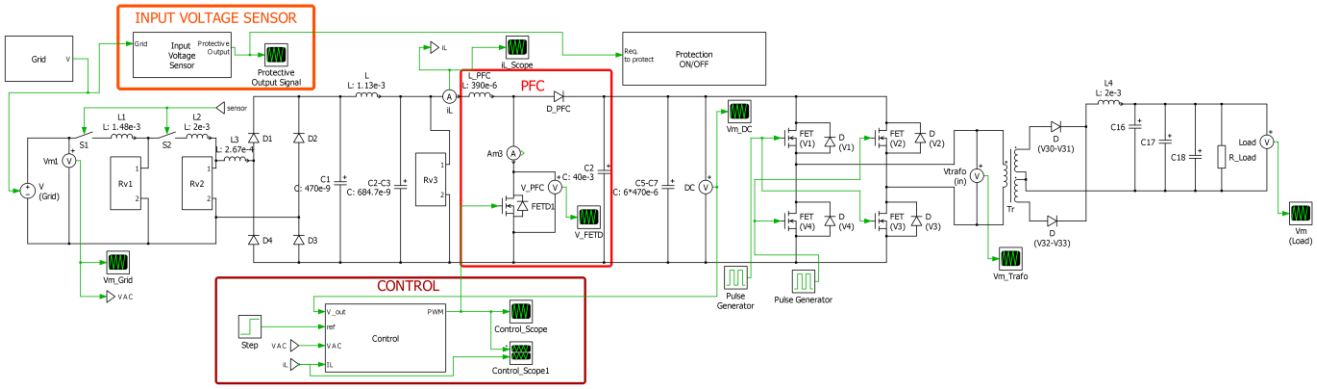


Figure 11: Complete model with PFC, PFC control and input voltage sensor

B. *Input voltage disturbance sensor model*

The PFC operate properly in the presence of nominal input voltage waveform. In the presence of distortions and overvoltages in the input voltage, the boost converter does not satisfactorily protect the circuits behind it. To achieve satisfactory protection, protection elements based on varistors and a detector of input voltage disturbance must be added. For this type of detection, a sensor has been developed whose only task is to detect the start of disturbances in the input voltage of the network and connect varistor-based protection elements before the boost converter. In this way, the rest of the system is protected from overvoltage and overcurrent. The block diagram of such a sensor is shown in Figure 12.

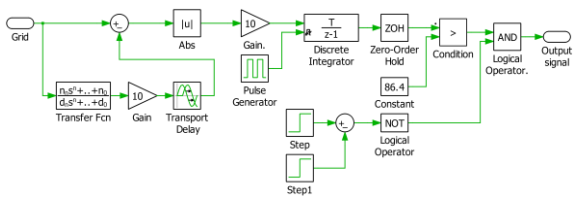


Figure 12: Input voltage disturbance sensor model

The output of the distortion sensor is based on the integral absolute error (IAE) between the actual input voltage and the first harmonic of the voltage. The first harmonic is obtained by a low-pass Butterworth filter with a cutoff frequency of 50 Hz (the frequency of the base harmonic), compensating gain and phase shift to maintain amplitude and phase equality with the original signal. The integrator is programmed to reset every 100 ms, which corresponds to 5 periods of input voltage. The 100 ms is the settling time of the IAE function and the Butterworth filter. This also means that the sensor operates with a sampling time of 100 ms. Comparing the IAE output signal to a threshold determines the presence of the distortion. If the output signal is greater than the threshold for the undisturbed signal, the input voltage is distorted for the last 5 periods and the sensor output signal changes from an initial 0, meaning no disturbance, to a 1. In the model, the sensor is disabled for the first 200 ms to prevent the sensor from responding to an initial transient response at power-up.

C. *Voltage controller and simulation results*

The controller in the outer loop of the boost converter cascade control is the voltage controller. The voltage controller of the PFC controls the output voltage of the PFC by influencing the sinusoidal current reference of the boost converter. It is designed as a PIT_1 controller with a time constant T_f equal to the period of the rectified input voltage of the boost converter to provide a smooth and undisturbed reference for the current controller during the input voltage period. The gain and integral time constant of the controller are determined by optimization using the Matlab/Simulink package and the optimization toolbox. The model for the optimization is based on the average boost converter model. Since the current model is much faster than the voltage part, the current part of the model is approximated by a current source controlled by a reference current signal. The gain K_r and the integral time constant T_i are determined by optimization using Newton based optimization function fmincon according to the ISE criterion, where the constraints are defined such that the rise time should be longer than 10 ms and the overshoot should be less than 1 %. These constraints and the filter of the controller guarantee the fastest disturbance compensation that does not affect the current shape required for the behavior of PFC.

TABLE 2: PARAMETERS OF THE EXISTING MODEL

Macromodel / Function	**Element**	**Value**
Control	K_r	2.2721
	T_i	0.2580 s
	T_f	0.01 s
	Current feedback filter	$13.33 \cdot 10^{-5}$ s

Using the above parameters and the full model shown in Figure 11 the following results were obtained. Figure 13 shows the input voltage at PFC and the corresponding input current. In the beginning, the input voltage signal is sinusoidal and PFC controls the input current so that it is also sinusoidal and in phase. The total harmonic distortion (THD) in this case is 4.22 %, which is acceptable for a system with PFC. At the 1 second mark, higher harmonics as well as white noise are introduced

into the base sine wave signal. At 1.1 seconds, the voltage sensor detects the disturbance in the input signal and responds by activating inductors L1 and L2, which were previously deactivated.

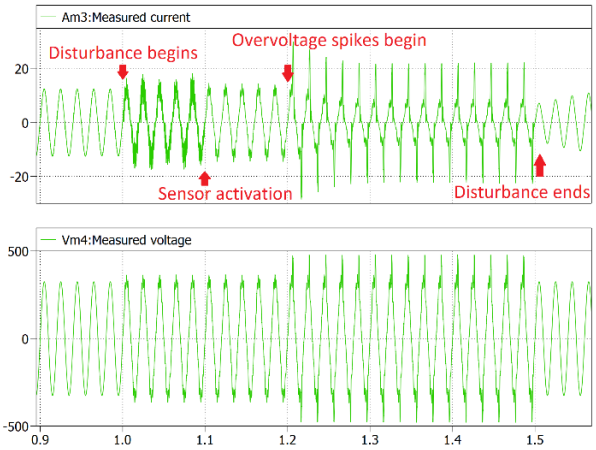


Figure 13: Input voltage and current with PFC

From the 1.2 to 1.5-second mark, overvoltage spikes are added to the higher harmonics, so the peak current is larger (about 30 A). After the input disturbance disappears, the system returns to its previous state, where it functions as PFC.

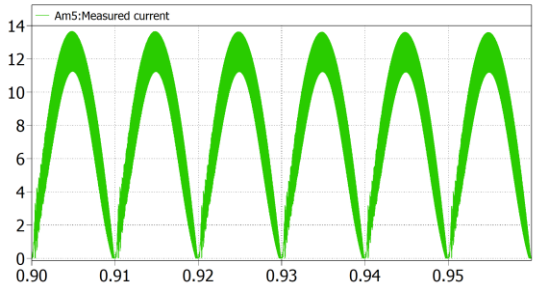


Figure 14: Current through inductor in the boost converter

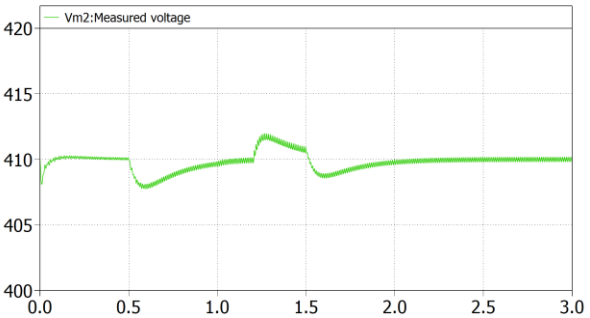


Figure 15: Output voltage of boost DC/DC converter

Figure 15 shows the output signal of the boost DC / DC converter. The set point for the output voltage is set to 410 V. At 0.5 s, a disturbance is applied in the form of a load change. The disturbance causes a current rise and an initial voltage drop. The controller compensates for the disturbance in about 0.5 seconds. At a time of 1.2 s the input voltage disturbance occurs (Figure 13) causing the output voltage to jump because the input voltage in the boost converter is greater than its output and the boost converter cannot operate under these conditions. As soon as the disturbance ends and the input voltage drops, the output voltage returns to its set point.

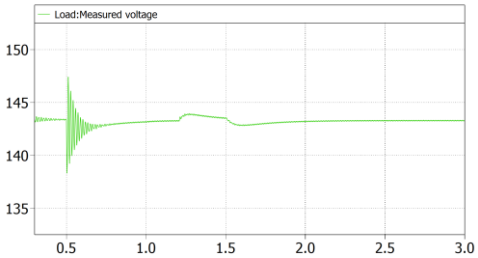


Figure 16: Output voltage from the system

Figure 16 shows the output voltage of the system, and as can be seen, the system functions as UPS with minor disturbances when an additional load is introduced at 0.5 seconds. Figure 14 shows that the inductor current follows the given reference.

V. Conclusion

The protection against overvoltage and higher harmonics, which consists of varistors and inductors, successfully reduces the voltage stress on the components of the UPS. However, the introduction of the boost DC/DC converter into the system presents a new challenge, as a voltage sensor must be incorporated to detect the occurrence of higher harmonics and spikes in the input voltage so that the protection can be activated at the right time. As long as the overvoltage protection isn't in effect, the boost converter acts as PFC and does its job very efficiently, with the input current distortion factor being around 4.2 %.

REFERENCES

[1] I. Uglešić, M. Mandić, “Napajanje električne vuče”, pp. 54, Graphis d.o.o., Zagreb, 2015.

[2] M. Štetić, „Problematika napajanja trošila električnom energijom iz kontaktnog voda 25 kV“, izvještaj s terena (MARETON d.o.o.), March 2019.

[3] A. Marcos-Pastor, E. Vidal-Idiarte, A. Cid-Pastor, L. Martinez-Salamero “Interleaved Digital Power Factor Correction Based on the Sliding-Mode Approach”, IEEE Transaction on Power Electronics, Vol. 31, no. 6, pp. 4641-4653, June 2016.

[4] A. Fernández, J. Sebastián, P.Villegas, M. M. Hernando, D. G. Lamar, “Dynamic Limits of a Power-Factor Preregulator”, IEEE Transactions on Industrial Electronics, Vol. 52, no. 51, pp. 77-87, February 2005.

[5] Šolc, I., Hećimović, A., Šunde V., Ban, Ž, “UPS with PFC input stage for railway applications with improved immunity on input overvoltage and energy strikes”, IEEE (EDPE2021), 2021.

[6] Abdel-Rahman, S., Stückler, F., Siu, K., “PFC Boost converter design guide”, Infineon, 2016

[7] Stepins, D., Huang, J., “Effects of Switching Frequency Modulation on Input Power Quality of Boost Power Factor Correction Converter”, International Journal of Power Electronics and Drive System (IJPEDS), Vol. 8, No. 2, pp. 882-899, June 2017.

[8] Moore, T., Schmid, F., Tricoli, P., “Voltage transient management for Alternating Current trains with vacuum circuit breakers”, IET Electrical Systems in Transportation, 2021.

[9] P. Marian Nicolae, M. Stefan Nicole, I. Diana Nicolae, A. Netoiu, “Overvoltages Induced in the Supplying Line by an Electric Railway Vehicle”, University of Craiova, Romania, 2021.

[10] J.Schoenberger – Modelling PFC controller with PLECS

[11] MATLAB documentation, fmincon function

Reduction of Current Harmonics in Data Center Power System using Single-Phase Shunt Active Power Filter

V. Zeleničić, M. Miletić, K. Raič Raguž, D. Sumina and I. Erceg
University of Zagreb/Faculty of Electrical Engineering and Computing
vinko.zelenicic@fer.hr, marinko.miletic@fer.hr, katica.raic-raguz@fer.hr, damir.sumina@fer.hr, igor.erceg@fer.hr

***Abstract*—With the significant increase of nonlinear loads in data center power systems, the losses due to the injection of current harmonics have also increased significantly. In order to reduce current harmonics, the use of Shunt Active Power Filters (SAPF) has become widespread. These filters use active components such as IGBT or MOSFET as well as R, L and C passive components. The SAPF power stage consists of a voltage source inverter (VSI) and an LCL filter, which is used to filter the signal at the switching frequency. The paper presents two simulatons of SAPF. The first simulation refers to a filter used only to reduce the current harmonics, while the second simulation uses a filter that also acts on the current reactive component. The control loop consists of a current controller and a PWM. In the first simulation, the grid side current and the converter side current are used as the current feedback signal. In the second simulation, only the grid side current is used as the current feedback signal. The results of two SAPF simulations are compared and the results show the influence of the control methods on the total current distortion as well as on the current ripple.**

***Keywords*—harmonic; active filter; voltage source inverters; reactive component**

I. INTRODUCTION

Recently, the need for data storage and processing has increased, and with it the need for data centres. Nonlinear loads in data centers, such as servers, IT equipment and cooling systems, are the most common sources of current harmonics. Harmonics not only increase the heating of equipment and shorten its life, but also have an impact on increasing overall losses in the system. The most common method for harmonic reduction is the use of active filters. Active filters compensate current harmonics by injecting equal but opposite harmonic compensating current. Type L and LCL filters are the most commonly used filters for reducing current harmonics. The LCL filter attenuates the high frequencies better than the L filter because the capacitor acts as a short circuit at high frequencies [1]. In addition, the LCL filter has a smaller volume for the same filtering effect and thus lower production costs. The LCL filter disadvantage is that they have current spikes at the resonant frequency [2]. It is necessary to provide damping at the resonant frequency to overcome this shortcoming. This attenuation can be passive or active. In the active method, fewer additional losses are introduced into the system, but the control loop is more complicated due to the additional control of active elements [3]. The passive damping method introduces additional losses into the system, but is the preferable option for more complex systems. In the simplest passive method, a resistor is connected in series with the capacitor. In contrast, in improved versions of passive methods, an inductor or a combination of inductor and capacitor is connected in parallel with the resistor. Due to their simplicity and good harmonic filtering characteristics, Shunt Active Power Filters with damping resistors have the widest application. This paper examines such a method.

In order to improve the properties of the LCL filter, the improved parameter design method of LCL SAPF is presented in [1]. In series with the capacitor, a resistor, an inductor, and a capacitor are connected in parallel. To design the inductor L_1 on the inverter side, which is used to suppress the inverter output current ripple, the peak ripple current method is used. For the best filtering effect, the two inductors ratio L_1/L_2 can be selected between 3 and 5. The filter's capacitance value (C_f) is designed to the systems rated power, and the reactive power generated by the capacitance is less than 5%. The damping inductor L_d short-circuits the damping resistor R_d at low frequencies, and the damping capacitor C_d short-circuits the resistor R_d at high frequencies. Conventional and improved LCL filters are compared in MATLAB/SIMULINK simulations, and the results show that the improved LCL filter has higher filter accuracy for higher harmonics.

A comparison of four different passive damping methods for LCL filter is given in [4]. Topologies, transfer functions and frequency characteristics are presented. Simulations for all methods were performed on an 80 kVA SAPF using the MATLAB/SIMULINK software package. Losses, damping efficiency, and performance were compared. The results show that the best attenuation of low frequencies up to resonance has the first method, where a resistor R is added in series with the capacitor of the LCL filter. Frequencies between the resonant and switching frequency are best attenuated by the second method, which involves connecting a parallel of the resistor and inductor in series with the capacitor. Frequencies above double switching frequency are best attenuated with third method which includes connecting a parallel resistor, an inductor, and a capacitor in series with the capacitor. Frequencies

This work has been fully supported by the European Regional Development Fund under the project "Development of a system for optimizing power consumption in data centers" (KK.01.2.1.02.0082).

approximately equal to the switching frequency are best attenuated by the fourth method, which adds a parallel of resistor and a series connection of inductor and capacitor in series with the capacitor. It is concluded that additional losses for all methods have a small effect on efficiency.

To improve the properties of the LCL filter the SAPF control system has two loops, the outer and the inner loop have feedback by grid current [5]. The inner loop is the PI control and the outer loop is the repetitive control, where the controller accumulates the error signal cycle by cycle and the error in this cycle occurs again in the next cycle. The PI control corrects the frequency characteristics and improves the dynamic response speed, while the repetitive control ensures the control system accuracy in a steady-state. A line voltage feedforward control has been used to suppress the influence of line fluctuations. The results show significant influence is in the low and medium frequency range. The voltage feedforward also reduces the line current *THD*. The experimental results show faster speed responses, good disturbance rejection, and higher accuracy in a steady-state compared to conventional SAPF control.

Another example of current control system is presented in [6]. In the LCL filter, there is no resistor in series with capacitor C. Resonance is prevented by an additional internal control loop. The internal circuit uses the capacitor current as a feedback signal. The coefficient (R) in the capacitor current feedback acts on the LCL filter attenuation like a physical resistor, which is why it is called a *virtual resistance*. The internal circuit stabilizes the control system, and the external circuit provides the required accuracy with a PI controller. Using the damping method with a virtual resistor improves the SAPF efficiency.

This paper compares the results of two Shunt Active Power Filters simulations when grid side and converter side currents are used as feedback signals. The paper is divided into four sections. Section I consists of an introduction and related work. In Section II, circuit and control diagrams of SAPF with LCL output filter are presented. Section III presents a comparison of the two simulations results. In the first simulation, measured values of the current harmonics were used, while in the second simulation the diode rectifier with an RL load on the DC side was used. Future work and conclusion are given in Section IV.

II. CONTROL MODELS

Fig. 1 shows the SAPF scheme with the LCL output filter. The grid side current (i_2) and the converter side current (i_1) can be used as the current feedback signal. To achieve the desired LCL damping factor resistor is connected in series with the capacitor. Transfer function, where ($L = L_1 \parallel L_2$) is given by (1). The resonant frequency of the second order term in the denominator is given by (2), and the damping factor ζ is given by (3).

$$G(s) = \frac{1}{s(L_1 + L_2)} \frac{RCs + 1}{LCs^2 + RCs + 1} \tag{1}$$

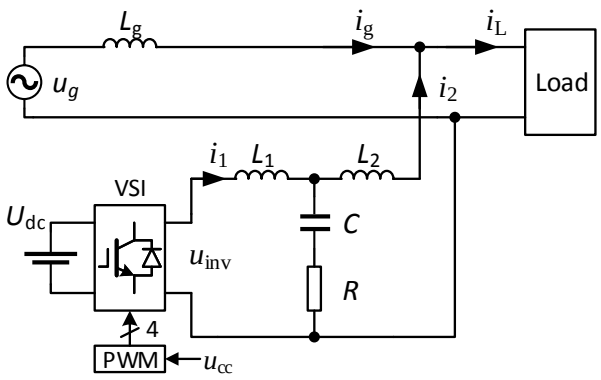


Fig. 1: Circuit diagram of SAPF with LCL output filter

$$\omega_n = \frac{1}{\sqrt{LC}} \tag{2}$$

$$\zeta = \frac{R}{2}\sqrt{\frac{C}{L}} \tag{3}$$

The inductance L_1 on the converter side has a value of 2 mH, the inductance L_2 on the grid side has a value of 0.5 mH and the capacitor has a value of 8 μF. The desired damping factor ζ is 1.41 and is achieved with a resistance value of 20 Ω.

The VSI consists of a single-phase transistor bridge with a constant DC voltage of 450 V. The transistors are controlled by a PWM circuit with a constant switching frequency of 16.7 kHz. Fig. 2 shows two different non-linear loads. Load 1 is a diode rectifier with RL load on the DC side, which is standardly used for obtaining current harmonics in experimental verification. When the RL circuit time constant is more than 30 ms (3 half cycles of the supply voltage), the RL circuit current pulsation is less than 7%, so that the current i_L on the AC side is almost rectangular. Such a waveform is rich in harmonics and is therefore suitable for SAPF experimental verification. The values for resistance and inductance are 13 Ω and 0.48 H, respectively. Load 2 is designed as a current source representing the actual current harmonics obtained by measurements in the data center. The current source consists of odd current harmonics up to the 25th harmonic.

Fig. 3 shows a control diagram of an APF filtering only the higher current harmonics, without compensation of the

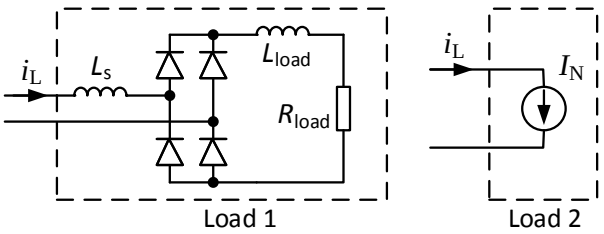


Fig. 2: Different SAPF Load

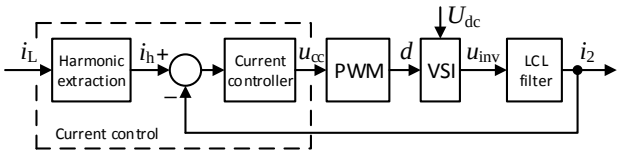


Fig. 3: SAPF Control diagram

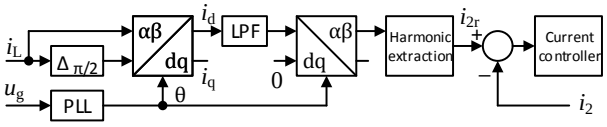


Fig. 4: dq Control diagram

fundamental harmonic reactive component. Fig. 4 shows a control diagram in the dq frame, with compensation of the higher harmonics and the fundamental harmonic reactive component. In the harmonic extraction block, the load current fundamental harmonic signal is subtracted from the load current signal i_L. The difference between these current signals is the harmonic current signal i_h and forms the reference for the current control loop. The current controller is a linear controller of the PI type. In dq frame a single-phase system can be translated directly into an α-β frame without matrix conversions by shifting β by 90° relative to α (i_L). To generate the fundamental harmonic, the active component (i_d) and the θ obtained from the PLL are needed [7]. After the generation of the α from the i_d, the rest of the control proceeds to harmonic extraction block previously described (Fig. 3).

III. Systems simulations

A. Simulations on nonlinear load

This subsection shows the responses of current harmonics filtering in a circuit with a nonlinear load "Load 1" (Fig. 2). Reason for choosing RL was the simplicity of its realisation in a laboratory environment to test the prototype active filter. Two cases have been considered. The converter side current i_1 and the grid side current i_2 (Fig. 1) are used as feedback signals.

Fig. 5 represents the first case, showing the waveform of the load current (blue), the active filter current (red) and the grid current (green). The load current RMS value is 15.35 A and the grid current RMS is 14.0 A. The load current *THD* is 42.3%, and the grid current *THD* after filtering is 2.76% (Fig. 6).

Fig. 7 shows the load, filter and grid current waveforms for the case where the current i_2 is used as feedback signal. The RMS value of the load current is 15.35 A and the grid current RMS is 14.1 A. The load current *THD* is 42.3%, and the grid current *THD* after filtering is 1.55%. With the same setting of the PI controller, the grid current *THD* is lower when current i_2 (Fig. 8) is used as feedback signal than when current i_1 (Fig. 6) is used as feedback signal.

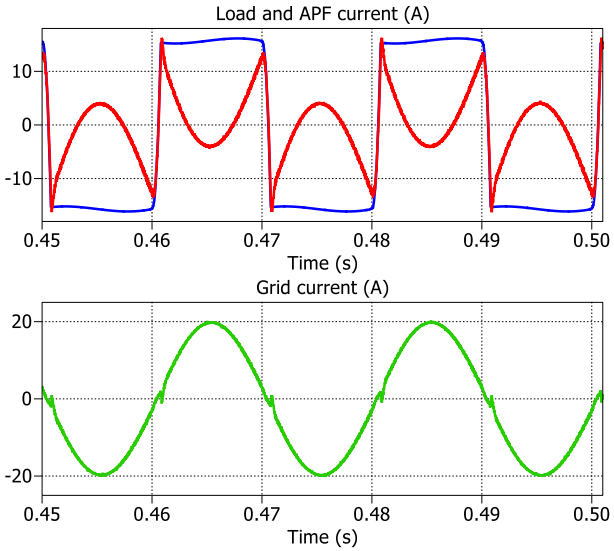


Fig. 5: Current waveforms using converter side current (i_1) as a feedback signal

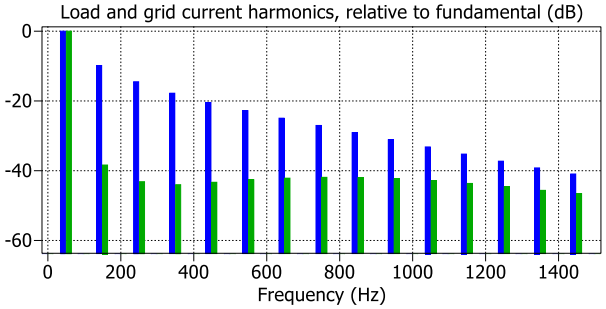


Fig. 6: Current harmonics spectrum using converter side current (i_1) as a feedback signal

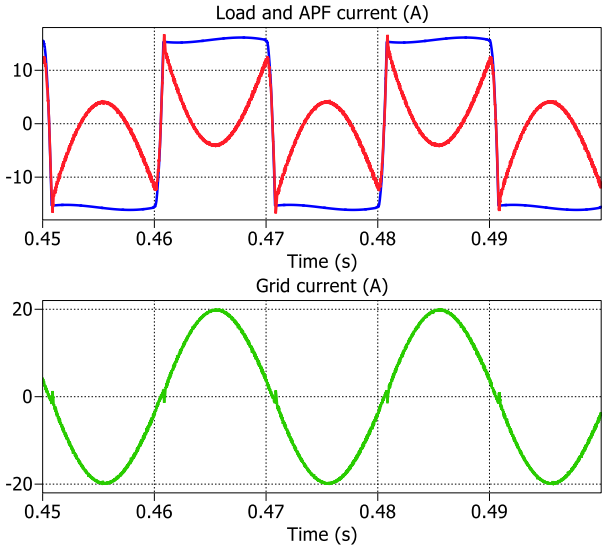


Fig. 7: Current waveforms using grid side current (i_2) as a feedback signal

For simulations in dq frame presented below, the current i_2 was used as feedback signal (Fig. 9).

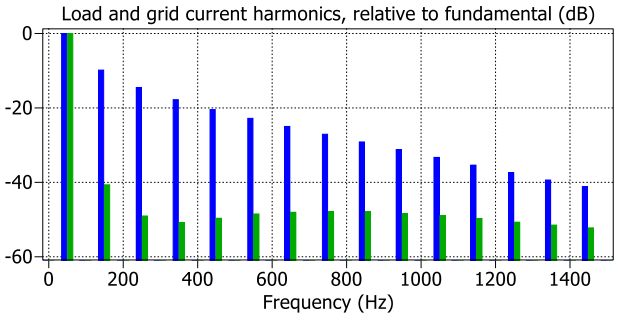


Fig. 8: Current harmonics spectrum using grid side current (i_2) as a feedback signal

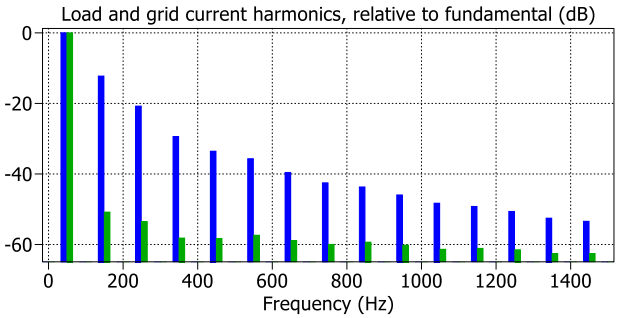


Fig. 10: Current harmonics spectrum using dq frame control and grid side current (i_2) as a feedback signal

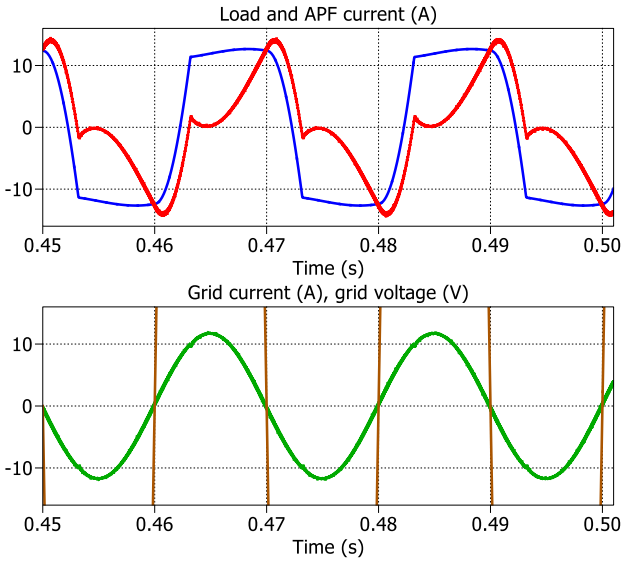


Fig. 9: Current waveforms using dq frame control and grid side current (i_2) as a feedback signal

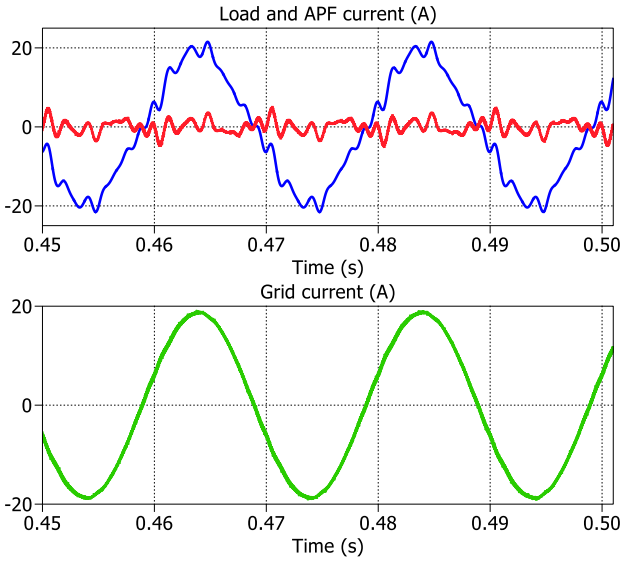


Fig. 11: Current waveforms using measured data rack harmonics as load current

In order to compensate for the fundamental current harmonic reactive component, a control is carried out in the dq frame. The inductance L_s value was increased from $1mH$ to $20mH$. The load current *THD* was decreased from 42.3% to 26.5%, due to the inductance L_s higher value. The load current waveform has a lower slope and a lower *THD*. The load current harmonics from the 15th harmonic upwards are smaller than -40 dB, whereas in the case with $1mH$ only the 29 harmonic is smaller than -40 dB. The grid current *THD* is 0.5% (Fig. 10), whereas in the first case it is 1.55%.

B. Simulations with measured load current harmonics

Fig. 11 and Fig. 12 show the load, APF and grid current waveforms as well as the harmonic spectrum for the case where the load was simulated with real current harmonics data measured in the data center. These measurements were made in data center rack with data servers. The load current RMS value is 13.34 A and the grid current RMS is 13.25 A. The load current *THD* is 11.94%, and the grid current *THD* after filtering is 0.66%.

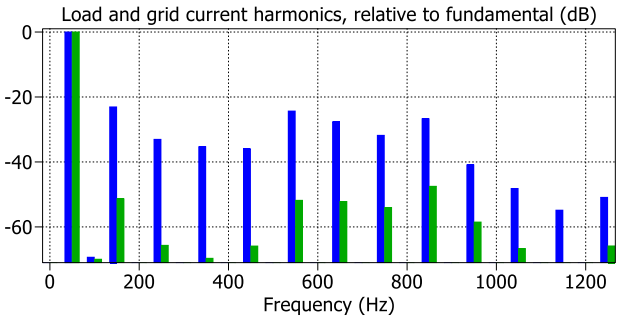


Fig. 12: Current harmonics spectrum using measured data rack harmonics as load current

Measurements on another rack with communication equipment show a higher current distortion (Fig. 13, blue). Since the equipment is underloaded, the load current *THD* is 47% with an RMS value of 1.82A. After filtering, the grid current *THD* is 2.11% and the RMS value of 1.56A (Fig. 14).

The case with real load current harmonics data from the data center was also simulated in the dq frame. The

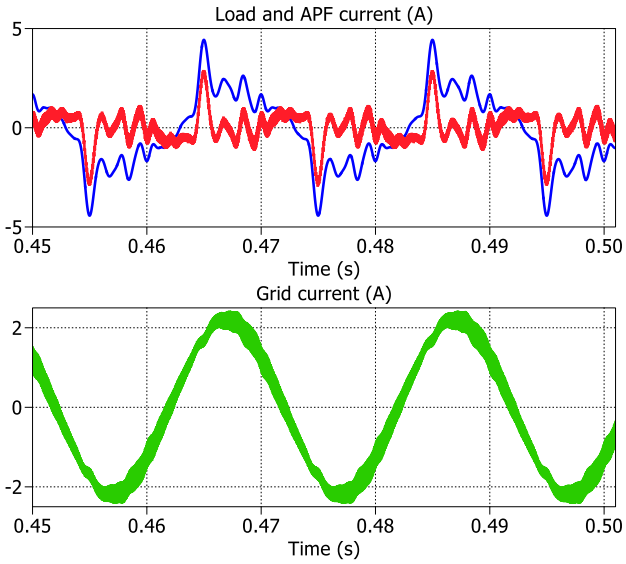


Fig. 13: Current waveforms using measured comunication rack harmonics as load current

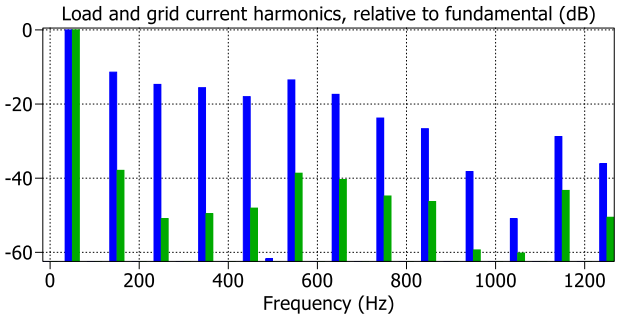


Fig. 14: Current harmonics spectrum using measured comunication rack harmonics as load current

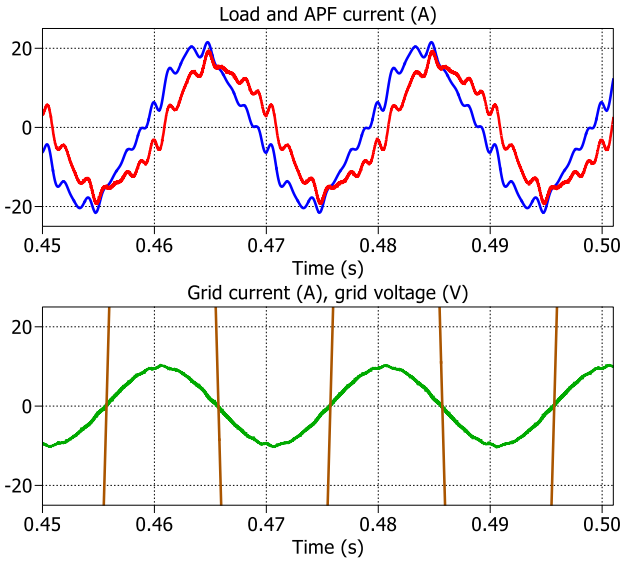


Fig. 15: Current waveforms using dq frame control and measured data rack harmonics as load current

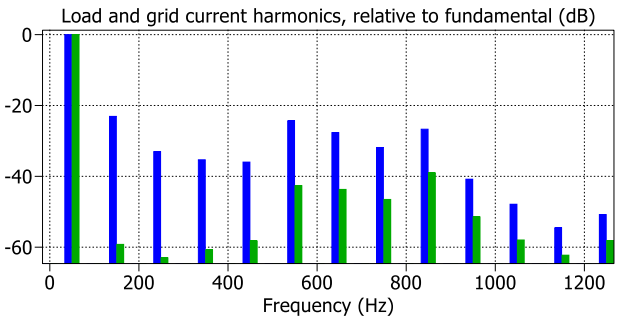


Fig. 16: Current harmonics spectrum using dq frame control and measured data rack harmonics as load current

load current fundamental harmonic lags the grid voltage by 60°. The load current RMS value is 13.34 A and the grid current RMS is 7.04 A (Fig. 15). The load current *THD* is 11.94%, and the grid current *THD* after filtering is 1.61% (Fig. 16).

IV. CONCLUSION

The paper presents two Shunt Active Power Filter simulations. The first simulation refers to a filter used only to reduce the current harmonics, while the second simulation uses a filter that also acts on the current reactive component. In the first simulation, the grid side and converter side currents were used as the current feedback signal. With the same settings of the PI controller, the grid current *THD* is lower when the grid side current is used as the feedback signal. For this reason, only the grid side current was used as the current feedback signal in the second simulation with control in the dq frame. Two types of loads were considered, an RL load and a load represented as a current source with current harmonic values measured in the data center. In both cases, the current *THD* after filtering is reduced to a value of less than one tenth of the original value. Future work will be based on experimental verification of the Shunt Active Power Filter prototype.

REFERENCES

[1] L. Zhou, Z. Liu, Y. Ji, D. Ma, J. Wang, and L. Li, "A improved parameter design method of lcl apf interface filter," in *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2020, pp. 948–952.

[2] H. Yuan and X. Jiang, "A simple active damping method for active power filters," in *2016 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2016, pp. 907–912.

[3] C. Xie, Z. Wang, and G. Chen, "A simple method of realization of low power loss and high compensation-precision active power filter," in *2009 International Conference on Sustainable Power Generation and Supply*, 2009, pp. 1–5.

[4] M. Büyük, A. Tan, K. Ç. Bayindir, and M. Tümay, "Analysis and comparison of passive damping methods for shunt active power filter with output lcl filter," in *2015 Intl Aegean Conference on Electrical Machines & Power Electronics (ACEMP), 2015 Intl Conference on Optimization of Electrical & Electronic Equipment (OPTIM) & 2015 Intl Symposium on Advanced Electromechanical Motion Systems (ELECTROMOTION)*. IEEE, 2015, pp. 434–440.

[5] Y. Zhang, Z. Dai, and Y. Fang, "Voltage feed forward shunt active power filter based on double loop control," in *2021 International Conference on Power System Technology (POWERCON)*, 2021, pp. 1251–1256.

[6] W. Zhao, Y. Li, and G. Chen, "A double-loop current control strategy for shunt active power filter with lcl filter," in *2009 IEEE International Symposium on Industrial Electronics*, 2009, pp. 1841–1845.

[7] V. Khadkikar, M. Singh, A. Chandra, and B. Singh, "Implementation of single-phase synchronous d-q reference frame controller for shunt active filter under distorted voltage condition," in *2010 Joint International Conference on Power Electronics, Drives and Energy Systems, 2010 Power India*, 2010, pp. 1–6.

OPHO

International Conference on

OPTOELECTRONICS AND PHOTONICS

Steering Committee

Chairs: Vera Gradišnik, University of Rijeka, Croatia
Irena Jurdana, University of Rijeka, Croatia
Marin Karuza, University of Rijeka, Croatia
Duško Čakara, University of Rijeka, Croatia
Mario Stipčević, Ruđer Bošković Institute, Zagreb, Croatia

Program Committee

Members: Domenico Caputo, Sapienza University of Rome, Italy
Andreja Gajović, Ruđer Bošković Institute, Zagreb, Croatia

INVITED PAPER

Polarisation Based Entanglement Distribution Quantum Networking

M.J. Clark*, R. Wang†, S. Bahrani†, M. Peranić‡, O. Alia†, M. Loncaric‡, Ž. Samec‡,
A. Radman‡, M. Stipcevic‡, R. Nejabati†, D. Simeonidou†, J. Rarity*, S.K. Joshi*
* Quantum Engineering Technology Labs, University of Bristol, Bristol, United Kingdom
† High Performance Networking, University of Bristol, Bristol, United Kingdom
‡ Center of Excellence for Advanced Materials and Sensing Devices, Ruđer Bošković Institute, Zagreb, Croatia
mj.clark@bristol.ac.uk

***Abstract*—Quantum networks based on entanglement distribution have shown promise for building scalable and fully connected systems that support quantum key distribution. This work aims to go beyond simply implementing quantum key distribution and explore the potential of such networks for implementing quantum photonic interconnects. Our research demonstrates the passive polarization stability of these networks for over a week and highlights the benefits of dynamic reconfiguration to remove redundant resources. We discuss recent advancements in quantum frequency conversion and quantum memory-based networks, and argue that the development of scalable, long-distance interconnects is crucial for advancing quantum technology. Our findings have important implications for the future of quantum networking and highlight the need for entanglement based photonic interconnect networks, such that quantum technology can scale beyond monolithic systems.**

***Keywords*—quantum communication, quantum network, entanglement, entanglement distribution**

I. INTRODUCTION

In the face of ever rapidly developing Quantum Technology (QT), a method of interconnecting QTs situated in separate locations[1] is required. The prevailing method is to connect QT nodes through photonic interconnects, such that quantum information is shared between locations, by building Quantum Networks (QNet). Recent demonstrations of QNets have been focused towards the single use case of Quantum Key Distribution (QKD) [1]–[3], and commercial implementations have been announced in recent years [4].

This focus on QNets for only a single QT application limits the future applications in Quantum Computation and Sensing [5]–[8]. Utilizing entanglement as a resource expands use cases of QNets into a regime of distributed QT [9], as shown in Fig. 1. One of the main obstacles to this approach is the variety of Quantum Memory (QM) platforms that are being developed. Each platform takes advantage of a different process to produce a quantum state, such as neutral atoms and trapped ions using atomic energy levels [10].

This produces a wide variation of the interaction wavelength by platform [11]–[14]. These photons are commonly produced in wavelength regimes that traditionally have not been used to transmit optical signals through fibre optic infrastructure, due to the high propagation loss [15]. Operating in this high loss regime prevents the separation of the clusters beyond a few hundred meters [11].

The technique of entanglement preserving Quantum Frequency Conversion (QFC) [17] can be used to overcome this limit. QFC is most commonly performed using the second order properties of non-linear crystals, through Difference Frequency Generation (DFG) and Sum-Frequency Generation (SFG) [18]–[21]. These processes allow photons in the $400 - 900\,\text{nm}$ regime to be converted into the C-Band for efficient transmission through deployed fiber telecommunications infrastructure, or for photons in the C-Band to be converted into the correct frequency regime to interact with a QM. By

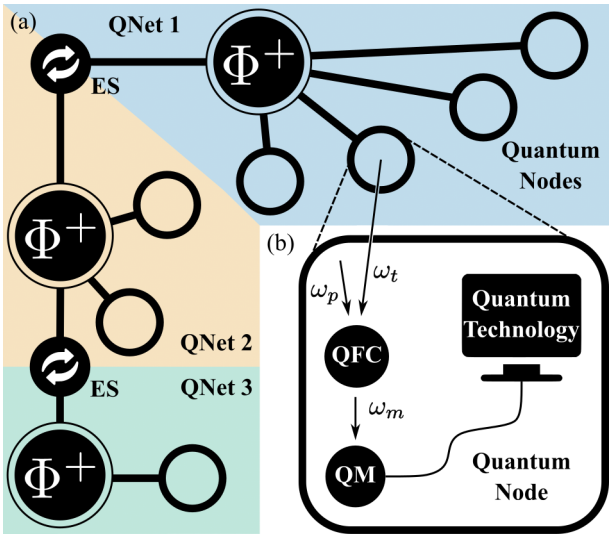


Fig. 1: Schematic of a Quantum Internet with Quantum Networks (QNet) providing photonic interconnects between Quantum Technologies (QT). Fig. **(a)** depicts the physical topology of QNets, and how they can be interconnected through Entanglement Swapping (ES). Fig. **(b)** shows the process of Quantum Frequency Confersion (QFC) into a Quantum Memory (QM), to connect QT with entanglement from the QNets. Here ω_p is the pump photon, ω_t is the entanglement transporting photon, and ω_m is the QM native photon.

[1]It is important to note that this does not mean exclusively space-like separated locations

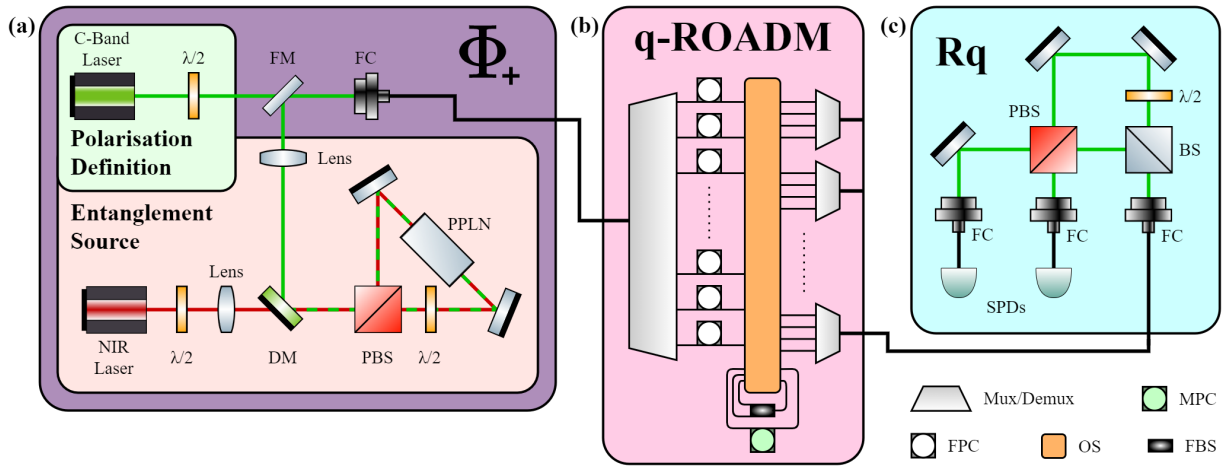


Fig. 2: The main components of the quantum network discussed [16]. Fig. **(a)** shows the photon pair source, comprised of a bulk optic entanglement source (ES) and the polarisation definition (PD). Fig. **(b)** is the quantum enabled Reconfigurable Optical Add Drop Multiplexer(q-ROADM). Fig. **(c)** is the quantum receiver (Rq). Here $\lambda/2$ is a half-wave plate, DM is a dichroic mirror, BS is a beam-splitter, PBS is a polarising beam-splitter, PPLN is a magnesium doped periodically poled lithium niobate crystal, FM is a motorised flip mirror, FC is a fibre collimater, SPD is a single-photon detector, FPC is a fibre polarisation controller, MPC is a motorised fibre polarisation controller, OS is an optical switch, FBS is a fibre beams splitter, and a (De)Mux is a wavelength (de)multiplexer.

careful selection of the pump wavelength in DFG and SFG, any desired wavelength within the C-Band can be achieved. If ω_p is the pump frequency, and ω_t is the frequency of the entanglement transporting photon, then the memory photon would have frequency $\omega_m = \omega_p + \omega_t$, as shown in Fig. 1 **(b)**. This process is reversible, so by performing the process on the memory(C-Band) photon the C-Band(memory) photon would be received out of the system. Development of these systems into high-efficiency processes is required for large scale QNets, where current systems reach an efficiency of about 30% [22].

With these QFC systems, entanglement distribution QNets will distribute entangled photons between nodes, without the associated loss of memory native wavelengths. Different entanglement resources may provide additional constraints on the QFC implementation. For example, when polarization entanglement is utilized the process of QFC must not apply a polarisation projection or measurement that produces a mixed quantum state. As such polarization-insensitive QFC is needed [23], [24].

By utilizing bipartite entanglement QNets, Greenberger-Horne-Zeilinger (GHZ) states can be generated in QM [11], [25] to share entanglement beyond two-node entanglement. QNets can distribute entangled photon pairs between many locations in a dynamic way. This allows these GHZ-style memory networks to be connected in a scalable way, and by leveraging QFC, with low-loss C-Band interconnects across Metropolitan areas.

Here we discuss an approach to building entanglement distribution QNets [26], [27], that connects many nodes with fully connected topologies. These QNets are designed as photonic interconnects that can facilitate entanglement between QT, without limiting the systems to small set of

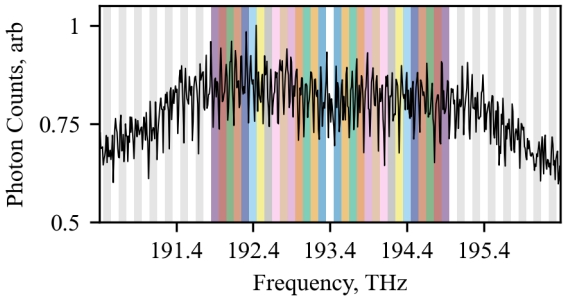


Fig. 3: The output spectrum of the entanglement source. Here the Photon counts are normalised to the maximum number seen in a frequency bin.

use cases. Entanglement is measured using the Bennett-Brassard-Mermin (BBM92) [28] method of QKD, and networks are assessed using a weighted average Secret Key Rate (SKR) [16], [29]. The quantum measurements taken to preform BBM92 QKD also allow other protocols to be preformed, without changes being made to the system. These include preforming secure anonymous protocols [30] and exchanging secure digital signatures [31].

II. Polarisation Entanglement Networks

Polarization entangled photon pairs can readily be produced through Spontaneous Parametric Down Conversion (SPDC). The quantum source [16], as shown in Fig. 2 **(a)**, used a magnesium doped periodically poled lithium niobate non-linear crystal to produce frequency entangled C-Band photon pairs centred at 1550.12 nm (193.4 THz), as shown in Fig. 3. A Sagnac interferometer induces polarization entanglement on to the photon pairs, producing a hyper-entangled state in polarization and frequency.

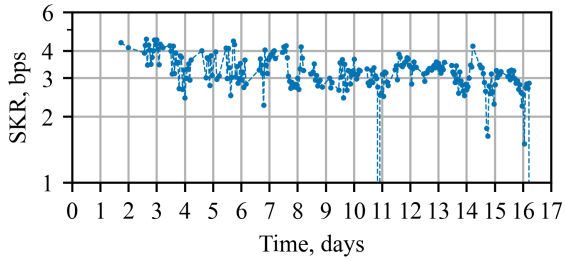


Fig. 4: Weighted average SKR of a 10-node fully connected entanglement based Quantum Networks [16].

Frequency entanglement is utilized to produce discrete bins of entanglement, through Dense Wavelength Division Multiplexing (DWDM) technology as shown in Fig. 2 **(b)**, in accordance to the telecommunications ITU grid. These 100 GHz bins are labeled communication channels (CC), where each CC shares entanglement with one other CC. This frequency splitting can be seen in Fig. 3, where the bars with matching color indicate the 15 CC pairs.

Once frequency demultiplexed, selected CCs are time multiplexed into 4 logical channels (LC), and the remaining CCs are relabelled as LCs. Taking advantage of time multiplexing and frequency multiplexing simultaneously, larger networks of nodes can be connected with a single source with a limited set of CCs. Each LC is connected to a port of an optical switch via a fibre polarisation controller. A series of frequency multiplexers are connected to the outputs of the optical switch, allowing many LCs to be combined into a single fiber, each of which is connected to a separate node. In such a setup, LCs can be assigned to any frequency multiplexer, allowing for a dynamically connected network of entanglement sharing [32].

Here we have shown how a quantum-enabled Reconfigurable Optical Add Drop Multiplexer (q-ROADM), as shown in Fig. 2 **(b)**, can be utilized to allow active switching between different topologies [32]. Changing the topology in a QNet allows for many different scenarios, beyond changing which nodes generate a secret key in a QKD network at that point in time. Reducing the number of simultaneous LCs reduces the perceived noise counts when measuring only a single LC [29], [32]. Due to the required switching between topologies required to produce a fully-connected QNet in this manner, this reduces the total time in which each LC can be measured. We have noted a trade-off between the quality of network components and the advantage that active switching gives. Notably, when the detector Jitter increases a dynamically switched set of non-fully connected network topologies see a reduced Quantum Bit Error Rate (QBER) and gain a boost in SKR across the combined fully-connected QNet [29].

With a set of 5 LCs per node, a QNet with 10 nodes can be constructed, as shown in Fig. 5 [16]. Using both time and frequency multiplexing creates duplicated connections in the system. These connections can be distributed to

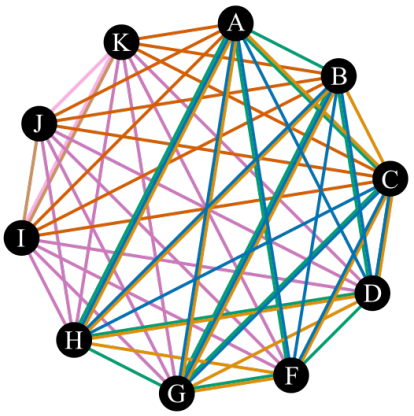


Fig. 5: Topology of the entanglement distribution Quantum Networks. Here color represents a frequency channels received by nodes.

nodes that require a higher bandwidth of entanglement sharing or can be utilised for different applications, by demultiplexing the channels at the node.

We demonstrated such a QNet, with deployed telecommunications fiber, and showed polarization stability upto 11 days of continuous operation with a weighted average SKR of 3.4 bps, improving on the previous record of 1 day [27]. In this QNet, all LCs were birefringence compensated before up-time began, such that the definition of polarization at the node matched that of the source location [33]. No further compensation was completed for the duration of data collection, showcasing that current fiber infrastructure has sufficiently low birefringence induced polarization rotation that a network is stable within expected poissonian fluctuation of the QBER for BBM92 QKD.

In the set of 10 nodes, 3 were connected through metropolitan scale deployed fiber with propagation loss ranging from 1.5 to 3.2 dB, and the remaining nodes had local connections. A further 2 nodes were connected, but failed in the network initialisation stage. This shows that such QNets are resilient to node downtime, and that extra nodes can be connected with no change to the current resource distribution.

III. Conclusion

Entanglement-based QNets have shown promise in connecting nodes for both QKD and QM entanglement [11], [16], [27]. While previous research has focused on either long-distance QKD networks [16], [26], [27] or small-scale QM networks [11], [13], [14], [25] , a critical challenge remains: developing low loss C-Band QNets to interconnect QMs across metropolitan areas. To address this challenge, we recommend focusing on the incorporation of bidirectional QFC [18] and developing new technologies to interface photons in the C-Band with current QM systems. By doing so, we can unlock the potential for QT to develop beyond single monolithic systems [5]–[8].

ACKNOWLEDGMENT

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grants EP/SO23607/1, for the Quantum Engineering Centre for Doctoral Training, Centre for Nanoscience & Quantum Information, University of Bristol, and EP/T001011/1 for the Quantum Communications Hub. This work was supported by the Croatian Science Foundation, HRZZ grant No. IPS-2020-1-2616 and the Croatian Ministry of Science and Education, MSE grant No. KK.01.1.1.01.0001.

REFERENCES

[1] M. Peev, C. Pacher, R. Alléaume, C. Barreiro, J. Bouda *et al.*, "The secoqc quantum key distribution network in vienna," *New Journal of Physics*, vol. 11, no. 7, p. 075001, jul 2009. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/11/7/075001

[2] D. Stucki, M. Legré, F. Buntschu, B. Clausen, N. Felber *et al.*, "Long-term performance of the swissquantum quantum key distribution network in a field environment," *New Journal of Physics*, vol. 13, no. 12, p. 123001, dec 2011. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/13/12/123001

[3] J. Dynes, A. Wonfor, W.-S. Tam, A. Sharpe, R. Takahashi *et al.*, "Cambridge quantum network," *npj Quantum Information*, vol. 5, no. 1, p. 101, 2019.

[4] "BT and Toshiba to build world's first commercial quantum-secured metro network across London," *BT Group Newsroom*, October 6 2021, Accessed: March 13 2023. [Online]. Available: https://newsroom.bt.com/bt-and-toshiba-to-build-worlds-first-commercial-quantum-secured-metro-network-across-london/

[5] R. Van Meter and S. J. Devitt, "The path to scalable distributed quantum computing," *Computer*, vol. 49, no. 9, pp. 31–42, 2016.

[6] Z. Zhang and Q. Zhuang, "Distributed quantum sensing," *Quantum Science and Technology*, vol. 6, no. 4, p. 043001, jul 2021. [Online]. Available: https://dx.doi.org/10.1088/2058-9565/abd4c3

[7] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini *et al.*, "Quantum internet: Networking challenges in distributed quantum computing," *IEEE Network*, vol. 34, no. 1, pp. 137–143, 2020.

[8] X. Guo, C. R. Breum, J. Borregaard, S. Izumi, M. V. Larsen *et al.*, "Distributed quantum sensing in a continuous-variable entangled network," *Nature Physics*, vol. 16, no. 3, pp. 281–284, 2020.

[9] D. Awschalom, K. K. Berggren, H. Bernien, S. Bhave, L. D. Carr *et al.*, "Development of quantum interconnects (quics) for next-generation information technologies," *PRX Quantum*, vol. 2, p. 017002, Feb 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PRXQuantum.2.017002

[10] D. S. Weiss and M. Saffman, "Quantum computing with neutral atoms," *Physics Today*, vol. 70, no. 7, 2017.

[11] M. Pompili, S. L. N. Hermans, S. Baier, H. K. C. Beukers, P. C. Humphreys *et al.*, "Realization of a multinode quantum network of remote solid-state qubits," *Science*, vol. 372, no. 6539, pp. 259–264, 2021. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.abg1919

[12] T. van Leent, M. Bock, F. Fertig, R. Garthoff, S. Eppelt *et al.*, "Entangling single atoms over 33 km telecom fibre," *Nature*, vol. 607, no. 7917, pp. 69–73, 2022.

[13] P. Drmota, D. Main, D. Nadlinger, B. Nichol, M. Weber, E. Ainley, A. Agrawal, R. Srinivas, G. Araneda, C. Ballance *et al.*, "Robust quantum memory in a trapped-ion quantum network node," *arXiv preprint arXiv:2210.11447*, 2022.

[14] R. Stockill, M. Stanley, L. Huthmacher, E. Clarke, M. Hugues *et al.*, "Phase-tuned entangled state generation between distant spin qubits," *Physical review letters*, vol. 119, no. 1, p. 010503, 2017.

[15] G. P. Agrawal, *Fiber-Optic Communication Systems, Fourth Edition*. John Wiley & Sons, Inc., 2010.

[16] M. J. Clark, O. Alia, R. Wang, S. Bahrani, M. Peranić *et al.*, "Entanglement distribution quantum networking within deployed telecommunications fibre-optic infrastructure," in *Quantum Technology: Driving Commercialisation of an Enabling Science III*, M. J. Padgett, K. Bongs, A. Fedrizzi, and A. Politi, Eds., vol. 12335, International Society for Optics and Photonics. SPIE, 2023, p. 123350E. [Online]. Available: https://doi.org/10.1117/12.2645095

[17] P. Kumar, "Quantum frequency conversion," *Opt. Lett.*, vol. 15, no. 24, pp. 1476–1478, Dec 1990. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?URI=ol-15-24-1476

[18] K. Zhang, J. He, and J. Wang, "Two-way single-photon-level frequency conversion between 852 nm and 1560 nm for connecting cesium d2 line with the telecom c-band," *Opt. Express*, vol. 28, no. 19, pp. 27 785–27 796, Sep 2020. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-19-27785

[19] J. S. Pelc, L. Yu, K. D. Greve, P. L. McMahon, C. M. Natarajan *et al.*, "Downconversion quantum interface for a single quantum dot spin and 1550-nm single-photon channel," *Opt. Express*, vol. 20, no. 25, pp. 27 510–27 519, Dec 2012. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-20-25-27510

[20] Z. Xie, K. H. Luo, K. C. Chang, N. C. Panoiu, H. Herrmann *et al.*, "Efficient c-band single-photon upconversion with chip-scale ti-indiffused pp-linbo3 waveguides," *Appl. Opt.*, vol. 58, no. 22, pp. 5910–5915, Aug 2019. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-58-22-5910

[21] T. Walker, K. Miyanishi, R. Ikuta, H. Takahashi, S. V. Kashanian *et al.*, "Long-distance single photon transmission from a trapped ion via quantum frequency conversion," *Physical review letters*, vol. 120, no. 20, p. 203601, 2018.

[22] M. Bock, P. Eich, S. Kucera, M. Kreis, A. Lenhard *et al.*, "High-fidelity entanglement between a trapped ion and a telecom photon via quantum frequency conversion," *Nature communications*, vol. 9, no. 1, p. 1998, 2018.

[23] R. Ikuta, T. Kobayashi, T. Kawakami, S. Miki, M. Yabuno *et al.*, "Polarization insensitive frequency conversion for an atom-photon entanglement distribution via a telecom network," *Nature communications*, vol. 9, no. 1, p. 1997, 2018. [Online]. Available: https://www.nature.com/articles/s41467-018-04338-x

[24] S. Ramelow, A. Fedrizzi, A. Poppe, N. K. Langford, and A. Zeilinger, "Polarization-entanglement-conserving frequency conversion of photons," *Phys. Rev. A*, vol. 85, p. 013845, Jan 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.85.013845

[25] G. Avis, F. Rozpędek, and S. Wehner, "Analysis of multipartite entanglement distribution using a central quantum-network node," *Phys. Rev. A*, vol. 107, p. 012609, Jan 2023. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.107.012609

[26] S. Wengerowsky, S. K. Joshi, F. Steinlechner, H. Hübel, and R. Ursin, "An entanglement-based wavelength-multiplexed quantum communication network," *Nature*, vol. 564, no. 7735, pp. 225–228, 2018. [Online]. Available: https://www.nature.com/articles/s41586-018-0766-y

[27] S. K. Joshi, D. Aktas, S. Wengerowsky, M. Lončarić, S. P. Neumann *et al.*, "A trusted node-free eight-user metropolitan quantum communication network," *Science Advances*, vol. 6, no. 36, p. eaba0959, 2020. [Online]. Available: https://www.science.org/doi/abs/10.1126/sciadv.aba0959

[28] C. H. Bennett, G. Brassard, and N. D. Mermin, "Quantum cryptography without bell's theorem," *Phys. Rev. Lett.*, vol. 68, pp. 557–559, Feb 1992. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.68.557

[29] R. Wang, M. J. Clark, S. K. Joshi, S. Bahrani, O. Alia *et al.*, "Optimum switching scenario analysis in a dynamic entanglement network," 2023 Optical Fiber Communications Conference and Exhibition (OFC), in press.

[30] Z. Huang, S. K. Joshi, D. Aktas, C. Lupo, A. O. Quintavalle *et al.*, "Experimental implementation of secure anonymous protocols on an eight-user quantum key distribution network," *npj Quantum Information*, vol. 8, no. 1, p. 25, 2022.

[31] Y. Pelet, I. V. Puthoor, N. Venkatachalam, S. Wengerowsky, M. Lončarić *et al.*, "Unconditionally secure digital signatures implemented in an eight-user quantum network," *New Journal of Physics*, vol. 24, no. 9, p. 093038, 2022.

[32] R. Wang, O. Alia, M. J. Clark, S. Bahrani, S. K. Joshi *et al.*, "A dynamic multi-protocol entanglement distribution quantum network," in *2022 Optical Fiber Communications Conference and Exhibition (OFC)*, 2022, pp. 1–3.

[33] M. Peranić, M. Clark, R. Wang, S. Bahrani, O. Alia *et al.*, "Polarization compensation methods for quantum communication networks," *arXiv preprint arXiv:2208.13584*, 2022. [Online]. Available: https://arxiv.org/abs/2208.13584

PAPERS

Monolithic Receiver with Nine Single-Photon Avalanche Diodes

A. Kuttner, S. Marschner, B. Steindl, K. Schneider-Hornstein, M. Hofbauer and H. Zimmermann
TU Wien, EMCE, Vienna, Austria
horst.zimmermann@tuwien.ac.at

***Abstract* – An optical receiver containing nine single-photon avalanche diodes (SPADs) and nine active quenching circuits is presented. This optical receiver was fabricated in a 0.35µm high-voltage CMOS technology without any process modifications. The SPADs avoid the excess noise of linear-mode avalanche diode receivers and allow for a better optical sensitivity. With 680 nm light, the 9-SPAD receiver achieves a sensitivity of -56.8 dBm for a bit error ratio of 2×10^{-3} at a data rate of 50 Mb/s. Compared to a 4-SPAD receiver fabricated in the same technology and introduced in 2020, the sensitivity is improved by 1.7 dB. The remaining gap to the quantum limit, which is at -73.7 dBm for 50 Mb/s, is reduced to 16.9 dB. A comparison to the state-of-the-art of SPAD receivers is included.**

Keywords – optical receiver; integrated optoelectronics; single-photon avalanche diode; high-voltage CMOS.

I. INTRODUCTION

Receivers with single-photon avalanche diodes (SPADs) are very interesting because of their very high gain [1], which eliminates the excess noise of avalanche photodiodes (APDs) in linear-mode [2]. A better sensitivity of SPAD receivers is therefore possible than for linear-mode APD receivers. Due to the very high gain of SPADs in the Geiger mode of larger 10^6, a digital output signal of the SPAD is possible and a digital gate (e. g. an inverter) can detect the absorption of a photon [3]. So, the excess noise of linear-mode APDs and electronic noise of an amplifier in a receiver is not relevant anymore.

However, SPADs are not ideal devices. Thermally generated charge carriers can trigger dark counts. During a Geiger mode event, traps can be filled and those can release charge carriers statistically with a certain time constant, which causes afterpulses. A long dead time and/or a fast active quenching of the SPADs have to be applied to reduce afterpulsing. So, there still can be Geiger mode events not caused by photon absorption. Already because of this non-ideal behavior of SPADs, one SPAD and absorption of one photon in a "1"-bit will not be sufficient to obtain a bit error ratio below 2×10^{-3}, which is needed to perform error correction [4]. Therefore, a kind of coincidence detection of at least two or more photons in a "1"-bit will be necessary.

If these issues can be handled, the final sensitivity is defined by the Poisson statistics of light sources like light emitting diodes and laser diodes, which causes the so-called quantum limit [5].

The first SPAD receiver containing 32×32 SPADs was designed for a high dynamic range and its sensitivity of −31.7 dBm at 100 Mbit/s and 450 nm wavelength [6], where the quantum limit is at −63.6 dBm, was quite moderate. A receiver with 100 SPADs was introduced for 20 Mbit/s in [7], but unfortunately no sensitivity was reported. We had investigated a 4-SPAD receiver following the simple idea that 4 SPADs with an afterpulsing probability of 10% each should result in a bit error ratio of 10^{-4}, assuming afterpulsing as the dominant effect [8]. This 4-SPAD receiver was fabricated in 0.35 µm PIN-photodiode CMOS technology. Later, we realized a 4-SPAD receiver [9] in the same 0.35 µm high-voltage CMOS technology as we use for the 9-SPAD receiver in this work.

In section II, the used SPAD and the receiver will be described. Section III reports the properties of the SPADs implemented in the 9-SPAD receiver. The receiver performance is introduced in section IV. In section V, the suggested 9-SPAD receiver is compared to the state of the art of SPAD receivers. Conclusions are drawn in section VI.

II. SPAD AND RECEIVER

The SPAD presented in Fig. 1 does not need any process modification. It can be fabricated in the XH035 (0.35 µm) high-voltage CMOS process of XFAB using the standard ASIC procedure.

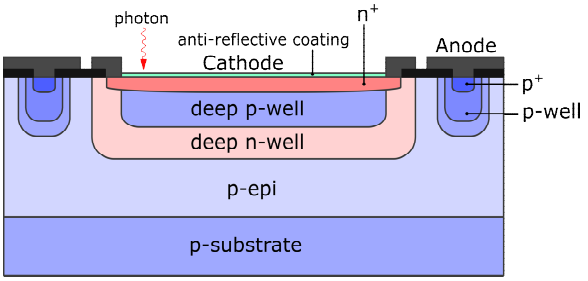


Fig. 1. Cross section of the integrated SPAD (not in scale).

The multiplication zone extends from the n+/deep-p-well junction into the deep p-well. The deep n-well compensates part of the deep p-well doping and part of the standard epitaxial layer doping. In such a way, this reachthrough avalanche photodiode allows for the depletion of the p-epi layer at an acceptable reverse voltage. The breakdown voltage of this SPAD is about 68 V [10], which is well within the isolation capabitlity of this process. So, this SPAD can be integrated together with transistors on the same chip. This SPAD was characterized in [11]. The photon detection probability (PDP) of the SPAD increases with the excess bias voltage. Thanks to the anti-reflection coating being available in the 0.35 µm high-voltage CMOS process used, the PDP for 635 nm and 6.6 V excess bias voltage is 45% [9].

The active quenching circuit described in detail in [12] was slightly improved. The principle of the quenching circuit is shown in Fig. 2. When the SPAD fires, the avalanche current increases and the voltage drop across R_C (actually realized as an active resistor MOSFET) grows. When the threshold V_{REF} is undercut, the comparator decides and the quenching switch is closed rapidly. To enable a large excess bias voltage of the SPAD of up to 6.6 V, the switches have to withstand this voltage. Therefore, the switches are realized as cascoded MOSFETs [12]. In such a way the fast 3.3 V transistors can be used although the excess bias of the SPADs can be up to 6.6 V. The closed quenching switch discharges the SPAD (slightly) below its breakdown voltage and the avalanche ends. After the dead time the quenching switch is opened and the reset switch is closed to charge the SPAD to the excess bias voltage again to make it sensitive for a new photon detection again. During the detection period, both switches are open. Via a bias input the dead time of the SPAD can be controlled in the range of 5.8 ns to 34 ns. It should be mentioned that the output pulses of the quencher have a shorter duration than the dead time [12].

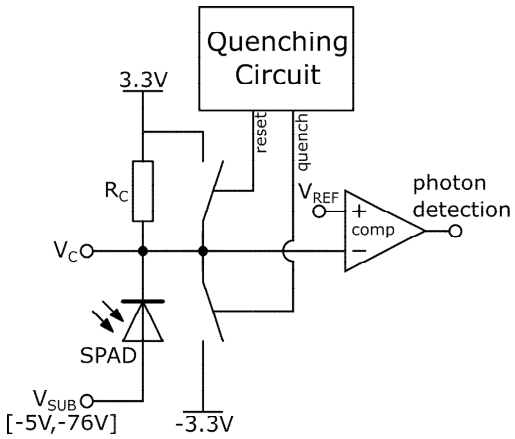


Fig. 2. Principle of the implemented active quenching circuit with comparator.

According to the supply voltages of -3.3 V and +3.3 V, which define the maximum excess bias of 6.6 V, the substrate potential of -71.3 V corresponds to a SPAD breakdown voltage of 68 V. If the excess bias of 4 V is desired, the substrate potential has to be set to -68.7 V.

In the presented 9-channel SPAD receiver, nine SPADs are equipped with one active quenching circuit each. The nine quencher outputs are connected to a bondpad each. The layout is introduced in Fig. 3 (upper part). The receiver chip was fabricated in 0.35 µm high-voltage CMOS (XH035) and the chip-photo is presented in Fig. 3 (bottom part). The size of this receiver chip is 2120 µm × 1640 µm. The diameter of the 9-segment SPAD detector is 490 µm and the optical fill factor is 49%. This large diameter was chosen because wide gaps between the concentric SPADs reduce the optical crosstalk. The concentrical segments were easier to layout than nine equal sectors or pieces of a tart. In addition, the distance between neighbouring pieces (especially closer to the center) would be too small concerning crosstalk. The SPAD segments are covered by an anti-reflection coating. The substrate contact ring surrounds the nine SPADs.

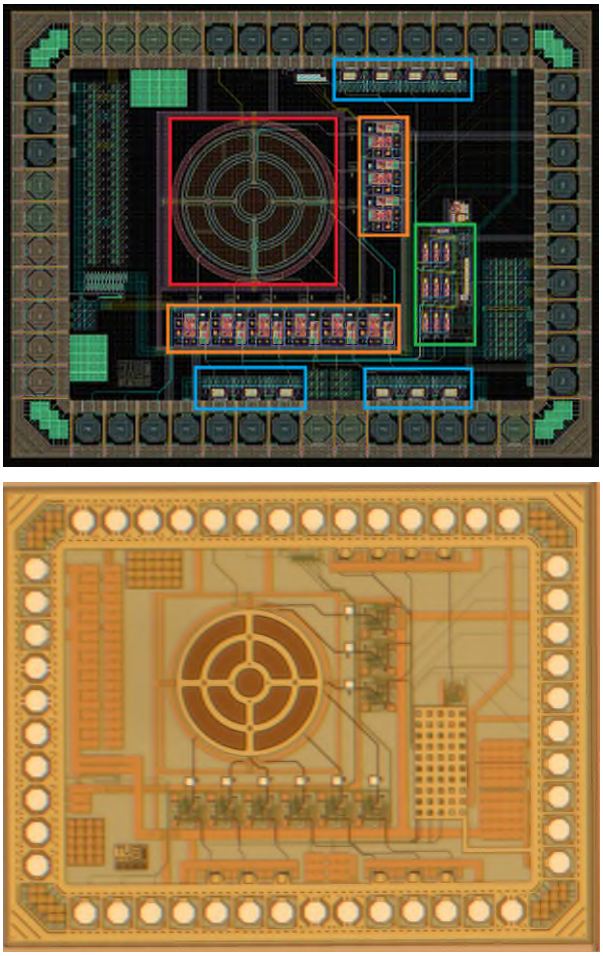


Fig. 3. Layout (top) and chip-photo (bottom) of receiver chip with 9 SPAD segments arranged in circular shape.

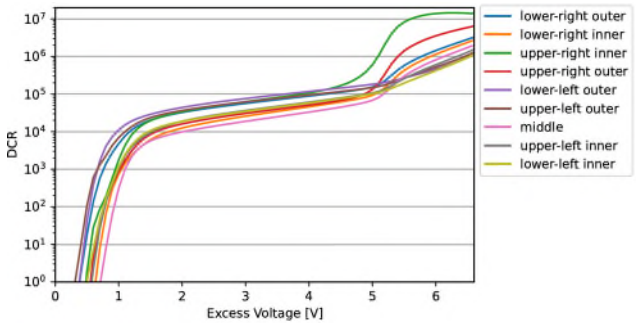


Fig. 4. Dark count rate of all 9 SPADs in the receiver.

I. PROPERTIES OF SPADS

Since all nine quencher outputs are available, each SPAD could be characterized concerning dark count rate and afterpulsing probability. A measurement time of 100 s was used. The measured dark count rates (DCR) are presented in Fig. 4 in dependence on the excess bias voltage. The curves indicate that the SPADs can be operated for acceptable BERs of the receiver up to an excess bias of about 4.5 V.

The measured afterpulsing probabilities (APPs) of all 9 SPADs are shown in Fig. 5. The upper-right inner SPAD shows the worst APP as the worst DCR. Therefore, the highest possible excess bias voltage of 6.6 V cannot be used in the receiver.

Another source of digital noise in SPAD arrays is optical crosstalk. This effect happens because during impact ionization photons are emitted. These photons can be detected in other SPADs and trigger avalanche effects there although no “true” or “signal” photons fell into these SPADs. Optical crosstalk in a SPAD receiver increases the bit error ratio [13]. Therefore, the circular 9-segment SPAD detector was also characterized with respect to its optical crosstalk probability (OCTP). Figure 6 shows the total OCTP, i. e. between all of the nine SPADs. The total OCTP is below about 20% in the excess bias range relevant for the BER measurements presented.

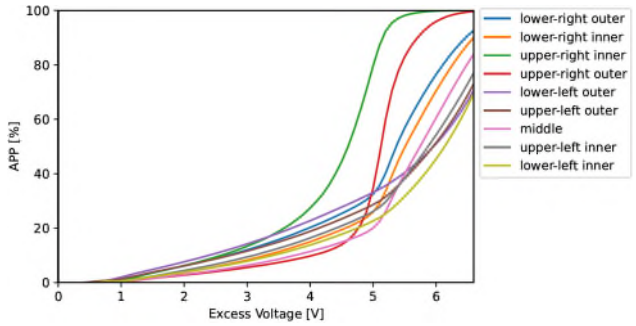


Fig. 5. Afterpulsing probability of all 9 SPADs in the receiver.

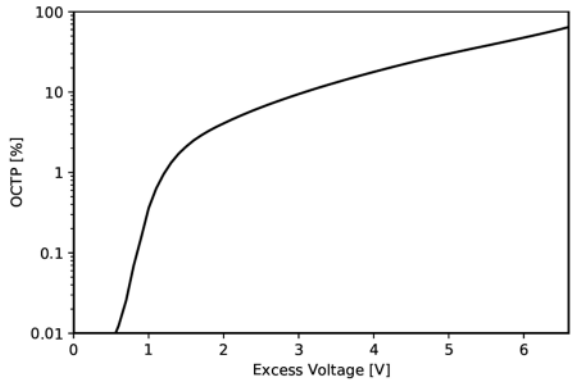


Fig. 6. Total crosstalk probability within the 9 SPADs of the receiver.

When having the large diameter of the 9-segment SPAD array in mind, the DCR, APP and OCTP seem to be reasonably low.

II. RECEIVER PERFORMANCE

As light source for the receiver experiments a high-extinction-ratio direct modulated laser diode with a wavelength of 680 nm was used. The receiver was placed inside a dark box and an optical fiber fed by the laser diode was adjusted to the center of the 9-segment SPAD array. The height of the fiber end over the reciver was set that a minimum BER was obtained. The optical power was measured with a Thorlabs PM200 power meter and the chip was held at constant temperature with a Thorlabs TC200C TEC controller.

The measurements were done via LabView. The program ran on an NI PXIe-8840 controller. The 9 outputs of the quenchers were connected via an interface board to a NI-6589 I/O module. A NI PXIe-7972R FlexRIO FPGA generated the 50 Mb/s return-to-zero (RZ) PRBS7 signal to modulate the laser source. The RZ signal had a duty ratio of 50%. Two source measure units (NI PXIe-4132 and NI PXI3-4145) delivered the substrate voltage (anode voltage of the SPADs), the supply voltages, the control voltage for the dead time and V_{REF} for the comparators to the SPAD receiver chip. The FPGA processed the 9 data streams of the SPAD receiver according to the principles of a latched and non-latched method (see Fig. 7). The latched method implies that a photon detection after the dead time but within the same bit is not taken into account. Whereas a photon detection after the dead time is considered with the non-latched method. The threshold of photon counts within each bit could be set in the range of 1 to 9 The obtained bits were compared by the FPGA with the PRBS stream used for laser modulation. In this way, the bit errors were determined efficiently directly within the FPGA. The dead time was optimized for minimum BER. The best performance was obtained with dead times of somewhat less than 10 ns. The measurement time was 1s.

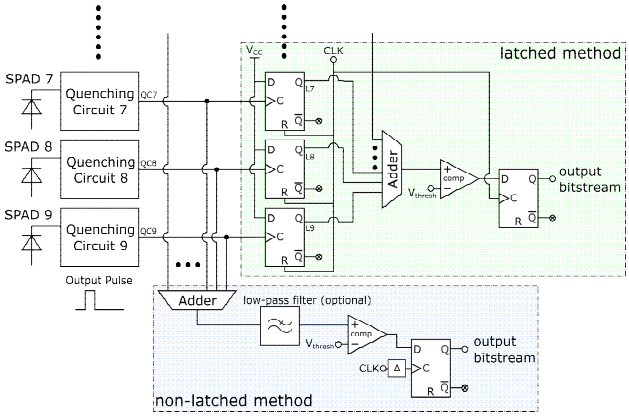


Fig. 7. SPAD receiver with 9 detectors, quenching circuits and postprocessing realized in an FPGA with a latched and non-latched method. The digital adder counts photons with equal weights.

Figure 8 shows the obtained BERs at 50 Mb/s for the non-latched processing for different thresholds in dependence on the average optical power. The sensitivity of the receiver is best for the thresholds of 2 and 3. The sensitivity for the threshold of 2 is -56.81 dBm for the non-latched processing method.

Figure 9 presents the BERs at 50 Mb/s for the latched processing for all thresholds that reached the BER limit of 2×10^{-3} for error correction. The best thresholds are 3 and 4 photon counts in a "1" bit. The sensitivity for the threshold of 3 is -56.26 dBm for the latched processing method. We can conclude that the non-latched processing method is better appropriate for determining the receiver sensitivity.

For the curves in Figs. 8 and 9, the substrate voltage (anode) voltage of the SPADs, i. e. the excess bias voltage, was held constant at the values for which the BER of 2×10^{-3} was achieved at the lowest optical power. The BER could be reduced (see Figs. 10 and 11), when the optimum substrate voltage, i. e. the optimum excess bias voltage, was used for each optical power.

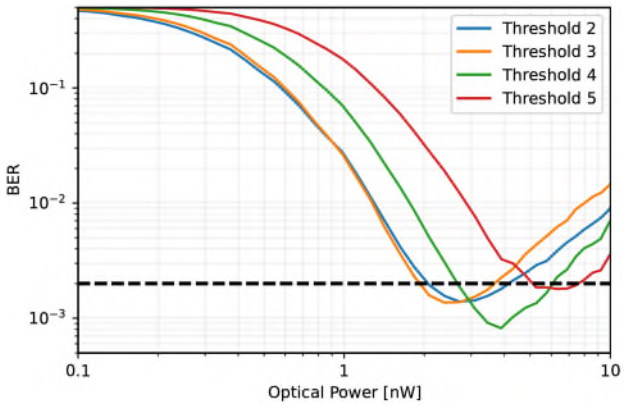


Fig. 8. Bit error ratio for the non-latched processing.

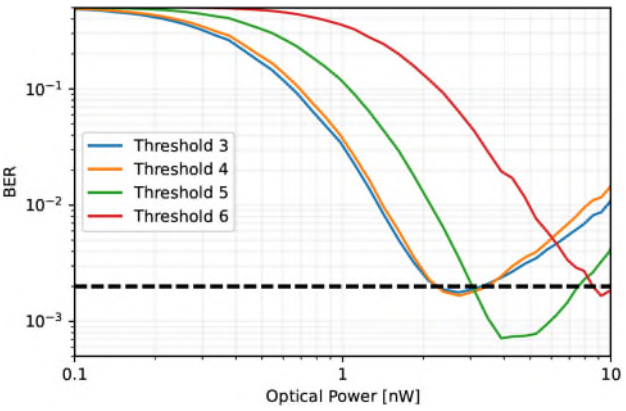


Fig. 9. Bit error ratio for the latched processing.

With this additional degree of freedom, the minimum BER of 10^{-4} is achieved at about 9 nW optical power with a threshold of three photon detections in the case of the non-latched processing. With latched processing the minimum BER of 2×10^{-4} is present for about 7 nW optical power with the threshold of four photon counts. The excess bias voltages for BER<2×10^{-3} are lower than about 4 V for both cases (non-latched and latched processing). For higher excess bias voltages the bit errors increase because of rising DCR, APP and OCTP.

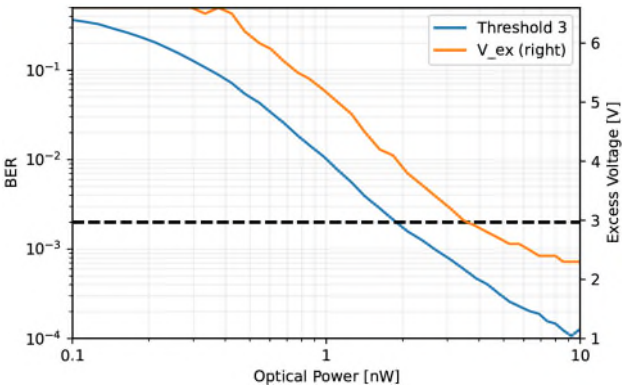


Fig. 10. Bit error ratio for the threshold of 3 photons (left) and excess bias voltage (right) for the non-latched processing.

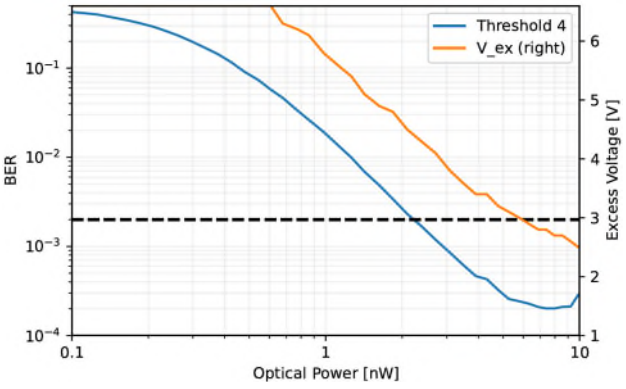


Fig. 11. Bit error ratio for the threshold of 4 photons (left) and excess bias voltage (right) for the latched processing.

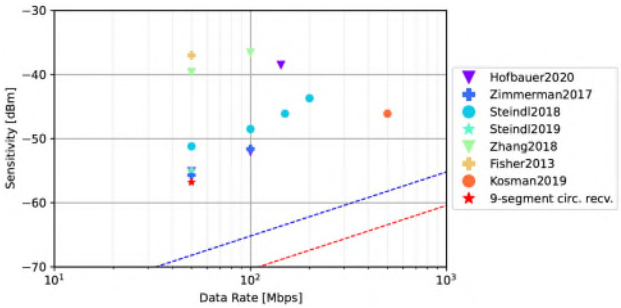


Fig. 12. Comparison of the sensitivity to the state of the art with quantum limits for BER=2×10^{-3} (blue line) and for BER=10^{-9} (red line).

I. COMPARISON

The 9-SPAD receiver (represented by the red star in Fig. 12) improves the sensitivity compared to a 4-SPAD receiver [9] (-55.1 dBm, represented by the violet triangle in Fig. 12) in the same 0.35 µm high-voltage CMOS technology at 50 Mb/s by 1.7 dB, which corresponds to a reduction of the necessary optical power for BER=2×10^{-3} by about 30%. The reason for this is that still a few SPADs from the 9 SPADs can detect photons if some are already dead after detection in the same bit. Therefore, less optical power is sufficient compared to the 4-SPAD receiver.

Compared to the 4-SPAD receiver in 0.35 µm PIN-photodiode CMOS [8] (represented in Fig. 12 by the blue +) with a sensitivity of -55.7 dBm at 50 Mb/s, the 9-SPAD receiver shows an improvement in sensitivity of 1.1 dB.

The 4-SPAD receiver of [14] (represented by the blue circle in Fig. 12) in 0.35 µm PIN-photodiode CMOS technology has a sensitivity of -51.2 dBm at 50 Mb/s. Another 4-SPAD receiver in 0.35 µm CMOS achieved a sensitivity of -54.9 dBm [15] (see the blue star in Fig. 12).

A receiver in 0.18 µm CMOS containing 60 SPADs achieved a sensitivity of -39.6 dBm at 50 Mb/s [16] (represented by the green triangle in Fig. 12). This sensitivity was mainly due to the low optical fill factor of 3.2%.

A 64×64 SPAD receiver in 0.13 µm CMOS showed a sensitivity of -41.6 dBm at 500 Mb/s [17]. This data point is the orange circle in Fig. 12.

II. CONCLUSION

There were several attempts to develop SPAD receivers reducing the gap to the quantum limit. Increasing the number of SPADs in a receiver improves the sensitivity and enables higher data rates. The receiver circuit and system complexity increase in turn. But dark counts, afterpulsing, optical crosstalk and the dead time prohibit to reach the quantum limit. The photon protection probability has to be increased to approach the quantum limit closer – however without increasing the dark count rate, afterpulsing probability and optical crosstalk. Finally, these issues will require higher purity in CMOS processing. We should, however, be aware that silicon technology and wafer fabrication are already at a very high level.

REFERENCES

[1] Cova, S.; Ghioni, M.; Lacaita, A.; Samori, C.; Zappa, F. (1996). "Avalanche photodiodes and quenching circuits for single-photon detection". *Applied Optics*. **35** (12): 1956–76.

[2] E. Säckinger, *Broadband Circuits for Optical Fiber Communication*, Hoboken, NJ, USA: Wiley, 2005, ch. 4, sec. 3, pp. 54–66.

[3] Y. Xu, J. Lu and Z. Wu, "A Compact High-Speed Active Quenching and Recharging Circuit for SPAD Detectors," in *IEEE Photonics Journal*, vol. 12, no. 5, pp. 1-8, Oct. 2020, Art no. 6803208, doi: 10.1109/JPHOT.2020.3015872.

[4] B. Sklar, *Digital Communications: Fundamentals and Applications*. Englewood Cliffs, NJ, USA: Prentice Hall, 2001.

[5] K. J. Ebeling, *Integrated Optoelectronics*, Berlin, Germany: Springer, 1993.

[6] E. Fisher, I. Underwood and R. Henderson, "A Reconfigurable Single-Photon-Counting Integrating Receiver for Optical Communications," in *IEEE Journal of Solid-State Circuits*, vol. 48, no. 7, pp. 1638-1650, July 2013, doi: 10.1109/JSSC.2013.2253222.

[7] D. Chitnis and S. Collins, "A SPAD-based photon detecting system for optical communications," *J. Lightw. Technol.* vol. 32, no. 10, pp. 2028–2034, May 2014.

[8] H. Zimmermann, B. Steindl, M. Hofbauer, " Integrated fiber optical receiver reducing the gap to the quantum limit", *Sci. Rep.* vol. 7, art. no. 2652, 2017, doi: 10.1038/s41598-017-02870-2.

[9] Hofbauer, B. Steindl, H. Zimmermann, "Fully Integrated Optical Receiver Using Single-Photon Avalanche Diodes in High-Voltage CMOS," SPIE Optical Engineering 59(7), 070502-1 -- 070502-6 (2020). DOI: 10.1117/1.OE.59.7.070502.

[10] B. Steindl, R. Enne, S. Schidl and H. Zimmermann, "Linear Mode Avalanche Photodiode with High Responsivity Integrated in High-Voltage CMOS," in *IEEE Electron Device Letters*, vol. 35, no. 9, pp. 897-899, Sept. 2014, doi: 10.1109/LED.2014.2336678.

[11] Hofbauer, B. Steindl, K. Schneider-Hornstein, H. Zimmermann, 'Performance of High-Voltage CMOS Single-Photon Avalanche Diodes with and without Well-Modulation Technique', *SPIE Opt. Eng.*, Vol. 59, No. 4, 040502-1 -- 040502-8 (2020). DOI: 10.1117/1.OE.59.4.040502.

[12] R. Enne, B. Steindl, M. Hofbauer and H. Zimmermann, "Fast Cascoded Quenching Circuit for Decreasing Afterpulsing Effects in 0.35- µ m CMOS," in *IEEE Solid-State Circuits Letters*, vol. 1, no. 3, pp. 62-65, March 2018, doi: 10.1109/LSSC.2018.2827881.

[13] H. Mahmoudi, M. Hofbauer, B. Steindl, K. Schneider-Hornstein and H. Zimmermann, "Modeling and Analysis of BER Performance in a SPAD-Based Integrated Fiber Optical Receiver," in *IEEE Photonics Journal*, vol. 10, no. 6, pp. 1-11, Dec. 2018, Art no. 7908411, doi: 10.1109/JPHOT.2018.2875519.

[14] B. Steindl, M. Hofbauer, K. Schneider-Hornstein, P. Brandl and H. Zimmermann, "Single-Photon Avalanche Photodiode Based Fiber Optic Receiver for Up to 200 Mb/s," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1-8, March-April 2018, Art no. 3801308, doi: 10.1109/JSTQE.2017.2764682.

[15] B. Steindl, „Einzelphotonen-Lawinendioden für integrierbare optische Empfänger", Ph. D. thesis, TU Wien, 2019.

[16] L. Zhang *et al.*, "A Comparison of APD- and SPAD-Based Receivers for Visible Light Communications," in *Journal of Lightwave Technology*, vol. 36, no. 12, pp. 2435-2442, 15 June15, 2018, doi: 10.1109/JLT.2018.2811180.

[17] J. Kosman *et al.*, "29.7 A 500Mb/s -46.1dBm CMOS SPAD Receiver for Laser Diode Visible-Light Communications," *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, San Francisco, CA, USA, 2019, pp. 468-470, doi: 10.1109/ISSCC.2019.8662427.

Occupancy Determination by Backscattered Visible Light Sensing

Omar El Ouahabi*, Christian Fragner *, Andreas P. Weiss* and Erich Leitgeb**
* Institute for Sensors, Photonics and Manufacturing Technologies – Smart Connected Lighting, JOANNEUM RESEARCH Forschungsges.mbH, Pinkafeld, Austria
** Institute of Microwave and Photonic Engineering, Technical University Graz, Austria
omar.elouahabi@joanneum.at
christian.fragner@joanneum.at
andreas-peter.weiss@joanneum.at
erich.leitgeb@tugraz.at

***Abstract* - Demand driven control of heating or cooling, has the potential to reduce the energy consumption of buildings to a significant portion. One of the main cornerstones of such a demand driven control is the accurate determination of occupancy in the building or sub-areas of the building. In recent years many different approaches, founded on various technologies and sensing principles have been used to establish an occupancy detection system. Especially in existing buildings, retrofitting such a system can be associated with high installation effort and cost. Furthermore, some technological approaches, mainly camera-based systems are inflicting concerns in regard to the privacy of the users. In this manuscript, we present and discuss an approach for a lighting fixture that performs person recognition and occupancy determination based on the technology of Visible Light Sensing. In our implemented system by solely capturing reflections of the visible light, caused by bypassing persons, the occupancy, the movement direction and the walking speed can be accurately determined. Based on comprehensive real world experiments we will show how our system, fully integrated into an off-the shelf luminaire, is able to operate in different scenarios and ambient light conditions and only requires minimal installation effort without causing any privacy concerns.**

***Keywords* - Visible Light Sensing, Occupancy detection, Photonic remote Sensing**

I. INTRODUCTION

One of the key elements to fight the energy crisis is to make buildings more energy efficient in their operation. Worldwide, buildings are responsible for up to 40% of total energy consumption. Therefore, many efforts are made to reduce the energy consumed by buildings. To achieve this without reducing the comfort of the people in the building, technological innovations are needed in various areas. Reliable occupancy information is one of the most important aspects when it comes to tailoring the operation of Heating, Ventilation and Air Conditioning (HVAC) systems of a building. In [1] it is stated that implementing occupancy detection systems has the potential to realize energy savings between 10% and 40%.

In regard to new buildings, in planning or under construction, such occupancy detection systems are often already integrated, but especially in existing buildings the subsequent installation is often not feasible due to high installation cost.

In terms of presence and occupancy detection, the main approaches can be divided into three main categories, where each category utilizes again various methods and underlying sensor technologies.

The first category are so called active systems, where the user has to carry an active/semi-active device. The main technologies used are Wi-Fi [2,3], Bluetooth [4] or RFID [5]. The main principle is that the corresponding access point (e.g. Wi-Fi Router) recognizes the device or tag worn by the user and consequently, by counting the "connected" users, the occupancy of an area can be inferred. One advantage of such systems is that the existing infrastructure (e.g. Wi-Fi) can be used to gather this data, but this also outlines one of the main disadvantages. If a user does not carry the device, he also cannot be "counted".

The second category is based on environmental sensors. This category includes CO2 sensors or smart-meter based approaches. [6,7] describe, how by monitoring the CO2 concentration the number of persons in room can be inferred, since every human emits CO2 by breathing. The main drawbacks of this approach are that the emission of CO2 strongly depends on the activity and condition of the person [8] and that such systems have a quite long reaction time of about 20 minutes [9]. [10] and [11] demonstrate how the energy consumption, monitored by a smart meter can be used to infer occupancy or determine the activity of persons in a building.

The third category can be called passive systems. In this context, passive refers to the fact that the person is recognized based on physical effects, such as heat emitted by the person, its physical presence etc. Popular approaches in this category are Passive Infrared Sensors (PIR) [12], Ultrasonic Sensors [13] or camera-based systems [14]. While PIR and Ultrasonic based approaches show good performance in occupancy detection and the components are low cost, but nevertheless are also highly relying on an optimal placement [15], which again can lead to higher installation effort in existing buildings. Camera-based systems [16] are capable of providing detailed occupancy data, but nevertheless there are always privacy concerns.

One approach that can also be assigned to this category is to use the lighting infrastructure in combination with Visible Light Sensing (VLS) for inferring the occupancy. This approach has the main advantages that artificial lighting is almost omnipresent throughout buildings, the utilized components such as LEDs and Photodiodes are low-cost, no privacy issues anticipated and that the obligatory lighting remains unchanged.

II. Background & Solution Approach

In the most general sense, the technology of Visible Light Sensing (VLS) can be described as the approach to infer information from light received at a photosensitive device. VLS can be distinguished in 4 main categories [17], whereas in this work we will focus on backscattered scenarios, also referred to as Backscattered Visible Light Sensing (BVLS). Such scenarios are defined by the essential approach that no Line-of-Sight is given between the light source and the receiving element and that consequently only reflections from the object or person of interest are acquired at the receiving element. This approach has the big advantage that the light source and the receiving elements can be integrated into the same module or housing, without the need to place components into the rooms infrastructure, as described in [18].

In [19] the approach of presence detection based on BVLS has been described. In this work, the authors show that presence detection of humans can be achieved by detecting changes in the received signal strength (RSS) at a photodiode (PD). Furthermore, comprehensive experiments have been conducted to show the influence of different clothes, walking speeds and placement of light sources and receivers. This work shows in so-called pass-by experiments that a person walking through a hallway can be detected with 100% correct results. In comparison to this work, the results show on the one hand the feasibility of utilizing BVLS, but on the other hand do not investigate further into the influence of ambient light or when more than one person needs to be detected. Furthermore, our work extends the scope to also infer the walking direction of the persons and the walking speed. Especially determining the walking direction is an essential parameter in order to infer occupancy, since by determining the direction, the people entering or leaving an area can be counted.

[20] presents a system where the occupancy of a room can be inferred based on 16 sensing units (luminaires) in a room of 5 m x 6 m. In this work, the LED luminaires on the ceiling serve as both transmitter and as receivers (by putting the LED in reverse bias mode). By cooperative analysis of these 16 luminaires, an occupancy accuracy of 90 % is described. In contrast to our work the approach is based on non-moving persons in a room, as well as other ambient light sources are not present. Our work focuses on settings where people are moving in a hallway towards or away from an area, which in turn gives the big advantage that the occupancy can be inferred by a single adapted luminaire. Furthermore, this has the additional advantage that our luminaire can work stand-alone, meaning that the luminaire itself infers the presence, the walking direction, walking speed and consequently the occupancy.

A. Solution Approach

Our solution approach is based on the following theorems and assumptions:

- A person entering the Field-of-View (FoV) of a photodiode causes a clear deviation in the RSS compared to the RSS in the absence of a person
- The RSS deviation increases during the passage of the person until the person is located directly under the photodiode
- The RSS deviation caused by the person has a clear peak when the person is directly under the photodiode
- The RSS deviation decreases from the peak value in the further passage until the person has left the FoV of the photodiode
- By utilizing more than one photodiode the sequence of the RSS deviation peaks occurring at the respective photodiodes can be clearly separated
- Utilizing the known distance between the photodiodes and the time difference between the detected peaks, the walking speed can be inferred
- Additional ambient light has similar effects on all the utilized photodiodes
- The floor (e.g. material) is assumed to have the same properties under the whole length of the luminaire

Based on these theorems and assumptions, the hardware laboratory sample have been built and the software algorithm implemented, described in the following subsections.

III. Luminaire

The designed and constructed luminaire in our work consists of two main parts. First, we designed and manufactured a Printed Circuit board that is equipped with a visible light sensitive photodiode TEMD5510FX01 from the vendor Vishay, a Lumileds LMZ7-QW50 LED alongside the necessary Transimpedance Amplifier (TIA) to convert the photocurrent of the photodiode into a voltage signal as well as the driver circuitry for the LED. The output of the TIA stage is connected to an in-built Analog-Digital-converter (ADC) of a NUCLEO-WB55RG microcontroller board. In order to achieve a better directionality of the LED and the photodiode, CA10716 reflectors from the vendor Ledil with a spot beam of ~20° were placed over both the LED and the photodiode. Fig. 1 shows a 3D representation of the designed and manufactured PCB.

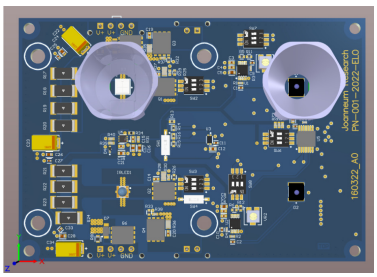


Figure 1: 3D representation of the designed and manufactured PCB

In the following, we integrated 3 of those PCB's into an off-the-shelf luminaire with the size of 120 cm length and 28 cm width. The placement of the PCB's and respectively the LED and photodiodes where chosen to be at the two outer borders of the luminaire and one PCB in the middle of the luminaire. Fig. 2 (left) shows the luminaire with the three integrated PCB's, with descriptive text added to the figure, marking the photodiodes the LEDs. Please note that the other visible reflectors in the picture are placed over other sensors, which are not used in this work and are consequently marked as such. Fig. 2 (right) shows the luminaire finally mounted in one of the laboratory rooms.

As mentioned earlier the output of the TIAs of the photodiodes (PD 1 to PD 3), see Fig. 2 (left), are connected to a NUCLEO-WB55RG microcontroller board also integrated in the luminaire. This microcontroller board runs the algorithm, described in the next subsection.

IV. ALGORITHM

Based on the theorems given in chapter II, the algorithm was implemented in the NUCLEO-WB55RG microcontroller.

Following the first theorem, that a person entering the Field-of-View (FoV) of a photodiode causes a clear deviation in the RSS compared to the RSS in the absence of a person a triggering mechanism was implemented, that is used to determine a "steady" state from a "non-steady" state. The "steady" state in this regard means that no person is currently passing through the luminaire, while the "non-steady" state means that the presence of a moving person is detected. The triggering mechanism is based on a rolling average calculation by an array of 1001 values at the chosen sample rate of 3.2 Ms/s by each photodiode. By comparing the previous value of the rolling average with the current value of the rolling average and applying a threshold level of +/- 1 % difference, the current state is determined. If the threshold is exceeded, the "non-steady" state is triggered by the corresponding photodiode and the next steps in the algorithm are performed. One thing we want to point out in this regard, is that by applying a relative threshold level already one big advantage of our approach can be highlighted, namely being independent from the static reflections caused by the flooring material or the presence or absence of ambient light. As we will show later in our experimental results, therefore no adjustments of changes had to be made to the algorithm although the scenarios differ in terms of present ambient light or flooring material.

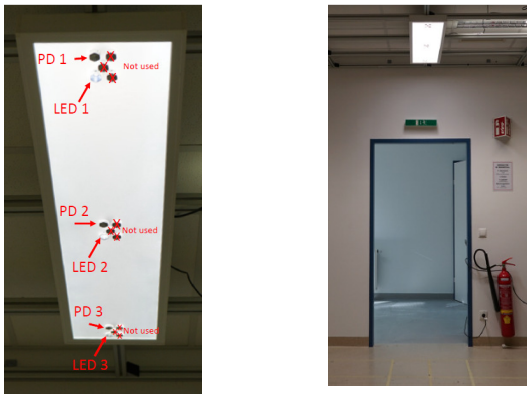


Figure 2: (left) Detailed view of the Luminaire with descriptive text (right) Luminaire mounted in the laboratory

When the "non-steady" state is triggered, by for example, PD 1, the algorithm starts computing the derivate from the acquired values from PD 1. By computing the derivate, a simple peak detection is achieved, since the value of the derivate equals 0 at the peaks of the underlying values. When a peak (derivate equals 0) has been detected the corresponding timestamp and photodiode (e.g. PD 1) is stored. Following the theorem that a person causes a peak at every photodiode it passes, the result will be three stored timestamps from the three utilized photodiodes. In order to illustrate this, please see Fig. 3, that exemplarily shows the passage of a person from PD1 (blue line) to PD 2 (red line) towards PD 3 (green line) under the luminaire with the digital "raw" values of RSS of each PD given on the x-axis and the time in seconds on the y-axis.

As can be easily seen from Fig. 3, by a simple comparison of the stored timestamps the walking direction consequently can be inferred.

Furthermore, Fig. 3 also shows how after the person has left the FoV of the corresponding PD the "steady" state is reached again. With the known length of the luminaire (120 cm) and the stored time between the start and the end of the "non-steady" state, finally the average walking speed of the person is calculated.

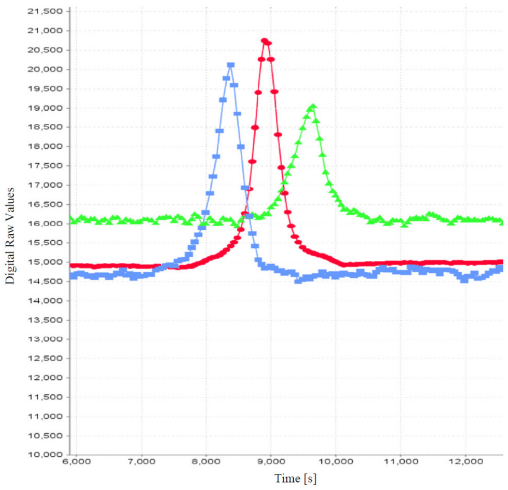


Figure 3: Exemplary data of a person walking under the luminaire

Given by the theorem that additional ambient light has similar effects on all the utilized photodiodes, the algorithm for mitigating the effects of sudden ambient light has been derived. Sudden ambient light changes, for example turning on or off the preexisting room lighting, will cause a change from the "steady" state to the "non-steady" state at all the three photodiodes at the same time. Our designed algorithm therefore monitors if a change from "steady" state to "non-steady" state occurs only at one PD or at all PD's at the same time. If all three PD's exhibit the change, the occupancy determination algorithm is paused until all three PD's have reached the "steady" state again. Please note that this mitigation is aimed at sudden light changes, whereas slow changes in ambient light, e.g. sunlight during the day, must not be mitigated separately, since the rolling average value will of course rise or fall over time, but will not exceed the threshold level in order to trigger a "non-steady" state.

With the described algorithm implemented in the microcontroller of the luminaire, several experiments in two settings where conducted in order to verify the validity of our solution approach. The achieved results are given in the following chapter V.

V. RESULTS

In order to verify our solution approach experimentally, two basic scenarios and different test cases were conducted. We would like to emphasize that during the different test cases and the scenarios, neither the luminaire, nor the implemented algorithm were changed or adapted.

The first scenario, which we further on call "room entrance", was built to resemble the determination of occupancy when persons enter a room, please see Fig. 2 (right). In the first test case, a person (male, 180 cm height, black curly hair), referred to as TP1, enters the room by walking in a straight line directly centered under the luminaire, entering the room and then leaving the room again. Consequently, the occupancy of the room should first be 0, then after TP1 passed under the luminaire should be increased to 1 and then after leaving the room go back to 0. We will refer to this sequence in the following as one Run. The occupancy determination is only considered correct when the numbers are correctly determined throughout one run.

In the first test case, TP1 conducted 40 runs with the ambient light off and walking in a straight line centered under the luminaire. With the ambient light (preexisting room lighting) off and shades of the laboratory closed to block sunlight, our luminaire was the only active light source. The achieved results show that for all 40 runs, the occupancy, the direction and the walking speed was determined correctly. In regard to the walking speed, please note that the speed was varied from walking slowly to running and the assessment was done on a subjective rating if the determined speed corresponds to the performed action.

In a second test case, the same test person repeated the same 40 runs with the ambient room light (fluorescent tubes) on. To illustrate how this affects the lighting situation, the Lux values were measured with a handheld spectrometer at the same points with and without the ambient light. For example under the left upper corner of the luminaire, the measured illumination on the floor was 511.2 Lux without the ambient light and 997.2 Lux with the ambient light on. Nevertheless, we can report that for the 40 conducted runs, again the occupancy, the direction and the walking speed were determined 100 % correct.

These results substantiate that one of our major goals in this work was successfully reached, namely that independent from the overall ambient lighting conditions, the occupancy can be inferred correctly. In order to further investigate the robustness of our approach, further test cases, where the test person is not passing centered under the luminaire were conducted. In these experiments, the test person was walking displaced by 40 cm to the left from the centered line for 20 runs and 40 cm to the right for 20 runs, with the ambient light on. The results show that also when the test person is not walking centered under the luminaire, again 100% correct results in terms of occupancy, direction and speed approximation was achieved.

In order to experimentally verify the applicability of our solution in regard to the second major goal, namely being independent from the persons in terms of hair, clothes and height, we introduced a second test person (male, height 175 cm, baldhead), further on referred to as TP2.

In these test cases, the two test persons were entering and leaving the room one after another with a distance between TP1 and TP 2 of 1 meter. First 20 runs were conducted in the absence of ambient light, and then 20 runs in the presence of the ambient room lighting. The achieved results show that in 17 of the 20 runs and respectively 18 of the 20 runs the occupancy, the direction and the walking speed were determined correctly. Investigating the incorrect runs shows that the direction was determined correctly, but that the occupancy was incorrect by 1 person. This means instead of 2 persons entering and leaving the room the occupancy count was only 1 person. This miscounting was caused by the circumstance that the persons were walking closer together as the targeted 1m. Nevertheless, the results still show that our occupancy determination renders good results with the two differing test persons and with and without ambient room light present.

For the second basic scenario, the luminaire was mounted in the corridor towards our laboratory rooms. Fig. 4 shows this setup, where we would like to point again that neither the hardware nor the algorithm was changed, only the luminaire was relocated.



Figure 4: Corridor experimental setup

In this second scenario, the test cases were repeated with the same procedure and by the same test persons TP1 and TP2. Please note that in this scenario the ambient light (sunlight + preexisting fluorescent tubes) was always present (see Fig. 4).

The achieved results are summarized in Table 1.

TABLE I. RESULTS CORRIDOR SCENARIO

Corridor scenario		
Test case	***Runs***	***Correct***
TP1, centered	40	40
TP1, 20 cm to the left	20	20
TP1, 20 cm to the right	20	20
TP1 and TP2, 1m distance	20	19

With almost identical results as compared to the "room entrance" scenario (see Fig. 2 (right)), we can show that our built occupancy determination successfully works also in much changed scenarios (flooring material, furniture, etc.) without the need for adaption or recalibration, etc.

One limitation that was encountered during the experiments, as mentioned earlier, is the miscounting of persons when the moving persons have a distance of below 1 meter. Overcoming this limitation is one of the main goals in our future work. The key starting point to solve this, is to narrow the FoV of the utilized photodiodes. Based on initial tests with adapted reflectors, we can report that the achieved results are very promising in overcoming this limitation.

VI. CONCLUSION

In this manuscript, we presented an approach for realizing an occupancy determination system based on the technology of Backscattered Visible Light Sensing. Based on low-cost off-the-shelf components we were able to fabricate a luminaire that with fully integrated photodiodes can be easily installed and consequently can reduce the installation effort to a minimum. The implemented algorithm was tested in two much different scenarios and test cases, showing that not only satisfactory results have been achieved, but that additional ambient light or differing test persons do not have a negative effect on the results. Furthermore, our solution approach of employing BVLS does have the big advantage that no concerns in regard to privacy is raised as compared to other technologies in the field of occupancy determination systems.

In combination with another aspect followed in the "Vilipa" project, namely realizing the communication of our luminaires by the means of Powerline Communication (PLC), we can outline our vision, where a multitude of connected BVLS enabled luminaires throughout a building delivers occupancy data to effectively reduce the energy consumption, by only causing minimal installation effort.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support by the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK), the FFG and the funding program "Stadt der Zukunft" - Project "Vilipa"

REFERENCES

[1] K. Sun, Q. Zhao and J Zou, "A review of building occupancy measurement systems", Energy and Buildings, Vol. 216, 2020, pp. 109965, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2020.109965..

[2] W. Wang, X. Xu, H. H. Wei, B. Ren and J. Chen, "Modeling occupancy distribution in large building spaces for HVAC energy efficiency", Energy Procedia, Vol. 152, 2018, pp. 1230-1235, ISSN 1876-6102, https://doi.org/10.1016/j.egypro.2018.09.174..

[3] J. Chen, H. Chen, X. Luo, "Collecting building occupancy data of high resolution based on WiFi and BLE network", Automation in Construction, Vol. 102, 2019, pp. 183-194, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2019.02.016.

[4] A. Filippoupolitis, W. Oli and G. Loukas,"Occupancy detection for building emergency management using BLE beacons", International Symposium on Computer and Information Sciences, 2016, pp. 233-240.

[5] J. Suk, "Enhancing Energy Efficiency in HVAC with RFID Technology Based Occupancy Detection Solution: A Thesis Submitted to Auckland University of Technology" , Auckland University of Technology, Master of Computer and Information Sciences (MCIS), 2017.

[6] M.S. Zuraimi, A. Pantazaras, K.A. Chaturvedi, J.J. Yang, K.W. Tham and S.E. Lee, "Predicting occupancy counts using physical and statistical Co2-based modeling methodologies", Building and Environment, Vol. 123, 2017, pp. 517-528, ISSN 0360-1323, https://doi.org/10.1016/j.buildenv.2017.07.027.

[7] C. Jiang, Z. Chen, R. Su, M. K. Masood, Y. C. Soh, "Bayesian filtering for building occupancy estimation from carbon dioxide concentration", Energy and Buildings, Vol. 206, 2020, 109566, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2019.109566

[8] C. Jiang, M. K. Masood, Y. C. Soh, H. Li, "Indoor occupancy estimation from carbon dioxide concentration", Energy and Buildings, Vol. 131, 2016, pp. 132- 141, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2016.09.002.

[9] C. Brennan, G. W. Taylor and P. Spachos, "Designing learned CO2 based occupancy estimation in smart buildings", IET Wireless Sensor Systems, Vol. 8, no. 6, 2018, pp. 249-255, DOI: 10.1049/iet-wss.2018.5027

[10] R. Razavi, A. Gharipour, M. Fleury and I. J. Akpan, "Occupancy detection of residential buildings using smart meter data: A large-scale study", Energy and Buildings, Vol. 183, 2019, pp. 195-208, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2018.11.025.

[11] G. Tang, K. Wu, J. Lei and W. Xiao, "The meter tells you are at home! nonintrusive occupancy detection via load curve data", IEEE International Conference on Smart Grid Communications (SmartGridComm), 2015, pp.897-902

[12] Y. P. Raykov, E. Ozer, G. Dasika, A. Boukouvalas and M. A. Little, "Predicting room occupancy with a single passive infrared

(PIR) sensor through behavior extraction", In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16), Association for Computing Machinery, New York, NY, USA, pp. 1016–1027 https://doi.org/10.1145/2971648.2971746

[13] A. Hammoud, M. Deriaz and D. Konstantas, "UltraSense: A Self-Calibrating Ultrasound-Based Room Occupancy Sensing System", Procedia Computer Science, Vol. 109, 2017, pp. 75-83, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.05.297

[14] J. Zou, Q. Zhao, W. Yang and F. Wang, "Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation", Energy and Buildings, Vol. 152, 2017, pp. 385-398, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2017.07.064

[15] N. Lasla, M. Doudou, D. Djenouri, A. Ouadjaout and C Zizoua, "Wireless energy efficient occupancy-monitoring system for smart buildings",Pervasive and Mobile Computing, Vol. 59, 2019, 101037, ISSN 1574-1192, https://doi.org/10.1016/j.pmcj.2019.101037

[16] S. Hu, P. Wang, C. Hoare and J. O'Donnell, "Building Occupancy Detection and Localization Using CCTV Camera and Deep Learning," in IEEE Internet of Things Journal, vol. 10, no. 1, pp. 597-608, 1 Jan.1, 2023, doi: 10.1109/JIOT.2022.3201877.

[17] Q. Wang and M. Zuniga, "Passive visible light networks: taxonomy and opportunities", In Proceedings of the Workshop on Light Up the IoT (LIOT '20), Association for Computing Machinery, New York, NY, USA, 2020, pp. 42–47. DOI:https://doi.org/10.1145/3412449.3412551

[18] M. A. Mestiraihi, H. Abuella and S. Ekin, "Indoor Occupancy Estimation Using Visible Light Sensing (VLS) System," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 2018, pp. 1-5.

[19] K. Deprez, S. Bastiaens, L. Martens, W. Joseph, and D. Plets, "Passive Visible Light Detection of Humans," Sensors, vol. 20, no. 7, p. 1902, Mar. 2020, doi: 10.3390/s20071902.

[20] Y. Yang, J. Hao, J. Luo and S. J. Pan, "CeilingSee: Device-free occupancy inference through lighting infrastructure based LED sensing," 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom), Kona, HI, USA, 2017, pp. 247-256, doi: 10.1109/PERCOM.2017.7917871.

An Experimental Investigation of Whispering Gallery Mode Microsphere Resonators for Sensing, Lasing, and Nonlinear Optical Applications

Snigdha Thekke Thalakkal*, Davor Ristic*, Daniil Zhivotkov*, Gualtiero Nunzi Conti**, Mile Ivanda*
*Institute Ruder Boskovic, Bijenička Cesta 54, 10000, Zagreb
**MiPLab, IFAC – CNR, 50019 Sesto Fiorentino, Italy.
sthekket@irb.hr

***Abstract* - Microsphere resonators are gaining popularity as a key component in photonic circuit devices, optical communications, sensing, lasing, and nonlinear optical applications. As one of the essential categories of optical microcavities, microsphere resonators based on whispering gallery modes (WGMs) have benefits of ultra-high Q factor, small mode volume, ease of fabrication, and low cost. In this paper, we will present experimental results achieved by our group in regards to different application of microsphere resonators:**

1) **We studied the refractometric gas sensing sensitivity and the effect of thickness of the sol-gel silica layer coated to the microsphere to detect small traces of ammonia.**

2) **Rare earth doped microsphere resonators are promising candidates for lasing applications. Tellurite, fluoride, phosphate, and silicate glasses are good hosts for doping with rare earth elements for broadband applications. Tellurite doping with various rare earth elements can provide a wide stimulated emission cross section in the communication band.**

3) **High Q factor and low mode volume of WGM microresonators is a perfect host for non-linear optics application. We will present some non-linear effects which are present in silica spherical microresonators which include stimulated Raman scattering and frequency comb generation via degenerative four wave mixing.**

Keywords - Whispering gallery mode, Microsphere, Microlaser

I. INTRODUCTION

Acoustical whispering gallery modes (WGMs) were first investigated by Lord Rayleigh in the dome of St Paul's cathedral in London [1]. Analogous optical WGMS have been found in micrometer sized spheres where light is spatially confined inside the microspheres surface by total internal reflection. Whispering gallery mode microresonators exhibit different geometries such as micropillars, toroids, spheres, disks and rings [2]. Among them, optical microspheres are the prime focus of several theoretical and experimental studies. Microsphere resonators acquired considerable attention due to their potential to become vital components in photonic devices, telecommunications, sensing, lasing and non-linear optical applications. Whispering gallery mode microspheres have unique properties such as a high-quality factor and a low mode volume, which make them an indispensable tool in various optical systems. Three indices, *p*, *l*, and *m*, which represent the mode orders in radial, azimuthal, and axial directions, respectively, define microresonators. Microsphere resonators can be fabricated using a variety of techniques. Melting the tip of a standard optical fiber is the most common and easiest method for the fabrication of microspheres. In our work we prepared microspheres using a commercial fiber splicer, plasma torch and micro heater method.

We mainly focused on the experimental investigation of microsphere resonators for sensing, lasing, and nonlinear optical applications. Initially, we studied the refractometric gas sensing sensitivity and effect of thickness of sol gel silica layer coated to the microsphere to detect small traces of ammonia. Silica microspheres were produced by melting the tip of a standard silica fiber using an electric arc from a commercial fiber splicer. The sol-gel method is employed to produce a porous silica layer, which is then dip coated onto the sphere. A gas molecule that diffused into the porous silica layer and interacted with the coating's whispering gallery mode changed the coating's refractive index, which induced a shift in wavelength. This mechanism is called refractometric sensing [3]. For ammonia gas sensing measurements, a signal from a tunable laser is coupled to the coated microsphere using a tapered optical fiber.

Furthermore, our work based on rare earth doped microsphere resonators for lasing applications. Tellurite, fluoride, phosphate, and silicate glasses are good hosts for doping with rare earth elements for broadband applications. The high refractive index and wide-range transparency tellurite glass offers in the communication band make it a viable material for the fabrication of microlasers. Rare earth

metals including erbium, thulium, holmium, and ytterbium attracted a lot of attention because of their powerful, and broad stimulated emission cross section in the tellurite host and are therefore frequently utilized in laser systems [4-7]. Tellurite glasses exhibit exceptional mechanical and thermal characteristics as well as excellent transparency from the visible to the mid-infrared spectrum. The plasma torch method is used for the fabrication of erbium doped tellurite microspheres. We worked on erbium doped tellurite microresonators and used pump laser of wavelength 980 nm to achieve laser emission at 1562.5 nm.

In addition, we also worked on non-linear effects which are present in silica spherical microresonators. The first frequency comb generation from a microresonator is reported by P. Del'Haye [8]. We achieved stimulated Raman scattering and frequency comb generation via degenerative four wave mixing.

II. EXPERIMENTAL

For gas sensing application, silica microsphere is produced by melting the tip of a standard silica fiber using an electric arc from a commercial fiber splicer. The fabricated microsphere is coated with a porous silica layer using the dip coating method. The coating solution is made using a standard sol gel technique. Tetraethyl orthosilicate (TEOS), distilled water, and HCl diluted in ethanol are mixed together to form the sol gel solution (procedure explained in detail [3]). Fig.1 depicts the sketch of a porous silica coated microsphere. Thickness and refractive index of the coating is estimated using ellipsometry.

Signal from a tunable laser is used for the gas sensing measurement which is coupled to the microsphere using a tapered fiber. The coupled microsphere system is placed in an isolated chamber. To remove the humidity, the chamber is flushed with argon before the measurements. Fig.2 represents the set up for gas sensing measurement. Ammonia in argon atmosphere is used for sensing. Two mass flow controllers control the concentration of Ar and ammonia in the chamber. The tunable laser is swept through a small wavelength range. The transmission is

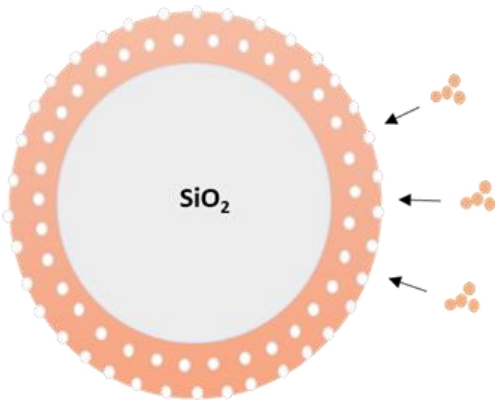


Figure 1. Sketch of microsphere coated with a porous silica layer.

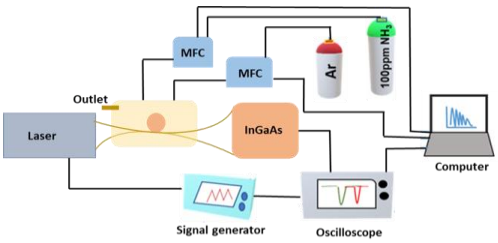


Figure 2. Experimental set up for ammonia gas sensing.

detected using an InGaAs detector (Thorlabs DET08CFC/M) which is connected to the oscilloscope.

Fig.3 depicts the schematic diagram of the plasma torch method for the fabrication of tellurite glass microspheres. To fabricate microlasers erbium doped tellurite glass is used. We used $15Na_2O_25WO_360TeO_2$ glass powder doped with 0.5 mol% Er^{3+} for the fabrication of microspheres. Plasma torch method is used to fabricate the microsphere. In the plasma torch method two electrodes are connected to a high voltage supply.

When the glass grains are passing through the electric arc from the plasma torch they melt and then acquire a spherical shape due to surface tension. Fabricated microsphere is glued to a fiber. A pump is coupled to the microsphere using a half taper. Fig. 4 shows the experimental set up for the microlaser experiment. An optical spectrum analyzer receives the counter propagating light from the microsphere. A pump laser of 980 nm is used to achieve the laser emission.

For the non-linear optical measurement, a tunable laser is connected to an erbium doped fiber amplifier which is connected to the coupled microsphere system. A tapered fiber is used for the coupling. Typical experimental set up for the nonlinear optical application used on a silica microsphere is shown in Fig.5 Signal from the microsphere is sent to the optical spectrum analyzer and an InGaAs detector using a 50:50 splitter. Output of the InGaAs detector is connected to an oscilloscope. Several non-linear effects were observed including Raman scattering and four wave mixing.

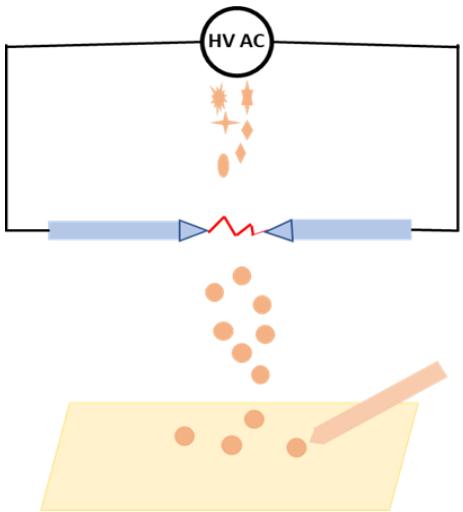


Figure 3. Plasma torch method for the fabrication of microspheres.

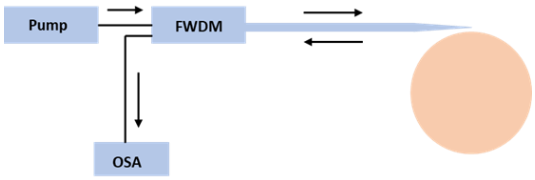


Figure 4. Experimental set up for the microlaser experiment.

III. RESULTS AND DISCUSSIONS

Sensing measurements were carried out for different cycles of ammonia ranging from 1 ppm to 5ppm (Fig. 6). We found that higher ammonia concentration in the chamber leads to a higher mode shift. According to Fig. 7 the sensitivity of our system is 113.2 pm/ppm, although for higher concentrations we observe some saturation. This is because while increasing the concentration of ammonia in the chamber all the bonding sites for ammonia bonding in the sol-gel coating is occupied. The sensitivity of different modes is found to be different. Fig. 8 depicts two modes that are shifting their positions due to different sensitivities. The different modes have different sensitivity because they have different radial orders (*p* values). The major problem we faced using our set up is the difficulty for selecting the mode with desired values of *p*. (detailed explanation is given in [3,9])

Fig. 9 shows the observed erbium emission at 1562.5 nm of an erbium doped tellurite microsphere doped with Er^{3+} ions. In order to get lasing at 1562.5 nm we adjusted the coupling point between the sphere and the half taper. We used a half-tapered fiber for the coupling of 980 nm pump light into the microsphere and observed single mode lasing. The Q factor of the sphere is also measured and it is found to be 10^6. Fig. 10 shows the Lorenz fitting of the peak for the Q factor measurements.

In Fig. 11 is shown the observed stimulated Raman scattering achieved using a pump laser at 1570 nm. The Raman lines are observed at 1692 nm and 1702.4 nm. The peak at 1692 is stronger than the pump laser peak. The conversion of energy from pump to Raman line is efficient. The Raman line shift itself is between 459 and 495 cm^{-1}, which corresponds to the energies of Raman transitions in quartz glass.

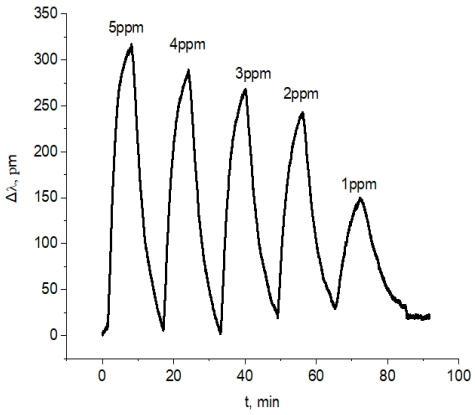


Figure 6. Resonance shift with time.

In Fig. 12 is shown a typical frequency comb extending from 1550 to 1750 nm including 46 lines. The initial 1560 nm pump serves as a seed from which different comb lines are generated. Two pump photons generate one signal and one idler photon by degenerative four wave mixing. All the other lines are then generated by non-degenerative four wave mixing between photons belonging to different lines. The lines correspond to the whispering gallery modes with subsequent *l* values. Energy and momentum conservation laws are satisfied in four wave mixing which means the comb lines have to be uniform. The lines are generated at the exact eigen frequencies of the whispering gallery modes. The Q-factors of the WGMs are extremely high so the eigen-modes must also be uniform. The uniformity of

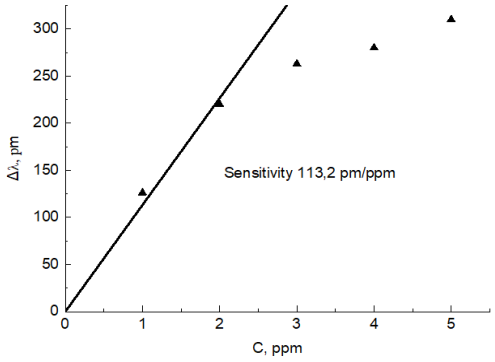


Figure 7. Resonance shift with concentration of ammonia.

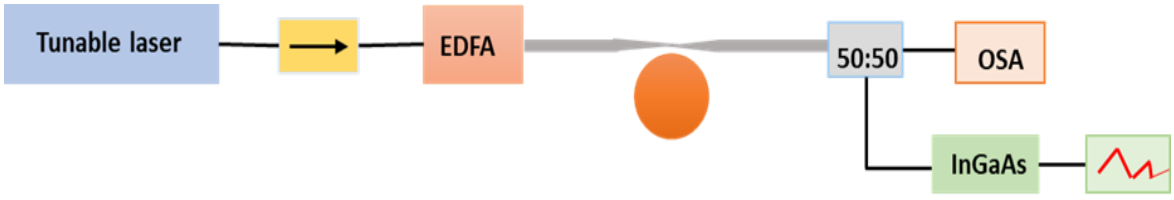


Figure 5. Experimental set up of nonlinear optical application.

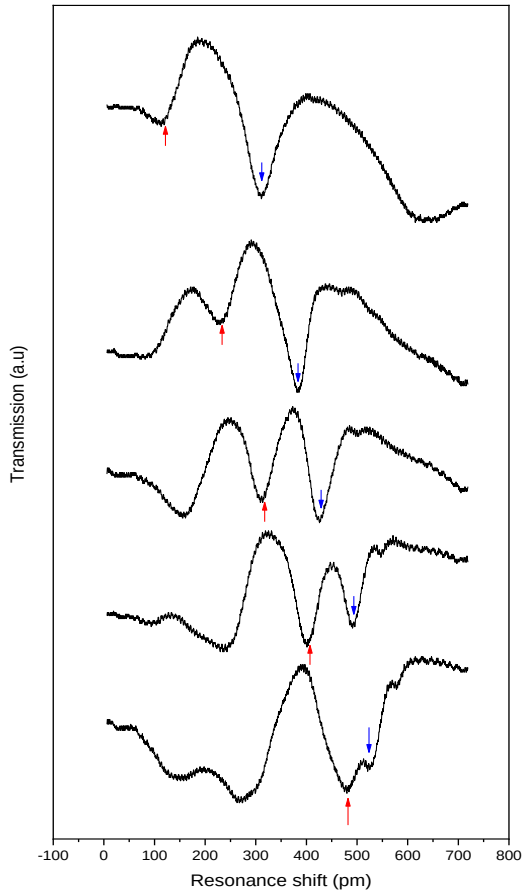


Figure 8. Transmittance of a sphere coated with a porous silica layer after the injection of 5ppm of NH_3 in the testing chamber.

the eigenmodes is determined by the dispersion of the resonator's material as well as the geometrical cavity dispersion of the sphere itself [10]. Additionally, Stimulated Raman scattering lines can also serve as an additional seed for the frequency comb.

IV. CONCLUSION

To conclude, an experimental study was made of the refractometric gas sensing sensitivities of a silica microsphere coated with a thick porous silica layer. We deduce from the observations that different modes can have various sensitivities for a given sphere and explain this by the fact that the observed whispering gallery modes had different *p* values. We demonstrated erbium doped tellurite microsphere as microlaser and achieved laser emission at 1562.5 nm. Furthermore, we observed several

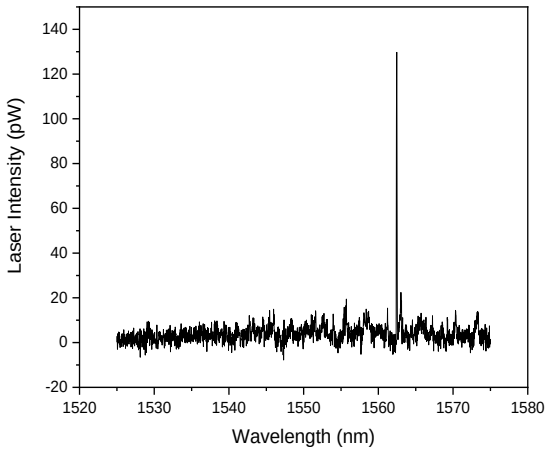


Figure 9. Laser emission of Er3+ doped tellurite microsphere at 1562.5 nm using 980 nm pump.

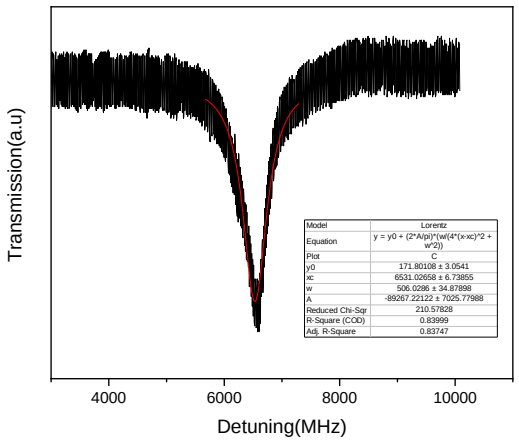


Figure 10. Transmittance of a sphere coated with a porous silica layer after the injection of 5ppm of NH_3 in the testing chamber.

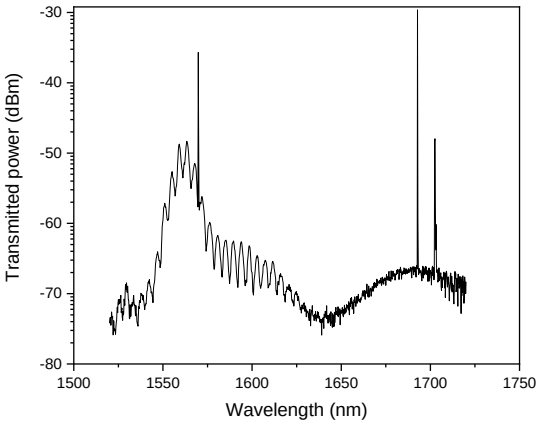


Figure 11. Stimulated Raman scattering.

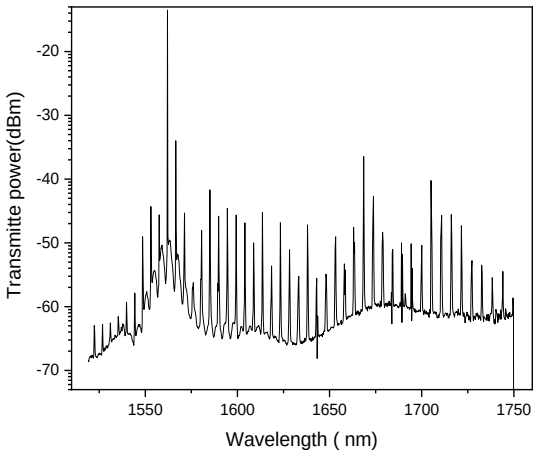


Figure 12. Frequency comb generation.

non-linear effects includes stimulated Raman scattering and frequency comb generation in a silica microsphere.

ACKNOWLEDGMENT

Croatian Science Foundation IP-2019-04-3045; Croatian Government and the European Union through the European Regional and Development Fund – the Competitiveness and Cohesion Operational Programme (KK.01.1.1.01.0001).

REFERENCES

[1] A. Chiasera et al, "Spherical whispering-gallery-mode microresonators," Laser & Photonics Reviews, vol. 4, pp. 457-482, 2010.

[2] VS. Ilchenko, S. Vladimir, and A. B. Matsko, "Optical resonators with whispering-gallery modes-part II: applications," IEEE Journal of selected topics in quantum electronics, vol. 12, pp. 15-32, 2006.

[3] D. Zhivotkov et al, "Radial order dependence of the gas sensing sensitivity of whispering gallery mode microspheres," Optical Materials, vol. 129, pp. 112544, 2022.

[4] X. Peng et al, "Fiber-taper-coupled L-band Er 3+-doped tellurite glass microsphere laser," Applied physics letters, vol. 82, pp. 1497-1499, 2003.

[5] J. Wu, S. Jiang, T. Qua, M. Kuwata-Gonokami, N. Peyghambarian, "2 μ m lasing from highly thulium doped tellurite glass microsphere," Applied Physics Letters, vol. 87, pp. 211118, 2005.

[6] J. Yu et al, "An experimental and theoretical investigation of a 2 μm wavelength low-threshold microsphere laser," Journal of Lightwave Technology, vol 38, pp. 1880-1886, 2020.

[7] A. Li et al, "Infrared-laser and upconversion luminescence in Ho3+-Yb3+ codoped tellurite glass microsphere," Journal of Luminescence, vol. 218, pp. 116826, 2020.

[8] P. Del'Haye et al, "Optical frequency comb generation from a monolithic microresonator," Nature, vol. 450, pp. 1214-1217, 2007.

[9] D. Ristić, et al. "Theoretical Analysis of the Refractometric Sensitivity of a Coated Whispering Gallery Mode Resonator for Gas Sensing Applications," *Sensors*, vol. 22.23, pp. 9155, 2022.

[10] D. Ristić, et. al. "Tailoring of the free spectral range and geometrical cavity dispersion of a microsphere by a coating layer," Optics Letters, vol. 39, pp 5173–5176, 2014.

DS-BE

International Conference on

DATA SCIENCE AND BIOMEDICAL ENGINEERING

Steering Committee

Chairs: Karolj Skala, Ruđer Bošković Institute, Zagreb, Croatia
Aleksandra Rashkovska Koceva, Jožef Stefan Institute, Ljubljana, Slovenia
Davor Davidović, Ruđer Bošković Institute, Zagreb, Croatia

Members: Marian Bubak, ACC Cyfronet AGH University of Science and Technology, Krakow, Poland
Jesús Carretero Pérez, University Carlos III, Madrid, Spain
Tiziana Ferrari, EGI Foundation, Amsterdam, Netherlands
Andrej Jokić, University of Zagreb, Croatia
Dieter Kranzlmüller, Ludwig-Maximilians-University, München, Germany
Ludek Matyska, Masaryk University, Brno, Czech Republic
Dana Petcu, West University of Timisoara, Romania
Uroš Janez Stanič, Biomedical Research Institute-BRIS, Ljubljana, Slovenia
Marko Švaco, University of Zagreb, Croatia
Matjaž Veselko, University Medical Centre Ljubljana, Slovenia
Yingwei Wang, University of Prince Edward Island, Charlottetown, Canada

Program Committee

Members: Enis Afgan, Ruđer Bošković Institute, Zagreb, Croatia

Viktor Avbelj, Jožef Stefan Institute, Ljubljana, Slovenia
Davor Davidović, Ruđer Bošković Institute, Zagreb, Croatia
Matjaž Depolli, Jožef Stefan Institute, Ljubljana, Slovenia
Simeon Grazio, University Clinical Centar Sestre milosrdnice, Zagreb, Croatia
Marjan Gusev, Ss. Cyril and Methodius University, Skopje, North Macedonia
Vojko Jazbinšek, University of Ljubljana, Slovenia
Jurij Matija Kališnik, Klinikum Nürnberg Süd, Germany
Zalika Klemenc-Ketiš, Angela Boškin Faculty of Health Care, Jesenice, Slovenia
Dragi Kocev, Jožef Stefan Institute, Ljubljana, Slovenia
Gregor Kosec, Jožef Stefan Institute, Ljubljana, Slovenia
Miklos Kozlovszky, Obuda University, Budapest, Hungary
Lene Krøl Andersen, Danish e-infrastructure Cooperation Denmark, Kgs. Lyngby, Denmark
Željka Mihajlović, University of Zagreb, Croatia
Panče Panov, Jožef Stefan Institute, Ljubljana, Slovenia
Tonka Poplas Susič, Health Centre Ljubljana, Slovenia
Aleksandra Rashkovska Koceva, Jožef Stefan Institute, Ljubljana, Slovenia
Karolj Skala, Ruđer Bošković Institute, Zagreb, Croatia
Viktor Švigelj, University Medical Centre Ljubljana, Slovenia
Ivan Tomašić, Mälardalen University, Västerås, Sweden
Roman Trobec, Jožef Stefan Institute, Ljubljana, Slovenia
Roman Wyrzykowski, Institute of Computer and Information Science, Częstochowa, Poland

PAPERS

DATA SCIENCE

Missing Values Interpolation in PurpleAir Sensor Data based on a Correlation with Neighboring Locations using KNIME Analytics Platform

Samir Omanovic*, Admir Midzic**, Zikrija Avdagic*, Damir Pozderac* and Amel Toroman**
* University of Sarajevo, Faculty of Electrical Engineering, Sarajevo, Bosnia & Herzegovina
** University of Bihac, Technical Faculty, Bihac, Bosnia & Herzegovina
samir.omanovic@etf.unsa.ba, admir.midzic@unbi.ba, zikrija.avdagic@etf.unsa.ba,
damir.pozderac@etf.unsa.ba, amel.toroman@unbi.ba

***Abstract* – Missing values handling in any collected data is one of the first issues that must be resolved to be able to use that data. This paper presents an approach used for missing values interpolation in PurpleAir particle pollution sensor data, based on a correlation of the measurements from the observed locations with the measurements from its neighboring locations, using KNIME Analytics Platform. Results of our experiments with data from five locations in Bosnia & Herzegovina, presented in this paper, show that this approach, which is relatively simple to implement, gives good results. All modeling and experiments were conducted using KNIME Analytics Platform.**

Keywords - interpolation; correlation; sensor data; KNIME; missing values; particle pollution (PM2.5)

I. INTRODUCTION

We are witnessing an increasing presence of data collection from many different sensors in various fields of application. These data should be analyzed and used to find certain relationships, rules, patterns, etc. which is not so easy, bearing in mind the volume and the quality of the data. One of the problems with the collected data is related to missing values. There is no straightforward answer how to deal with missing values in the collected data and various approaches are used [1].

Nowadays, there is more and more awareness of environmental pollution, so there are more and more activities in that field with an emphasis on pollution measurements using sensors like PurpleAir sensors. PurpleAir sensors are low-cost sensors for measurement of air quality monitoring. They use laser particle counters to provide real time measurement of temperature, humidity, PM1.0, PM2.5 and PM10. This paper is only about particle pollution (PM2.5) data. These sensors can be connected to a WiFi network. That way, the measurements from the sensor are accessible on that network. Additionally, users of PurpleAir sensors can register their sensors on the PurpleAir real-time map. The PurpleAir real-time map is a web application that displays a network of community owned PurpleAir sensors. It enables downloading and use of data for various research projects [2,3].

This research is supported by the Ministry of Education and Science of entity Federation of Bosnia & Herzegovina of the state Bosnia & Herzegovina.

One of the major issues related to the use of the collected data is that there are missing values (missing measurements). To be able to use data in a good way, it is important to solve that issue at the beginning of any research. Measurement errors and sensor failure are not so rare in data collection and are one of major reasons for missing measurement values [4].

The data collected from PurpleAir sensors are time-series data. There are many methods to interpolate time-series data and that way to solve missing values issue. In [5] is given good overview of the interpolation methods:

- deterministic like nearest-neighbor, polynomial, based on distance weighting, based on Fourier's theory, and others; and
- stochastic like regression methods, autoregressive methods, machine learning methods, methods based on data dynamics, and others.

This paper is focused on resolving the issue of missing PM2.5 values in data from PurpleAir sensor, using existing data from neighboring locations. Neighboring locations usually share the same climatic and geographical properties which means that they mostly share the same causes of particle pollution. Logically, some causes will always be related to some local properties, but if neighboring locations are close enough then we can conclude that they share many common causes of particle pollution and that they share common climatic and geographical conditions. Based on that, we can say that there should be a significant correlation between PM2.5 data from these locations. Generally, correlation should not be used as a proof of mutual causality, but it can be an indicator of sharing common causes, which is important in this case. So, we can say that common causality is proved by the proximity of locations. In this case we can conclude that the correlation between locations is a good indicator of sharing common causes of particle pollution.

The main idea is to use correlation as a base for calculating weights, instead of distance weighting. This way, missing values handling is not so complicated process, while the obtained results are very good.

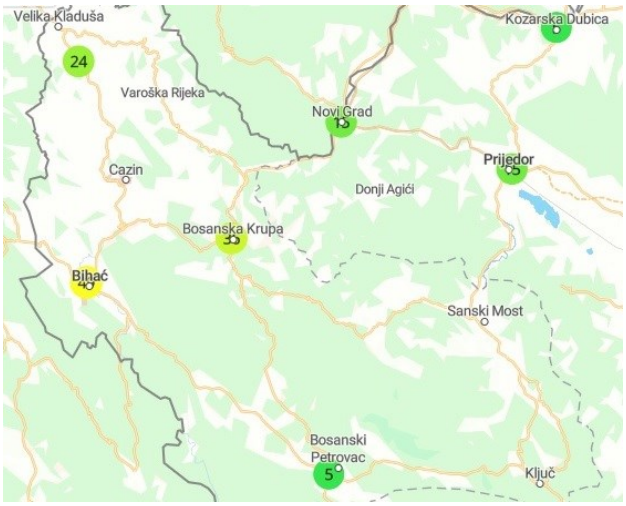


Figure 1. Map of locations with PurpleAir Sensors (image taken from the PurpleAir real-time map on 30.09.2022)

II. DATA

In our project we analyze data publicly available on the PurpuleAir real-time map [6], collected on several PurpleAir sensors located in Bosnia and Herzegovina. This paper uses data from sensors located in cities: Bihac, Bosanski Petrovac, Bosanska Krupa, Prijedor, and Velika Kladusa. Fig. 1 shows these locations on the map. Results and conclusions in this paper are based on the data from listed locations, for the period 01.01.2021 – 29.09.2022, and 60 minutes averages were used. Data was downloaded as CSV files on 30.09.2022.

If we observe listed locations on the map on Fig. 1 we can note that city Bosanska Krupa is surrounded by cities Bihać (air distance around 33 km), Velika Kladuša (air distance around 33 km), Prijedor (air distance around 47 km), and Bosanski Petrovac (air distance around 35 km). Based on the geographical location of these cities we decided to analyze and interpolate missing values for the city Bosanska Krupa, based on data from cities: Velika Kladuša, Bihać, Bosanski Petrovac, and Prijedor.

Table 1. shows locations list with data counts for the observed period 01.01.2021-29.09.2022 (primary, 60 minutes average) It is expected to have 15288 measurements for 636 days x 24 hours/day, but each of the sets do not have all entries. For the observed location there are 13970 entries, and 1318 entries are missing.

There are 787 entries in the data sets of all neighboring locations with entry time (*created_at* column) that do not exists in the data set for the observed location. That means that these values can be interpolated based on the existing data from the neighboring locations. 10060 entries in the data sets of all 5 locations are with the same entry time (created_at column) and they can be used for finding correlation between the observed location and their neighboring locations.

III. MODELING USING KNIME

KNIME Analytics Platform [7] is a very popular tool for various research projects. In this project, related to PurpleAir data, we use it for modeling different kind of workflows including those related to handling missing values in data.

TABLE I. LOCATIONS AND THEIR DATA COUNTS

ID	*Location name*	*Location Type*	*Data Counts*
#1	Velika Kladuša	neighboring	Number of entries: 12519 Missing entries: 2769
#2	Bihać	neighboring	Number of entries: 14050 Missing entries: 1238
#3	Bosanski Petrovac	neighboring	Number of entries: 15275 Missing entries: 13
#4	Prijedor	neighboring	Number of entries: 13992 Missing entries: 1296
#5	Bosanska Krupa	**observed**	Number of entries: 13970 **Missing entries: 1318**

At the beginning of the workflow, we use CSV Reader input nodes to read data from CSV files. Then manipulation nodes: Row Filter, Duplicate Row Filter, and Sorter are used to remove records with missing PM2.5 values, remove duplicated rows and sort data by date and time. We do not need rows with missing PM2.5 values in further processing.

By using Joiner nodes and the inner join setting for these nodes we create a new set composed of joined measurements for all five locations involved (four neighboring and one observed – target). That set contains only existing measured values on all locations for the same date and hour. If measurement is missing for any of five locations, then that date and hour is not included in this set. After that outliers are removed by using Numeric Outliers node. This new set is used to calculate the correlation between the four neighboring locations (Velika Kladuša, Bihać, Bosanski Petrovac, and Prijedor) and the observed location (Bosanska Krupa), using the Linear Correlation node. This node will calculate the correlation for all combinations of input columns, so it is necessary to filter the values of interest for the further processing. The results shown on Fig. 2.a indicate that there is a moderate to high positive correlation of PM2.5 values between the neighboring locations and the observed location. p-value stands for 'probability value' and it indicates how likely is that a result occurred by chance alone. Fig. 2.b shows zero p-values which indicates zero probability that this result is by chance.

S First column name	S Second column name	D Correlation value
PM2.5_ATM_ug/m3 (L #1)	PM2.5_ATM_ug/m3 (L #5)	0.7081996022459...
PM2.5_ATM_ug/m3 (L #2)	PM2.5_ATM_ug/m3 (L #5)	0.663273428193733
PM2.5_ATM_ug/m3 (L #3)	PM2.5_ATM_ug/m3 (L #5)	0.5906837388301...
PM2.5_ATM_ug/m3 (L #4)	PM2.5_ATM_ug/m3 (L #5)	0.6766711805903...

(a) Correlation values

D Correlation value	D p value	I Degrees of freedom
0.7081996022459994	0.0	8153
0.663273428193733	0.0	8153
0.5906837388301037	0.0	8153
0.6766711805903904	0.0	8153

(b) Two-sided p-values

Figure 2. Filtered result of the Linear Correlation node.

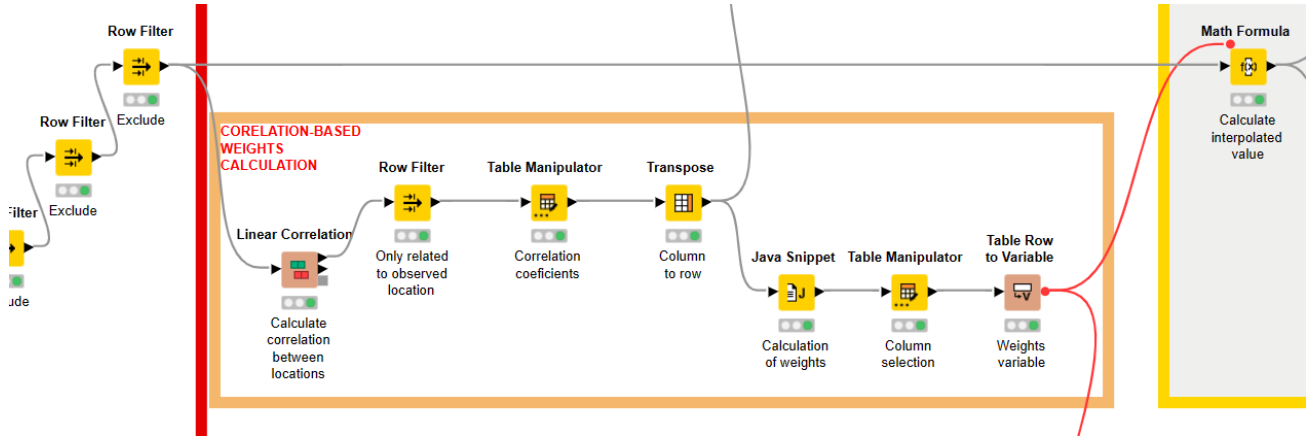


Figure 3. Part of the KNIME workflow where the correlation-based weight coeficients were calculated

In general, the Linear Correlation node calculates, for each pair of selected columns, a correlation coefficient – a measure of the correlation of the two columns in data (two variables). When columns contain numeric values then Pearson's product-moment coefficient (1) is calculated [7].

$$\rho_{X,Y} = cov(X,Y) / (\sigma_X \sigma_Y) \quad (1)$$

$cov(X,Y)$ is a covariance of two columns X and Y. σ_X is a standard deviation for the column X, and σ_Y is a standard deviation for the column Y.

The Fig. 3 shows a snippet of the workflow with nodes related to weights calculation based on the correlation values. In the Java Snippet node are calculated four weight values (w_i values) based on the correlation values (cor_i) using (2).

$$w_i = cor_i / \Sigma cor_i \text{ , } i = 1,2,3,4. \quad (2)$$

Weights are then transformed to flow variables using Table Row to Variable node. These weights variables are then used to calculate interpolated values as a weighted sum of measurements from the neighboring locations. In a further analysis of the results, it is concluded that it would be good to correct the interpolated result by multiplying it with the average ratio between the expected and the calculated value. The calculated multiplication factor has a value of 0.896. Fig. 4 shows quality indicators generated by the Numeric Scorer node.

Calculated weights and the multiplication factor are then used to interpolate missing values for the observed location. For that it is necessary to prepare the dataset that contains all rows with missing PM2.5 values in the dataset for the observed locations, but only for those date and time for which there are entries in datasets on all four neighboring locations. That way we ensure that we can calculate the interpolated value. For joining datasets, we again use Joiner nodes.

Statistics - 0:166 - Numeric Scorer (Quality)

File Edit Hilite Navigation View

Table "Scores" - Rows: 7 | Spec - Column: 1 | Properties | Flow Variables

Row ID	New_calculated_value
R^2	0.574
mean absolute error	9.652
mean squared error	218.999
root mean squared error	14.799
mean signed difference	1.11
mean absolute percentage error	0.407
adjusted R^2	0.574

Figure 4. Numeric Scorer node shows some quality indicators for the results.

In the final phase we joined the interpolated data and the existing data for the observed location to be able to make further analysis of the results. Our observations are presented in the next section of the paper.

So, the interpolation is based on a weighted sum, where the weighting factors are not calculated on the basis of distance but on the basis of correlation between data. The reason for this is that climatic influences, which are a very important part of this type of problem, cannot be connected only to the distance of locations. There are a lot of factors that come into play, from terrain configuration to air flow, and all of these factors are reflected through the correlation factor, which is why the correlation factor was taken into account when calculating weights, instead of distance.

IV. Discussion

Fig. 5 presents comparison of the final calculated

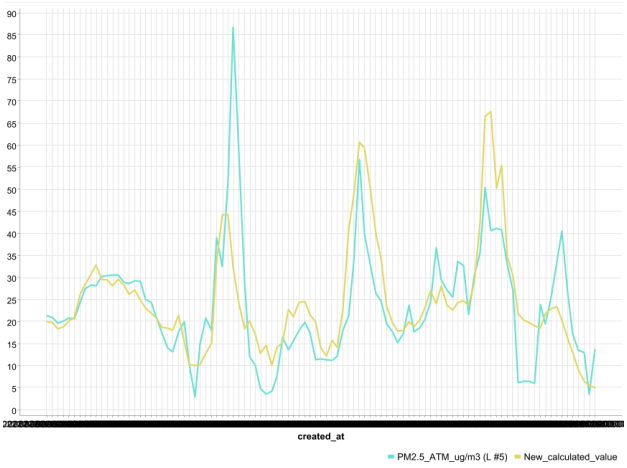


Figure 5. Checking results by comparing calculated value (after applying the correction factor) and known values.

(interpolated) values and the real (expected) values. We can note that the calculated values are near the real values in most of the points (60 minutes averages). It can be noted that there are differences and that they are mostly related to higher peaks in real values. That can be explained as a local influence (local factors related to the observed location) that cannot be calculated (predicted) based on the data from the neighboring locations.

Fig. 6 shows how interpolated values filled a larger gap of missing values. It can be noted that the transition from the interpolated to the existing values doesn't deviate much. For a few hours where there are interpolated values and for a few hours where there are original (real) values it can be noted that numbers are close to each other – around 20. Also, it can be noted that values decrease up to 18.44 and then they rise, which means that the trend is captured too. Also, it can be noted that the average of three marked interpolated values on Fig. 6 is 21,11 which is close to the average of three marked original values that are following them in a sequence, which is 20,18. If we bear in mind the fact that pollution influences between locations are not transmitted instantly but with a certain delay, then these small deviations (10-20%) are expected. All that is an argument that this approach is good enough, especially when there are larger gaps of missing values, like in the presented example.

Fig. 7 shows how interpolated values filled a small gap of missing values. It can be noted that the transition from the original to the interpolated values and back to the original values again is quite smooth.

The presented model can be seen as a simple model that can at least partially fill in the missing values in the data. Since it relies on existing data from neighboring locations, it has limited possibilities because it depends on them. In practice, this means that missing values cannot be interpolated if only a single value from one of the four neighboring sites is missing.

Of course, extensions of the model are possible in such a way that values are predicted based on more than four or less than four neighboring locations. In the case of an increase in the number of locations, it would be logical to expect an improvement in accuracy, but this increase cannot be significant, and it is difficult to verify it, because it would be very complicated to try to model microclimatic influences only for the observed location. Using fewer locations would result in less accurate values. The decision to use the four listed locations was more aimed at roughly covering the four sides (north, south, east, west) and thus covering the effects of climatic factors for those sides. That is why is less aimed at taking the closest locations.

Creating a model for something involves making a number of decisions related to the complexity of the model. The introduction of additional parameters usually leads to greater precision, but this may not be of vital interest for solving a particular problem. Therefore, in the modeling process, it is very important to balance between the practical usefulness of the model and the ideal model. Although the presented model has its shortcomings, it is still good enough for practical application and allows us to make good enough estimations of missing PM2.5 values for the observed location.

Regardless of all the shortcomings, this shows that, due to the fact that the locations share similar climatic characteristics, such a model allows us not only to fill in the missing values but also to roughly predict the values at the observed location, in case we don't have a sensor there, or the sensor doesn't work.

The number of missing values at the neighboring locations does not necessarily mean that the presented model will not be able to produce many interpolated values. The result depends on overlapping missing periods. If for the missing period on the observed location we have data on the neighboring location then the total number of missing values on the neighboring locations is not important. For example, although for the location Velika Kladuša are missing 2769 out of totally possible 15288 values which is 18,11%, the number of missing values on the observed location for which there are values on neighboring locations is 787, out of 1318 that are missing for the observed location. That means that for the observed location there are 8,62% missing values before applying the model and after applying the model number of missing values decreased to 3,47% which is an improvement of 5,15%. It would be possible to decrease this further by repeating the same process with other neighboring locations.

V. CONCLUSION

Our experiments showed that interpolating PM2.5 values in data from PurpleAir sensor, using existing data from neighboring locations and weighted sum based on a historical correlation of data has many advantages including: (1) easy implementation, and (2) possibility to interpolate large gaps in the data.

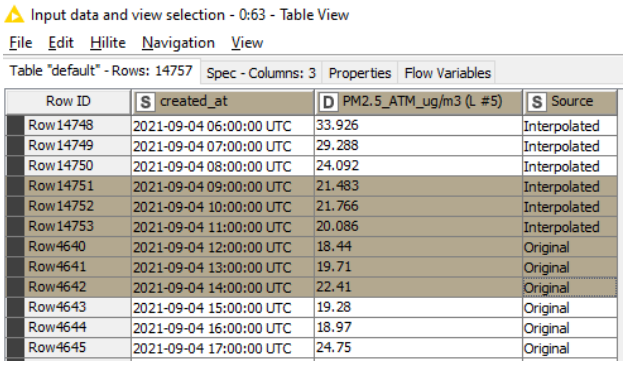
Comparing interpolated values and known (measured) values is possible to conclude that the interpolation results are very good. Discussion about results, and figures presented, argument our conclusions.

KNIME Analytical Platform tool enables very effective work during data analysis and modelling so that research focus is on modeling and model representation in the form of a workflow, and not on programming of a solution.

Our goal is to continue this research by calculating missing data using approaches based on the data for the observed location only, which can be a problem for larger gaps in data, so that we can compare approaches. Our final goal is to work on predictions of PM2.5 values for locations where PurpleAir sensor is temporary not working and not showing results on PurpleAir real-time map.

ACKNOWLEDGMENT

The authors thanks to the Ministry of Education and Science of entity Federation of Bosnia & Herzegovina of the state Bosnia & Herzegovina for the financial support to the project entitled: “Analysis of data from PurpleAir



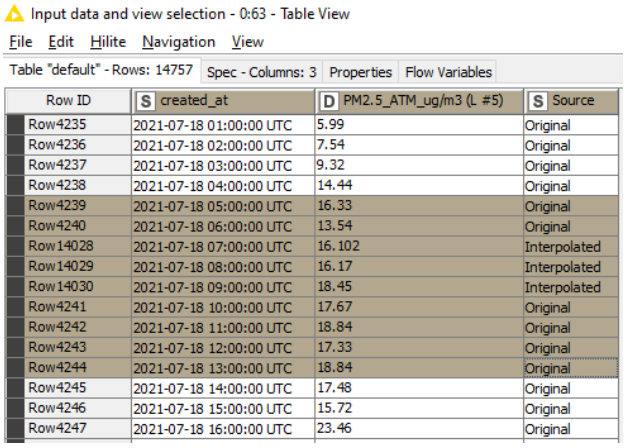
Input data and view selection - 0:63 - Table View

File Edit Hilite Navigation View

Table "default" - Rows: 14757 | Spec - Columns: 3 | Properties | Flow Variables

Row ID	created_at	PM2.5_ATM_ug/m3 (L #5)	Source
Row14748	2021-09-04 06:00:00 UTC	33.926	Interpolated
Row14749	2021-09-04 07:00:00 UTC	29.288	Interpolated
Row14750	2021-09-04 08:00:00 UTC	24.092	Interpolated
Row14751	2021-09-04 09:00:00 UTC	21.483	Interpolated
Row14752	2021-09-04 10:00:00 UTC	21.766	Interpolated
Row14753	2021-09-04 11:00:00 UTC	20.086	Interpolated
Row4640	2021-09-04 12:00:00 UTC	18.44	Original
Row4641	2021-09-04 13:00:00 UTC	19.71	Original
Row4642	2021-09-04 14:00:00 UTC	22.41	Original
Row4643	2021-09-04 15:00:00 UTC	19.28	Original
Row4644	2021-09-04 16:00:00 UTC	18.97	Original
Row4645	2021-09-04 17:00:00 UTC	24.75	Original

Figure 6. Part of time-series where interpolated values filled a larger gap of missing values. Note the transition from the interpolated to the existing values.



Input data and view selection - 0:63 - Table View

File Edit Hilite Navigation View

Table "default" - Rows: 14757 | Spec - Columns: 3 | Properties | Flow Variables

Row ID	created_at	PM2.5_ATM_ug/m3 (L #5)	Source
Row4235	2021-07-18 01:00:00 UTC	5.99	Original
Row4236	2021-07-18 02:00:00 UTC	7.54	Original
Row4237	2021-07-18 03:00:00 UTC	9.32	Original
Row4238	2021-07-18 04:00:00 UTC	14.44	Original
Row4239	2021-07-18 05:00:00 UTC	16.33	Original
Row4240	2021-07-18 06:00:00 UTC	13.54	Original
Row14028	2021-07-18 07:00:00 UTC	16.102	Interpolated
Row14029	2021-07-18 08:00:00 UTC	16.17	Interpolated
Row14030	2021-07-18 09:00:00 UTC	18.45	Interpolated
Row4241	2021-07-18 10:00:00 UTC	17.67	Original
Row4242	2021-07-18 11:00:00 UTC	18.84	Original
Row4243	2021-07-18 12:00:00 UTC	17.33	Original
Row4244	2021-07-18 13:00:00 UTC	18.84	Original
Row4245	2021-07-18 14:00:00 UTC	17.48	Original
Row4246	2021-07-18 15:00:00 UTC	15.72	Original
Row4247	2021-07-18 16:00:00 UTC	23.46	Original

Figure 7. Part of time-series where interpolated values filled a small gap of missing values. Note the transition from the original to the interpolated values and back to the original again.

sensors" approved by the contract no. 05-35-2135-1/22 from 27.10.2022.

REFERENCES

[1] Struminskaya, B., Lugtig, P., Keusch, F., & Höhne, J. K. (2020). Augmenting Surveys With Data From Sensors and Apps: Opportunities and Challenges. Social Science Computer Review, 0(0). https://doi.org/10.1177/0894439320979951

[2] Lu T, Liu Y, Garcia A, Wang M, Li Y, Bravo-Villasenor G, Campos K, Xu J, Han B. Leveraging Citizen Science and Low-Cost Sensors to Characterize Air Pollution Exposure of Disadvantaged Communities in Southern California. Int J Environ Res Public Health. 2022 Jul 19;19(14):8777. doi: 10.3390/ijerph19148777. PMID: 35886628; PMCID: PMC9322770.

[3] A. Caseiro, S. Schmitz, G. Villena, J. V. Jagatha, and E. von Schneidemesser, "Ambient characterisation of PurpleAir particulate matter monitors for measurements to be considered as indicative," *Environ. Sci.: Atmos.*, p., 2022, doi: 10.1039/D2EA00085G.

[4] Kasam, A.A., Lee, B.D. & Paredis, C.J.J. Statistical methods for interpolating missing meteorological data for use in building simulation. Build. Simul. 7, 455–465 (2014). https://doi.org/10.1007/s12273-014-0174-7

[5] Lepot, M.; Aubin, J.-B.; Clemens, F.H.L.R. Interpolation in Time Series: An Introductive Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. Water 2017, 9, 796. https://doi.org/10.3390/w9100796

[6] PurpleAir, "Map Start-up Guide," PurpleAir Community. [Online]. Available: https://community.purpleair.com/t/map-start-up-guide/90. [Accessed: Dec. 06, 2022].

[7] KNIME AG: "KNIME Analytics Platform – Allowing anyone to build and upskill on data science". [Online]. Available: https://www.knime.com/knime-analytics-platform. [Accessed: Dec. 07, 2022].

[8] KNIME AG: "Linear Correlation". [Online]. Available: https://hub.knime.com/knime/extensions/org.knime.features.base/latest/org.knime.base.node.preproc.correlation.compute2.CorrelationCompute2NodeFactory [Accessed: Dec. 07, 2022].

Spectral Analysis, Agglomerative, Mean Shift and Affinity Propagation Algorithms, Use on the Content from Social Media for Low-Resource Languages

Mërgim H. HOTI, Jaumin Ajdari, Xhemal Zenuni, Mentor Hamiti
South East European University (SEEU)/ Computer Science, Tetovo, Republic of North Macedonia
{mh28356, j.ajdari, xh.zenuni, m.hamiti}@seeu.edu.mk

***Abstract* - Social networks, as part of our daily life, affect our behavior and lifestyle in various ways, making it important for each of us to be aware of their impact. Posts on social media platforms can have a profound effect on our mood, depending on our personal interpretation and opinion of them. Therefore, it is crucial to correctly classify these textual data to gain a better understanding of their impact. However, this task can be challenging, particularly when dealing with unlabeled data such as social media posts. An added challenge is working with low-resource languages.**

In this research, we investigate four unsupervised text clustering methods by testing them on a low-resource language, such as Albanian. The investigated algorithms are Spectral, Agglomerative, Mean Shift and Affinity Propagation, and by adjusting the working parameters, we tried to find a more appropriate application of them. Methods are applied to pre-processed data, textual posts, by use of different preprocessing techniques, and the results are presented and interpreted.

This research aims to assist other researchers in the same field who have a specific focus on working with low-resource languages.

Keywords - text classification, low-resource language, Agglomerative, Spectral Analysis, Affinity Propagation, Mean Shift..

I. INTRODUCTION

Nowadays, there is a huge increase in social content, data that is distributed on many social networks such as Facebook, Instagram, Twitter, LinkedIn, Tik-Tok, and so on. Those data are very extensive and require proper tools for effective management. The proper management and processing of these data make today's life easier, especially in the process of decision-making. Knowing the opinions of customers is of great value in business. Therefore, the correct classification of data is a step in helping companies to make changes in services and customer satisfaction.

To the best of our knowledge, there is no literature available that has worked specifically with unsupervised clustering algorithms using low-resource languages, excepted to some researchers which have used Albanian as low resources language on their papers such as [1], [2], [3], therefore, we decided to test several unsupervised clustering algorithms on a low-resource language such as Albanian. The main idea of the research was to analyze the correctness and accuracy of these algorithms in the dataset of posts, on the social network (Facebook), written in the Albanian language.

After analyzing some existing literature on this issue, we decided to focus our research on unsupervised [4] clustering methods and used clustering methods such as spectral [5, 6, 7], agglomerative [8, 9], mean shift [10] and affinity propagation [11, 12, 13, 14] while other resources has been identified what kind of parameters are included [15, 16]. We tried to find and demonstrate their behavior when used in a dataset composed of posts written in a low - resources language. For all methods, appropriate data preprocessing is performed, and then the method is used, and the results are analyzed. As a result, a more appropriate data classification and clustering is sought for critical words or terms that frequently appear, as cluster centers. Results and clusters were also manually checked.

The paper starts with an overall introduction, explaining the reasons for selecting these algorithms then, in section II presents a brief literature review. Section III presents the methodology used and steps performed for generating the research results. While section IV presents obtained results with selected algorithms. In the final section of this paper, specifically section V, the conclusions are extracted, and future work is discussed.

II. LITERATURE REVIEW

Collecting the ideas and thoughts from random users, clients of specific services, expressed on social media, in order to use them and improve the company's services, is something very important and should be appreciated by everyone.

After reviewing the relevant literature, we have not found any materials that discuss or implement

experiments similar to ours for low-resource languages. Therefore, we have not pursued this direction of research.

The opinion of each client is very important to the company in order to improve the quality of its services. But sometimes it's hard to use those opinions because they're usually text-based and unstructured. These text responses, customer feedback, are important to the company, but difficult to analyze and extract real meaning from. The known methods used require input of numerical data and provide quantitative estimates, which shows another difficulty such as numerical expression of textual data [17].

Therefore, in this way we need to emphasize that machine learning classifiers seem to be effective tools, methods for detecting clusters in a huge dataset. In terms of classification, text classification methods work in one of three data types such as supervised, semi-supervised and unsupervised data. The selected methods, for this research, have a wide range of use for all three data types and allow implementation to solve different problems and objectives.

According to [5], authors have operated the first mathematical solution of using flexible spectral algorithm in single and multi-dimensional issues using general linear Hermite processes. Also, they propose techniques to use three different adaptive spectral techniques such as p-adaptive, the moving procedure and scaling one. Also, they have tested by using three keys which are associated with scaling, separation and expansion of spectral order. Using these sub-techniques of spectral in explicitly form makes it possible to control error efficiently especially when it is used spectral algorithm.

There are several research papers which discuss how spectral and other algorithms operate in different issues. Authors in [18], propose a new form of implementation which is based on spectral algorithm and tries to detect communities. This means that, by using nodes inside of a system and by learning continuously exceed all forms proposed so far. Authors explain that nodes learn space features of each of them around in low dimensionality and calculate the similarity to finish the community detection.

Also, agglomerative algorithm is used by [9] and represent a framework implementation which has significant impact especially on the accuracy of traditional agglomerative grouping algorithms. This is divided into two implementation perspectives such as single, complete, group linkage and Ward's which are equal with hierarchical model-based method. While the second perspective shows how variants such as complete-link, Mahalanobis-link, and line-link could be used in an extended form in a case of agglomerative algorithm.

Charles Kumah et al. [19], used printed fabric pattern into full color to make image segmentation by using Mean Shift. The dataset contains 11 plain waives and 1 twill waive cotton fabrics. Each printed pattern consists of 600 x 600 pixels while the results have shown that this algorithm is so good in clustering of segmentation for printed fabric patterns even containing texture and illuminations. Approximately in each execution they have gained 6 to 7 clusters where each of them identifies separately object in that figure while at the end it collects all of segments.

Javier Fumanal-Idocin et al [20] propose a construction for novel affinity functions. Nonetheless, they have measured the performance of several datasets using less than-convex combinations of Affinity functions in different communities. At the same time, a proposed method is designated to collect different social mechanisms in a network which has interaction between used functions. Also, they found very good results in the case of modularity measurements for all datasets and algorithms.

III. METHODOLOGY

In this section, we present the way we applied all the methods, including the implementation parameters as well customization. This helped to execute the selected algorithms in the most appropriate way, according to the input data and data preprocessing. In order to analyze the behavior of the above algorithms for the case of low-resource languages, we used datasets consisting of Facebook posts written in Albanian, and the first dataset refers to the company Vala [1] (a telecommunications company) with 1325 posts, and the second to the company Art Motion[2] (a company that provides TV services), with 550 posts, both from the Republic of Kosovo. In the explanation below are used use X and Y, respectively, for those two companies. To convert the posts from text to vector of numbers, TF-IDF is used. The results below, on the left side correspond to dataset X and on the right to dataset Y.

To increase accuracy, we applied stop words removal and word lemmatization. Lemmatization was applied by manually creating a long list of common words which correspond to the root of specific words such as mbushje, mbushjen, mbushjes, mbushjev and these are turned to mbushje. This is applied for both of datasets. Then, we compiled our set of stop words and included a repository of stop words from research [21]. At the same time, we also used data filtering, removal of outliers (which was created a cluster with minimal comments) as well as principal component analysis (PCA) to reduce some dimensions and focus on more significant dimensions and post's words.

To verify the accuracy of the most suitable number of clusters, we used Silhouette which shows the higher accuracy of certain number of clusters. This has helped us to identify specific values for each algorithm.

In the following part, are shown the experimental results of each of the selected methods, applied on both datasets and for different parameter adjustments.

[1] https://www.facebook.com/valamobile

[2] https://www.facebook.com/artmotion.net

IV. EXPERIMENTAL RESULTS

A. *Spectral algorithm*

Spectral, as one of the unsupervised clustering algorithms and which was used in the study [22], and the analyzing of literature review we have decided to use it and see how much it will be suitable in our case.

By analyzing the method, adjusting the necessary parameters, preparing the input data (data preprocessing) and applying it to both datasets, we have seen that the method gives very good results, and we consider it suitable for languages with low resources. The classification was made into five main clusters (labeled as 0, 1, 2, 3 and 4) applied to dataset X and three main clusters (labeled as 0, 1 and 2) to dataset Y.

In addition, in the case of both datasets, we used specific parameters such as min and max values which means that these are the ranges which adapt the minimum and maximum values used on the datasets, while appropriate value of PCA is tested before to see how they operate in our cases. In our cases, the PCA method has contributed to determine the correct values within the range of values and facilitate the process for algorithm to cluster appropriate clusters.

Therefore, to do it in the right way, we have identified an interval of values that are applied, but in this way, we must be careful not to exceed this interval. So, the interval of values varies approximately 0.00325343 to 0.96325343 for the case of X set of data and 3.17385658 to 8.15436154 for Y set of data. Hence, just a prototype of values for P1 and P2 columns has been presented in table I.

TABLE I. WEIGHT OF COMMENTS SEPARATED IN 2 PRINCIPAL COMPONENTS (P1 AND P2) FOR X AND Y DATASETS.

No of comments	X dataset		No of comments	Y dataset	
N/A	P1	P2	**N/A**	P1	P2
0	0.00325343	0.05183993	0	3.1738565	1.40960012
1	0.00965092	0.07244956	1	4.8683066	3.66278086
2	0.04361599	0.08094468	2	1.1499502	1.47490599
…	…	…	…	…	…
1323	0.02807612	0.06707625	421	3.37407994	4.86895000
1324	0.21093181	0.16381852	422	2.8240604	8.15436154
1325	0.12649087	0.18273131	423	1.8556728	4.11241329

From the obtained results, we recognized the biased results, and in the case of set of data X, most posts were classified in cluster 1, while for set of data Y were in cluster 0. These deviations have increased our dilemmas, but we have explanations for this occurrence. And the explanation is, since how the discussion takes place on social network for example, when someone writes about specific topic, such as price of services, then others comment a lot because it is not suitable for all users and they could agree or disagree, and we needed to check if all comments have been grouped in a specific cluster or algorithm has mis-grouped in any other cluster. For this reason, we manually checked all printed results, while we identified a higher accuracy of clustering each of comment in certain cluster.

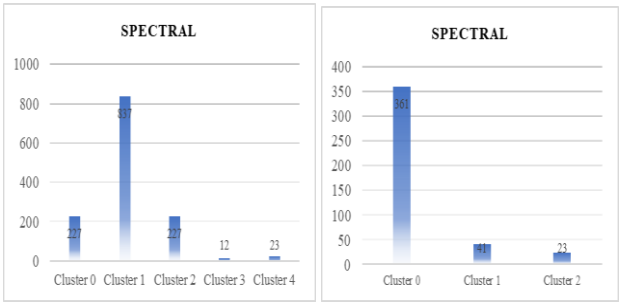


Figure 1. Spectral results using on X and Y datasets.

In Figure 1 we have presented a distribution of comments for both data sets.

To see and prove if the preprocessing is done in the right form, we have used the Silhouette technique to see which cluster has the highest accuracy. By this, we see how many topics are discussed mostly in each dataset.

In the case of dataset X, with the highest accuracy are generated 6 clusters with 69.2%, while for dataset Y are generated 3 clusters with 79.3%.

In table II we have presented in detail for each cluster

TABLE II. CLUSTER ACCURACY OF BOTH DATASETS.

Number of clusters:	Silhouette clusters accuracy of Spectral Clustering	
2	0. 6087469	0.7368314
3	0. 6514980	0.7945238
4	0. 6648371	0.6134574
5	0. 6887702	0.6750648
6	0. 6921797	0.6458574
7	0. 41424305	0.2363203
8	0. 41424305	0.1218392
9	0. 1674287	0.0653652

the accuracy for X and Y datasets while the maximum number of clusters shown is 9. This is because more than 9 clusters have very low percentage of accuracy and we did not present them.

After a whole process of generating clusters, in Figure 2 we have presented visualization of results for certain clusters, while on the left side is for X dataset, and on the right side is for Y dataset.

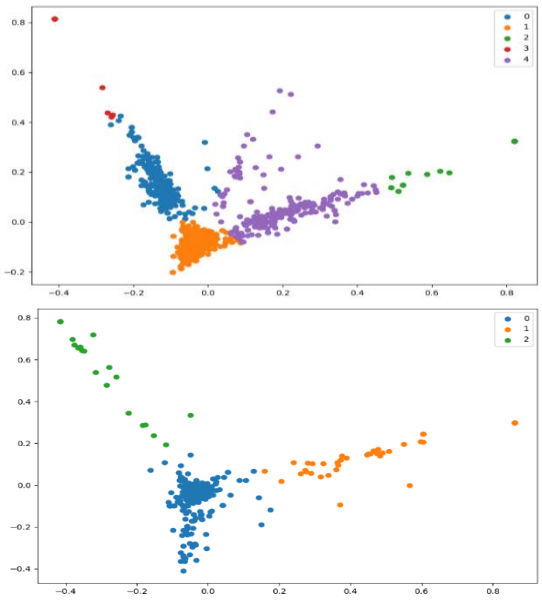


Figure 2. Visualization of data clustering for X and Y datasets.

B. *Agglomerative algorithm*

Agglomeration is hierarchical and works from the bottom up, creating clusters by gathering posts and growing clusters. To analyze its behavior in the case of low-resource languages, we used it to our datasets.

We have applied cosine similarity unlike from other selected algorithms which use PCA. This has contributed directly to increase the performance of visualization at the end stage. While this algorithm has shown a very good ability to link comment threads with others by creating a tree as a whole from the content.

In our case, we measure distance between the comment threads by using Euclidian method because for our dimensionality of data this is the most appropriate. This form of application has helped the algorithm to achieve accuracy and create the most suitable clusters, which means that, from all comments, only 5% have been categorized as outliers (and the conditions for outliers were when they could not link to any cluster while they create cluster by them self's). To be more precise in generating results, we tested different parameters, which correspond to having different results. Some of the parameters we used are in the case of clustering without using lemmatization, Stopwords, vectorization and N-Grams.

In the case when we use all of these parameters, it has generated the best result than in comparison with other tested cases. In Figure 3, we have presented the case, where we use all mentioned parameters for both datasets.

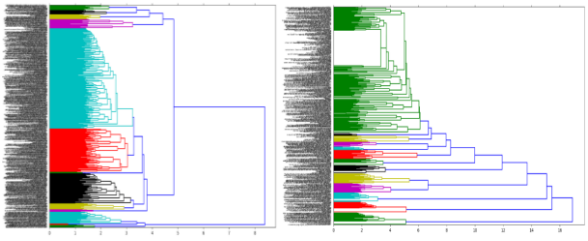


Figure 3. Agglomerative clustering using X and Y sets of data.

Visualizations have differences between two used datasets because grouping is applied into two different contents as it is mentioned in methodology part for whole of the paper. So, in the case of dataset X, a group with teal color has approximately more than 45% of all comments and the other part of the comments thread is on the other groups of the tree linkage. In total, have been generated 6 clusters where each group collect the most appropriate comment in context of sentiment by moving forward to create a tree with the same content as much as it is possible.

In addition, in the case when we use the dataset Y, it has generated 6 groups, but in this case, approximately more than 55% of comment threads are in groups with green color. Grouping of comments in this case Agglomerative algorithm achieved higher accuracy than in the case of X dataset.

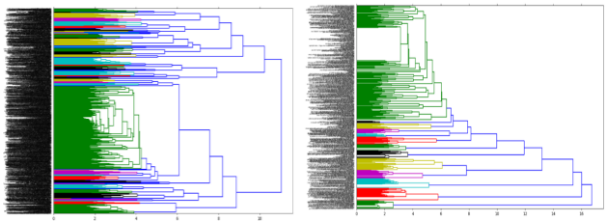


Figure 4. Agglomerative grouping of X and Y sets of data without lemmatization feature.

Results in figure 4, are shown by using other parameters such as Stopwords, N-grams, vectorization, max and min filtering form. So, in this figure, it is not included lemmatization for both cases. Also, here are incorporated N-Grams which it divides terms from 1 to 3, and this has affected by increasing the number of clusters in comparison with other case. This has complicated the process of identifying an appropriate group for each comment thread.

This is because a very large number of groups have been built and determined which one it should be in due to the smallest value from the vectorization. After a manual check of printed results, we identified that in any case within the set of data X and Y, it has clustered in a wrong group any thread of comment while the same comment it shouldn't be there. And this we called as outliers while it hasn't achieved to link in any cluster. So, in this way we have seen that the algorithm has generate a cluster only because they have a term that makes it stand out in terms of writing, while lemma parameter wasn't applied. If a term was there, it has transformed just like "aktivizimi", " aktivizon", and "aktivizu", into a single expression such as "aktivizohet".

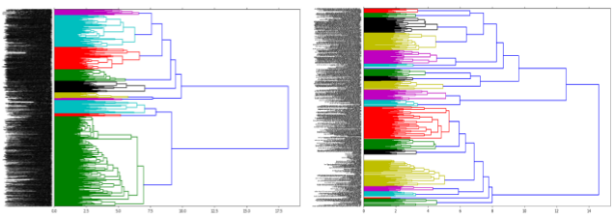


Figure 5. Agglomerative grouping of X and Y sets of data without vectorization, Stopwords and lemmatization.

Figure 5, depicts the final one as commitment in order, wherein we really do not use vectorization, lemma, and the collection of Stopwords. Even though the finding isn't really accurate, the methodology inside the pattern of dendrograms intends to group comment threads far greater than when just lemmas or a list of Stopwords have been used.

In the case of X dataset, most comment threads are classified in a cluster with green color, while in Y dataset, comment threads are grouped more in a cluster with yellow and red color.

C. *Affinity propagation algorithm*

Affinity propagation is based on the implementation of propaganda for each group of content discussed in a dataset. The algorithm works based on the similarity of a comment with others, and this continues until the cluster itself is made. This algorithm is used in different fields and some of them are explained in the research [12]. This algorithm deals with special attributes, analyses and uses datapoints from the dataset to group them by the similarity calculated and found. So, it analyses the comment's content and the relationship between comments using preferences for themselves and other comments.

In the following, we use the same datasets, as in the previous cases, and try to perform parameter tuning on the affinity propagation algorithm and test it in these datasets.

Each comment begins on its own, and to form a cluster, other comments must display the greatest similarity with the given comment. Then, this comment will be joined with the previous one, and so on, until the clusters are formed. This algorithm is based on a preference value on how similar each comment shows itself to others.

Based on the ways that this algorithm operates, the best values for both datasets have been found. This helps to determine for each comment in which cluster it should be, therefore, the values vary from -1, -2, and -3. However, the third case is shown to be the best group by generating better results compared to other cases.

In the case in which the value -2 has been applied, sic clusters generation for X set of data have been affected. Consequently, the distribution of comments is determined based on the number of optimal groups generated that means that the most comments were generated in group 4 (a total of 724), then we have group 5 (213 comments) and so on. Further, the formation of the centroid was calculated based on the highest density of comments at that point, and here the considered terms are “internet, packages and problems, offers/services”.

On the other hand, in the case of dataset Y, the value -1 was applied and 5 groups were generated in total in which 320 comments were grouped in group 2, then in group 0 with 43 comments, and so on. These results for both set of data are shown in figure 6.

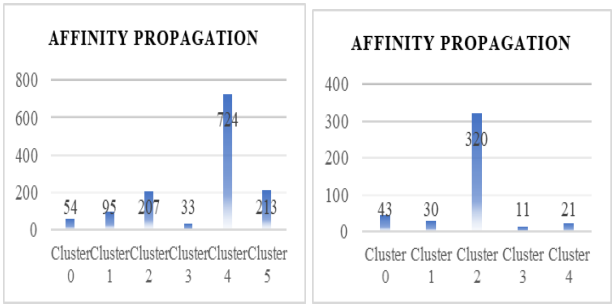


Figure 6. Affinity propagation clustering using X and Y dataset.

The illustration of this is grouping for both cases are shown in figure 7.

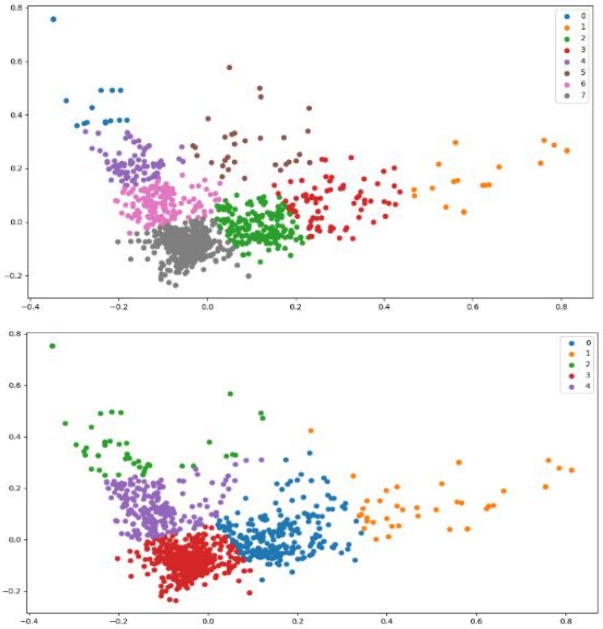
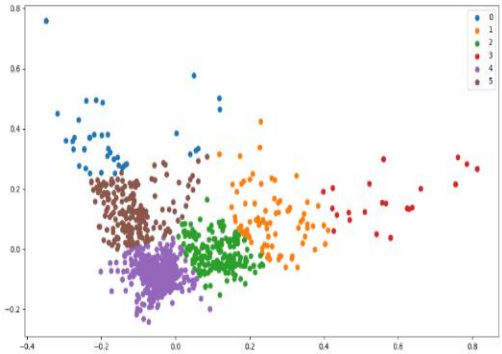


Figure 7. Visualization of results of Affinity propagation clustering using preferences = -1 (in the left side) and -3 (on the right side) of X dataset.

For "preferences = -3" we got 5 clusters, the clusters were created based on keywords such as: problems, Vala, internet, offers and 3G & 4G services, and most of the comments were collected in the same cluster, i.e., 0. Comparing based on the results obtained from all three conducted experiments, we conclude that preferences -2 show better and more accurate results. The “preference = -2” was also shown to be the best choice for the Y dataset as well. The results are shown in Figure 8.



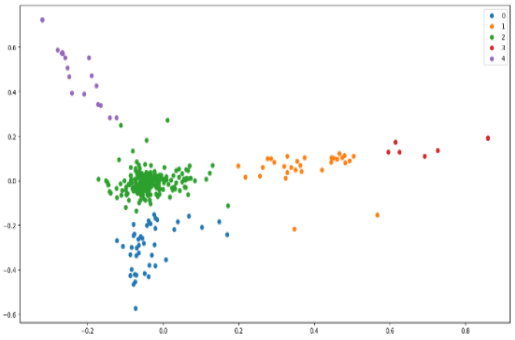


Figure 8. *Visualization of results of Affinity propagation clustering using preferences = -2 for X and Y dataset.*

D. Mean Shift algorithm

The focus of Mean Shift is solely based on the widely held belief that averaging nearly location information results in a greater concentration, and therefore more traditional areas. This algorithm was also applied to find solutions including graphics processing according to [23], [24], in addition to the exhaust system and surface noise removal [25]. A specific used parameter for classifying comment threads by this algorithm is quantile, which is a value that assists to define the data rate to produce the findings. Quantile values have been defined for both datasets and these are applied according to the content of each of them.

At first, we discovered the boundaries, where the classifier works effectively, while it is used to have a specific value that is available to determine the best specific instance of grouping, so this boundary seems to be 0.2 to 0.8 in the scenario of set of data X. In comparison, for set of data Y, the boundary seems to be from 0.321 to 0.621 which produces exactly 8 groups. In this case, when we attempt to increase the value to 0.721, we get 6 groups. Even so, when we increase the value once more, we get only two groups. Considering the above-mentioned range of possibilities (0.02, 0.4, and 0.6 for X set of data), we prefer throughput significance 0.451 as the best optimum value, as well as throughput significance of 0.421 for Y set of data. The number of groups formed by the applied values is presented in figure 9 for both sets of data.

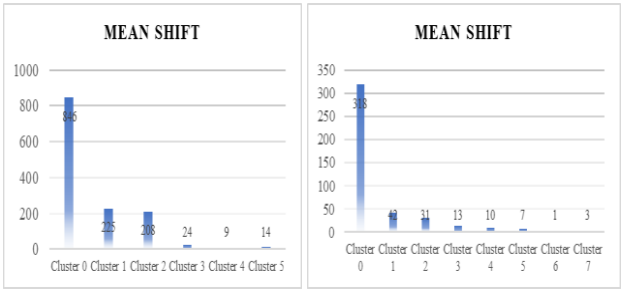


Figure 9. Mean Shift grouping using X and Y datasets.

In figure 10, are shown illustrative forms of how the classification has been generated for all groups for both datasets, applying the parameters mentioned above.

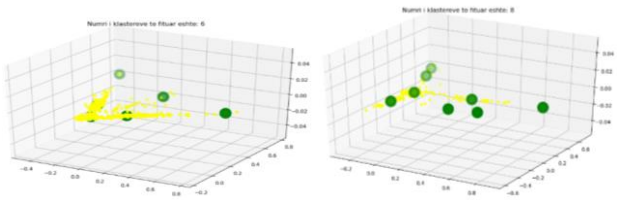


Figure 10. Mean Shift grouping using X and Y datasets.

Then, after several tests we did with values from 0.2 to 0.8, as a constant, it is obtained the following results, with the value of 0.2 are obtained 8 clusters, for 0.3 = 7 clusters, 0.4 = 6 clusters, 0.5 = 7 clusters, 0.6 = 6 clusters, 0.7 = 3 clusters.

After several tests with the increase of the quantile value, in which it exceeds the value of 0.8, then the number of clusters decreased extremely. This proves that the ideal value/limit is between 0.2 and 0.8. Meanwhile, the values with the best results are 0.2, 0.4, and 0.6, all within the above range.

The above quantile values help a lot in identifying the bandwidth of the algorithm (since this algorithm works in that form), and from the value 0.3, it is obtained this value: 0.134282, which subsequently generates 13 clusters. Then we got a 0.4 value from the bandwidth: 0.142802 and it generates 12 clusters, then 0.5 with bandwidth: 0.152266 generates 10 clusters. From this, it seems that as the quantile/bandwidth values increase, the number of clusters decreases.

Further, below we have presented the values obtained during the process of identifying the most suitable cluster:

- 0.182 generates 9 clusters.
- 0.192, 0.211 generate 8 clusters,
- 0.251, 0.281 generate 7 clusters,
- 0.481, 0.451, 0.581 generate 6 clusters,
- 0.681 generates 3 clusters.

Considering the value of 0.251 of bandwidth, the distribution of the clusters was done mostly in cluster 0 with 846 comments. In cluster 1, 224 comments have been collected and in cluster 2 there are 201 comments. Meanwhile, in the other clusters there is a much smaller number of comments, which in cluster 4 and 5 can also be seen as outliers of these results in order not to lose accuracy due to some comments.

V. CONCLUSIONS

The stages of pre-processing and clustering of comments extracted from social networks by analyzing them is now very important for many users, especially companies, as it facilitates and guides them in fulfilling the requests of their users. Therefore, the use of appropriate algorithms to classify each opinion correctly is now essential.

In this paper, we have used two datasets with different contents of low-resource language (in the Albanian language), which, through algorithms, aim to present the groups that are ideal for the topics being discussed. By combining these elements with other algorithms, we aim to determine their perception of the services offered, which is a crucial aspect of our future work. Therefore, to contribute

even a little, through this research, we have selected a total of four algorithms in which, in each of them, the preprocessing stages and the most appropriate parameters are applied, depending on the form of the output that we received from the previous stage.

For example, in the Spectral, Mean Shift, and Affinity Propagation algorithms, the PCA technique is applied, while in Agglomerative cosine similarity. So far, TF-IDF vectorizer, Stopwords list, lemmatization of words, and min and max filtering have been used for all algorithms. Also, for all algorithms, we used Silhouette methods that predict and verify if the results are generated correct from the algorithms used.

The algorithm that generated the best content grouping was shown to be Agglomerative, followed by Affinity Propagation, which once again prove that they are very good in cases where data density is applied. While Spectral generates a very good and suitable number of groups and ideal visualization against the number of groups. This work, in addition to its contribution, also has its limitations, and something like that was access to the most prestigious databases (mainly paid ones) from the literature review part.

Furthermore, based on our thorough research and analysis of existing literature, we have not encountered any prior implementation of similar algorithms in low-resource languages. This posed a significant challenge for us in terms of making comparisons with our own work.

VI. REFERENCES

1 A. Kadriu and L. AbaziA comparison of algorithms for text classification of Albanian news articles,"*Entrenova-Enterprise Research Innovation Conference*Vol. 3 No. 1p. pp. 62–682017

2 A. Kadriu, L. Abazi, and H. AbaziAlbanian text classification: Bag of words model and word analogies*Business Systems Research Journal, doi: 10.2478/bsrj-2019-0006*Vol. 10No. 1p. pp. 74–87Apr. 2019

3 E. Trandafili, N. Kote, and M. Biba,Performance evaluation of text categorization algorithms using an Albanian corpus*Advances in Internet, Data & Web Technologies*p. pp. 537–5472018

4 Stewart, G. Al-Khassaweneh, M. AnAn Implementation of the HDBSCAN* Clustering Algorithm*Appl. Sci., https://doi.org/10.3390/app12052405,* 125pp. 1-212022

5 Khader, M.M.; Adel, M,Modeling and Numerical Simulation for Covering the Fractional COVID-19 Model Using Spectral Collocation-Optimization Algorithm*Fractal and Fractional, https://doi.org/10.3390/fractalfract6070363,* 6363pp. 1-192022

6 Tom Chou, Sihong Shao, Mingtao XiaAdaptive Hermite spectral methods in unbounded domains *Applied Numerical Mathematics* vol. 183p. 201–2202023

7 Kenneth Joseph, Ryan J. Gallagher, Brooke Foucault WellesWho Says What with Whom: Using Bi-Spectral Clustering to Organize and Analyze Social Media Protest Networks*Computational Communication Research*22pp. 153-1742020

8 Arshia Naeem et al.Development of an efficient hierarchical clustering analysis using an agglomerative clustering algorithm*CURRENT SCIENCE*1176pp. 1045-10532019

9 Sepandar D. Kamvar, Dan Klein, Christopher D. ManningInterpreting and Extending Classical Agglomerative Clustering Algorithms using a Model-Based Approach*ICML '02: Proceedings of the Nineteenth International Conference on Machine Learning*pp. 283-2902002

10 S S Pavithra, S Chitrakala, C M BhattSpectral Clustering of Events in Social Media Flood Images based on Multimodal Analysis*7th International Conference on Electrical Energy Systems (ICEES), DOI: 10.1109/ICEES51510.2021.9383727*Chennai, India2021

11 Jia, H., Ding, S., Meng, L. et al.A density-adaptive affinity propagation clustering algorithm based on spectral dimension reduction*Neural Comput & Applic, Springer, https://doi.org/10.1007/s00521-014-1628-7*25p. 1557–15672014

12 Limin WANG, Zhiyuan HAO, Xuming HAN, Ruihong ZHOUGravity Theory-Based Affinity Propagation Clustering Algorithm and Its Applications*Tehnički vjesnik*254pp. 1125-11352018

13 Ge, H. Wang, L. Pan, H. Zhu, Y. Zhao, X. Liu, M.Affinity Propagation Based on Structural Similarity Index and Local Outlier Factor for Hyperspectral Image Clustering*Remote Sensing, MDPI, https://doi.org/10.3390/*141195p. 20222022

14 Limin Wang, Kaiyue Zheng, Xing Tao & Xuming HanAffinity propagation clustering algorithm based on large-scale data-set*InternatIonal Journal of Computers and applICatIons, DOI: 10.1080/1206212X.2018.1425184,* 403pp. 1-72018

15 Campello, R.J.; Moulavi, D.; Sander, J.Density-based Clustering Based on Hierarchical Density Estimates,"*Advances in Knowledge Discovery and Data Mining, Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer*pp. 160-1722013

16 Melvin et al,Visualizing correlated motion with HDBSCAN clustering,"*Protein Science*271pp. 62-752017

17 Singh, U., Saraswat, A., Azad, H.K. et al.Towards improving e-commerce customer review analysis for sentiment detection*Sci Rep 12, https://doi.org/10.1038/s41598-022-26432-3,* p. 022022

18 F. Hu, J. Liu, L. Li et al. "Community Detection in Complex Networks using Node2vec with Spectral Clustering*PHYSA A, Elsevier,* pp. 1-302019

19 Charles Kumah et al.,Unsupervised segmentation of printed fabric patterns based on mean shift algorithm*The Journal of The Textile Institute, Taylor and Francis, DOI: 10.1080/00405000.2020.1867413,* pp. 1-92021

20 Javier Fumanal-Idocin et al.Combinations of Affinity Functions for Different Community Detection Algorithms in Social Networks*Conference: Hawai International Conference on System Sciences,* Hawai: Social and Information Networks2022

21 A. Dine GithubGitHub05072018Available https://github.com/ arditdine /albanian-nlp/blob/master/corpus /stopwords/ albanian. 02062022

22 Mërgim H. HOTI & Jaumin AJDARIUnsupervised Clustering of Comments Written in Albanian Language*International Journal of Advanced Computer Science and Applications (IJACSA*128pp. 287-2922021

23 D. DeMenthonSpatio-temporal segmentation of video by hierarchical mean shift analysis*In Statistical Methods in Video Processing Workshop (SMVP)*Copenhagen, Denmark2002

24 S. Paris and F. DurandA topological approach to hierarchical segmentation using mean shift*in Proc. of the 2007 IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR'07)*Minneapolis, MN, USA2007

25 Miguel A. Carreira-PerpinanA review of mean-shift algorithms for clustering*ArXiv, abs/1503.00687*pp. 1-282015

Interactive redescription set mining and exploration

Iva Kozjak* and Matej Mihelčić*
* University of Zagreb, Faculty of Science/Department of Mathematics, Zagreb, Croatia
kozjak.iva@gmail.com, matmih@math.hr

***Abstract*—InterSet is a client-server web application that allows targeted and contextual redescription set exploration. The main drawback of this tool is that it uses fixed, precomputed sets of redescriptions to allow obtaining novel knowledge. In this work, we significantly extend the capabilities of InterSet by adding the possibility to create new redescriptions on predefined data. The main advantage of this is that the user can create new redescriptions in any step of exploration depending on the context and current research hypothesis, utilizing entity or attribute constraint-based redescription mining and enriching the existing set of redescriptions with newly obtained knowledge. This procedure further enhances the potential of forming and exploring new research hypotheses using large sets of redescriptions. The redescription mining ability of the InterSet tool is achieved by utilizing the CLUS-RM algorithm. Usefulness of the obtained environment for interactive redescription set creation, targeted and contextual exploration is demonstrated on real-world use case datasets.**

Keywords*—*InterSet, interactive redescription mining, redescription set exploration, world countries, dblp authorship network, phenotypes

I. INTRODUCTION

Redescription mining [1] is a field of data mining that has two main goals: a) discover subsets of entities in the data that can be described in multiple ways, b) construct appropriate redescriptions of discovered subsets of entities. Finding multiple descriptions (called redescriptions) of the same or very similar subsets of entities, in an unsupervised manner, is the main distinguishing factor of this approach compared to the related fields of clustering [2], [3], conceptual clustering [4], [5], rule learning [6] and subgroup discovery [7]–[9]. In contrast to the association rule mining [10]–[12], where discovered rules form implications, descriptions obtained by redescription mining are in strong equivalence relations.

Redescription mining approaches either return all redescriptions that satisfy some predefined quality criteria (e.g. [1], [13]–[15]) or use redescription set optimization to obtain a smaller set of diverse redescriptions that offer favourable trade-off between many different quality criteria (e.g. [16], [17]). Approaches that utilize filtering, often create a large set of results. This set is very hard to explore and extremely difficult to reason about without the use of additional tools.

There exist two main tools that allow obtaining additional knowledge from sets of redescriptions: the Siren [18] and the InterSet [19]. The Siren is an interactive redescription mining framework that allows mining redescriptions, inspecting and expending individual redescriptions, filtering, embedding instances described by a redescription into some lower-dimensional plane etc. InterSet [19] allows targeted and conceptual redescription set exploration. The main idea was to allow exploration of large sets of redescriptions by allowing redescription subset creation and selection, based on entities described by redescriptions, attributes contained in their queries and general redescription quality measures. The tool also allows statistical analyses of individual redescriptions and relating different redescriptions. The main drawback of the InterSet tool, prior to this work, was that it worked with the precomputed sets of redescriptions.

In this work, we present a way to remove the aforementioned disadvantage of the InterSet by incorporating interactive and constrained-based redescription mining into this tool. The resulting tool, among others, has the ability to perform interactive redescription mining, potentially using entity, attribute-based constraints or the combination thereof and to allow interactive targeted and contextual redescription set exploration.

II. DEFINITIONS AND NOTATION

The input data $\mathcal{D}$ is in a form of one or more tables $D_i,\ i = 1, \ldots, n$, where each table contains values of Boolean, categorical or numerical attributes for one, shared set of entities. We denote a set of all attributes with $\mathcal{A}$ and a set of entities with E. Data tables D_i form views of the input data $\mathcal{D}$. $\mathcal{W} : \mathcal{A} \mapsto \mathbb{N}$ is a mapping that returns the view of some attribute $a \in \mathcal{A}$. Views can, for example, represent repeated experiments, experiments performed in different conditions or by different institutions, these can be natural groupings of attributes based on domain knowledge or the current research problem.

Knowledge obtainable from the data is represented in a form of logical formulae, where each formula consists of variables (attributes with corresponding conditions) and logical connectives (conjunction, disjunction and negation). Such formulae are called queries (q). The choice of allowed logical operators and the definition of rules of generative grammar define the set of allowed queries (Q), also called a query language. A set of entities described by a query q (satisfying the logical formula) form the support set of this query $supp(q)$. Redescriptions are tuples of queries $R = (q_1, q_2, \ldots q_n)$ where $supp(q_1) \sim$

$supp(q_2) \sim \cdots \sim supp(q_n)$ (queries forming redescriptions should have similar support sets). The set of all attributes contained within queries of a redescription R is denoted by $attrs(R)$. Similarity of support sets of queries is measured using the Jaccard index [20]. The corresponding similarity score $J(R) = (|\cap_{i=1}^{n} supp(q_i)|)/(|\cup_{i=1}^{n} supp(q_i)|)$ represents the accuracy of a redescription R. Statistical significance of a redescription is measured as: $p(R) = \sum_{i=|supp(R)|}^{|E|} \binom{|E|}{i} \prod_{j=1}^{n}(|supp(p_j)|/|E|) \cdot (1 - \prod_{j=1}^{n}(|supp(p_j)|/|E|))$. The complexity of a redescription is measured as $comp(R) = min((|attrs(R)|/k), 1)$, where k represents the maximal desired size of $|attrs(R)|$.

A set of redescriptions is denoted by $\mathcal{R}$. The average redundancy of a redescription R, contained in a set of redescriptions $\mathcal{R}$, based on attributes contained in its queries is denoted as: $AAJ(R, \mathcal{R}) = (\sum_{R_i \in \mathcal{R},\ R_i \neq R} J(attrs(R), attrs(R_i)))/(|\mathcal{R}| - 1)$. The average redundancy of a redescription based on entities it describes is denoted as $AEJ(R, \mathcal{R}) = (\sum_{R_i \in \mathcal{R},\ R_i \neq R} J(supp(R), supp(R_i)))/(|\mathcal{R}| - 1)$.

We say that a redescription R is associated with the entity cluster $\mathcal{G}$ if $J(supp(R), G) \geq \delta$, where δ denotes the required similarity level. The average homogeneity of an entity cluster $\mathcal{G}$ is defined as $avHom(\mathcal{G}) = \sum_{R \in \mathcal{G}} J(supp(R), \mathcal{G})/|\mathcal{G}|$.

Given a set of user-defined constraints on redescriptions $\mathcal{C}$, the input data $\mathcal{D}$ and the query language Q, the goal of redescription mining is to discover all redescriptions that satisfy constraints defined in a set $\mathcal{C}$.

III. Enabling Interactive Redescription Mining with the InterSet

InterSet is a web application, where the client part was created using standard technologies: `HTML`, `CSS`, `JavaScript` and the `angular.js`. The server part of the InterSet tool was originally implemented using the `node.js` environment (see Figure 1).

In this work, we replace the `node.js` backend with the `Java Spring Boot`. This allows stronger integration of the CLUS-RM algorithm [17], extended with the ability to perform constraint-based redescription mining [21], implemented in the `Java` programming language, with the InterSet tool [19].

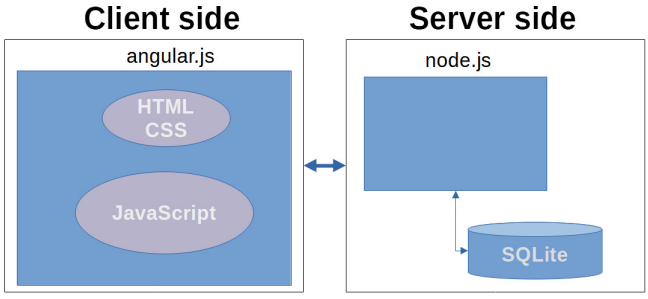


Fig. 1: Architecture of the InterSet tool prior to the proposed extensions.

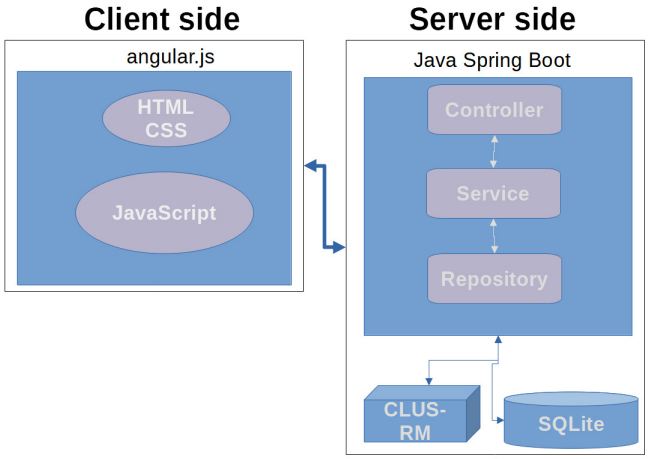


Fig. 2: Architecture of the InterSet tool after the proposed extensions.

The newly obtained structure of the InterSet can be seen in Figure 2. The *controller* package of the InterSet tool contains one class per dataset: the *DBLPRedescriptionsController*, *PhenotypeRedescriptionsController* and *TradeRedescriptions-Controller*. The *service* package contains a class *RedescriptionsService* and the *repository* package contains a class *RedescriptionsRepository*.

The CLUS-RM algorithm [21] has been adopted for integration with the InterSet tool. The main modification is to allow adding joined entity-attribute constraints into the constrained-based redescription mining process. The constraints have been modified to contain three main categories: *none* (no constraints, perform regular redescription mining), *soft* (at least one entity and at least one attribute from the predefined set of constraints must be contained in the support and attribute sets of the newly created redescriptions), *hard* (all entities and attributes from the predefined set of constraints must be contained in the support and attribute sets of the newly created redescriptions).

The InterSet tool [19] contains three main exploration views: the Entity view, the Attribute view and the Redescription quality view. The first two are modified to allow interactive redescription mining, but the newly added redescriptions in any of these views cause updates of the statistics reported in the Redescription quality view.

A. Newly added functionality to the InterSet

When the user selects one cluster from the SOM (Self-organizing map) [22], contained in the *Entity - based exploration* tab, a table of entities contained in this cluster is displayed. The main goal of enabling the use of constrained-based redescription mining is to allow discovering new redescriptions that describe all or large subsets of entities contained in the selected cluster. This allows expanding the knowledge base and can lead to proving, disproving or forming new research hypotheses.

The most important settings of the CLUS-RM algorithm can be adjusted through the user interface, and

the required constraints and type can be selected using corresponding radio-buttons. When both the SOM cluster and the constraint type are selected, the CLUS-RM algorithm can be invoked. By clicking the run button, the selected properties are sent to the server side of the application where the CLUS-RM algorithm is executed, and the result file is created. For each of the resulting redescriptions R, the Jaccard similarity index of entities with the selected cluster G is being calculated as $J(supp(R), G) = (|supp(R) \cap E|)/(|supp(R) \cup E|)$. Redescriptions and the aforementioned Jaccard similarity are saved and sent to the client side, where they are displayed in a table. This table allows determining the support sets of the newly created redescriptions, which redescriptions describe entities from the selected cluster in the most suitable manner, and it provides information about the most similar redescription, to each of the newly created redescriptions, already contained in the database. Based on the information provided in this table, the user can save appropriate redescriptions into the database.

In the *Attribute - based exploration* tab, the user can select pairs of attributes from the heatmap that visualizes statistics about frequency and co-occurrence of attributes contained in the queries of redescriptions from the database. The main advantage of incorporating constraint-based redescription mining in this view is to allow discovering new redescriptions containing interesting set of attributes in their queries. This allows confirming, exploring, expanding or forming hypothesis about associations of attributes in the dataset.

Similarly as in the *Entity - based exploration* tab, depending on the choice from the heatmap, the appropriate set of attributes is saved and displayed as a potential set of attribute constraints. The appropriate radio-buttons allow selecting constraint type, and the CLUS-RM algorithm can be invoked using the run button. Again, a table is created containing all newly created redescriptions with information about the most similar redescription already contained in the database. An additional table provides description of the attributes contained in the queries of a selected newly discovered redescription. Based on the provided information, the user can save appropriate redescriptions in the database.

Running the CLUS-RM algorithm with conditions on both entities and attributes is possible if the selection (set of entities or attributes) is made and saved in one of the views and then loaded in the other view. In that case, the set of selected attributes is displayed alongside entities from the selected cluster.

B. Saving newly created redescriptions into the database

The user can save subsets of newly found redescriptions from the table by marking the appropriate checkboxes in the table. The save action causes updates of multiple tables in the database, necessitates computation of redescription set based quality scores (AEJ,AAJ) of all newly added redescriptions and re-computation of these quality scores for all redescriptions previously contained in the database.

For a set of redescriptions $\mathcal{R}_{db}$ and a set of redescriptions $\mathcal{R}_{new}$, the AEJ and AAJ for redescriptions contained in $\mathcal{R}_{new}$ are computed as:
$AEJ(R, \mathcal{R}_{db} \cup \mathcal{R}_{new}) = (\sum_{R_i \in \mathcal{R}_{db} \cup \mathcal{R}_{new},\ R_i \neq R} J(supp(R), supp(R_i)))/(|\mathcal{R}_{db} \cup \mathcal{R}_{new}| - 1)$.
$AAJ(R, \mathcal{R}_{db} \cup \mathcal{R}_{new}) = (\sum_{R_i \in \mathcal{R}_{db} \cup \mathcal{R}_{new},\ R_i \neq R} J(attrs(R), attrs(R_i)))/(|\mathcal{R}_{db} \cup \mathcal{R}_{new}| - 1)$. The updates for redescriptions contained in $\mathcal{R}_{db}$ are computed as:
$AEJ(R, \mathcal{R}_{db} \cup \mathcal{R}_{new}) = (AEJ(R, \mathcal{R}_{db}) \cdot |\mathcal{R}_{db} - 1| + \sum_{R_i \in \mathcal{R}_{new}} J(supp(R), supp(R_i)))/(|\mathcal{R}_{db} \cup \mathcal{R}_{new}| - 1)$.
$AAJ(R, \mathcal{R}_{db} \cup \mathcal{R}_{new}) = (AAJ(R, \mathcal{R}_{db}) \cdot |\mathcal{R}_{db} - 1| + \sum_{R_i \in \mathcal{R}_{new}} J(attrs(R), attrs(R_i)))/(|\mathcal{R}_{db} \cup \mathcal{R}_{new}| - 1)$.

Complete code of the current version of the tool and links to the older versions of the InterSet tool and the CLUS-RM algorithm can be found on: https://github.com/ivakozjak/MIPRO2023.

IV. Applying InterSet for scientific discovery

We apply InterSet to the three well-known use-case datasets: the Country dataset, the Phenotype dataset and the DBLP dataset (see [19] for more details). For all presented redescriptions $p(R) \leq 0.01$.

A. Discovery on the Country dataset

We perform constraint-based redescription mining with entity constraints on 3 different clusters from the SOM map on the Country dataset, as demonstrated in Figure 3.

The top left cluster has the highest homogeneity 0.3374 and 2737 redescriptions describing at least one Country contained in this cluster. This cluster contains 11 highly developed Western European countries that share many common trading and socio-demographic patterns. Using the presented extensions of the InterSet, we managed to find redescription R_1 (first in Table I) with entity Jaccard of 0.91 with this cluster. After adding this redescription to the database, the overall homogeneity of the top left cluster increased to 0.3376. The newly discovered redescription describes that small percentage of population of these countries was employed in agriculture, there was high trade in Stocks and that these countries exhibited significantly larger export than import of specialized machinery and had high import of manufactured goods. Although there already exists redescriptions with equal support set in the database, the newly discovered redescription has an attribute redundancy of 17% with the most similar redescription, thus it provides novel knowledge about the entities contained in the first cluster.

The neighbouring cluster (first in the second row) mostly contains Central and Eastern European countries. This cluster originally had homogeneity of 0.2163 and there existed 3002 redescriptions describing at least one country from this cluster. The newly discovered redescription R_2 (second in Table I) with entity Jaccard of 0.72 with the aforementioned cluster when added to the database

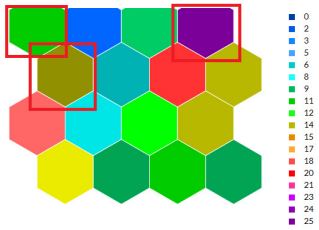


Fig. 3: Studied clusters on the Country dataset.

increases the overall homogeneity of this cluster to 0.2165. The most important features described by R_2 include higher unemployment of females ($2.5 - 18\%$), larger percentage of bad employment ($5 - 35\%$), import of non-metallic manufactures and labour-intensive manufactures, pronounced export of metal manufactures and mixed ratio of export to import of fuels. The newly discovered redescription describing the second selected cluster has a maximal redundancy of 90% with a redescription already contained in the database, but the attribute redundancy of only 9% with a redescription having most similar support.

The third selected cluster contains a diverse set of 28 countries (e.g. Afghanistan, Algeria, Angola, Cameron, Congo, Iraq, Mali, Niger, Tajikistan, Yemen). This cluster originally had homogeneity of 0.2396 and had 2019 redescriptions describing at least one country from this set. For this cluster, we discovered 7 redescriptions with entity Jaccard ≥ 0.4 with the selected cluster. After adding all 7 discovered redescriptions, the homogeneity of this cluster increased to 0.2403. Redescription R_3 (third in Table I) with entity Jaccard of 0.48 with the selected cluster is the most related redescription obtained in our analyses. This redescription describes 27 countries from the selected cluster and 18 countries from the neighbouring 3 clusters. Thus, it contains relevant information for many related countries. The main characteristics of these countries are relatively small percentage of population aged 64 or above ($2.2 - 5.0\%$), high percentage of rural population ($22.1 - 80\%$) and higher import than export of medicinal and pharmaceutical products. Level of export of beverages and plastics in non-primary form, level of export to import ratio of footwear and plastics in primary form and export to import ratio of fuels varies between countries. Newly discovered redescriptions associated with the third selected cluster have a maximal entity redundancy with other redescriptions contained in the database of $58 - 69\%$ and attribute redundancy of $0 - 17\%$ with redescriptions having the most similar support. Thus, newly discovered redescriptions contain useful knowledge, previously unavailable in the database.

TABLE I: Redescription examples obtained on the Country dataset.

q_1	q_2	J	$\lvert supp \rvert$
$1.2 \leq E_{AG} \leq 4.9 \wedge$ $27.0 \leq ST \leq 166.6$	$1.1 \leq E/I_{SM} \leq 4.3$ $\wedge\ 58 \leq I_{MG} \leq 78$	0.67	12
$2.5 \leq U_F \leq 28.0 \wedge$ $5.0 \leq E_{bad} \leq 32.1$	$1 \leq I_{NMM} \leq 2 \wedge$ $0.1 \leq E/I_F \leq 1.2 \wedge$ $8 \leq I_{LIRIM} \leq 16 \wedge$ $2 \leq E_{MM} \leq 6 \wedge$	0.53	18
$2.2 \leq P_{64} \leq 5.0 \wedge$ $22.1 \leq R_P \leq 88.8$	$0 \leq E/I_{MP} \leq 0.19 \wedge$ $0 \leq E_B \leq 2 \wedge$ $0 \leq E/I_{FW} \leq 9.2 \wedge$ $0 \leq E_{Pl} \leq 1 \wedge$ $0 \leq E/I_F \leq 161.2 \wedge$ $0 \leq I_{PPSN} \leq 1 \wedge$ $0 \leq E/I_{PPF} \leq 9.2 \wedge$	0.54	55
$0 \leq C_{Cv} \leq 53.5 \wedge$ $2.2 \leq P_{64} \leq 11 \wedge$ $50 \leq P_{15,64} \leq 73$ $\wedge\ 18 \leq P_{14} \leq 47.6$	$0 \leq E/I_{SM} \leq 0.64$	0.63	103
$10.6 \leq P_{64} \leq 24.4$	$0.65 \leq E/I_{SM} \leq 61.9$	0.5	29
$0.3 \leq W_R \leq 1.9 \wedge$ $59 \leq L_M \leq 67.8 \wedge$ $30 \leq E_{IM} \leq 50 \wedge$ $11 \leq ST \leq 167 \wedge$ $13 \leq C_{Cv} \leq 100 \wedge$ $-0.2 \leq P_G \leq 0.5$ $\wedge\ 22.4 \leq R_P \leq 47.2$	$1 \leq E/I_{MT} \leq 1.8 \wedge$ $2 \leq E_{MM} \leq 5 \wedge$ $0.18 \leq E/I_{DP} \leq 3.1$ $\wedge\ 0 \leq E/I_{CR} \leq 0.95$	0.75	9

We test two attribute-based constraints. First, we search redescriptions with hard constraints on the most-frequently co-occuring attributes P_{64}, E/I_{SM}. Using these constraints, we managed to obtain new redescriptions containing both attributes in their queries. Redescriptions R_4 and R_5 (fourth and fifth in Table I) are examples of such redescriptions. Both redescriptions have very small attribute redundancy ($1 - 4\%$) with respect to used attributes to redescriptions with the most similar support contained in the database. Thus, they provide novel associations of the selected pair of attributes.

Soft attribute constraints C_{Cv}, E/I_{MP} allow discovering redescriptions that contain either of these attributes in their queries. Redescription R_6 (sixth redescription in Table I), discovered using soft constraints on the aforementioned attributes, contains C_{Cv} but not E/I_{MP}.

B. Discovery on the Phenotype dataset

Using the proposed extensions of the InterSet tool, we discovered a redescription R_7 (first in Table II) with entity Jaccard of 0.97 with the first (top left) cluster in the SOM map. Although there existed a redescription with equal support set in the database, the newly discovered redescription has an attribute redundancy of 10% with this redescription, providing novel information about the target subset of bacteria. This redescription describes unicellular photosystem bacteria residing in the water. After adding this redescription to the database, the homogeneity of the first cluster increased from 0.5160 to 0.5163. Redescrip-

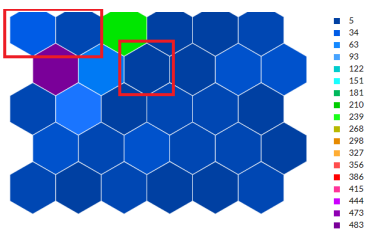


Fig. 4: Studied clusters on the Phenotype dataset.

tion R_8 (second in Table II) has the entity Jaccard of 0.79 with the second selected cluster in the first row.

TABLE II: Redescription examples obtained on the Phenotype dataset.

q_1	q_2	J	$\|supp\|$
$unicPhotosyOF$	$C_{260} \wedge C_{349} \wedge C_{321} \wedge C_{1624} \wedge C_{5474} \wedge C_{1530} \wedge C_{99}$	0.84	32
$hlHsArSFNaCl \wedge \neg TRHypTP \wedge \neg GSP$	$C_{4749} \wedge C_{442} \wedge \neg C_{3890} \wedge \neg C_{1481} \wedge \neg C_{4412} \wedge \neg C_{1753}$	0.61	17
$cytFeHCDis \wedge \neg unicPhotosysOF \neg OF \wedge \neg TPHypTP \wedge \neg GSP \wedge oxSRA$	$(\neg C_{3102} \wedge \neg C_{4129} \wedge \neg C_{3764} \wedge \neg C_{3373} \wedge \neg C_{419} \wedge C_{442} \wedge C_{3012} \wedge C_{2033}) \vee (\neg C_{3580} \wedge \neg C_{3611} \wedge \neg C_{1205} \wedge \neg C_{2926} \wedge \neg C_{1844} \wedge C_{1183} \wedge C_{4881} \wedge C_{1379})$	0.67	12
$oxSA \wedge shCoc \wedge \neg metSulf \wedge \neg metMeth \wedge hypAHdrVHet$	$C_{3354} \wedge C_{4025}$	0.67	10
$brChUlMyEnd \wedge \neg khabSl \wedge \neg ecsyEnv$	$C_{1853} \wedge C_{5468} \wedge \neg C_{123} \wedge \neg C_{4286} \wedge \neg C_{3373}$	0.64	7

Although there exists a redescription with very similar support set with this redescription (entity Jaccard of 0.94), their attributes are different, thus we obtained novel knowledge about bacteria contained in the target cluster. Bacteria described by this redescription live in extreme (hypersaline) environments. The homogeneity of the second cluster increased from 0.7130 to 0.7136 after adding the newly discovered cluster to the database. Given that only 127 redescriptions were originally associated with this cluster, the importance of the newly discovered redescription is substantial. Redescription R_9 has entity Jaccard of 0.54 with the third cluster (third in the second row). This redescription describes all bacteria from the target cluster except Geobacter uraniireducens. Additionally, it describes 5 bacterial species from the genus Desulfovibrio contained in the neighbouring cluster (third row, fourth cluster). Given that the database contains only 41 redescriptions that describe at least one species of the bacteria contained in the third cluster and its homogeneity is 0.41, the discovered redescription provides a very useful additional knowledge. Addition of the presented redescription increases homogeneity of the third cluster to 0.413.

We performed constrained-based redescription mining with soft attribute constraints $shCoc$, C_{3373} and present two obtained redescriptions R_{10} and R_{11} (fourth and fifth redescription in Table II). Redescription R_{10} contains positive information about bacteria that are Coccus shaped whereas the second redescription contains information about a set of bacteria that does not contain COG_{3373}. The described set of bacteria is connected to the terms buruli ulcer and chancroid.

C. Discovery on the DBLP dataset

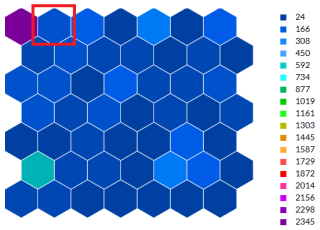


Fig. 5: Studied clusters on the DBLP dataset.

On the DBLP dataset, we focus on the second cluster in the first row. This cluster contains 109 different authors of scientific manuscripts, where 144 redescriptions describe at least one author from the cluster. Redescriptions associated with this cluster are very general, mostly containing negations of attributes (saying what is not true for many authors contained in their support sets). As a result, the homogeneity of this cluster is very small 0.0213. Using authors contained in this cluster as soft constraints, we discovered 17 new redescriptions with entity Jaccard $\geq$ 0.03 with the selected cluster. The important aspect of the discovery is that newly obtained redescriptions contain some positive attributes, revealing conferences in which a part of the authors contained in the cluster published their manuscripts and their co-authors. After adding newly discovered redescriptions to the database, the homogeneity of this cluster increased to 0.0243. Redescription R_{12} presented in Table III has the largest entity Jaccard (0.0365) with the selected cluster. It describes a group of authors that co-authored their manuscripts with either Ravishankar K. Iyer, Amin Vahdat, Robbert van Renesse, Robert J. Stroud, E. N. Elnozahy, Michele Garetto, Michel Raynal or Maria L. Gini. The described authors published manuscripts in the ICDS (International Conference on Digital Society) and the Symposium on Reliable Distributed Systems. R_{12} describes 5 authors from the selected cluster, revealing information about their scientific activity and 15 authors from the neighbouring cluster (second in the second row).

V. Conclusions

We presented the extension of the InterSet tool that allows performing interactive and constrained-based redescription mining from the user interface of the InterSet tool. This important extension transforms InterSet

TABLE III: Redescription examples obtained on the DBLP dataset.

q_1	q_2	J	$\|supp\|$
$ICDS \wedge SRDS$	$R.K.Iyer \vee A.Vahdat \vee R.vanRenesse \vee R.J.Stroud \vee E.N.Elnozahy \vee M.Garetto \vee M.Raynal \vee M.L.Gini$	0.17	33

into targeted and contextual redescription mining and redescription set exploration tool.

Scientific merit of the approach is demonstrated on three different use-case datasets. The authors obtained novel redescriptions, not previously contained in the database, with significantly higher entity Jaccard index with the targeted SOM cluster than the homogeneity score of this cluster on all three datasets and demonstrated meaningfulness of the created redescriptions for the target cluster and the neighbouring clusters. Analyses were performed both for clusters with high and with low homogeneity score. As was expected, clusters with low homogeneity score contain heterogeneous entities, thus newly created redescriptions, obtained using constraint-based redescription mining, describe only subsets of the targeted clusters, potentially including neighbouring related entities. The authors also demonstrate functionality of constrained-based redescription mining with attribute-based constraints, where novel previously unknown associations were discovered for the targeted attributes both using hard and soft constraints. It should be noted that obtaining novel results is not always possible (e.g. with to strict constraints). In such cases, it is necessary to relax the constraints and re-run constraint-based redescription mining.

The obtained tool is one of only two interface-based interactive redescription mining tools and the only tool capable of incorporating information about redescription sets into constrained-based redescription mining and targeted and contextual redescription set exploration.

REFERENCES

[1] N. Ramakrishnan, D. Kumar, B. Mishra, M. Potts, and R. F. Helm, "Turning CARTwheels: an alternating algorithm for mining redescriptions," in *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'04), Seattle, Washington, USA*. ACM, 2004, pp. 266–275.

[2] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comp. Surv.*, vol. 31, no. 3, pp. 264–323, 1999.

[3] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *An. Dat. Sci.*, vol. 2, no. 2, pp. 165–193, 2015.

[4] R. S. Michalski, "Knowledge acquisition through conceptual clustering: A theoretical framework and an algorithm for partitioning data into conjunctive concepts," *J. Pol. Anal. Inf. Syst.*, vol. 4, no. 3, pp. 219–244, September 1980.

[5] D. H. Fisher, "Knowledge acquisition via incremental conceptual clustering," *Mach. Learn.*, vol. 2, no. 2, pp. 139–172, 1987.

[6] J. Fürnkranz, D. Gamberger, and N. Lavrac, *Foundations of Rule Learning*, ser. Cognitive Technologies. Springer, 2012.

[7] S. Wrobel, "An algorithm for multi-relational discovery of subgroups," in *Procedings of first European Conference on Principles of Data Mining and Knowledge Discovery (PKDD'97), Trondheim, Norway*, ser. Lecture Notes in Computer Science, vol. 1263. Springer, 1997, pp. 78–87.

[8] N. Lavrač, B. Kavšek, P. A. Flach, and L. Todorovski, "Subgroup discovery with CN2-SD," *J. Mach. Learn. Res.*, vol. 5, pp. 153–188, 2004.

[9] F. Herrera, C. J. Carmona, P. González, and M. J. del Jesus, "An overview on subgroup discovery: foundations and applications," *Know. Inf. Syst.*, vol. 29, no. 3, pp. 495–525, 2011.

[10] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," in *Adv. Know. Disc. Dat. Min.* AAAI/MIT Press, 1996, pp. 307–328.

[11] J. Hipp, U. Güntzer, and G. Nakhaeizadeh, "Algorithms for association rule mining - A general survey and comparison," *SIGKDD Expl.*, vol. 2, no. 1, pp. 58–64, 2000.

[12] M. Zhang and C. He, *Survey on Association Rules Mining Algorithms*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 111–118.

[13] A. Gallo, P. Miettinen, and H. Mannila, "Finding subgroups having several descriptions: Algorithms for redescription mining," in *Proceedings of the SIAM International Conference on Data Mining, (SDM'08), Atlanta, Georgia, USA*. SIAM, 2008, pp. 334–345.

[14] E. Galbrun and P. Miettinen, "From black and white to full color: extending redescription mining outside the boolean world," *Stat. Anal. Dat. Min.*, vol. 5, no. 4, pp. 284–303, 2012.

[15] T. Zinchenko, E. Galbrun, and P. Miettinen, "Mining predictive redescriptions with trees," in *IEEE International Conference on Data Mining Workshop, (ICDMW'15), Atlantic City, NJ, USA*. IEEE Computer Society, 2015, pp. 1672–1675.

[16] M. Mihelčić, S. Džeroski, N. Lavrač, and T. Šmuc, "A framework for redescription set construction," *Expert Syst. Appl.*, vol. 68, pp. 196–215, 2017.

[17] ——, "Redescription mining augmented with random forest of multi-target predictive clustering trees," *J. Intell. Inf. Syst.*, vol. 50, no. 1, pp. 63–96, 2018.

[18] E. Galbrun and P. Miettinen, "Mining redescriptions with siren," *ACM Trans. Knowl. Discov. Data*, vol. 12, no. 1, jan 2018.

[19] M. Mihelčić and T. Šmuc, "Targeted and contextual redescription set exploration," *Mach. Learn.*, vol. 107, no. 11, pp. 1809–1846, 2018.

[20] P. Jaccard, "The distribution of the flora of the alpine zone," in *New Phytologist*, vol. 11, 1912, pp. 37–50.

[21] M. Mihelčić, G. Šimić, M. Babić Leko, N. Lavrač, S. Džeroski, T. Šmuc, and for the Alzheimer's Disease Neuroimaging Initiative, "Using redescription mining to relate clinical and biological characteristics of cognitively impaired and alzheimer's disease patients," *PLOS ONE*, vol. 12, no. 10, pp. 1–35, 10 2017.

[22] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, Jan. 1982.

Fake News Detection: A Comprehensive Survey

Elton Tata*, Jaumin Ajdari * and Nuhi Besimi*
Faculty of Contemporary Sciences and Technologies, South East European University, Tetovo, North Macedonia
et29825@seeu.edu.mk, j.ajdari@seeu.edu.mk, n.besimi@seeu.edu.mk

***Abstract* - The growth and increasing use of digital information platforms have dramatically changed the way news is produced, disseminated, and consumed in our society. Fake news can be found everywhere through popular platforms like social media and the internet. Efforts to develop an effective system for identifying fake news are numerous. Artificial intelligent tools are included to address this difficult issue. Fake news appears in different forms based on the features of their content. The aim of this research is to provide a comprehensive understanding of the various techniques within the domain of fake news detection through a systematic literature review of the existing work. This literature will demonstrate the most significant and relevant models to provide orientations in future research.**

Keywords - fake news, machine learning, deep learning

I. INTRODUCTION

Nowadays all over the world, mainly in developing countries, the main source of information where people get the news have become social media platforms on the Internet (Facebook, YouTube, Twitter, etc.). Thus, fake news has become a problem in society, being used to spread false information in order to change people's attitudes. Fake news refers to unrealistic information disguised as authentic news.

Detecting fake news is a difficult and challenging task, as it is associated with a number of characteristics of the news, such as its authenticity, author's intention, content and form.

Mostly fake news is in the categories: satirical news, fraud, completely fabricated news, and government (political) propaganda [1]. But when we talk about the categorization or types of fake news, we encountered different opinions about identifying their types. But looking at their content evaluation on the Internet, there are different types of fake or misleading news that the reader is faced with. A more detailed fake news categorization is as follows:

- Rumor: A rumor is a story or statement that is spread without being verified or confirmed by facts. It can also include innuendos or unverified information.
- Clickbait: these are contents whose main purpose is to attract the attention of visitors to increase their number on websites. To attract attention, they use sensational titles.
- Satire/Parody: fake news published on social media for entertainment, not to cause harm, but have the potential to deceive.
- Propaganda: news that deliberately deceives the audience to give a one-sided view of a certain cause or political agenda.
- Conspiracy Theory: news sources that promote conspiracy theories.
- Manipulated content: real content or images manipulated to deceive.
- Fabricated content: completely fake news created with the purpose of deception and causing harm.
- Hate News: content that actively promotes racism, homophobia, and other forms of discrimination.
- Misinformation: false content spread without malicious intent.
- Dis-information: false content created and distributed with malicious intent. Research in this area is oriented towards the development of automated techniques to discover the credibility of news on social media. Considering the different characteristics of news, according to the researchers of this field, the methods for evaluating fake news are divided in two directions, where the first is content-based, and the second one is social-based. Content-based methods discussed in these papers [2, 3] evaluate fake news by analyzing the textual or visual content of the news, or external data about the news subjects. The authors use these methods to detect fake news at an early stage, but these methods lack mechanisms for modeling important auxiliary data such as the news propagation model presented by Nguyen et al [4] which may limit their performance.

Methods based on social contexts are divided into two categories: stance-based and propagation-based methods. These methods integrate additional data as comments of social network users about certain news given by papers [4, 5], user profiles by Shu et al [6], and user behavior features given by Rajabi et al [7] such as "posting " and "re-posting", and user attitudes given by Van-Hoang et al [8]. Recent studies show that incorporating the structure of news propagation in social networks into the model plays a crucial role in improving the performance of fake news evaluation.

Although methods based on social contexts have shown efficacy in improving accuracy, the best way to improve explainability is unexplored. These methods do not have detailed information about topological connections. This paper presents a systematic review of existing literature, with the aim of understanding different techniques and models based on the domain of fake news, to identify challenges for future research. The rest of the paper is structured as follows: Section II focuses on the

Methodology. Section III presents the result and analysis. And in section IV are presented some conclusions.

II. Research Methodology and Questions

This study conducts a comprehensive search of published research papers investigating the field of automated ML and DL from a computer science perspective on fake news evaluation and the largest databases, such are Google Scholar, IEEE, Springer, ACM, Elsevier, etc., using the strategy based on keywords are used. The target keywords used during the search process are: "fake.", "automated.", "automatic.", "deep learning.", "DL.", "machine learning.", "ML." and "news.". The analyzed articles were selected in such a way that they meet the following condition:

• The analyzed study has a significant number of citations (10 < citations)

• The research paper is written in English.

In order to focus on our objective, we also proposed few research questions and based on these research questions, all the selected papers were analyzed and as well all the useful insights were extracted. The proposed research questions are as follows:

RQ I: What features influence the evaluation of a news item?

RQ II: Which models can be used in the field of ML and DL in the evaluation of news reliability?

III. Result and Analysis

Evaluation of news reliability could be based on automated solutions provided by artificial intelligence tools with Machine Learning (ML) and Deep Learning (DL) techniques. So having this idea, the following literature review will be based on these tools for reliable news assessment.

A. ML Techniques

Machine learning has found wide use in many areas of everyday life and in solving various problems. Bali et al [9] have applied machine learning algorithms to classify news and to assess their reliability. A stratified 10-fold cross-validation based on linguistic features such as ease of readability and lexical diversity was used to evaluate the model. The characteristics used for the given model are: n-grams Count Feature, TF-IDF which shows how many times a word appears in a certain document, Word Embedding, Sentiment Polarity Score, and linguistic features. Three datasets from Kaggle and GitHub were used to test the model. The used algorithms are Random Forest (RF), Support Vector Classifier (SVC), Gaussian Naive Bayes (GNB), AdaBoost (AB), K-Nearest Neighbor (KNN), Multi-Layer Perceptron (MLP) and Gradient Boosting (XGB). The comparison showed that the lowest average accuracy recorded is 62.5%, 72.54%, and 64.8% when using KNN whereas Gradient Boost (XGB) achieved 87.2%, 92.0%, and 87.3% average accuracy for these used datasets. Pedro et al., in their paper [10] present a model for evaluating news using only textual features which can be generated independently of the news publishing platform regardless of the language used. Authors show that lexical size and sentiment are helpful in all datasets. They used five datasets to obtain test results using four different Natural Language Processing (NLP) techniques. From the conducted experiments, it was shown that the best result was achieved with Random Forest and SVM algorithms, combined with Bag-of-Words (BoW) as a text representation technique with an accuracy of 95% and 94% respectively, for news in the analyzed language groups. Another study in this field, by Ahmed et al [11], presents news evaluation techniques based on the extraction of different text features. Two datasets were used to test them, and it was shown that Passive Aggressive (PA) gives a higher accuracy of 93% while Naive Bayes is 85% and 84% for SVM. While Ott et al., in the paper [12] applying the SVM with LIWC + Bigrams have achieved an accuracy level of up to 89.8%. In the publication [13], the authors have designed a model for automatically identifying fake news based on machine learning. The authors trained several machine learning models suitable for binary classification using textual and graphic features. From the obtained experimental results, the study showed that the highest accuracy of 80% was achieved by SVM. M. E. Tacchini et al., in their paper [14], have used two models based on logistic regression and on Boolean crowdsourcing algorithms. Authors classify the news into hoax or non-hoax based on their features for both classifiers, based on what users liked. The datasets consist of data obtained from the posts of various users on the social platform Facebook. From the evaluation of the models, it results that for the complete dataset, the BLC harmonic algorithm shows better performance and with accuracy above 90%. Another paper, presented by W. Han and V. Mehta [15], presents an assessment and comparison of the different approaches used to classify fake news. They present some traditional machine learning models such as Naive Bayes and advanced deep learning models such as the hybrid CNN and RNN. Evaluation of their performance was done using TF-IDF bi-gram, PCFG, and a combination of both. From the obtained results, authors have shown that the Naive Bayes classifier performs well in fake news detection, but the hybrid model CNN and RNN using TF-IDF results in better performance.

B. DL Techniques

N Ahlem et al., in their paper [16], present a model where CNN and LSTM are combined, which is a type of Recurrent Neural Network as shown in figure 1.

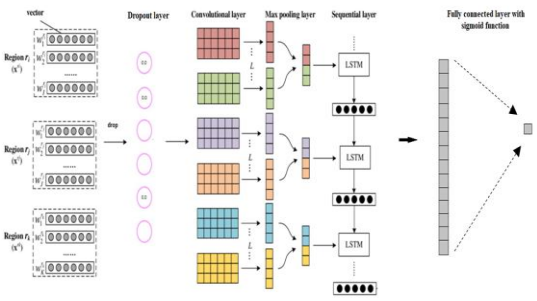


Figure 1. The architecture of the proposed fake news detection model Ahlem et al [16]

The first layer represents each statement (text) as a series of vectors. Then drop-out layers are added to reduce overfitting and to determine the probabilities of the vectored data during the training phase. Then comes the CNN layer that serves to extract features from local inputs. Some of the features used are total words, characters per word, frequency of large words, unique words, number of followers, number of tweets, etc. The LSTM layer measures the long-term dependencies of feature sequences. The outputs from this layer are merged and passed to a fully connected layer which converts the array into a single output in the range {0,1} using the sigmoid function. The overall performance of the CNN-LSTM model is much better than the baselines in terms of accuracy. What is evident is that this model is more efficient to apply to a large dataset to improve the detection of fake news compared to small datasets. Gallego et al., in the paper [17], present some techniques based on ML and DL for evaluating news in the Spanish language. For comparison purposes, in the case of ML were used the techniques of Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Tree (GBT), and Multi-Layer Perceptron (MLP) while for the DL were used LSTM-RNN using a many-to-one architecture and CNN. For testing, two datasets, in Spanish and English, that correspond to fake and real data for both languages were used. Accuracy was used as the meter parameter and it resulted that the highest accuracy, up to 80%, for DL was achieved using the LSTM model with 25 epochs. Where in the group of ML techniques, RF has shown the best performance getting an accuracy of 80%, using TF-IDF, and a rough accuracy has shown the SVM. What was noticed was that the application of these techniques to the automatically translated dataset had a significant decrease in accuracy. There were also no significant differences in model performance when applying Stemming or removing stop-words.

The authors, in the paper [18], present a hybrid model for evaluating different news. The proposed model called CSI consists of three modules: Capture, Score, and Integrate. The model uses three characteristics of news such as: the content of an article, the response of the users who receive and the source user who distributes it. The first module analyzes the response and the text, while it uses a Neural Recurrent Network to analyze the user activity. The second module learns the characteristic of the source based on user behaviors and then integrates with the third module for news classification. For the comparison and testing of the model are used two datasets that were obtained from the social networks Twitter and Weibo. From the obtained results the proposed model performs better than the others respectively with an accuracy of 89.2% and an F-score of 89.4% for Twitter, and accuracy of 95.3%, and an F-score of 95.4% for Weibo. As the text representation technique used, the paper shows the benefit of using doc2vec over simple TF-IDF.

Another study in this field is presented in the research work [19], where the authors present an ensemble-based deep learning model for news evaluation. LIAR datasets were used to test and evaluate the model. For the statement attributes, NLP techniques like tokenization, lemmatization, and stop word removal are applied. The first model is built with 9 neural network layers, while the second model, was designed with 10 fully connected dense layers and with 9 feature variables as input data. The obtained results of this research, show an accuracy of 89.8% using the statement feature. Another work in this area that uses neural networks is given in research [20], where the proposed model accurately predicts the attitude between headlines and the body of a given article. The FNC-I datasets are used to test and evaluate the model which includes the body and the headline of the news article, and the label for relatedness (stance) of an article and headline. The preprocessing techniques used are Stop Word Removal, Punctuation Removal, and Stemming. Bag of Words and TF-IDF techniques are applied to transform the raw text and extract the features. The presented model estimates the relative importance of a word present in article-title pairs locally (importance of that word for that specific headline-article pair) and globally (how often that specific word is found in relation to all words in the corpus). For the similarity measure between the headline and article pair the Cosine similarity between Headline- Article TF-IDF pairs are calculated. Authors have tested three different architectures: TF-IDF Vectors with Dense Neural Network, Bag of Words Vector with Dense Neural Network, and Pre-trained word embedding with Neural Networks. The architecture with the use of a fine-tuned TF-IDF - Dense neural network (DNN) has shown the best result with 94.21% accuracy. K. Kaliyar et al., in research [21], present a Deep Hybrid Neural Network (DeepNet) model for evaluating fake news using the BuzzFeed and Fakeddit datasets. Initially, they present a model based on the Artificial Neural Network (ANN) with four dense hidden layers and activation function Leaky ReLU and Adam optimizer. The implemented Deep Hybrid Neural Network architecture, whereas the first layer has an embedding layer that accepts the input as a vector and continues with three convolutional layers where the vector is transformed and passed to the LSTM layer responsible to handle the nature of sequential data. After the LSTM layer are implemented seven dense layers. The proposed model uses a combination of content and context features. From the obtained results, DeepNet outperformed existing fake news detection methods with an accuracy of 95.2% for BuzzFeed and 86.4% for Fakeddit.

In the research [22], Singhal et al., present a model called SpotFake - a multi-modal framework for the classification of fake news. This model evaluates fake news based on the textual and visual features of an article. It uses the BERT to extract text features and VGG-19 to extract image features pre-trained on the ImageNet dataset. The proposed model consists of three sub-modules. The first sub-module is the extraction of textual features composed of 12 encoding layers, where the vectors of these features are marked as Tf. The second sub-module serves to extract visual features from VGG-

19 pre-trained and implemented through Vf vectors. The third sub-module serves to join two feature vectors obtained through different modes Tf and Vf fused together using a simple concatenation technique to form the news. The authors have compared the performance of SpotFake with the models presented in EANN [3] and MVAE [23] tested with the same datasets and the results are shown in the following table (Table1).

TABLE I. PERFORMANCE COMPARISON OF EANN, MVAE WITH SPOTFAKE

	Accuracy %	
Model	***Twitter***	***Weibo***
EANN	71.5	82.7
MVAE	74.5	82.4
SpotFake	77.77	89.23

The model SpotFake proposed by Singhal et al [22] performs better compared with both above mentioned models on Twitter and Weibo datasets.

Farokhian, M. et al., in the research [24], present a model where parallel BERT deep neural networks are used. The proposed model is named MWPBert and uses two parallel BERT networks, where the first one is used to encode news headlines, and the second on to encode news bodies. As an algorithm for fact-checking, they used the MaxWorth algorithm, by which it is possible to get the most valuable part of the news text. The outputs from the two BERT networks are encoded into a single output network to classify the news. To evaluate the model, they used Fakenewsnet datasets. To show the effectiveness of using two BERT networks, the result is compared with the result of MWBERT model, which uses only one BERT. From the obtained results, it appears that MWBERT shows an accuracy of approximately 84.7% while MWPBert is 85.4%. Thus, the MWPBert model, with parallelization, shows that the use of two parallel BERTs was beneficial.

C. Answer of the research questions

By analyzing the above research, we came up with the following answers to the above-mentioned questions.

RQ I: What features influence the evaluation of a news item?

Considering the different characteristics of news, the methods for evaluating fake news can be divided into two main categories, content-based, and social-based features.

Content-based methods evaluate fake news by analyzing the textual or visual content of the news, or external data about the news subjects. The set of features used captures the structure of the article, the similarity between the headline and the body of the news, and readability. Features from these data are grouped into stylometric, semantic, and syntactic. Some stylometric features that are used in techniques for detecting fake news are the average number of words, characters, punctuation marks, and the number of words, sentences, and unique words used. Semantic features are anger, anxiety, feelings, polarity, and informality. Syntactic features can be: word impact, certainty, inconsistencies in the text, self-references, questions, disfluencies, numbers, and tentative language use.

Social-based methods are categorized into two subtypes: stance-based and propagation-based methods. These methods, for detecting fake news, integrate additional data used in social media, such as extracted features about social network users' comments on a given news and user profiles. Other features used are the number of words in a user's self-description, posts, likes, shares, replies, followers–followees, and the time difference between the source tweet's post time and my retweet.

RQ II: Which models can be used in the field of ML and DL in the evaluation of news reliability?

There are various methods and techniques based on deep learning and machine learning, used to design, and implement different models for fake news detection, where we mention, and

ML techniques: Logistic Regression (LR) [13, 14], Gaussian Naive Bayes (GNB) [11, 12, 15], Gradient Boosting [2,9, 13, 17], Random Forest (RF) [9, 13, 17], Passive Aggressive (PA) [11], K-Nearest Neighbor (KNN) [9], and Support Vector Machine (SVM) [9, 11, 12, 13, 17], and

DL techniques: Convolutional Neural Network (C.N.N.) [3, 15, 16, 17], Recurrent Neural Network (R.N.N.) [15, 17, 18], Long Short-Term Memory (L.S.T.M.) [16, 17, 21] and Transformer (BERT, ALBERT) [22, 23, 24].

From the analyzed research papers, we collected some data, regarding the techniques used, the features analyzed and the relevant datasets to evaluate the performance of the algorithms used. This is shown in the below table (Table 2).

As can be seen from the table, different techniques give different performances depending on the features analyzed and on the datasets used for model evaluation.

TABLE II. PERFORMANCE COMPARISON OF SOME WORKS BASED ON ML AND DL TECHNIQUES

Reference	***Dataset***	***Classifier***	***Represen-tation technique***	***Features***	***Accuracy Metrics***
Pedro et al [10]	*5 datasets*	*RF*	*Bag-of-Words (BoW)*	*text features*	*95%*
Ahmed et al [11]	*Kagle datasets*	*PA/SVM*	*tf-idf*	*text features*	*93% / 84%*
Ott et al [12]	*spam detected*	*SVM*	*LIWC+ Bigrams*	*text features*	*89.8%*
Janze et al [13]	*BuzzFeed*	*SVM*	*binary classify-cation*	*textual and graphic*	*80%*
Gallego et al [17]	*four open datasets*	*LSTM-RNN/RF*	*tf-idf*	*Text Norma-lization*	*80%*
Ruchansky et al [18]	*Twitter/ Weibo*	*CSI*	*doc2vec*	*textual features*	*89,2% / 95,3%*
Aslam et al [19]	*LIAR*	*Bi-LSTM-GRU*	*Adam" optimizer*	*statement feature*	*89.8%*
Thota et al [20]	*FNC-I*	*Dense neural network (DNN)*	*Bag of Words and TF-IDF*	*text features*	*94.21%*
Kaliyar et al [21]	*BuzzFeed and Fakeddit*	*CNN-LSTM*	*Adam optimizer*	*content and context*	*95.2% / 86.4%*
Farokhian et al [24]	*Fake-newsnet*	*parallel BERT*	*MaxWorth*	*text-based features*	*0.855 %*

IV. CONCLUSION

This study has tried to collect and analyse different ML and DL methods used in model design for fake news detection. The research found that fake news can be spread in different forms and ways. Some of the categories encountered may be in the form of rumor, clickbait, satire/parody, propaganda, manipulated content, or news that is purposely fabricated for various benefits such as economic, political, etc. Based on the analyzed features, the news is divided into content-based and social-based categories.

Content-based methods evaluate fake news by analyzing the textual or visual content of the news. Social-based methods are categorized into two subtypes: stance-based and propagation-based methods. Based on the algorithms used for ML mention Logistic Regression (LR), Gaussian Naive Bayes (GNB), Random Forest (RF), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), while for DL techniques are used Convolutional Neural Network (C.N.N.), Recurrent Neural Network (R.N.N.), Long Short-Term Memory (L.S.T.M.) and recently Transformer (BERT, ALBERT) are also being used, leading to an increase in the accuracy of the models implemented for detecting fake news. The implementation of deep learning models overcame the limitation of machine learning models since the implementations by use of ML technique perform weakly when there is a large dataset. The techniques for identifying fake news are always in development since the methods for spreading fake news have a dynamic spread.

REFERENCES

[1] Shubham Bauskar, Vijay Badole, Prajal Jain, Meenu Chawla, " Natural Language Processing based Hybrid Model for Detecting Fake News Using Content-Based Features and Social Features", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.11, No.4, pp. 1-10, 2019. DOI: 10.5815/ijieeb.2019.04.01

[2] Saeid Sheikhi, an effective fake news detection method using WOA-xgbTree algorithm and content-based features, Applied Soft Computing, Volume 109,2021,107559, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107559.

[3] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. 2018. Eann: Event adversarial neural networks for multi-modal fake news detection. In KDD.

[4] Dat Quoc Nguyen, Thanh Vu, and Anh Tuan Nguyen. 2020. BERTweet: A pre-trained language model for English Tweets. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pages 9–14, Online. Association for Computational Linguistics.

[5] Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. 2019. DEFEND: Explainable Fake News Detection. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). Association for Computing Machinery, New York, NY, USA, 395–405. https://doi.org/10.1145/3292500.3330935

[6] Kai Shu, Xinyi Zhou, Suhang Wang, Reza Zafarani, and Huan Liu. 2019. The role of user profiles for fake news detection. In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '19). Association for Computing Machinery, New York, NY, USA, 436–439. https://doi.org/10.1145/3341161.3342927

[7] Rajabi, Z., Shehu, A., Purohit, H. (2019). User Behavior Modelling for Fake Information Mitigation on Social Web. In: Thomson, R., Bisgin, H., Dancy, C., Hyder, A. (eds) Social, Cultural, and Behavioral Modeling. SBP-BRiMS 2019. Lecture Notes in Computer Science (), vol 11549. Springer, Cham. https://doi.org/10.1007/978-3-030-21741-9_24

[8] Van-Hoang Nguyen, Kazunari Sugiyama, Preslav Nakov, and Min-Yen Kan. 2020. FANG: Leveraging Social Context for Fake News Detection Using Graph Representation. In The 29th ACM International Conference on Information and Knowledge Management (CIKM '20), October 19–23, 2020, Virtual Event, Ireland. ACM, New York, NY, USA, 10 pages. https: //doi.org/10.1145/3340531.3412046

[9] Bali, A.P.S., Fernandes, M., Choubey, S., Goel, M. (2019). Comparative Performance of Machine Learning Algorithms for Fake News Detection. In: Singh, M., Gupta, P., Tyagi, V., Flusser, J., Ören, T., Kashyap, R. (eds) Advances in Computing and Data Sciences. ICACDS 2019. Communications in Computer and Information Science, vol 1046. Springer, Singapore. https://doi.org/10.1007/978-981-13-9942-8_40

[10] Pedro Henrique Arruda Faustini, Thiago Ferreira Covões, Fake news detection in multiple platforms and languages, Expert Systems with Applications, Volume 158, 2020, 113503, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2020.113503.

[11] Ahmed, Sajjad and Hinkelmann, Knut and Corradini, Flavio, Development of Fake News Model using Machine Learning through Natural Language Processing, Computation and Language (cs.CL), Artificial Intelligence (cs.AI), FOS: Computer and information sciences, arXiv, 2022, doi: 10.48550/ARXIV.2201.07489

[12] Ott, Myle and Choi, Yejin and Cardie, Claire and Hancock, Jeffrey T., Finding Deceptive Opinion Spam by Any Stretch of the Imagination Computation and Language, Computer and information sciences, arXiv, 2011, Proceedings of ACL 2011: HLT, pp. 309-319, doi:10.48550/ARXIV.1107.4557

[13] Janze, Christian and Risius, Marten, "Automatic Detection of Fake News on Social Media Platforms" (2017). PACIS 2017 Proceedings. 261. https://aisel.aisnet.org/pacis2017/261

[14] E. Tacchini, G. Ballarin, M. Della Vedova, S. Moret, and L. de Alfaro. Some like it hoax: Automated fake news detection in social networks. doi 10.48550/ARXIV.1704.07506, 2017.

[15] W. Han and V. Mehta, "Fake News Detection in Social Networks Using Machine Learning and Deep Learning: Performance Evaluation," 2019 IEEE International Conference on Industrial Internet (ICII), 2019, pp. 375-380, doi: 10.1109/ICII.2019.00070.

[16] Ahlem Drif, Ahlem & Ferhat Hamida, Zineb & Giordano, Silvia. (2019). Fake News Detection Method Based on Text-Features. The Ninth International Conference on Advances in Information Mining and Management

[17] artínez-Gallego, Kevin and Álvarez-Ortiz, Andrés M. and Arias-Londoño, Julián D. Fake News Detection in Spanish Using Deep Learning Techniques, arXiv, 2021, doi 10.48550/ARXIV.2110.06461

[18] Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. CSI: A Hybrid Deep Model for Fake News Detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17). Association for Computing Machinery, New York, NY, USA, 797–806. DOI:https://doi.org/10.1145/3132847.3132877

[19] Nida Aslam, Irfan Ullah Khan, Farah Salem Alotaibi, Lama Abdulaziz Aldaej, Asma Khaled Aldubaikil, "Fake Detect: A Deep Learning Ensemble Model for Fake News Detection", Complexity, vol. 2021, Article ID 5557784, 8 pages, 2021. https://doi.org/10.1155/2021/5557784

[20] Thota, Aswini; Tilak, Priyanka; Ahluwalia, Simrat; and Lohia, Nibrat (2018) "Fake News Detection: A Deep Learning Approach," SMU Data Science Review: Vol. 1: No. 3, Article 10. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss3/10

[21] R. K. Kaliyar, P. Kumar, M. Kumar, M. Narkhede, S. Namboodiri and S. Mishra, "DeepNet: An Efficient Neural Network for Fake News Detection using News-User Engagements," 2020 5th International Conference on Computing, Communication and Security (ICCCS), 2020, pp. 1-6, doi: 10.1109/ICCCS49678.2020.9277353.

[22] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru and S. Satoh, "SpotFake: A Multi-modal Framework for Fake News Detection," 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 2019, pp. 39-47, doi: 10.1109/BigMM.2019.00-44.

[23] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, “Mvae: Multimodal variational autoencoder for fake news detection,” in The World Wide Web Conference, ser. WWW ’19. New York, NY, USA: ACM, 2019, pp. 2915–2921. [Online]. Available: http://doi.acm.org/10.1145/3308558.3313552

[24] Farokhian, M., Rafe, V. and Veisi, H., 2022. Fake news detection using parallel BERT deep neural networks. arXiv preprint arXiv:2204.04793.

Graphical user interface to perform glacier simulations with PISM

M. Urbanč*, M. Depolli*
* Jožef Stefan Institute/E6, Ljubljana, Slovenia
urbancmarko1@gmail.com

***Abstract*—Computer simulations of glacial dynamics help us understand how glaciers respond to changes in climate, such as warming temperatures and increased precipitation. Simulations are performed with PISM (Parallel Ice Sheet Model). To tackle the complexities of ice dynamics, PISM employs several sub-models for various aspects and thus exposes a plethora of parameters. There are also several options for setting parameter values, i.e. through config files, command line options or input files. To get a good handle on all the provided buttons and knobs, we developed a user interface for setting parameters of out simulations and for executing PISM with the selected set of parameters on a remote workstation.**

In this paper we present the GUI developed in Jupyter and deployed in a local JupyterLab installation. The environment contains DEM files, Python scripts and other inputs locally but connects to a 128-core workstation for simulation execution. The simulation inputs and outputs are transferred between the computers over an SSH connection. The GUI is publicly exposed in a single-user environment, only for use, while development is performed on a git repository which is then synchronized to the public through git. This approach is general enough to be used on other software as well.

***Keywords*—PISM, glacier simulations, user interface, jupyter, distributed execution, remote management**

I. INTRODUCTION

A. Motivation

PISM (Parallel Ice Sheet Model) [1], [2] is a complex simulation tool that can be used in wide range of applications that require simulation of ice sheets or glaciers. Performing glacier dynamics simulations is far from trivial, PISM is designed to executed primarily through scripts; even the tutorial available in the official documentation presents it through scripts. A scripting approach makes in silico experiments highly reproducible, since the whole experiment can be defined by a script, and the script can be archived. However, handling multiple geographical areas of interest and multiple modelling choices eventually leads the relatively simple execution scripts to evolve into a complex ecosystem of scripts. Such an ecosystem is difficult to manage and presents a formidable barrier for new researchers that are assigned to such an experimentation task.

Within our research, we found that our typical workflow for conducting in silico experiments on glaciers can be defined as follows.

- Select the geographical area to perform the experiment on from a predefined set of options. The geographical area uniquely defines a set of variables: the DEM (Digital Elevation Map), climatological data and expected glaciation extent.
- Select other PISM modelling parameters and parameters of the custom climatological models to fully define the experiments.
- Format all the inputs appropriately for the use in PISM.
- Execute PISM, either locally or on a remote workstation. In the latter case, communicate the input and output files between the local and remote machines.
- Analyze PISM outputs to evaluate the experiment. Some of the analysis can be fully automated while some is done interactively or on demand.

To aid in executing the workflow swiftly, correctly, and with repeatable results, we avoided scripts and rather designed a graphical user interface (GUI) for the experimenter. It is written in a combination of Python, Jupyter notebooks, and helper bash scripts. The presented approach could be viewed as an extension of the scripting approach with the majority of functionality migrated from bash to Python and from command line arguments to graphical elements. Python has been chosen because of its flexibility and capability of implementing complementary models to those included in the PISM itself. The visualization is implemented in Jupyter notebooks [3], [4] for ease of use and beauty of presentation. The workflow based on these two parts of GUI is presented in Figure 1. The presented approach was very successful within the context it was created for and it seems to be general enough to be applied more broadly. Therefore we present it as a concept that could be copied by researchers facing a different complex simulation software.

In this paper we focus only on the *main* GUI part/Jupyter notebook which is responsible for parameter setup, execution with execution monitoring, and some automated result analysis.

B. Abbreviations and Acronyms

Abbreviations	Definitions
PISM	Parallel Ice Sheet Model
DEM	Digital Elevation Map
JSON	JavaScript Object Notation
MPI	Message Passing Interface

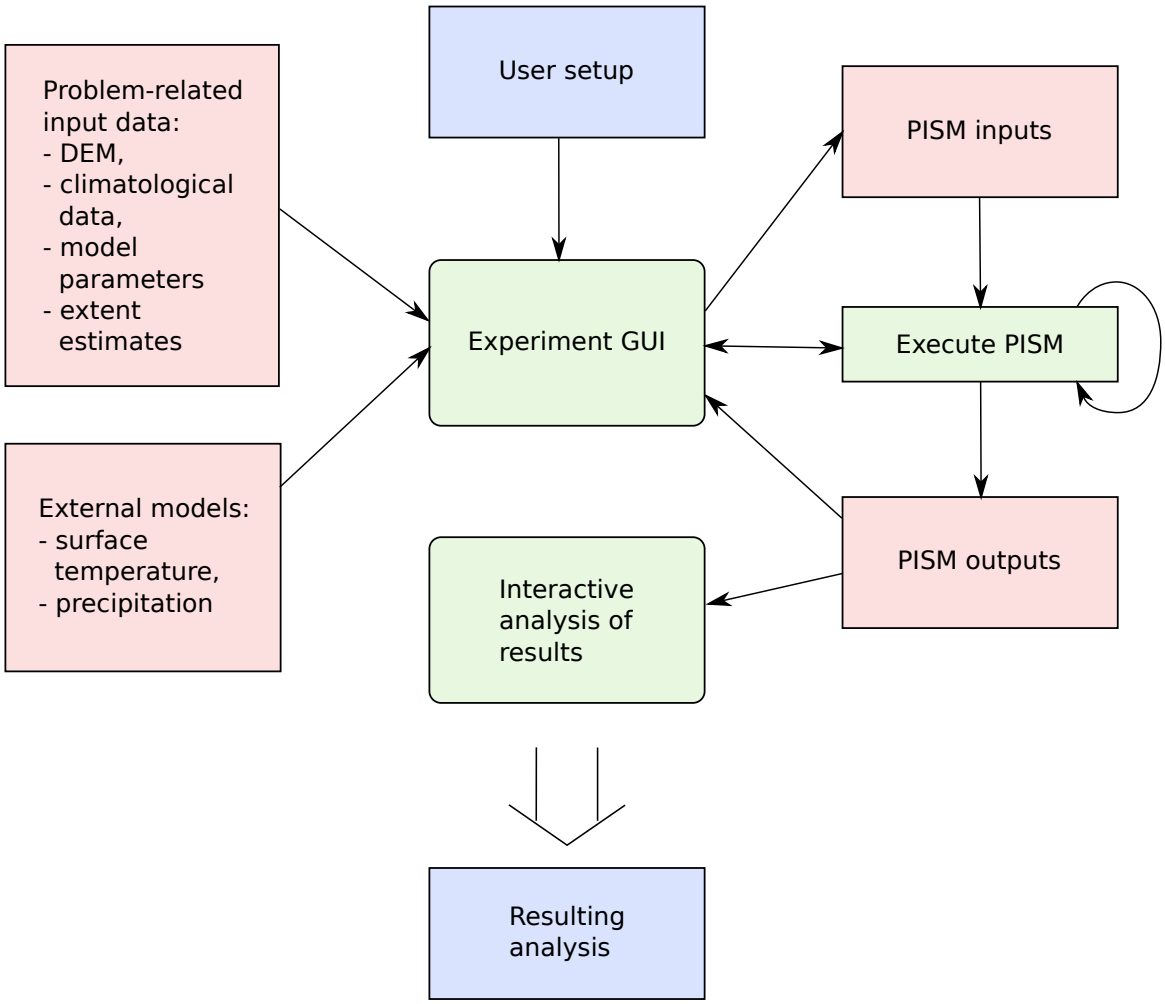


Fig. 1: The workflow of experimentation using the presented GUI. The user is aided by a selection of supported experimental parameters – the relevant subset of the PISM functionality. Several models are also implemented as Python modules and are seamlessly integrated in PISM execution.

II. METHODS

A. Hardware system

There are three different computers involved in the workflow. One is the experimenter's *personal computer*, on which they physically work and on which they interact with the GUI.

Second is the *Jupyter server* that serves the web presentation of Jupyter notebooks and executes their code locally. Note that the Python processing required for the custom (external to PISM) models are executed here. Communication with the user is done through the web server service that runs on the same computer.

Third is what we shall call the *PISM execution server*, but could in reality be a whole cluster of computers. PISM execution server is a number crunching system that only executes PISM as per the input arguments given. It communicates with the web server through the SSH protocol. Internal communication between the multiple workstations of this server is handled by MPI – a platform for distributed computing often used in scientific contexts.

B. Software system

Jupyter notebooks are documents produced by the *Jupyter Notebook App*. They contain both computer code (in our case Python) and rich text elements, organised in so called cells. Notebook documents are meant to be human-readable as well as executable. Each cell has a type - executable or rich text and can be executed or rendered on its own. A cell is executed by its *kernel* which is the "computational engine" that executes code contained within a notebook document. For example the *ipython kernel* [5] executes Python code and *markdown kernel* [6] renders rich text cells. Kernels for many other languages also exist. Cells typically contain analysis descriptions, analysis code and results (figures, tables, multimedia). The Jupyter Notebook App is a server-client application which allows editing and running notebook documents via a web browser. It can be executed on a local desktop where it is hosted locally or it can be installed on a remote server and accessed through the internet.

Several Jupyter notebooks are used to present the GUI to the user, with the central one designed to perform the

majority of tasks: setting of the parameters, selection of the host workstation for PISM execution, submission of a job for execution, monitoring of execution, performance of automated analysis and presentation of the results to the user. A further set of Jupyter notebooks is used for development and testing of additional models, and for interactive or specialist analysis of simulation results.

To keep the interface of Jupyter notebooks clean, the majority of its source code is organised into Python source files. In addition, several Python scripts are provided that are able to bypass the GUI and allow the execution of pre-prepared experiments from the console. The main use case of such scripts is the execution of parameter sweeps for parameter studies, which require simulation of multiple experiments and an automated analysis of results. For these scripts, the experiments are also designed in the aforementioned GUI, however, their setups are stored in a JSON (JavaScript Object Notation) [7], [8] files. Python scripts read the experiment setup from JSON files, but are also able to modify several parameter values on their own, e.g. a parameter sweep script generates experiments from an experiment setup, a set of parameters and a set of values for each of the parameters.

For the GUI we leverage the high level interface elements from *Jupyter Widgets* [9] – interactive browser controls for Jupyter notebooks, which can be added to Jupyter as a Python package. The framework has two components; a package in the kernel which provides an interface for the widgets and an extension for the browser Jupyter frontend to manage Jupyter Widgets. The `ipywidgets` package provides Jupyter Widgets for the *ipython kernel*. Installing `ipywidgets` automatically installs extensions for JupyterLab and Jupyter Notebook (the `jupyterlab-widgets` and `widgetsnbextension` packages). The package provides a basic and lightweight set of core form controls that use this framework. These include but are not limited to: text area, select/multiselect controls, sliders etc.

Lastly, several bash scripts are used for file synchronization and PISM execution. These are generated from templates by the Python code when PISM execution is requested, either from GUI or from scripts. When executing PISM, bash scripts redirect PISM standard outputs to temporary files and to the Python monitoring functions, providing the option of limiting PISM execution time and handling external signals to terminate execution if so requested by the user.

C. Remote PISM execution

The first panel (Execution host settings, shown in Fig 2) of the GUI is responsible for configuring the remote compute node's parameters. For ease of use and convenience the user can save to, and load the parameters from JSON files. The interface has built-in tests which check for a working SSH connection, a working PISM executable on the remote host and the existence of the specified working directory.

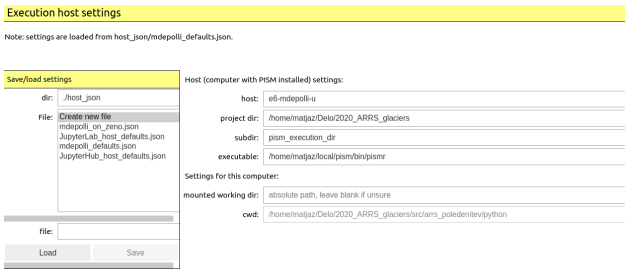


Fig. 2: Execution host settings presented on the GUI.

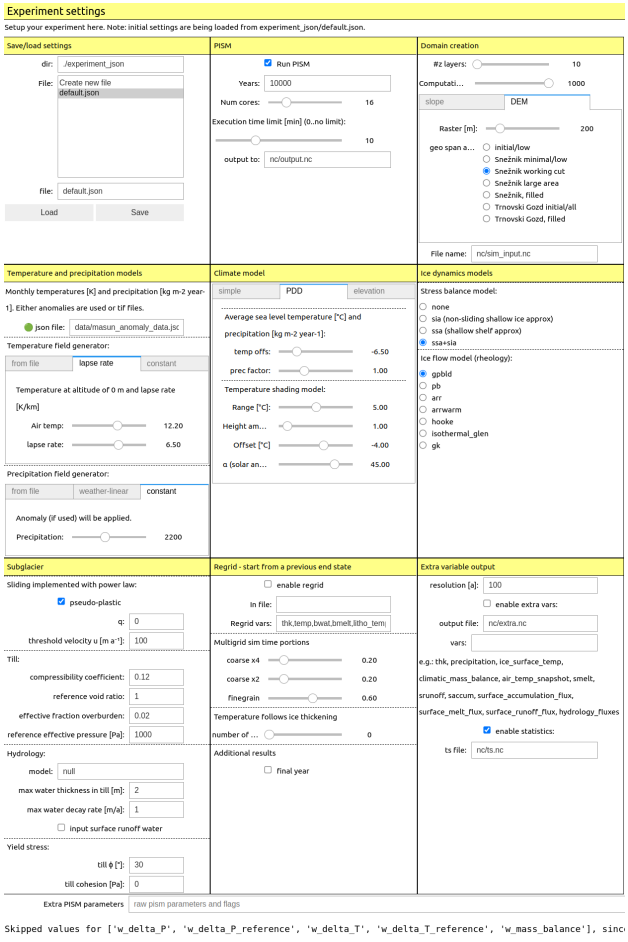


Fig. 3: Experiment settings presented on the GUI.

After passing the built-in tests the user moves onto the second panel (Experiment settings) of the GUI. Here lie the meat and potatoes of the interface. As before, the user can conveniently load or save configuration settings from JSON files. There are a plethora of parameters that can be set here, an example is shown in Fig 3.

Once the experiment parameters have been set, the interface creates the necessary input files for PISM. To verify the set parameters, some of the more important model inputs are plotted; see Fig 4. Then a cell can be executed that causes the interface to connect to the configured compute node via SSH. It first copies input files to the compute node, then iteratively copies and executes the bash scripts (as the simulation is often comprised of

several consecutive PISM executions). Finally a script is executed to copy the results from the execution server back to the experimenter's personal computer (where the Jupyter notebook is open).

After the output files are available on the experimenter's personal computer, they can be analysed. The Notebook with GUI provides some automated analysis that can be modified according to one's needs. An example of the visualization that follows the analysis is shown in Fig 5. Researchers can quickly overview the plotted results of simulations and decide whether the experimentation is progressing well and plan further experiments accordingly.

Each PISM execution requires the use of two scripts to enable all the code required for setting up PISM, including a timeout and handles for user-requested termination. The timeout is responsible for gracefully stopping the process if it were to take longer than allowed by the user. Note that the results of the incomplete simulation are nevertheless stored and can be useful to the user. After their operation, bash scripts are deleted. During the simulation the user is presented with progress indicators, as shown in Fig 6.

D. Test deployment

The presented GUI was developed iteratively. In the first iteration, the GUI was designed, developed and used only locally on a the main developer's personal computer. This was sufficient and also very efficient since only one person was working on it. In the second iteration, as of the time of writing this article, the GUI was made available for demonstration to other researchers from the areas of computer science and geography. At this stage, in a form of a demonstration of functionality, it is best served as a single user environment to avoid the overhead of user registration. Note that multiple users can nevertheless work with the notebooks at the same time, but cannot store their changes which would be a cause of race conditions. In its final stage, it will be migrated to a multi-user Jupyter environment, assisted with version control.

The test deployment consists of two computer workstations: a Jupyter node and a compute node. Ubuntu Server 22.04 was used on both machines. Installation on the compute node includes Spack [10] and PISM. Spack is a package management tool designed to support multiple versions and configurations of software on a wide variety of platforms. It was designed with large super computing centers in mind. It makes PISM's install a whole lot easier at the cost that it contains an outdated version of PISM in their repositories. This could of course change in the future. For our purposes this older version was fine.

On the Jupyter node, JupyterLab is installed within Docker. This gives us the advantages of *repeatability* and *ease of deployment* as Docker containers are easy to work with while they provide a predictable environment and a user-defined level of isolation between applications. Since the presented GUI has package dependencies we've created a custom Docker image using a Dockerfile based on the Jupyter Datascience Notebook. The required Ubuntu packages to run the GUI are: `libgdal-dev`, `gdal-bin`, `python3-dev`, `build-essential` and `rsync`.

Python dependencies are also installed via the Dockerfile so that they are permanent and persistent even if the Docker container is restarted (e.g. through `docker-compose down`). Installing additional packages is possible with `pip install ...` from inside the container but the installations will not persist. The required Python dependencies are: `ipywidgets`, `netCDF4`, `numpy`, `matplotlib`, `scipy`, `shapely`, `pyproj` and `gdal`. Some of these are already provided by the Jupyter Datascience Notebook source image but are listed here for the sake of completeness. It is crucial that `gdal` from `pip` and `apt` are of the same version.

All the required code for the GUI is stored on the Jupyter node in the form of a `git` repository. This repository is setup as read-only for the Docker container because it is not meant to be modified inside it. This test deployment is configured to be synchronized with a central code repository through `git`. The repository is mounted into the Jupyter Docker container as a Docker volume at path `/home/jovyan/work/`, which is hard-coded into the official Jupyter Datascience Notebook Docker image.

Additionaly we also mount the `ssh_config` and key files to their appropriate locations to enable Jupyter node connect to the compute node with non-interactive SSH sessions. In our environment an additional hop through a middle-man node was required due to our network configuration.

E. Development

The intention is that the GUI maintainer develops the GUI locally on his computer. The maintainer runs the same Jupyter notebook along with some helper notebooks that are development oriented. Changes are then synchronized across to the Jupyter node via `git` and the GUI is updated.

III. RESULTS

In this section, we shall list several features that we found invaluable during the development and use of the proposed GUI. Then we list the positive and negative aspects of the developed system.

A useful feature of PISM is its ability to stop the simulation at any time and force its unfinished result to be stored for inspection. We used this feature to implement a time limit for simulations, which we found to be very useful for experimenting with parameters. On a related note, the ability to monitor the ongoing simulation is important to experimenters and while we did implement it through Jupyter supported widgets, there is room for improvement over our approach. Using the PyQt framework could be one betterment of our approach as PyQt gives the GUI the flexibility to take on what ever form is needed, all the while not being limited by the selection of widgets in Jupyter. This all comes at a great cost though

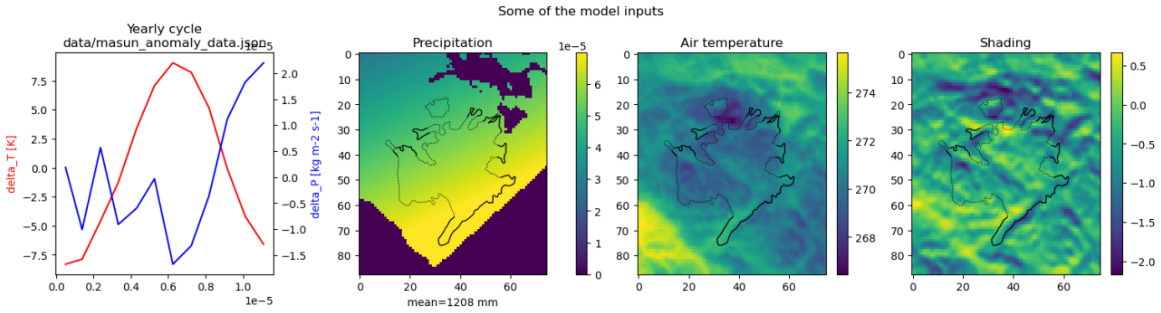


Fig. 4: An example of the overview of experiment settings, which the user can verify to match their desired settings before confirming to execute the experiment.

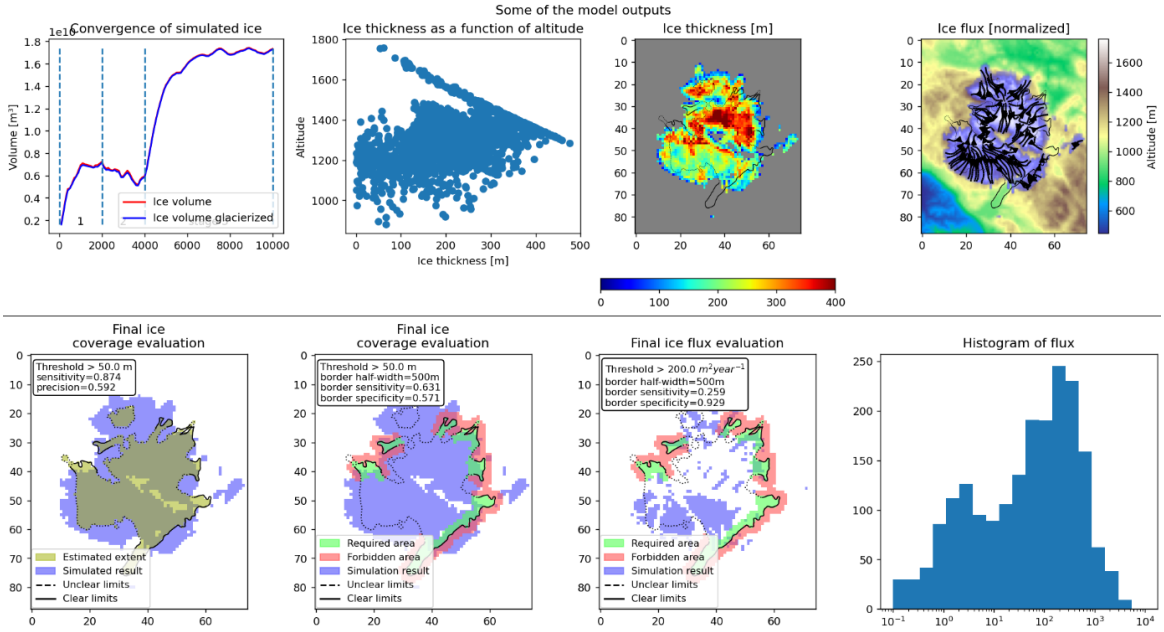


Fig. 5: An example of the visualization presented to the user as an experiment completes.

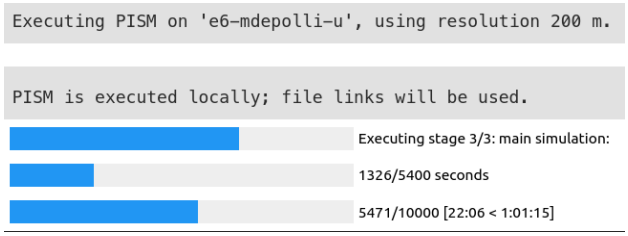


Fig. 6: While a simulation is in progress, GUI visualizes its progress to the user.

as development of a GUI in PyQt is much more complex and time consuming than using Jupyter widgets.

The most notable positive aspects of providing a custom interface towards PISM are listed below.

- Experiments are reproducible from a minimal set of input parameters.
- Experiment can comprise of an arbitrary number of consecutive PISM executions, e.g. to improve execution speed, multigrid simulations are often used, which are presented as multiple PISM calls with careful handling of input/output files and modifying command line arguments between the consecutive calls.
- Experiments do not need to be executed locally, workstations with ample computational power can be used instead of user's laptops or PCs. This in turn enables one to design and execute experiments that run for several days due to their computational complexity.
- The presented environment is easy to learn for newcomers.
- The presented environment is fully extensible, since it primarily comprises of Python scripts and Python code within Jupyter notebooks.
- GUI is straight-forward to develop since all the required Python widgets are supported by the framework. All the required input arguments of PISM were converted into widget-based inputs, by a fast and efficient development process, facilitated by the simplicity of widgets. Note that the simplicity of

widgets can be a two-edged sword if a more complex interface is required.

The presented GUI also has its set of problems.

- Execution is slow, for example, rendering the *Experiment setup* cell can take 10 seconds.
- Working with the Jupyter-based GUI is not the most intuitive. Selected parameters in a cell that renders the GUI are not immediately available – another cell with the code for collecting the parameter values must be executed, to make them available to the rest of the notebook. Moreover, while any cell of the notebook is being executed, input widgets from all the cells are not responsive.
- The GUI is limited by the visual elements provided by the *ipywidgets* library. Storing and loading settings from a file, for example, is a user-created widget and is not as responsive as the first-time users might expect from it.
- Multi-user support is only partial on PISM execution server. Multiple users could, for example, run simulations concurrently and inadvertently write over other user's input or output files. Multiple concurrent users could also overload the server, since there is no load-balancing implemented, nor are users aware of the existing load of the PISM execution server.
- Only a subset of PISM options is supported by the GUI – those that were used within our experiments so far.
- The GUI is tightly decoupled to the PISM simulations, therefore, Jupyter notebook must be kept open for the duration of the simulations, or the results will be lost. Therefore, running complex experiments that take hours or days to complete is unrealistic in GUI.

There is room for improvement in the deployment too. We had learned of several quirks while working on the test deployment. Multi-user support could be bettered by running Jupyter notebooks in an instance of JupyterHub instead of just JupyterLab, which will be attempted in the next phase of development.

A. Generalization

Given the flexibility of Python, Jupyter and especially Jupyter widgets our approach is general enough that it could be applied to other software as well. The condition is, that such other software supports either command-line or configuration file interface – there is no sense in creating another graphical interface if one already exists. We also see no hurdles in implementing execution on a remote host over an SSH connection for other software. The interactive part of the proposed GUI implementation is not great though (slow and with limited interface), and the source code that implements it is notably more complex than the rest. Therefore, we would not recommend our approach where real-time user interaction is required or where user interaction is complex, e.g. selection of a region on a plot, 3D rotation of plots, interaction with sound or animation.

IV. Conclusions

We present an attempt at making the interface towards a complex simulation software more manageable for experimenters and easier to learn for new users. The presented approach separates user interface from the simulation software and is even able to put it on another computer entirely. Currently, the presented GUI is meant to be only used online with the editing disabled. All the GUI development is performed on personal computers with Jupyter locally installed and is propagated to the web through `git`. There are ongoing efforts to make GUI and the external models also editable from the web interface. Although the presented GUI serves to handle PISM, the approach is general enough to be used on other software as well.

Acknowledgment

This work was funded by the Slovenian Research Agency, research core funding No. P2-0095 and project funding No. J1-2479.

References

[1] E. Bueler and J. Brown, "Shallow shelf approximation as a "sliding law" in a thermodynamically coupled ice sheet model," *J. Geophys. Res.*, vol. 114, p. F03008, 2009. [Online]. Available: https://doi.org/10.1029/2008JF001179

[2] R. Winkelmann, M. A. Martin, M. Haseloff, T. Albrecht, E. Bueler, C. Khroulev, and A. Levermann, "The potsdam parallel ice sheet model (pism-pik) part 1: Model description," *The Cryosphere*, vol. 5, pp. 715–726, 2011. [Online]. Available: https://doi.org/10.5194/tc-5-715-2011

[3] T. Kluyver and B. R.-K. et al., "Jupyter notebooks – a publishing format for reproducible computational workflows," in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides and B. Schmidt, Eds. IOS Press, 2016, pp. 87 – 90.

[4] J. M. Perkel, "Why jupyter is data scientists' computational notebook of choice," *Nature*, vol. 563, no. 7732, pp. 145–147, 2018.

[5] F. Pérez and B. E. Granger, "Ipython: a system for interactive scientific computing," *Computing in Science & Engineering*, vol. 9, no. 3, 2007.

[6] J. Allaire, J. Horner, V. Marti, and N. Porte, "The markdown package: Markdown rendering for r," 2014.

[7] F. Pezoa, J. L. Reutter, F. Suarez, M. Ugarte, and D. Vrgoč, "Foundations of json schema," in *Proceedings of the 25th international conference on World Wide Web*, 2016, pp. 263–273.

[8] T. Bray, "The javascript object notation (json) data interchange format," Tech. Rep., 2014.

[9] J.-W. maintainers, "Jupyter-widgets/ipywidgets: Interactive widgets for the jupyter notebook." [Online]. Available: https://github.com/jupyter-widgets/ipywidgets

[10] T. Gamblin, M. P. LeGendre, M. R. Collette, G. L. Lee, A. Moody, B. R. de Supinski, and W. S. Futral, "The Spack Package Manager: Bringing order to HPC software chaos," in *Supercomputing 2015 (SC'15)*, Austin, Texas, November 15-20 2015. [Online]. Available: http://tgamblin.github.io/pubs/spack-sc15.pdf

Computational Performance Aspects of CROCO-BFM Coupling

Filip Strniša[a], Martin Vodopivec[b], Gregor Kosec[a]
[a] Jožef Stefan Institute, Parallel and Distributed Systems Laboratory, Ljubljana, Slovenia
[b] National Institute of Biology, Marine Biology Station, Piran, Slovenia
filip.strnisa@ijs.si, martin.vodopivec@nib.si, gregor.kosec@ijs.si

***Abstract*—Coastal and Regional Ocean COmmunity model (CROCO) is a modeling system used in oceanographic simulations. CROCO solves primitive hydrodynamic equations for momentum, heat and mass transport on a three dimensional, terrain following grid. Several biogeochemical models are already implemented in CROCO, including the PISCES model, which it borrows from NEMO, a different oceanography modeling system. While PISCES is a highly complex model, used in many oceanographic applications, it lacks variable stoichiometry and some plankton functional types needed in certain applications. On the other hand, Biogeochemical Flux Model (BFM) is a dedicated biogeochemistry modeling system, customizable and of user-defined complexity. It can be run as a standalone model or in conjunction with an oceanographic simulation. Previously we have reported on our work regarding the incorporation of BFM into CROCO. In this work we address the performance aspects of this coupling. Namely, compared to e.g. PISCES, BFM is a much more complex model with more tracers, which makes it more computationally expensive. It is therefore of the essence to investigate, and to highlight possible areas of improvement in this coupling's computational performance.**

***Keywords*—CROCO; BFM; coupling; oceanogrphy; biogeochemistry; modelling; computational performance**

I. Introduction

Coupling specialist models with each other is important, as it merges different branches of ocean science. Often circulation models are coupled with marine biogeochemical models [1]–[4]. The circulation models bring the spatial aspect into biogeochemical calculations, while on the other hand, the biogecohemical models can provide spatialy and temporally variable light absorption which influences the temperature profile in the circulation models [5]. Although some circulation models come prepackaged with some of the biogeochemical models [6], [7], comprehensive studies may sometimes require the use of a specialist model such as the Biogeochemical Flux Model (BFM) [8]. The biogeochemical models that come coupled with the oceanographic models, can serve as a template on how to implement a different biogeochemical model in the existing code.

CROCO is an oceanographic model, which incorporates momentum, mass (tracer), and heat transport [6]. CROCO code includes coupling with several biogeochemical models and the most sophisticated among these is PISCES biogeochemical model. The latter was originally coupled with NEMO oceanographic model [7]. Although PISCES is widely used in biogeochemical simulations, we have decided that our needs would be better served with BFM. The latter was chosen because it allows for variable stoichiometry (nutrient ratios in living functional groups — LFG), because the addition of new LFGs is relatively straightforward, and because the BFM is the model of choice in Copernicus Mediterranean biogeochemical forecasts and reanalyses, which will provide the boundary conditions in simulations.

In this work we present some aspects of our coupling of CROCO and BFM. First we briefly introduce both modeling environments, and proceed to highlight some key aspects of the coupling. The main focus of the paper is the assessment of the computational performance of this coupling, as compared to the original CROCO-PISCES model. We then additionally briefly explore the accuracy of a performance enhancing solution, and discuss possible areas of improvement for the current state the coupled codes.

II. CROCO

CROCO is a branch of the widely used Regional Ocean Modeling System (ROMS) [9], with an additional non-hydrostatic core and several other features including an extensive MATLAB toolbox for model configuration and results analysis [10]. Although written in Fortran, CROCO's design enables the user to toggle between different model complexities through C-preprocessor by defining or undefining its components at compilation [11]. In its essence CROCO solves primitive transport equations. Besides the oceanographic physics, it also enables the user to employ the included biogeochemical models. This is done by introducing tracer variables which represent the concentrations of biochemical components in the ocean. By enabling ocean biology, the user may further make their pick of biology-dynamics model of varying complexity (5-24 tracer components). The most sophisticated of the pre-included models is PISCES with 24 tracer variables. PISCES is borrowed from a different oceanographic model — NEMO [7].

The authors would like to acknowledge the financial support of Slovenian Research Agency (ARRS) in the framework of the research core funding No. P2-0095 and projects J7-2599, and Z7-1884.

III. BFM

BFM is an independently developed biogeochemical model [8]. It allows the user to specify its complexity, first at build time, and also at run time. The complexity involves the choice to compute pelagic, and benthic variables, to include sea ice effects, etc. By varying the complexity, the user not only chooses the amount of tracer variables that are used in computations, but also how and to what degree these variables interact with each other. When building BFM the user can pick one of the preset configurations or they can create their own configuration. In this study the "PELAGOS2" preset was used, as it is appropriate for use in coupled models.

IV. Coupling

For the purpose of this research CROCO 1.1, and BFM 5.2.0 were used. BFM was previously successfully coupled with NEMO [7]. Therefore PISCES is a common link between CROCO and BFM. PISCES' implementation in CROCO, as well as BFM's implementation in NEMO were both used as blueprints for coupling of BFM and CROCO. BFM's coupling with NEMO is achieved by almost completely rewriting NEMO's Tracers in Ocean Paradigm (TOP) component. All of this was not necessary in CROCO, as CROCO has its own subroutines for tracer transport. However, BFM's coupling with NEMO was instrumental to understanding the initiation processes, as well as main BFM tracer dynamics calls.

BFM's "PELAGOS2" preset introduces 58 tracer variables. These tracers are then initialized in CROCO's `init_arrays` subroutine. Tracers' initial values can either be obtained from input NetCDF files or from analytical initialization through `ana_initial` subroutine, both of which are taken care of by CROCO. Analytical boundary condition settings need to be manipulated in the coupled source code. The `main` program then calls a BFM subroutine `bfm_ini_tile` which initializes the appropriate BFM arrays and parameters, including reading the BFM namelist files, and initializing the parameters found therein. Also in this step, masks for 3D-to-1D and *vice versa* array transformations are built. An additional call is made to the `trc_sbc_bfm` subroutine, which reads the input files and initializes the surface boundary conditions, such as irradiation, and dust deposition. The latter subroutine has been mostly rewritten by following the preexisting PISCES' subroutine `trc_sbc_pisces`'s example. During timestepping in the `main` program, the 3D tracer arrays are updated via the `step3d_t` subroutine, which also includes a biology update (by calling `biology_tile`). Therein the main BFM call is made. CROCO's 3D tracer arrays (`t`) are transformed to BFM's 1D tracer arrays (`D3STATE`). These are then passed to BFM's main subroutines (`trc_bfm`) to calculate their respective mass fluxes. BFM's tracer arrays are then updated with these mass fluxes through Euler's method. Finally the values from BFM's diagnostic arrays, as well as tracer arrays, are copied back to CROCO's arrays. The output is done entirely through CROCO's subroutines, as BFM's output was disabled, with the exception of a BFM's separate log file. Some of the key communications between different subroutines of the two codes, are laid out in Fig. 1.

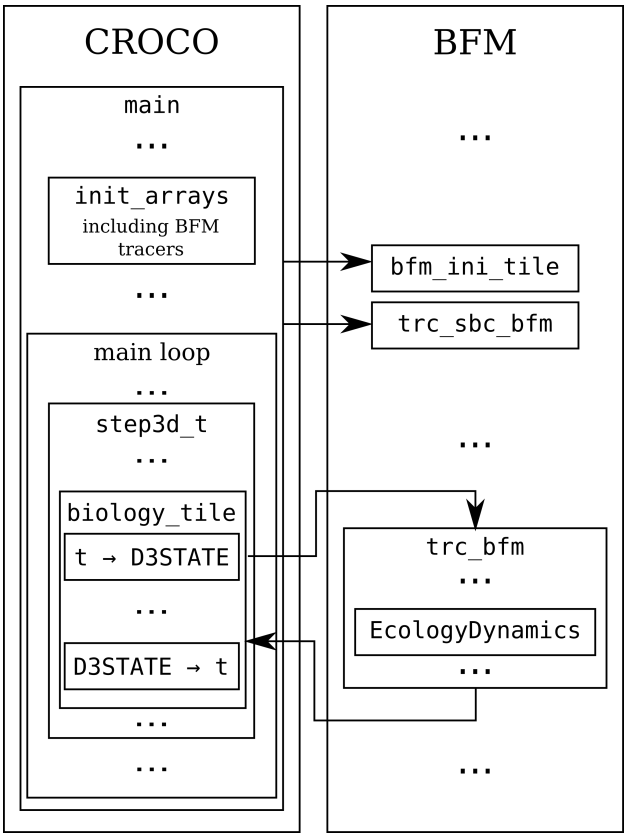


Figure 1. A schematic illustrating some of the key communications made between the two codes.

V. Test case

The CROCO-BFM coupled model was verified by comparing it to the coupling of MITgcm and BFM [2]. A test case was designed where a wind-induced gyre is formed within a closed square oceanic pool: 41.0°N – 43.0°N, -20.5°E – -17.5°E, with even 280 m depth. The grid dimensions were $95 \times 86 \times 32$. Model verification is not the purpose of this study, and these results were presented previously [12]. While conducting these tests, however, very slow performance of the coupled model was observed.

Consequently, to mitigate the slowdown by BFM, the biology call can be made only every n iterations, an approach also explored in [2]. The biological processes are expected to be slower than tracer physics which justifies the longer biology time step. Thus the test case is set up to simulate roughly 14 days of oceanic circulation (2000 time steps of 600 s) and biology in separate runs with PISCES, and BFM. In the PISCES run the biology call is made on every update, whereas in BFM runs it is made every n updates, $n \in [1, 8] \cap \mathbb{N}$. The PISCES run serves merely as an example of what "optimal" computation times could be, and BFM runs are compared to the $n = 1$ run to check for possible accuracy

concerns when decreasing the biology update frequency. To measure the accuracy of $n \in [2,8] \cap \mathbb{N}$ solutions the normalized ℓ^2-norm is calculated as follows:

$$\ell^2-\mathrm{norm}(c_n) = \frac{1}{\max(c_1)} \sqrt{\frac{\sum_i^N (c_{1,i} - c_{n,i})^2}{N}}, \quad n \in [2,8] \cap \mathbb{N}, \tag{1}$$

where $c_{n,i}$ represents the tracer concentration for solution n at i-th node, and N is the total number of discrete grid nodes.

VI. Results and Discussion

The results presented in this section were obtained by running the test case on a laptop PC equipped with AMD Ryzen 7 6800HS CPU, and 40 GB DDR5 RAM. Although the computational performance of the code can be significantly improved by running it in parallel mode (MPI), these tests were performed in single-core mode to better highlight the execution time ratios of the biology vs tracer physics step. Some aspects of the speed-up achieved through parallel execution were presented in [12].

First we present a profile break-down of the `step3d_t` subroutine, which lays out the total time spent at a certain part of the aforementioned subroutine. Results are given in Fig. 2, and Table I.

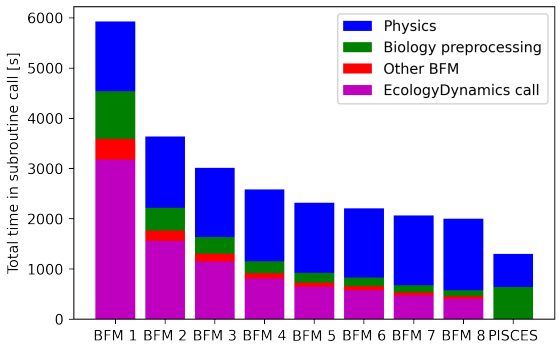


Figure 2. Total time spent on different subroutines. "Physics" refers to the `step3d_t` subroutine, excluding the biology update; "Biology preprocessing" refers to the transformation of `t` to `D3STATE`, and vice versa, and in PISCES's case the full biology update; "Other BFM" are diverse BFM calls; "EcologyDynamics call" is the main BFM call to `EcologyDynamics`. The number following "BFM" is indicating the period of biology update in time steps, n.

The slight deviation in average times between the cases, presented in Table I, are due to the PC not having locked frequencies, and the computations were being carried out, while the PC was used for other tasks as well. However, for the purpose of this study, the quality of these data should be sufficient, as expected trends are shown. Some performance hits of BFM compared to PISCES can be expected as the number of tracers used in the simulation is more than doubled (58 as opposed to 24). Consequently the average execution time of the tracer physics part of `step3d_t` subroutine increases from 0.33 s to 0.70 s. Therefore, equal execution time for BFM and PISCES cannot be expected. However, the average `biology_tile` call goes from 0.32 s for PISCES to 2.34 s for BFM, which is about 7-times slower. With BFM being a more complex model, some degree of slowdown is expected here as well, but this result seems excessive. The transformation between BFM's 1D and CROCO's 3D arrays appears to be the weak point of our coupling's current implementation, as it takes about 20% of `biology_tile` call time on average, and more work should be dedicated to alleviate this issue. Fortran supports vector remapping, which could be a solution here. In this case `t` (of CROCO) should be declared as target, and `D3STATE` (of BFM) as pointer. However, the issue is that `t` is structured as `(positional id, species)`, and `D3STATE` as `(species, positional id)`, which prevents the use of the proposed solution. As Fortran arrays are stored in memory in column-major order, this means that values of `t` for the same tracer species at different coordinates are stored close to each other. This is optimal for computing mass transport — the main goal of CROCO. In `D3STATE` different species at the same coordinate are stored close, which is optimal for computing interplay of these species — the main goal of BFM. Thus to solve the issue of restructuring the basic tracers' data structures on every biology call, the BFM code would need to be rewritten. All instances of `D3STATE`, and other arrays that it interacts with, would need to be changed to `(positional id, species)`, and this may bring upon a performance hit in the `EcologyDynamics` call. By increasing n (the period of biology update in time steps) the total time spent computing biology is brought down significantly, which is shown in Fig. 2. If one's aim was to make the time computing biology about equal to the time computing tracer physics (as is the case in PISCES), n should be between 3, and 4. At $n = 8$ the total time spent computing biology is even lower for BFM than for PISCES, which is expected, as, as it was mentioned above, BFM is about 7-times slower than PISCES. How the value of n affects the results accuracy is presented in Fig. 3.

Some data in Fig. 3 appear distorted, because the output was done every 10 time steps. Consequently, a tracer's concentration may not have been updated at a certain output and the deviation from $n = 1$ solution may be greater than at an output when the concentration was updated. Thus $n = 2$, and $n = 5$ lines appear smooth, as 2 and 5 are divisors of 10. Some tracers are more affected by increasing n than others. Out of 58 those, which suffer from accuracy issues the most, are presented in Fig. 3. These are the concetrations of: N1p - phosphate; N4n - ammonium; P1c - carbon in diatoms; P3c - carbon in picophytoplakton; R1c - carbon in labile dissolved organic matter; R6c - carbon in particulate organic matter; Z4c - carbon in omnivorous mesozooplankton; Z5c - carbon in microzooplankton. As one would expect the deviation from $n = 1$ solution increases with increasing n, which effectively increases the time interval in Euler's

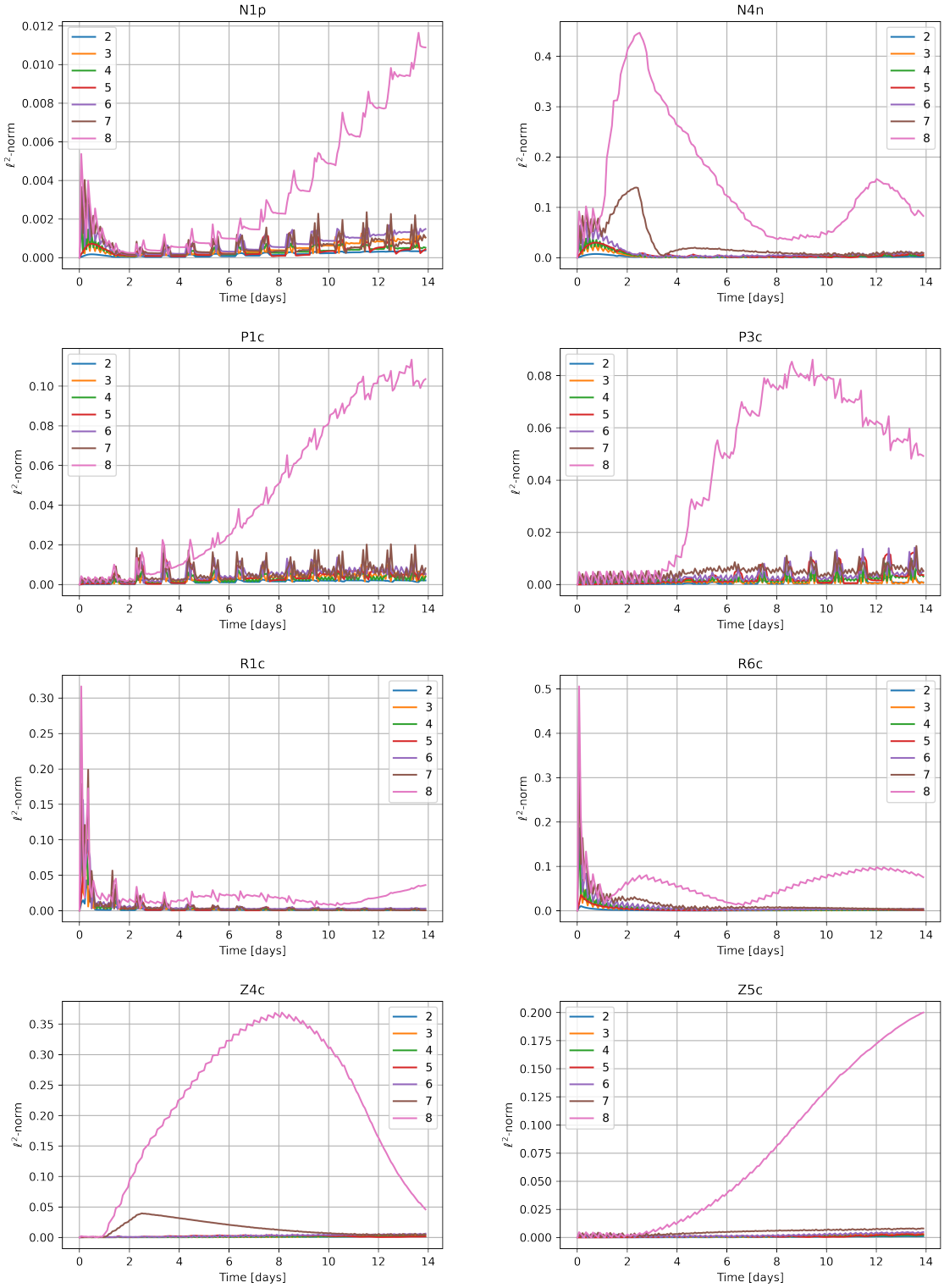


Figure 3. The discrepancy of tracer concentrations for different biology update frequencies. A case where biology is solved on every time step ($n = 1$) is the base result, to which we compare the dimensionless ℓ^2-norm of other cases. Displayed are only the tracers where the ℓ^2-norm exceeds 0.01. The labels represent the concentrations of the following species: N1p - phosphate; N4n - ammonium; P1c - diatoms; P3c - picophytoplakton; R1c - labile dissolved organic matter; R6c - particulate organic matter; Z4c - omnivorous mesozooplankton; Z5c - microzooplankton. The label numbers represent the biology update period in expressed in time steps (n).

TABLE I. AVERAGE AND TOTAL TIME SPENT COMPUTING TRACER PHYSICS AND BIOLOGY IN STEP3D_T.

Case	Average tracer physics time [s]	Average biology time [s]	Total tracer physics time [s]	Total biology time [s]
BFM ($n = 1$)	0.69	2.27	1388	4538
BFM ($n = 2$)	0.71	2.22	1418	2220
BFM ($n = 3$)	0.69	2.46	1373	1639
BFM ($n = 4$)	0.71	2.31	1430	1156
BFM ($n = 5$)	0.70	2.30	1398	923
BFM ($n = 6$)	0.69	2.48	1378	829
BFM ($n = 7$)	0.70	2.36	1392	676
BFM ($n = 8$)	0.71	2.30	1421	578
PISCES	0.33	0.32	661	641

integration. However, these deviations can be over 40% (N4n, R6c, when $n = 8$). Although the computational performance is increased, there can be a significant impact on accuracy, and in going forward a balance between the two should be found. Implementing a different time-integration scheme, such as 2nd order Euler method, or even 4th order Runge-Kutta may improve the integration accuracy, but would inherently increase the amount of `EcologyDynmaics` calls in a given time step, effectively further slowing down the computation. Specifically for the presented case, $n = 3$ or $n = 4$ seem like a good compromise, as at these n the total biology time is about the same, as total tracer physics time, and the deviation from the $n = 1$ solution is not so big. However, one should note that this computations were done in the period of vigorous model dynamics due to the adaptation to the initial conditions. After the initial growth, many discrepancies start to fall as the model reaches a more stable state. It would perhaps be interesting to explore the possibility of an adaptive biology update period, as in oceanographic computations not all of the nodes are equally biologically active.

VII. CONCLUSION

Performance aspects of the CROCO-BFM coupling were presented, i.e. the computational performance, a possible measure to improve the latter, and a brief numerical accuracy analysis of these measures. We have considered a small-scale closed basin as a test case, solving it for approximately 14 days of simulated time. The coupled model performs significantly slower than the reference model, which is CROCO with PISCES biology. One reason behind slower performance is BFM simulating 58 tracers vs PISCES' 24. This results in longer computation times for both tracer physics, and biology. However, while the tracer physics computation time merely doubles, the biology computation time increases sevenfold. Some of these issues come from having to transform between CROCO's, and BFM's data structures in their communication, but there is no simple solution to this problem. The issue of poor performance can be solved by performing the biology update at a lower frequency, i.e every $n-\mathrm{th}$ step. In doing so, one can achieve satisfactory performance results, but there is noticeable penalty to the accuracy. Our tests show that $n = 3$ or $n = 4$ produce good results in test case conditions, while keeping the slowdown at a reasonable amount. The effects of the biology call on reduction of accuracy in large-scale simulations, as well as for longer time periods, have not been explored in this study.

REFERENCES

[1] J. S. Lessels, D. Tetzlaff, S. K. Carey, P. Smith, and C. Soulsby, "A coupled hydrology–biogeochemistry model to simulate dissolved organic carbon exports from a permafrost-influenced catchment," *Hydrological Processes*, vol. 29, no. 26, pp. 5383–5396, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10566

[2] G. Cossarini, S. Querin, C. Solidoro, G. Sannino, P. Lazzari, V. Di Biagio, and G. Bolzon, "Development of bfmcoupler (v1.0), the coupling scheme that links the mitgcm and bfm models for ocean biogeochemistry simulations," *Geoscientific Model Development*, vol. 10, no. 4, pp. 1423–1445, 2017. [Online]. Available: https://gmd.copernicus.org/articles/10/1423/2017/

[3] R.-H. Zhang, F. Tian, and X. Wang, "A new hybrid coupled model of atmosphere, ocean physics, and ocean biogeochemistry to represent biogeophysical feedback effects in the tropical pacific," *Journal of Advances in Modeling Earth Systems*, vol. 10, no. 8, pp. 1901–1923, 2018. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017MS001250

[4] H.-C. Jung, B.-K. Moon, H. Lee, J.-H. Choi, H.-K. Kim, J.-Y. Park, Y.-H. Byun, Y.-J. Lim, and J. Lee, "Development and assessment of nemo(v3.6)-topaz(v2), a coupled global ocean biogeochemistry model," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 56, no. 3, pp. 411–428, Aug 2020. [Online]. Available: https://doi.org/10.1007/s13143-019-00147-4

[5] K. Fennel, J. P. Mattern, S. C. Doney, L. Bopp, A. M. Moore, B. Wang, and L. Yu, "Ocean biogeochemical modelling," *Nature Reviews Methods Primers*, vol. 2, no. 1, p. 76, 2022.

[6] "Croco – coastal and regional ocean community model," 10.5281/zenodo.7415055, accessed: 16 Jan 2023.

[7] "Nemo," nemo-ocean.eu, accessed: 17 Feb 2023.

[8] "Biogeochemical flux model — bfm," bfm-community.eu, accessed: 16 Jan 2023.

[9] "Roms," myroms.org, accessed: 17 Feb 2023.

[10] "Croco tools," 10.5281/zenodo.7431960, accessed: 16 Jan 2023.

[11] "Croco – coastal and regional ocean community model documentation," 10.5281/zenodo.7400758, accessed: 16 Jan 2023.

[12] M. Vodopivec, F. Strniša, and G. Kosec, "Coupling the coastal and regional ocean community model (croco) with the biogeochemical flux model (bfm)," in *EGU General Assembly 2022, Vienna, Austria*, 23–27 May 2022.

OpenMP offloading and OpenACC programming model approach for object-oriented plasma device algorithms

Ezhilmathi Krishnasamy*, Ivona Vasileska†, Leon Kos†, Pascal Bouvry*,
* University of Luxembourg/PCOG, Belval, Luxembourg
† University of Ljubljana/LECAD, Ljubljana, Slovenia
ezhilmathi.krishnasamy@uni.lu, ivona.vasileska@lecad.fs.uni-lj.si, leon.kos@lecad.fs.uni-lj.si, pascal.bouvry@uni.lu

***Abstract*—Plasma physics is becoming more important due to its applications in clean energy production (using fusion technology) and other fields, such as chemical and material science. Even recently, Lawrence Livermore National Laboratory (LLNL) has demonstrated the capability of producing more energy through fusion compared to laser energy. Therefore, in the future, we might need to do more computational simulations for further understanding and explore the advancement in plasma physics. Furthermore, this could be possible with the help of supercomputers. In this work, we parallelise a one-dimensional object-oriented plasma device algorithm, Object Oriented Plasma Device 1d (oopd1), on a multicore CPU and GPU. We use the OpenMP programming model for the CPU version, and for the GPU, we use OpenMP offloading and OpenACC offloading. All of these approaches are compared to each other. Thus, it provides further suitable programming models with parallel capabilities for the existing oopd1 to explore the available parallel architectures.**

***Keywords*—GPU, OpenACC, OpenMP Offloading, Plasma Physics**

I. INTRODUCTION

Plasma simulation is the computational modelling of the interaction of charged particles with electric and magnetic fields. The plasma may exhibit complex nonlinear behaviour, and the fields and particles may interact with time-dependent boundary conditions. Fluid codes model the plasma using moments of a distribution function at discrete grid points. In contrast, particle-in-cell (PIC) codes to model the plasma using discrete particles, each representing many charged particles. The particle-in-cell Monte Carlo (PIC-MC) method is a self-consistent kinetic approach capable of predicting the electron and ion energy distribution function (EEDF and IEDF), respectively. Essentially, the PIC simulation employs thousands of simulated particles, known as superparticles, to represent a significantly larger number of real particles (10^{14} - 10^{18} m^{-3}). In a PIC simulation, the motion of each particle is simulated, and macro-quantities (such as particle density, particle flux, current density, etc.) are calculated from the position and velocity of these particles. The macro force acting on the particles is calculated from the field equations.

The **o**bject **o**riented **p**lasma **d**evice 1d (oopd1) is a one-dimensional object-oriented PIC program that pushes particle positions and velocity to advance. The main idea of using particles is to represent space collocation, which is much computationally cheaper than the Vlasov/Boltzmann methods [1]. Furthermore, at the same time, Vlasov/Boltzmann represents six-dimensional space and enables arbitrary plasma systems, similar to magnetohydrodynamic methods. It contains a model for planar geometry and includes models for cylindrical and spherical geometries. The oopd1 also allows for different weights of simulation particles and relativistic treatment of electrons. The object-oriented methodology is used for plasma simulations in a physical device with field-emitted electrons in the presence of ion current in a one-dimensional diode. The configuration for a one-dimensional bounded plasma system is shown in Fig. 1.

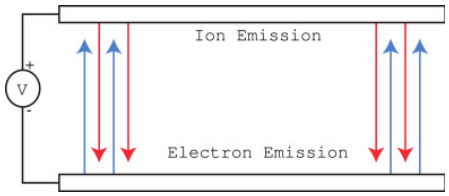


Fig. 1: Plasma one-diode system.

The algorithm of the oopd1 is common with the other PIC codes [2]–[4]. It contains a particle mover, which updates the position and velocities of the simulated particles according to the famous Newton's laws of motion, and a field solver, which calculates the fields inside the simulated spatial region at some grid points.

A. Contribution

For the oopd1 code, we have implemented the following approaches:

- OpenMP implementation that uses the available threads in the CPU
- OpenMP offloading directives targeting GPU
- OpenACC directive programming model for GPU enabled parallelisation on GPU
- Performance analysis comparison between all of these implementations

B. previous study

As the oopd1 is based on the PIC code, many studies have been conducted to parallelise similar code on the CPU and GPU [5]–[7]. Moreover, the simplified version of the oopd1 is called SIMPIC, which is also well studied with various parallel implementation approaches [2]–[4]. However, oopd1 involves complex physics and many options to simulate, such as argon, hydrogen, oxygen and chlorine. As far as we know, no study has been explored on parallel implementation of the oopd1 code on GPU.

II. Parallel Programming

Parallel programming exists to achieve parallel computation on parallel computers. Parallel architecture exists in the form of distributed memory, shared memory and hybrid memory (compute node is connected with accelerators, e.g., Graphics Processing Unit (GPU) and Field Programmable Gate Array (FPGA)). For the past few decades, parallel programming exists, notably Message Passing Interface (MPI) and Open Multi-Processing (OpenMP) programming models. For example, MPI is for parallelisation on distributed memory, and OpenMP is for parallelisation on shared memory parallel architecture.

In particular, MPI and OpenMP have version standards which are being updated regularly; the initial MPI standard was created in 1992 [8]. Similarly, the OpenMP version standard was set in 1997 for the FORTRAN language [9]; since then, OpenMP standards have been regularly updated, and it supports C/C++ as well. But recently, many parallel programming models have been emerging, e.g., Compute Unified Device Architecture (CUDA), Heterogeneous Interface for Portability (HIP), and Open Computing Language (OpenCL). These are being used for explicitly targeting the accelerators, such as GPUs and FPGAs.

As the MPI and OpenMP have existed for many decades, most scientific applications have been written using those programming models. However, to make use of the new parallel heterogeneous architecture (compute nodes with accelerators), scientific codes or programmers have to adapt to vendor-specific (e.g., CUDA and HIP) or low-level programming (OpenCL) to target these parallel heterogeneous architectures.

Although these vendor-specific or low-level programming languages might be employed to port the existing code to accelerators, it might need more effort in terms of time spent on implementing and understanding these programming languages. In order to avoid those obstacles, one could use OpenACC and OpenMP offloading. Both of this OpenACC and OpenMP offloading are very easy to implement with minor changes in the existing code and are directive-based programming languages. This is what we address in this work.

III. Testing Platform

Presently supercomputers have achieved exascale computing with the help of heterogeneous architecture, e.g., Frontier [10] has AMD CPU and GPU. Similarly, most supercomputers worldwide have accelerators connected to compute nodes (CPU) such as Nvidia GPU, Intel GPU and FPGA. This work considers IBM POWER9 AC922 at 3.1 GHz with Nvidia Volta V100 GPU. Each compute node has 2 CPU sockets, which have in total of 32 cores. Each core can have up to 4 threads, bringing a total of 128 [11]. Although both OpenACC and OpenMP offloading can be executed on AMD and Nvidia GPU, we consider only Nvidia GPU here.

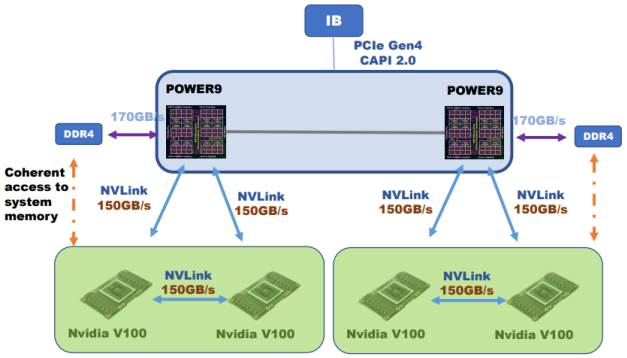


Fig. 2: IBM: Power9 (CPU) with Nvidia Volta V100;source [13].

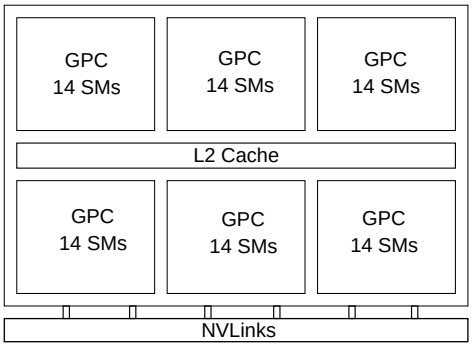


Fig. 3: Nvidia: Volta V100 GPU with 84 streaming multiprocessors.

Figure 2 shows the technical hardware configuration of IBM POWER AC922 with Nvidia GPU compute node. The important perspective of this architecture is that two GPUs are connected to each CPU socket, thus providing a bandwidth of up to 150 GB/s between CPU and GPU. Furthermore, Figure 3 shows the Nvidia Volta architecture GPU that has 6 GPU Processing Cluster (GPC), and each GPC has 14 Streaming Multiprocessors (SMs) with a total of 80 SMs [12]. Nvidia V100 has 2688 double precision cores running at a maximum speed of 1530 MHz [12].

IV. Implementation

This section focuses on the parallel implementations of using OpenMP, OpenACC and OpenMP offloading simulation for argon gas in oopd1.

A. OpenMP

We have introduced the OpenMP directives wherever it is needed. We have identified the parallelisation blocks, where there is no data dependency and data race. This ensures the correctness of the solution and parallelisation in oopd1. We have also introduced `private` and `shared` to identify the variables and avoid computational errors. Listing 1 shows the simple syntax of the OpenMP implementation that we have used in the oopd1 code. We have used the GNU compiler for the compilation.

Listing 1: OpenMP.

```
#pragma omp parallel for
for (int i=0; i<N; i++)
   {
   // computation
   }
```

B. OpenACC

OpenACC is very similar to OpenMP, which is based on the directive programming model. Where adding the directive clauses above the loops will parallelise the loop. Furthermore, OpenACC has three concepts; they are incremental (original or existing code can be maintained), single source (some code can be executed on different architecture), and low learning curve (easy to learn). Listing 2 shows the pseudo-code implementation and its output, where we can see that the loop is already parallelised with 128 GPU threads. An `acc` will instruct the compiler to execute the code block on the device, and `parallel` will instruct the code block to be executed in parallel on the device. Since the code has 1000 lines and more than 100 files, we opted to use the unified memory option in OpenACC. This option will take care of the memory transfer wherever it is necessary. Even if we do it with explicit memory allocation, we would not expect to get a large difference in performance. Therefore, we have enabled the unified memory option flag `-ta=tesla:managed` during the compilation.

Listing 2: OpenACC.

```
#pragma acc parallel loop
for (int i=0; i<N; i++)
   {
   // computation
   }

/***compilation output***/
Fields::Fields(SpatialRegion *):
     81, Generating NVIDIA GPU code
         83, #pragma acc loop gang, vector(128)
         /* blockIdx.x threadIdx.x */
```

OpenACC is supported by both commercial and open-source compilers, such as HPE, Nvidia HPC SDK, GNU, and OpenARC. However, we chose the Nvidia HPC SDK compiler, and moreover, this is the only option we got from the testing HPC platform. For compilation, we used flags as follows `nvc++ -fast -Minfo=accel -ta=tesla:managed -acc`.

C. OpenMP offloading

The OpenMP API standard started to support the OpenMP offloading from 4.0, and thereafter the standard included more advanced features to support the GPU [14]. Although OpenMP offloading is supported by many compilers, such as AMD, GNU, HPE, IBM, and Intel, we use the Nvidia HPC SDK compiler in this work [15]. Here in OpenMP offloading, `target` will execute the code on the device, `teams` will create teams of thread blocks, and finally, `distribute parallel` will create a number of threads to be executed in parallel on each thread blocks. Listing 3 shows the pseudocode example used in the oopd1 code and the compilation output of how the device code block is parallelised using the default thread block.

Listing 3: OpenMP offloading.

```
#pragma omp target teams distribute parallel for
for (int i=0; i<N; i++)
   {
   // computation
   }

/***compilation output***/
Fields::Fields(SpatialRegion *):
#pragma omp target teams distribute parallel for
         82, Generating "nvkernel__ZN6FieldsC1EP13
         SpatialRegion_F1L82_2" GPU kernel
         85, Loop parallelized across teams and
         threads(128), schedule(static)
```

We used the Nvidia HPC SDK compiler for compilation: `nvc++ -fast -mp=gpu -gpu=managed -Minfo=mp,accel`.

V. Code optimisation

We have chosen the default thread blocks for both OpenACC and OpenMP directive clauses; however, as a programmer, we can also choose to set the threads blocks in OpenACC and OpenMP to get maximum performance. Table I shows how the CUDA terminologies `thread blocks`, `threads`, and `warps` are interpreted in OpenACC (for both `kernels` and `parallel`) and OpenMP offloading. OpenACC offers two options for creating a parallel code block, using `kernels` and `parallel`; however, we have used `parallel` (explicit and creates just one device kernel) for the entire oopd1 code [16]. According to OpenACC, `kernels` is safer (implicit and also creates more device kernels) to use when a programmer does not know much about data dependency [16], [17]. However, since we knew the data flow in oopd1, we have used `parallel`.

Listing 4 shows how the loop can be optimised to use more thread blocks within the parallel code block. As can be seen in Listing 2, we do not know exactly how many gangs will be created, and for each and every gang, we will have up to 128 threads; this can be seen in Figure 4. However, on the Nvidia device, threads in the thread blocks are executed as `warp`, and each `warp` will have up to 32 threads. Each thread block can be executed on the GPU on the SMs; therefore, the equivalent amount of

TABLE I: OpenACC and OpenMP offloading loop mapping in terms of CUDA terminologies [19], [20].

CUDA Terminologies	OpenACC Kernels	OpenACC Parallel	OpenMP Offloading
Thread Blocks	gang	num_gangs	teams
Threads	vector	vector_length	parallel
Warps	worker	num_workers	simd

SMs with threads block will have maximum performance. But this is only possible when we know the problem size; otherwise, it is better to let the compiler choose the default thread blocks for a specific problem. Furthermore, GPU architecture is based on Single Instruction Multiple Threads (SIMT) which execute the threads based on the Single Instruction Multiple Data (SIMD), and it uses a set of 32 threads (warp) [18]. So, warps of 32 threads will be in a queue depending on the SMs resource available. Figure 5 shows these phenomena, where each thread block has just 32 threads and will be executed as one warp on the SM.

Similarly, we have done it in OpenMP offloading as well; that is, setting the thread blocks and the number of threads on each thread block. Listing 5 shows the num_teams and thread_limit; however, we do not see this as a compiler output similar to OpenACC; perhaps using the GNU compiler might give more information for compilation output. We can also see the visualisation phenomena in Figure 5, where threads are grouped as warps and will be executed one by one. Moreover, Nvidia Volta V100 GPU has 32 double-precision floating-point cores (FP64).

Listing 4: OpenACC optimisation.

```
#pragma acc parallel loop
num_gangs(32) vector_length(32)
for (int i=0; i<N; i++)
   {
   // computation
   }

/***compilation output***/
Fields::Fields(SpatialRegion *):
     12, Generating NVIDIA GPU code
         17, #pragma acc loop gang(32), vector(32)
         /* blockIdx.x threadIdx.x */
```

Listing 5: OpenMP offloading optimisation.

```
#pragma omp target teams num_teams(32)
distribute parallel for thread_limit(32)
for (int i=0; i<N; i++)
   {
   // computation
   }

/***compilation output***/
Fields::Fields(SpatialRegion *):
     82, #omp target teams distribute parallel for
     num_teams(32) thread_limit(32)
         82, Generating "nvkernel__ZN6FieldsC1EP13
         SpatialRegion_F1L82_2" GPU kernel
         85, Loop parallelized across teams and
         threads(128), schedule(static)
```

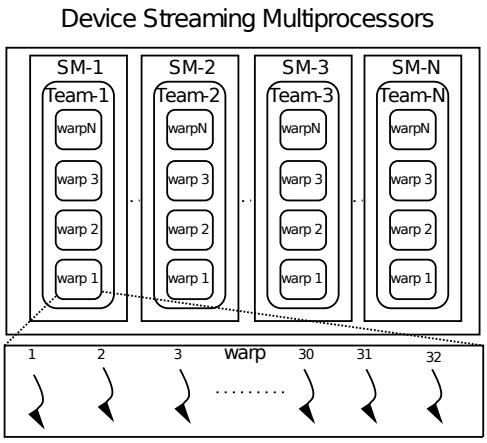


Fig. 4: Schematic workflow of OpenACC and OpenMP offloading threads (default) on the device.

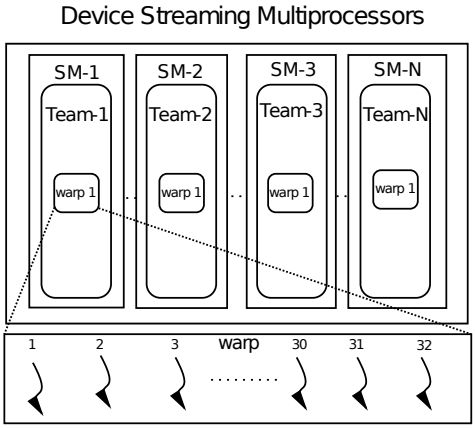


Fig. 5: Schematic workflow of manually setting threads on a device for OpenACC and OpenMP offloading.

VI. Discussion

In oopd1, we have tried the argon gas model for testing with 1000 iterations. Figure 6 shows the time difference between serial and various parallel implementations. Since our main goal is to offload the computational task to GPU, we used all the threads(128) from the compute node for the OpenMP version. As we can notice, OpenACC and OpenMP offloading show similar behaviour and also with optimized thread blocks of these two approaches. This is mainly due to two reasons: the compiler already chooses the optimised thread blocks for the given problem, and the computational task within the device code is not very intensive. However, this phenomenon might look different where the device code needs lots of arithmetic computations, that is, if each warp execution takes more time. In that scenario choosing the thread blocks manually would give little more performance benefit than choosing the default thread block.

At the same time, the OpenMP version with 128 threads does not compete with the other two options, OpenACC and OpenMP offloading; this is mainly because the com-

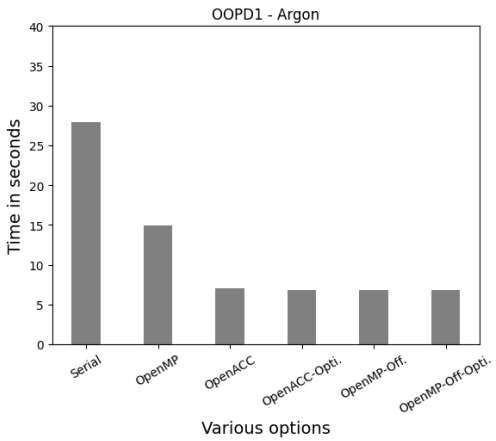


Fig. 6: Schematic workflow of OpenACC threads on the device.

putational domain has more than 128 spatial domains (spatial domain is discretised with 1001). Therefore, the OpenMP version takes a little longer time to complete the task compared to GPU (OpenACC and OpenMP offloading). Because GPU (OpenACC and OpenMP offloading) can have up to 1001 threads, these thread blocks can be executed on the SMs on the GPU. In addition, V100 GPU has up to 84 SMs [12], and each warp can be executed on each SM and adding up to 32 SMs (total, we get 1024 threads (32*32)).

VII. Conclusion

We have ported the oopd1 code to GPU using OpenACC and OpenMP offloading. We notice that both OpenACC and OpenMP offloading show good speedup compared to serial and OpenMP versions. Although optimised option thread blocks for both OpenACC and OpenMP offloading do not show any difference due to the problem (argon gas) nature in oopd1; however, it was good to try this option in oopd1 and notice the behaviour.

We would like to continue with MPI implementation with GPUs in the future. At the same time, we also want to study further OpenMP offloading on the AMD GPU and, if possible, with various compilers for OpenMP offloading. Moreover, we do not see any huge difference between OpenACC and OpenMP offloading; therefore, in the future, we would like to focus on OpenMP offloading, considering the code can be easily run on both Nvidia and AMD GPUs. Finally, we would also like to try various physics parameters in oopd1 and check if the code can be further optimised for the GPU (including heterogeneous computing). From this work, we believe that in the future, a directive-based programming model for the device would completely eliminate the rewriting of any existing scientific code for GPU.

ACKNOWLEDGMENT

We are also grateful to CINECA for letting us to use their GPU cluster computational resources. The first author is funded by the EuroCC, and the second author is funded by PRACE.

References

[1] T. M. Tyranowski and M. Kraus, "Symplectic model reduction methods for the vlasov equation," *Contributions to Plasma Physics*, p. e202200046, 2022.

[2] I. Vasileska, P. Tomšiè, and L. Kos, "Modernization of the pic codes for exascale plasma simulation," in *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2020, pp. 209–213.

[3] I. Vasileska, L. Bogdanović, and L. Kos, "Particle-in-cell code for gpu systems," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2021, pp. 291–294.

[4] E. Krishnasamy, I. Vasileska, L. Kos, and P. Bouvry, "Openmp offloading and openacc programming model approach for object-oriented plasma device algorithms," in *2023 The 8th International Conference on Computer and Communication Systems*. IEEE, 2023.

[5] Z. Juhasz, J. Ďurian, A. Derzsi, Š. Matejčík, Z. Donkó, and P. Hartmann, "Efficient gpu implementation of the particle-in-cell/monte-carlo collisions method for 1d simulation of low-pressure capacitively coupled plasmas," *Computer Physics Communications*, vol. 263, p. 107913, 2021.

[6] S. W. Chien, J. Nylund, G. Bengtsson, I. B. Peng, A. Podobas, and S. Markidis, "sputnipic: an implicit particle-in-cell code for multi-gpu systems," in *2020 IEEE 32nd International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*. IEEE, 2020, pp. 149–156.

[7] H. Shah, S. Kamaria, R. Markandeya, M. Shah, and B. Chaudhury, "A novel implementation of 2d3v particle-in-cell (pic) algorithm for kepler gpu architecture," in *2017 IEEE 24th International Conference on High Performance Computing (HiPC)*. IEEE, 2017, pp. 378–387.

[8] D. W. Walker and J. J. Dongarra, "Mpi: a standard message passing interface," *Supercomputer*, vol. 12, pp. 56–68, 1996.

[9] C. university, "Openmp history," 2022. [Online]. Available: https://cvw.cac.cornell.edu/openmp/openmp-history-and-evolution

[10] Frontier, "Frontier," 2022. [Online]. Available: https://www.olcf.ornl.gov/frontier/

[11] CINECA, "Marconi100 userguide," 2023. [Online]. Available: https://wiki.u-gov.it/confluence/display/SCAIUS/UG3.2%3A+MARCONI100+UserGuide

[12] Nvidia, "Volta gpu white paper," 2023. [Online]. Available: https://images.nvidia.com/content/volta-architecture/pdf/volta-architecture-whitepaper.pdf

[13] Nvidia and IBM, "Ibm ac922," 2023. [Online]. Available: https://on-demand.gputechconf.com/gtc/2018/presentation/s8309-nvidia-v100-performance-characteristics-on-ibm-power-\9-system-and-hpc-application-performance.pdf

[14] OpenMP, "Openmp 5.1 api syntax reference guide," 2020. [Online]. Available: https://www.openmp.org/wp-content/uploads/OpenMPRefCard-5.1-web.pdf

[15] OpenMP.org, "Openmp compilers & tools," 2023. [Online]. Available: https://www.openmp.org/resources/openmp-compilers-tools/

[16] S. Chen, "Introduction to openacc," 2017. [Online]. Available: https://www.bu.edu/tech/files/2017/04/OpenACC-2017Spring.pdf

[17] R. Farber, *Parallel programming with OpenACC*. Newnes, 2016.

[18] A. Papakonstantinou, K. Gururaj, J. A. Stratton, D. Chen, J. Cong, and W.-M. W. Hwu, "Efficient compilation of cuda kernels for high-performance computing on fpgas," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 13, no. 2, pp. 1–26, 2013.

[19] C. Shen, X. Tian, D. Khaldi, and B. Chapman, "Assessing one-to-one parallelism levels mapping for openmp offloading to gpus," in *Proceedings of the 8th International Workshop on Programming Models and Applications for Multicores and Manycores*, 2017, pp. 68–73.

[20] J. Bispo, J. G. Barbosa, P. F. Silva, C. Morales, M. Myllykoski, P. Ojeda-May, M. Bialczak, M. Uchronski, A. Wlodarczyk, P. Wauligmann *et al.*, "Best practice guide: Modern accelerators," Technical report, PRACE aisbl, June 2021. URL: https://prace-ri.eu/training ..., Tech. Rep., 2021.

Benchmark DPC++ code and performance portability on heterogeneous architectures

Nenad Mijić*, Davor Davidović*
* Centre for Informatics and Computing, Ruđer Bošković Institute, Zagreb, Croatia
{nenad.mijic, davor.davidovic}@irb.hr

***Abstract*—Source code portability is becoming increasingly important in the development of new solutions in HPC due to the wide diversification of hardware and heterogeneity of systems. With Intel's oneAPI suite of programming tools and the Data Parallel C++ compiler, a single source code containing both host and device code can leverage hardware architectures from different vendors. Using the compiler's interoperability, it can be linked to existing libraries such as MPI to run the program on a distributed memory system.**

In this paper we benchmark and analyze the performance that can be achieved with the Intel DPC++ compiler, using the distributed Cholesky QR2 algorithm as an example and comparing it with the native CUDA and C++ implementation. The analysis shows that the performance degradation when using SYCL is negligible when a smaller number of nodes are used, but with the cost that some additional self-made optimizations are required in SYCL code.

***Keywords*—code portability, DPC++, GPU, MPI, HPC, Choleksy QR2**

I. Introduction

HPC is currently based on extremely heterogeneous systems, and there is no reason to believe that this situation will change in the near future. The most powerful systems in the TOP500 [1] supercomputers are dominated by such systems, and it is enough to look at the list to see the diversity of combinations of CPUs and accelerators. Intel, AMD and Arm dominate in the field of central processing units, while NVIDIA had a leading role for a decade in the domain of GPU accelerators, and more recently AMD has stepped forward by shipping systems with its GPU graphic processors. AMD has taken the first step forward by delivering the first exascale systems. In addition, thanks to the rise of artificial intelligence, specialised accelerators for AI are being developed and appearing on the scene. For developers and researchers, writing portable code that can run on a number of these architectures and maintain performance is a real challenge.

The diversity of hardware brings along a variety of programming models that achieve high performance using different approaches. Low-level programming models for heterogeneous systems are often completely vendor locked and therefore do not provide portability. The low-level architectural details are exposed directly by the programming language that can exploit them explicitly, often resulting in higher performance. For example, CUDA is exclusive to NVIDIA GPUs while ROCm is exclusive to AMD GPUs. The HIP programming model [2], provided by AMD, is an abstraction layer for CUDA and ROCm, targeting both GPU vendors. Opposite from previously described low-level programming models, OpenCL stands out as a portable solution but relies on the vendor to provide and maintain the backend for runtime execution.

The models based on directives provide direct portability by using high-level *pragma* directives to instruct the compiler to use different architectural implementations for exposing parallelism and provide solutions for offloading computation to the accelerator. They also handle memory management and data transfer between the host and accelerator. Notable examples include OpenMP, which was originally developed for multicore architectures but now supports offloading to the accelerator, and OpenACC [3]. All hardware vendors provide libraries and codes tuned for their respective hardware.

High-level programming models are often abstraction layers that enable the development of cross-platform portability, preferably from a single source code. They are intended as C++ portable APIs mapped to the specific backend for providing portability and features are often overlapped among these models. Such models take advantage of different libraries, directive programming approaches, and other languages to enable portability and leverage hardware specifics providing abstractions for parallel execution and data structures. The best known examples are Kokkos [4], [5], Raja [6], and SYCL [7] targeting shared memory systems. Kokkos and Raja include OpenMP, CUDA, Hip and SYCL backends. Celerity [8] extends SYCL with MPI to distributed memory systems and implements a task-based runtime system for distributed memory. Celerity creates task-based graphs and designates a master node to execute the schedule graph for the entire cluster and oversees how data and compute load is distributed.

This paper presents the code and performance portability of oneAPI solutions with respect to direct native implementation on distributed HPC systems. MPI is used for parallelization on a distributed memory system. The original contributions of this research are as follows:

- achieved single source code portability across different hardware architectures (CPUs and GPUs from different vendors),
- achieved performance portability for simple linear algebra algorithms,
- implements inter-node communication library for distributed memory systems,

- code benchmarked and analyzed on different architectures: AMD CPUs, Intel CPUs and NVIDIA GPU,
- detected several performance issues in Intel DPC++ compiler and libraries, and workarounds on how to resolve them were proposed.

The paper is divided into the following logical order. Section II presents related benchmarks for linear algebra algorithms implemented using the SYCL programming model. The distributed CholeslyQR2 algorithm used for benchmarking is briefly described in Section III with an overview of its implementation. Numerical results and achieved performance portability are presented in section IV, and the paper concludes with comments and an outlook on future work in section V.

II. SYCL and Related work

SYCL is an open standard published by the Khronos group [7] to provide a unified programming model for all heterogeneous architectures. The model is based on modern C++ with a single-source code approach that includes both host and device parts. Parallel part of code is scheduled and executed with kernels object written with lambda functions, offering several choices for parallel construction. During compilation, kernels intended to run on accelerators are extracted from code, compiled with a compiler suitable for that architecture and embedded in a single executable. It was developed with the intention of providing a plugin interface for various programming models. It is assumed that in the future the plugin interface will allow easy integration and support of new programming models that have not yet been announced and thus provide long-term HPC support. The SYCL runtime manages task scheduling, task dependencies and synchronisation by creating task graphs through knowledge of memory access while also being responsible for memory and data management.

The interoperability of SYCL with built-in *host_task* kernels allows the use of a common, unifying interface for accessing non-SYCL libraries. Sycl has been implemented by the following entities among others: Codeplay's ComputeCpp, hipSYCL [9] and Intel's Data Parallel C++ (DPC++) [10]. ComputeCpp supports any OpenCL platform that accepts SPIR/SPIR-V, and additionally NVIDIA by using the PTX backend from LLVM. hipSYCL is an open-source implementation led by the University of Heidelberg based on existing compiler toolchains for heterogeneous computing with Clang plugin via HIP or CUDA toolchain for GPUs, Intel Level Zero for Intels GPUs and OpenMP for CPUs. DPC++ is an open-source implementation based on the LLVM compiler that adopts the SYCL 2020 standard. Codeplay contributed to the Intel project by providing an interface to CUDA that allows direct programming for the NVIDIA GPUs. Lately, a beta version interface for AMD GPU has been announced.

The SYCL standard provides 2 containers for controlling data during code execution. A buffer is a memory object responsible for allocating or controlling the allocated memory on the host or device and implicitly takes care of transferring the memory between the devices and the host during the computation. Unified Shared Memory (USM) provides explicit control of memory allocation and transfer with traditional pointers. USM distinguishes two memory containers, one for the device and one for the host. Buffers provide modern C++ style memory management, while USM provides C style with fine-grained controls that give developers precise control over memory management, which can lead to better-optimized memory transfers.

Code portability has been tested in recent years due to various occurring computing architectures, especially in the area of high-performance computing. Most programming models that guarantee portability work well on different platforms, but often show unsatisfactory performance portability. Extensive testings of various programming models conducted in ([11]) showed at least good code portability but have performance problems with Kokkos and OpenMP models.

First benchmark suite [12] written in SYCL has 71 tests. To test all the capabilities of SYCL, the authors divided the benchmark into microbenchmarks, runtime tests, and application tests with ComputeCPP and hipSYCL. The microbenchmarks tested the performance of code on GPU devices targeting arithmetic or the memory subsystem. Application tested were mainly linear algebra algorithms *gemm*, *syrk*, *syr2k*, Gram-Schmidt *QR*, and *axpy*. For testing the runtime performance the benchmark includes tests that explicitly create complex inter-task dependencies. They showed close to the peak performance of microbenchmarks and divergent results for other tests with different toolchains. The paper [13] extended the work of SYCL-bench to 4 toolchain implementations showing the difference in performance between them but indicating somewhat problems within runtime systems targeting devices with non-optimal configuration.

In [14], the authors studied the performance of kernels from the RAJA performance suite written in the SYCL standard using the hipSYCL toolchain. The authors observed the competitive performance of their implementation against the native CUDA implementation.

To test the performance portability on distributed systems, the authors of Celerity tested the runtime system with a synthetic benchmark to evaluate performance portability. They observed negligible or slightly worse performance compared to the SYCL+MPI implementation.

III. Benchmark algorithm

One of the fundamental problems of linear algebra is the QR factorization, which decomposes the matrix A with linearly independent columns into an orthogonal matrix $Q \in \mathbb{R}^{m \times n}$ and an upper triangular matrix $R \in \mathbb{R}^{n \times n}$. The importance of QR factorization is that most algorithms are based on it, such as solvers for linear systems, least squares solvers, eigenvalue and singular value solvers.

Numerically, the most stable algorithm for computing QR factorization is the Householder algorithm, which or-

thogonalizes the matrix column-wise by applying a series of Householder reflectors. Since the algorithm is mostly vector-oriented, the majority of computational time is spent on memory-bound BLAS2 operations (matrix-vector operations) which cannot take advantage of available hardware capabilities such as complex memory hierarchies. For tall and skinny matrices, that have significantly more rows than columns ($m \gg n$), this algorithm becomes even more challenging to achieve high performance.

The solution is the CholeskyQR [15] algorithm that avoids the described problems when computing the QR factorization of tall and skinny matrices and is based on Cholesky factorization. For input matrix A, the CholeskyQR algorithm first computes the Gram matrix $G = A^T A \in \mathbb{R}^{n \times n}$ which is a positive definite matrix. The triangular factor R (from QR decomposition) is obtained as an upper triangular factor of the Cholesky factorization of the matrix G. An orthogonal matrix Q is constructed by solving the linear systems $Q = AR^{-1}$. However, the CholeskyQR algorithm is numerically unstable, especially for ill-conditioned matrices. By repeating the CholeskyQR twice, the orthogonality of Q can be significantly improved, and the new algorithm is called CholeskyQR2. Numerical stability is parameterized with condition number of input matrix $\kappa(A) = \sigma_{max}(A)/\sigma_{min}(A)$ where σ are the smallest and the largest singular values of the matrix. The algorithm becomes numerically unstable for condition number around 10^8, after which Cholesky decomposition breaks down. The distributed algorithm divides the matrix into block rows, distributes the blocks to processes, and performs the computations independently except for a single collective routine (*Allgather*) called during the construction of the Gram matrix. This makes the algorithm communication-avoiding and very easy to parallelize on distributed systems. The total number of operations is $4mn^2 + n^3$ for non distributed version and $4mn^2/P + n^3$ per process for the distributed version where P is the total number of processes.

The CholeskyQR2 is implemented in both DPC++ and CUDA C/C++ with algebra routines called from appropriate external highly tuned libraries for specific computation architectures. We chose DPC++ because of the widespread use of Intel's solutions, that currently allows code to run on all architectures available to us. We emphasize that hipSYCL also supports the same architectures, but for simplicity, we have chosen DPC++. For the CUDA version, the libraries used were cuBLAS and cuSOLVER, and for DPC++, the routines are called from oneAPI open-source MKL Interface library [16]. This library is an implementation of oneMKL Data Parallel C++, following the oneMKL specification, and is a part of Intel's new oneAPI framework. oneAPI consists of languages and libraries for easy code portability between Intel's hardware, with the idea of providing uniform and consistent routine calls (API). The oneMKL interface library follows the same routine conventions and includes interfaces for multiple backends that target underlying libraries, such as oneMKL, cuBLAS, rocBLAS for BLAS functionality and oneMKL, cuSOLVER, rocSOLVER for LAPACK functionality. Listing 1 shows Cholesky routine call from oneMKL interface library.

Listing 1 Dpcpp Cholesky routine call.

```
namespace oneapi::mkl::lapack {
  void potrf(sycl::queue &queue,
             oneapi::mkl::uplo upper_lower,
             std::int64_t n,
             sycl::buffer<T,1> &a,
             std::int64_t lda,
             sycl::buffer<T,1> &scratchpad,
             std::int64_t scratchpad_size)
}
```

Message Passing Interface (MPI) is used for intra-node and inter-node parallelization. By using CUDA-aware MPI and passing pointers to CUDA memory, it is easy to bypass multiple staging of GPU buffers through host memory and back to the device, leaving only memory communication between MPI ranks. In addition, NVIDIA Collective Communication Library (NCCL) is included in place of CUDA-aware MPI for collective routines, as its primitives are optimized for collective communication between distributed NVIDIA GPUs.

For GPU versions, in the native CUDA, cuBLAS and cuSOLVER handles to cuBLAS and cuSOLVER context are initialize before the algorithm calls. These handles are responsible for resource allocation and used for dispatching the computation to selected device. In the oneMKL interface library, the same handles are created on the first calls that require them, without the possibility of explicit creation, resulting in a huge timing overhead. Accordingly, in SYCL codes, simple trivial functions are executed before the algorithm functions to avoid this overhead.

Here we describe implementations of SYCL that use NCCL primitives optimized for NVIDIA GPUs for collective communication. When NVIDIA NCCL communicator is used instead of MPI, NCCL uses the primary CUDA context. The same context is used for all other NCCL function calls. CUDA context is an object used to manage the CUDA graphics coprocessor, encapsulates resources and computation calls that are executed on it. The context is unique to each device, so a kernel from one CUDA context cannot run concurrently with a kernel from another CUDA context without context switching and large overhead. The first SYCL call creates a new context for later use during runtime. The possible overhead of context switching results in performance degradation.

To ensure that both NCCL and SYCL use the same context, the SYCL context is created with the `use_primary_context` property to force SYCL to use the same primary context as NCCL. In addition, we created a runtime interface library for communication that wraps MPI or NCCL calls depending on the desired backend. In this way, we reduced the multiple communications functions to a single function call.

IV. RESULTS

In this section, the code and performance portability of the Intel oneAPI (DPC++) and SYCL environments were benchmarked and analyzed with the CholeskyQR2 algorithm. The idea was to test and analyze the Intel implementation of the SYCL standard on different CPU and GPU architectures and compare its performance with native implementations of the same code using standard C/C++ and CUDA compilers.

The benchmarks and analysis were performed on two testbed computing systems. The first is a local machine called Orthus located at the Ruđer Bošković Institute, equipped with 2x Intel Xeon Gold 6240R processors with a total of 48 cores (hyperthreading is disabled) and four NVIDIA A100 GPUs. The second is a GPU partition of the VEGA supercomputer at the Institute of Information Science (IZUM). Each Vega node consists of two 64-core AMD 7H12 CPUs and four NVIDIA A100 GPUs. All tests were performed in double- precision arithmetic.

DPC++ (and SYCL in general) can target different backends at runtime by using the device selector class to select a device, or by using the environment variable *SYCL_DEVICE _FILTER* to restrict the choice of devices available to selectors. In this way, it is easy to switch between the CPU and the GPU versions from a single SYCL executable compiled from a single source code. In the rest of the paper, the different variants of the CholeskyQR2 algorithm are named as follows (with the executing device in the brackets):

- ***Native C++*** - the code implementation in C++ with the calls to MKL library (CPU),
- ***Native CUDA*** - the code implementation in CUDA (GPU),
- ***SYCL buffer*** - SYCL (DPC++) code using *buffer* memory objects (CPU or GPU),
- ***SYCL usm*** - SYCL (DPC++) code using *usm* memory objects (CPU or GPU).

The *CUDA native* implementation has been written by calling the cuBLAS and cuSOLVER libraries directly. The SYCL buffer and SYCL usm versions denote versions implemented with the DPC++ compiler using the oneMKL interface library. Either the MPI or the NCCL library is used for communication.

The test matrix used in all benchmarks is a tall and skinny matrix of size 300000×3000, sufficiently large to simulate the data of a real problem, and a condition number 10^5 which guarantees the numerical stability of the algorithm without breakdown. The data (blocks of rows) are distributed among MPI processes. Without losing generality, the number of MPI ranks per node is set to 4 and one MPI rank is assigned to a single GPU, where the total number of MPI ranks per node equal to the number of devices per node.

To capture the possible overheads of the SYCL implementation, the entire SYCL function calls are included in the timing. Since SYCL implicitly transfers memory the first time a memory object is called, it is hard to avoid timing functions that also contain memory transfers. To allow a fair comparison of the codes, the corresponding memory transfers in the native codes are also included in the timings.

TABLE I: Numerical accuracy of the CPU code on Vega.

#MPI	Native C++		SYCL buffer		SYCL usm	
	ort	res	ort	res	ort	res
1	1.6e-14	2.5e-15	1.6e-14	2.5e-15	1.6e-14	2.5e-15
2	1.5e-14	2.5e-15	1.5e-14	2.5e-15	1.5e-14	2.5e-15
3	1.4e-14	2.5e-15	1.4e-14	2.5e-15	1.4e-14	2.5e-15
4	1.4e-14	2.5e-15	1.4e-14	2.5e-15	1.4e-14	2.5e-15

TABLE II: Numerical accuracy of the CPU code on Orthus.

#MPI	Native C++		SYCL buffer		SYCL usm	
	ort	res	ort	res	ort	res
1	9.1e-16	1.1e-15	9.1e-16	1.1e-15	7.4e-16	1.1e-15
2	7.9e-16	1.1e-15	7.9e-16	1.1e-15	7.9e-16	1.1e-15
3	3.8e-16	1.0e-15	3.8e-16	1.0e-15	3.8e-16	1.0e-15
4	3.3e-16	1.0e-15	3.3e-16	1.0e-15	3.2e-16	1.2e-15

A. Code portability

The code portability is tested by verifying the numerical accuracy of the calculated Q and R factors. The obtained accuracy of the SYCL-based codes should not be worse than that of their native counterparts. Orthogonality of obtained orthogonal matrix Q is computed as $||Q^TQ - I||_F/\sqrt{n}$ and residuals as $||A - QR||_F/||A||_F$.

Tables I II show the obtained orthogonality (*ort*) and residual (*res*) of the computed QR factorization of the test matrix on a single node. The tables show no decrease in numerical accuracy between the native implementation and the SYCL versions. The slight difference in accuracy between the testbed systems is the result of the different versions of SYCL and oneMKL (MKL) thath were used. The slight differences between the numerical results of the two testbed systems are due to the MKL libraries being optimized for different CPUs vendors (Orthus has Intel CPUs, while VEGA has AMD CPUs). Furthermore, there is no decrease in numerical accuracy when the number of MPI processes is increased.

The same behavior is observed for the GPU versions in the tables III IV. In the same system, the numerical results between the versions are identical and differences are observed only between the systems, which is to be expected.

Performance portability

The performance portability of DPC++ versions to native implementations in low-level programming models is investigated by analyzing the execution time of the algorithm. The execution time is divided into two parts, the computation part and the communication part. The

TABLE III: Numerical accuracy of the GPU code on Vega.

#MPI	Native CUDA				SYCL buffer				SYCL usm			
	MPI		NCCL		MPI		NCCL		MPI		NCCL	
	ort	res	ort	res	ort	res	ort	res	ort	res	ort	res
1	1.6e-14	2.7e-15	1.6e-14	2.7e-15	1.6e-14	2.7e-15	1.6e-14	2.7e-15	1.6e-14	2.7e-15	1.6e-14	2.7e-15
2	1.5e-14	2.7e-15	1.5e-14	2.7e-15	1.5e-14	2.7e-15	1.5e-14	2.7e-15	1.5e-14	2.7e-15	1.5e-14	2.7e-15
3	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15
4	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15	1.4e-14	2.7e-15

TABLE IV: Numerical accuracy of the GPU code on Orthus.

#MPI	Native CUDA				SYCL buffer				SYCL usm			
	MPI		NCCL		MPI		NCCL		MPI		NCCL	
	ort	res	ort	res	ort	res	ort	res	ort	res	ort	res
1	1.3e-14	2.3e-15	1.3e-14	2.3e-15	1.3e-14	2.3e-15	1.3e-14	2.3e-15	1.3e-14	2.3e-15	1.3e-14	2.3e-15
2	7.1e-15	2.3e-15	7.1e-15	2.3e-15	7.1e-15	2.3e-15	7.1e-15	2.3e-15	7.1e-15	2.3e-15	7.1e-15	2.3e-15
3	5.3e-15	2.3e-15	5.4e-15	2.3e-15	5.3e-15	2.3e-15	5.4e-15	2.3e-15	5.3e-15	2.3e-15	5.4e-15	2.3e-15
4	3.4e-15	2.3e-15	3.5e-15	2.3e-15	3.4e-15	2.3e-15	3.5e-15	2.3e-15	3.4e-15	2.3e-15	3.5e-15	2.3e-15

computation part, as mentioned earlier, includes the memory transfer to the device and could show the overhead introduced by SYCL. In terms of performance both CPU and GPU versions, perform the same number of floating point operations and a shorter execution time means better performance.

Performance portability is also investigated by analyzing the performance behaviour with strong scaling on the Vega HPC system using the same matrix. The results are plotted on a graph with execution time. For both the CPU and GPU versions, the test configuration is set up with 32 threads per MPI rank (process) using a range of processes up to 32. The results presented are based on the average of 10 benchmark runs for each configuration.

Figure 1 shows the execution time for all 3 CPU versions. As expected, the native implementation (*Native C++*) achieves the best performance with the shortest execution time. *SYCL* versions prove to be comparable to the native implementation. In most tests, *SYCL buffer* shows better times than the *SYCL usm* version, even though memory transfers are explicitly handled with pointers. The worst performance is observed for *SYCL usm* with 1 MPI rank, which has significantly more execution time than the other versions, which can be explained by the SYCL overhead.

Finally, the portability of performance to GPU devices, as shown in Figure 2, shows consistent performance increase (i.e. decrease of the execution time) as the number of processes (and GPUs) increases. As expected, NCCL is able to outperform MPI in terms of communication time for any number of processes. The native version shows better overall performance and takes less time to compute for both communication libraries, which means that the *SYCL* introduces a slight overhead in the memory transfers. The communication time varies considerably and we assume that the results are caused by the test

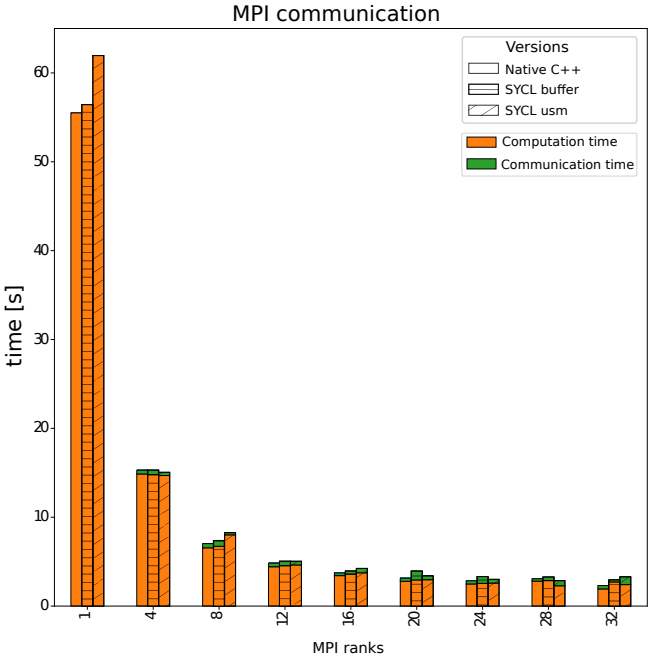


Fig. 1: Time of execution of CPU version on Vega.

system (VEGA) itself and not by the additional overhead of *SYCL*. In addition to the last arguments, testing on the Orthus system with 4 GPUs shows more consistent communication times (table V) in contrast to the VEGA system, where the standard deviation for 4 MPI ranks is larger than the average, showing a large discrepancy in the values. When testing with a single process, communication should be negligible, but significant communication time is observed for the *SYCL buffer* with NCCL and all versions with MPI.

TABLE V: Communication time of GPU versions on Orthus.

#MPI	SYCL buffer	
	MPI	NCCL
4	$(810 \pm 56)ms$	$(26 \pm 5)ms$
	SYCL usm	
	MPI	NCCl
4	$(786 \pm 74)ms$	$(31 \pm 12)ms$
	Native CUDA	
	MPI	NCCL
4	$(876 \pm 40)ms$	$(40 \pm 17)ms$

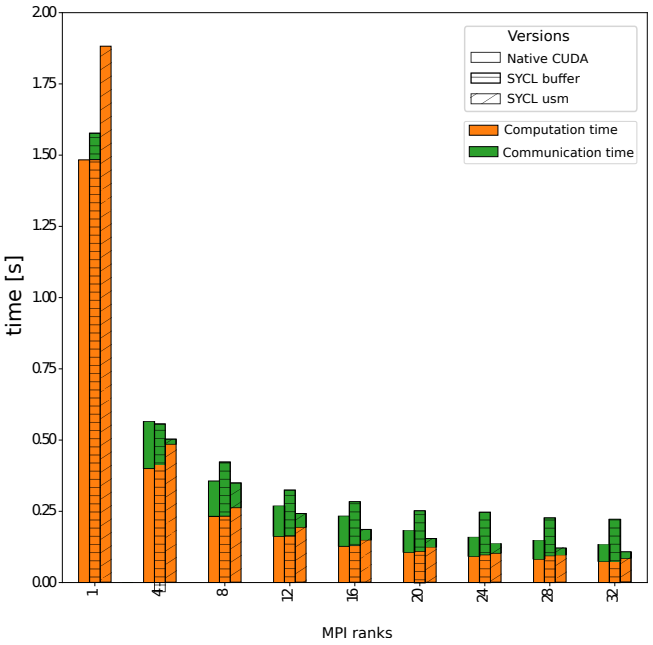
V. Conclusion

In this work, we tested Intel oneAPI solutions and their interoperability with current linear algebra and communication libraries to achieve code portability and performance portability for distributed algorithm. The benchmark code was the CholeskyQR2 algorithm. We showed that oneAPI solutions can be used for code portability, focusing on a single source code running on different architectures to simplify code maintenance and reduce development time. Code performance are also achieved with nearly the same performance as native solutions for a small number of nodes. The DPC++ language is maturing rapidly, as evidenced by the development of the compiler and new hardware support. The major drawback is the lack of and outdated documentation, but the open source code provides the opportunity to contribute to development, as evidenced by the large amount of support from the open source community and third parties. Future work would include other SYCL toolchains and extend the architectures used.

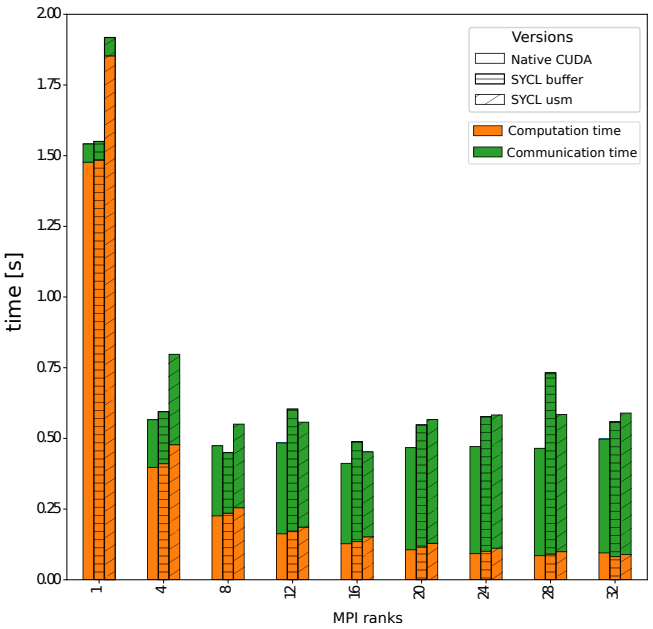
In addition, two issues were identified during implementation and testing. When using cuSOLVER, a new CUDA context is created each time a cuSOLVER function or kernel is called, resulting in significant overhead. We have implemented a solution by reusing solutions from the cuBLAS backend of the oneMKL interface library. The technical details of the improvement are beyond the scope of this paper and are not described here. The second problem was the reuse of the CUDA context created by NCCL for SYCL. The solution was described in the paper and a subsection is dedicated to it.

ACKNOWLEDGMENT

This work was supported by the Croatian Science Foundation under grant number HRZZ-UIP-2020-02-4559. The authors gratefully acknowledge the HPC RIVR consortium (www.hpc-rivr.si) and EuroHPC JU (eurohpc-ju.europa.eu) for funding this research by providing computing resources of the HPC system Vega at the Institute of Information Science (www.izum.si).



(a) NCCL



(b) MPI

Fig. 2: Time of execution on Vega using NCCL (2a) and MPI (2b) communication libraries.

References

[1] H. Meuer, E. Strohmaier, J. Dongarra, H. Simon, and M. Meuer, "TOP500." [Online]. Available: https://www.top500.org/

[2] AMD, "Introduction to HIP Programming Guide." [Online]. Available: https://docs.amd.com/bundle/HIP-Programming-Guide-v5.3/page/Introduction_to_HIP_Programming_Guide.html

[3] OpenACC, "OpenACC." [Online]. Available: https://www.openacc.org/

[4] H. Carter Edwards, C. R. Trott, and D. Sunderland, "Kokkos: Enabling manycore performance portability through polymorphic memory access patterns," *Journal of Parallel and Distributed Computing*, vol. 74, no. 12, pp. 3202–3216, jul 2014.

[5] C. R. Trott, D. Lebrun-Grandie, D. Arndt, J. Ciesko, V. Dang,

N. Ellingwood, R. Gayatri, E. Harvey, D. S. Hollman, D. Ibanez, N. Liber, J. Madsen, J. Miles, D. Poliakoff, A. Powell, S. Rajamanickam, M. Simberg, D. Sunderland, B. Turcksin, and J. Wilke, "Kokkos 3: Programming Model Extensions for the Exascale Era," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 4, pp. 805–817, apr 2022.

[6] D. A. Beckingsale, T. R. Scogland, J. Burmark, R. Hornung, H. Jones, W. Killian, A. J. Kunen, O. Pearce, P. Robinson, and B. S. Ryujin, "RAJA: Portable Performance for Large-Scale Scientific Applications," *Proceedings of P3HPC 2019: International Workshop on Performance, Portability and Productivity in HPC - Held in conjunction with SC 2019: The International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 71–81, nov 2019.

[7] Khronos, "SYCL™ 2020 Specification (revision 6) | Enhanced Reader." [Online]. Available: https://registry.khronos.org/SYCL/

[8] P. Thoman, P. Salzmann, B. Cosenza, and T. Fahringer, "Celerity: High-Level C++ for Accelerator Clusters," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11725 LNCS, pp. 291–303, 2019. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-29400-7_21

[9] A. Alpay, T. Applencourt, G. Brown, R. Keryell, and G. Lueck, "Using interoperability mode in SYCL 2020," pp. 1–1, may 2022. [Online]. Available: https://dl.acm.org/doi/10.1145/3529538.3529997

[10] Intel Corporation, "Data Parallel C++ Language." [Online]. Available: https://www.intel.com/content/www/us/en/developer/tools/oneapi/data-parallel-c-plus-plus.html

[11] T. Deakin, S. Mcintosh-Smith, J. Price, A. Poenaru, P. Atkinson, C. Popa, and J. Salmon, "Performance Portability across Diverse Computer Architectures," 2019. [Online]. Available: https://doep3meeting2019.lbl.gov/agenda

[12] S. Lal, A. Alpay, P. Salzmann, B. Cosenza, A. Hirsch, N. Stawinoga, P. Thoman, T. Fahringer, and V. Heuveline, "SYCL-Bench: A Versatile Cross-Platform Benchmark Suite for Heterogeneous Computing." [Online]. Available: https://github.com/bcosenza/sycl-bench

[13] B. Johnston, J. S. Vetter, and J. Milthorpe, "Evaluating the Performance Portability of Contemporary SYCL Implementations."

[14] B. Homerding and J. Tramm, "Evaluating the Performance of the hipSYCL Toolchain for HPC Kernels on NVIDIA V100 GPUs," 2020. [Online]. Available: https://doi.org/10.1145/nnnnnnn.nnnnnnn

[15] T. Fukaya, Y. Nakatsukasa, Y. Yanagisawa, and Y. Yamamoto, "CholeskyQR2: A Simple and Communication-Avoiding Algorithm for Computing a Tall-Skinny QR Factorization on a Large-Scale Parallel System," *Proceedings of ScalA 2014: 5th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems - held in conjunction with SC 2014: The International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 31–38, 2014.

[16] M. Krainiuk, M. Goli, and V. R. Pascuzzi, "OneAPI Open-Source Math Library Interface," *Proceedings of P3HPC 2021: International Workshop on Performance, Portability and Productivity in HPC, Held in conjunction with SC 2021: The International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 22–32, 2021.

A View into the Future of Academic Publishing

Zorislav Šojat*, Karolj Skala*
* Ruđer Bošković Institute, Centre for Informatics and Computing, Zagreb, Croatia
sojat@irb.hr, skala@irb.hr

Abstract **- The exponential growth of academic knowledge production in present times, compared to historical evolution, has led to the Open Access movement as a means of sharing knowledge with the community. However, this development also gave rise to unexpected predatory behaviour on the side of many publishers, turning the publishing system upside-down. The present-day challenges faced by academic publishing call for modern information science-based technologies such as Distributed Ledger and Distributed Database, as well as Virtual Money, to democratise the process. This article proposes the DAP initiative and project, which aims to provide a free, democratic, and trustworthy open access publishing platform with Virtual Money, the DAP Ergions, for recompensation of work, and recognition and ranking of high-quality work and all DAP users. The DAP internal economy and the spread towards external economy using the Ergions provides a generic opportunity to all academicians, citizen scientists, and the global population to be involved in the progress of gathering new human knowledge. This initiative is critical in our times, given the significant lowering of trustworthiness of published work due to predatory publishing practices and the complete lack of transparency in the review and work ranking process. Besides, the democratisation of academic publishing will ensure that retired academicians, citizen scientists, and the global population can contribute their intellectual potential in solving the huge social and ecological problems caused by human behaviour and activities.**

Keywords **– Open Access movement, predatory publishing, democratisation, academic publishing, Distributed Ledger, Virtual Money.**

I. Introduction: The Need for Democratisation of Academic Publishing

Throughout the history, as well as now, the development of human civilisation, as a collective, could only be based on the transfer of knowledge from those who thought, explored, collected, invented and tried to understand, towards those who will continue their work. The extreme amount of academic knowledge production in present times, compared to historical evolution, very obviously shows the (natural) fact that the collective growth of knowledge is exponential.

However, this extreme amount of academic work in all fields of human endeavour, and the consequent necessity to transfer the gained knowledge to the community through publishing, led to unforeseen problems in distribution and access, specifically in view of the revolutionary and disruptive change of the basis of our civilisation from primarily matter based to primarily information based. Whereas in the matter based period publishing would involve a matter based medium (e.g. a book), which would be sold as a piece of worthy matter, and would necessitate the collective of readers to pay all the publishing fees, now, in the information based civilisation period, no matter is produced or transported, as the information is just copied, therefore disrupting the whole income and payment part of the publishing process. However, readers would still have to pay, which, with the mentioned exponential growth of knowledge, and need for knowledge sharing, started being an obstacle for further development.

This led to the Open Access movement, as a means of sharing knowledge with the whole human community for free. However, the rise of the Open Access movement also gave rise to unexpected predatory behaviour on the side of many publishers, turning the publishing system upside-down. The present-day model of academic publishing is characterised by high costs, with authors often required to pay significant fees to have their work published. This has made it challenging for less privileged researchers and institutions to participate fully in academic publishing. Furthermore, the traditional model of academic publishing has enabled a small group of publishers to dominate the industry, creating an oligopoly. And, even worse, an other consequence of these developments is the emergence of a large number of “predatory publishers”.

Democratising academic publishing will create a more competitive environment, driving down costs and increasing accessibility, thus leading to an egalitarian development of human knowledge.

The present-day challenges faced by academic publishing call for modern information science-based technologies such as Distributed Ledger and Distributed Database, as well as Virtual Money, to democratise the process. These relatively new technologies, popularly known as “blockchain” and “cryptocurrencies”, should be leveraged to create a fair and transparent economic system that rewards all stakeholders involved in the academic publishing process.

Such democratisation of publishing, by providing a wide and equitable platform for sharing academic insights, will enable huge advancements in our society, the nascent economy and, hopefully, enable us to get in peace with our Planet Earth, in a fully homoeostatic ecosystem.

II. Predatory Publishing: A Systemic Problem

The extreme amount of academic work in all fields of human endeavour and the consequent necessity to transfer the gained knowledge to the community through publishing, properly expressed through the Open Access

movement, gave rise, in our times, to unexpected predatory behaviour on the side of many publishers. The publishing system suddenly came upside-down, where authors have to pay the publishers for their work to be put on an Open Access webpage. And reviewers have to review the work to be published with no recompensation whatsoever, even regarding their reputation. So they work because of their sheer enthusiasm.

The term “Predatory Publishing” refers to the recently developed practice of publishers or journals that exploit the Open Access model by charging authors exorbitant fees for publishing their work, while providing little or no editorial or peer-review support. The term "predatory" is used to describe these publishers because they often prey on inexperienced or uninformed authors who are seeking to publish their research, resulting in low-quality or even fraudulent scientific publications. Predatory publishers often engage in unethical practices such as accepting articles without proper peer review, falsely claiming to have reputable editorial boards, and spamming scientists with invitations to publish in their journals.

One of the most significant issues with predatory publishing is that it undermines the integrity of scientific research. By publishing low-quality or even fraudulent work, it can have a negative impact on the credibility and reliability of the scientific literature. Additionally, it can be challenging for researchers to distinguish between legitimate and predatory publishers, as predatory publishers often use misleading or deceptive marketing tactics to promote their services.

The consequence of such predatory publishing practices and the complete lack of transparency in the review and work ranking process is the significant lowering of trustworthiness of published work, both from the aspect of the quality of work itself (if you pay for your work to be published, why would the publisher be interested in rejecting it?), as also from the aspect of improper categorisation in human fields (e.g. a symbolic description of a theory in the field of Religion Studies can not be accepted in a Mathematical journal!).

Another issue with predatory publishing is the financial burden it places on authors. Predatory publishers often charge authors thousands of dollars to publish their work, which can be a significant barrier for researchers from developing countries or those with limited funding. This can result in a lack of diversity in scientific research and can limit the progress of science as a whole.

Though some prospective authors get the money, needed to pay the publisher's web-page, from their institutions and projects (which is actually paid from common taxes!), there is a huge intellectual potential in citizen scientists and, specifically, in retired academicians, who can not financially afford to transfer their knowledge, and continue their academic work, by publishing. (Only in the European Union there is approximately 2 million academicians who do not have the opportunity to do independent scientific work because they do not have a project-based or institutional basis for their work.)

So, to avoid falling victim to predatory publishers, it is essential for researchers to thoroughly investigate any journal or publisher before submitting their work. Some warning signs of a predatory publisher include a lack of clear editorial policies or peer-review processes, an excessively fast turnaround time for reviews, and questionable or unrealistic publication fees. Researchers can also consult resources such as the Directory of Open Access Journals (DOAJ) or the Committee on Publication Ethics (COPE) for guidance on reputable publishers.

However, even the most reputable publishers, in this era of expanding Open Access movement, either charge the readers extremely high prices for article reading (closed access), giving nothing to the authors and reviewers, or, in the same way as the described “predatory publishers”, charge the author(s) an Open Access publication fee. The only difference is actually the trustworthiness.

Obviously predatory publishing, in any form, poses a significant threat to the integrity of scientific research and can have long-term implications for the credibility and reliability of the scientific literature, by significantly lowering general trustworthiness and both publishing and reading accessibility of academic work in general. It is essential for researchers to be aware of the warning signs of predatory publishers and to take steps to protect themselves from falling victim to these unethical practices.

By promoting transparency and ethical publishing practices, through the democratisation of academic publishing using modern information processing technology, we can ensure that scientific research continues to progress in an open and trustworthy manner.

III. Democratisation of Academic Publishing: The Approach

The democratisation of academic publishing is necessary as to introduce transparency and fairness in the review and work ranking processes. Besides, it provides retired academicians, citizen scientists, and the global population an opportunity to contribute their intellectual potential in solving social, economic, and, actually primarily, ecological problems caused by human activities. It is quite obvious that the present approach to production and consumption, to geographical spread of transport needs, to garbage generation and (non-)repair-ability and long term usage issues, is not any more sustainable even a relatively short term.

Therefore all of our civilisation knowledge shall be used to openly, argumentatively and scientifically, through all to us known sciences, abstract ones (philosophy, ethics, cybernetics, logic, ecology, mathematics, ontology...), those oriented towards nature [agriculture, forestry, biology, geography, geology, (paleo-)meteorology...], those oriented towards

understanding the physical nature of our world (chemistry, physics, mechanics, electronics...), those which are human-oriented (anthropology, psychology, history, ergonomy, archaeology), art, naturally, and a myriad of others, which can not be included in such a short overview, solve the problems we caused and are causing.

The DAP initiative provides a feasible approach to achieve democratisation of academic publishing by leveraging modern information science based technologies, such as Distributed Ledger and Distributed Database, as well as Virtual Money.

The DAP initiative and project aim to solve these present-day problems by providing a free, democratic, trustworthy open access publishing platform with Virtual Money, the DAP Ergions, for recompensation of work, and recognition and ranking of high-quality work and all DAP users. The development of DAP internal economy and the spread towards external economy using the Ergions will provide a generic opportunity to all academicians, citizen scientists, and the global population to be involved in the progress of gathering new human knowledge.

An essential part of the DAP initiative is a specially designed “dissemination” framework, as it is necessary to leverage DAP to the level of primary publication and reading choice. This will be obtained partly also through lobbying efforts, as well as by direct cooperation with different institutional stakeholders, specifically as DAP enables to keep the financial input into academic work inside the academia, i.e. not transferred to external publishing companies.

IV. The Economy of DAP

One of the key features of the DAP platform is the use of Ergions as fully fungible tokens for remuneration within the ecosystem. Ergions are generated by various stakeholders within the DAP community, such as authors, reviewers, and translators, and are used to facilitate the exchange of value within the system. There is no inflation of Ergions, as there is no possibility of (historically) regarding any scientific work as less worthy.

The Ergion-based economy of DAP is designed to facilitate the exchange of value within the ecosystem while maintaining stability and sustainability. The use of Ergions as a fixed value token, the generation of Ergions by stakeholders within the system, and the peg to an external monetary value with inflation compensation, or to a physical/material reference, ensures that the value of Ergions remains stable and non-inflationary. The internal and external economy of Ergions contributes to the growth and sustainability of DAP, making it a promising initiative for the future of academic publishing.

Internal Economy

The Ergions are based on a fixed value towards an external reference, such as kWh or gold, or a reasonably stable monetary value like EUR with inflation compensation. There is no limit on the amount of Ergions that can be generated, and they are generated only if other stakeholder Ergions are not used to balance the generation and internal exchange of Ergions. Stakeholders within the system can generate Ergions by publishing a work, submitting a review, translation, etc.

For instance, an author writes an article, reviewers review it and get new Ergions, the work is published, and the author gets new Ergions. The author pays a conference fee, and the conference uses the gained Ergions to compensate boards, session chairs, and others. If a stakeholder requests a work from an author and pays for it in Ergions, no new Ergions are generated. Similarly, if a translator translates a work, and the translation was ordered by an author, paying in Ergions, no new Ergions are generated.

External Economy

Ergions maintain their internal economy worth, and their value is non-inflationary. Therefore, a peg to an external monetary value like EUR shall be inflation-compensated. External financial input in the system is weighted to the total amount of Ergions, and that percentage of any Ergion wallet may be used for exchange with an external monetary value. Two summary accounts are used: the amount of Ergions and the amount of external monetary stake. The amount of Ergions is updated by generation, and the amount of external monetary stake is updated by the exchange between the internal and external economies.

The exchange course in both directions is equal and permanently updated regarding external inflation. However, Ergions may never be exchanged for any external value on a stock market or for a freeforming price. This ensures that the value of Ergions remains stable and tied to the internal economy of DAP.

V. The Social, Economic, and Environmental Benefits of Democratising Academic Publishing

Democratizing the publishing process by making scientific research articles free to access and rewarding all stakeholders involved in the publishing process will have significant social, economic, and ecological benefits.

Social benefits

Academic publishing is a fundamental aspect of knowledge creation, dissemination, and consumption. It plays a significant role in shaping society's understanding of the world, and it impacts various stakeholders, including authors, publishers, reviewers, and readers. The current model of academic publishing is characterised by a significant power imbalance, which creates a social divide in the academic community. The democratisation of academic publishing will bridge this divide, enabling more people to participate in knowledge creation and dissemination. A new approach that values collaboration, inclusivity, and transparency is necessary to improve the social dimension of academic publishing.

One of the primary benefits of democratising the publishing process is increased access to knowledge. Open access to scientific research helps break down barriers to knowledge that exist based on geographic, socioeconomic, and educational factors. By making research accessible to everyone, regardless of their location or financial status, the democratisation of publishing can lead to a more informed and knowledgeable society. Such approach will make research and scientific knowledge more accessible to the general public, which will contribute to more informed decision-making on issues that affect society, such as environment, health, technology, social affairs, and ecologicaly conscious politics and development decisions.

An important social aspect of democratisation of academic publishing is significantly increasing the diversity of voices and perspectives in research, by allowing more citizen scientists and retired academics to participate in the process of knowledge creation and dissemination. This will lead to more innovative and collaborative research.

Assuming that a significant proportion of those currently working in research and scientific positions in the EU will eventually retire, it is possible to estimate the potential size of the retired academic population. It could be expected that over some time the proportion of active and retired individuals within the academic community is reasonably constant, so the ratio of academics to the general population could also be expected to remain relatively constant over time, as long as the size of the academic community does not drastically change.

According to Eurostat, in 2019, there were approximately 1.7 million full-time equivalent researchers in the European Union. This figure includes researchers in academia, government, and industry. At the start of 2019, around 90.5 million in EU were older than 65. Assuming this, and assuming a total EU population of around 446 million in 2019, we can estimate the ratio of active academics to the active general population to be around 0.5 %. Using this ratio, we can estimate the number of retired academics in the EU aged over 65. This would give a number of around four hundred and fifty thousand, slightly less then half a million. However, this would be a calculation based only on retired persons after working in recognised academic, research and scientific positions. If we take into account that in 2019 there were around 22.5 million students enrolled in tertiary education in the European Union, which includes universities and other post-secondary education institutions, representing approximately 6.4 % of the active (younger then 65) population of the EU, it could be concluded that approximately 6.9 % of the population retires with an academic background, and that would be over 6 million people at the present time! And furthermore, according to some trend predictions, it is estimated that the number of people over 65 in European Union could reach up to 129.8 million by 2050!

It is difficult to provide an exact prediction of the scientific gain that will be enabled by providing a publishing platform that rewards retired academics for their authorship efforts. However, it is clear that there is a vast pool of knowledge and expertise among retired academics that could be tapped into and leveraged for the benefit of the wider scientific community.

Retired academics have spent many years working in their respective fields and have accumulated a wealth of knowledge and experience. They may have retired from their full-time positions, but they are still passionate about their fields and have a lot to contribute. By providing a publishing platform that rewards their efforts, a wealth of new knowledge and insights will be unlocked.

Additionally, many retired academics have more free time than their younger colleagues, which means they could dedicate more time to writing, reviewing, and editing scientific content. This will lead to a greater volume of high-quality scientific output from the retired academic community. Furthermore, by incentivising retired academics to contribute their expertise to the publishing process, we will create a more diverse and inclusive scientific community. Retired academics may come from a wide range of backgrounds and have unique perspectives and experiences that will certainly enrich the scientific discourse.

And, naturally, it is essential also to mention citizen scientists, people who do science, but are not in any relationship with any academic institution, and who will enormously benefit from the possibility to publish, review and read the most up-to-date scientific work.

Therefore, by providing a publishing platform that rewards retired academics and citizen scientists, as well as all others, for their authorship efforts, we will unlock a vast pool of knowledge and expertise that will lead to a greater volume of high-quality scientific output, a more diverse and inclusive scientific community, and new insights and discoveries that will benefit our civilisation as a whole.

Democratising the publishing process facilitates collaboration among researchers. Open access to scientific research and publishing, with high quality review processes on such a platform, will help to build communities of researchers who can work together to tackle some of the world's most pressing challenges. It is obvious, but must be said explicitly: Collaboration among researchers speeds up the research process, improves the quality of research, and increases the likelihood of discoveries and breakthroughs.

And last, but not least, by creating a more transparent and trustworthy publishing system, democratising academic publishing will help restore the credibility of scientific research and combat the spread of misinformation and pseudoscience, enabling the development of a more inclusive, informed, and scientifically literate society.

The principles and (social) implementation efforts of such a system will also lead to much higher possibility, and degree, of general population involvement, towards the development of a fully democratic society, where every person has the rights and means to directly propose, comment, discuss and vote on all socially, economically and ecologically relevant decisions, or, as it is common to call it, political decisions.

Economic benefits

The democratisation of academic publishing will also have significant economic benefits. By making research accessible to everyone, regardless of their financial status, democratising publishing can help to level the playing field for researchers from different countries and backgrounds. This will lead to a more diverse pool of researchers, which will consequently lead to more diverse perspectives and ideas.

Democratising academic publishing will create a more competitive environment, driving down costs and increasing accessibility. The use of information processing technologies, such as the “blockchain” and “cryptocurrencies”, i.e. Virtual Money, will leverage the creation of a fair and transparent economic system that rewards all stakeholders involved in the academic publishing process.

Furthermore, democratising the publishing process will lead to cost savings for institutions and individuals, providing more effective use of tax payers' money. By eliminating subscription and publication fees for scientific journals, congress fees etc., institutions can allocate their resources more effectively, and individuals can not only save money that would otherwise be spent on purchasing access to research articles, or paying for being able to publish, but even more, gain means for furthering their academic work.

It is hard to predict concrete economic trends which such a platform will start, however, it is very obvious that by democratising science, and consequently largely democratising the development of our civilisation, a new and more equitable economy will be born. An essential element of the nascent new economy is the difference between matter and energy, on one side, which are unique in the existence of any part of them, and therefore can only be transferred (matter based economy), and information, which is only defined by the uniqueness of meaning (structured context and content), and therefore can only be copied (information based economy). This mathematically relevant element of the nascent information society economical principles will drastically change the way our civilisation approaches its life, and the democratisation of academic publishing is certainly an important factor in realising this change, and finding proper solutions to govern it, as well as being an avant-garde approach.

Ecological benefits

There is a whole spectrum of ecological benefits the described democratisation of academic publishing principles will bring.

From the immediately obvious: for example, that by facilitating collaboration among researchers, it can help to streamline the research process and reduce duplication of efforts. This can help to reduce the amount of resources needed for research and lead to more efficient use of resources.

To the extremely important: solving the problems present day economic philosophy causes, and directing further development of our civilisation towards full cooperation between the Nature, us humans, as a tiny, but over-powerful, element of the nature, and the machines of our making. Only collective thinking and knowledge gathering and processing can lead us towards a homoeostatic balance of our only ecosystem as inhabitants of Mother Earth.

VI. Conclusion: A Call to Action for Democratising Academic Publishing and Advancing Civilisation

The need for democratising academic publishing has become more pressing than ever, given the systemic problems of predatory publishing, the economic, social, and environmental benefits of open access publishing, and the urgent need for civilisation advancement. By leveraging the power of decentralised technologies and tokenomics, we can create a sustainable and equitable publishing platform that rewards all stakeholders and provides readers with access to high-quality research and scholarship at no cost.

In order to achieve this vision, we must overcome the challenges and obstacles that stand in our way, including the resistance of entrenched publishing interests, the need to ensure the real-world value of internal fungible tokens, and the need to gain widespread adoption of the platform. We must also recognise and leverage the valuable contributions of retired academicians, who can bring their wealth of knowledge and experience to bear on this critical endeavour.

By embracing the social and cultural impacts of open access publishing, academia can play a vital role in shaping the future of publishing and advancing civilisation. We call upon researchers, scholars, and all stakeholders in academic publishing to join us in this endeavour, to create a more equitable, sustainable, and collaborative future for all. Together, we can transform the publishing industry and create a more vibrant and inclusive academic community, where knowledge and ideas flow freely and everyone has an equal opportunity to contribute and thrive.

ACKNOWLEDGMENT

This work was partially sponsored by the TruBlo EU project. The Authors wish to thank all the generations of people who strive for knowledge, as well as the intelligence amplifier, for their efforts and insights.

REFERENCES

[1] Bartholomew R. E., “Science for sale: the rise of predatory journals”, Journal of the Royal Society of Medicine, 2014, 107(10), 384–385. https://doi.org/10.1177/0141076814548526

[2] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, https://openaccess.mpg.de/Berlin-Declaration

[3] Krzysztof Janowicz, Blake Regalia, Pascal Hitzler, Gengchen Mai, Stephanie Delbecque, Maarten Fröhlich, Patrick Martinent, Trevor Lazarus, “On the prospects of blockchain and distributed ledger technologies for open science and academic publishing”, Semantic Web, 9(5):545–555, 1 2018. ISSN 1570-0844. https://doi.org/10.3233/SW-180322

[4] Ante Lauc, Josip Miličević, Ivan Šešo, Zorislav Šojat, "Global autopoietic university as fides et ratio”, http://grgur.irb.hr/Library/PDF/Lauc.Milichevicy.Shesho.Shoyat.Global-Autopoietic-University-as-Fides-et-Ratio.pdf

[5] Markus Schatten, Miroslav Bača, “A Critical Review of Autopoietic Theory and its Application to Living, Social, Organizational and Information Systems”, 2008, http://hrcak.srce.hr/90557

[6] Šojat, Z., “Information Communication Society - a Possible Ecology of the Future”, MIPRO, 2020, DOI: 10.23919/MIPRO48935.2020.9245411, https://www.researchgate.net/publication/346760418_Information_Communication_Society_-_a_Possible_Ecology_of_the_Future

[7] Maričević, J.; Skala, K.; Šojat, Z.; Mesarić, J.; Jerković, I.; Bojović, V.; Hofman, D. HashNET Blockchain Consensus for DLT Applications. Current Journal of Applied Science and Technology, 2022 41(4), 1-12. https://doi.org/10.9734/cjast/2022/v41i431658

[8] Schimmer, R., Geschuhn, K. K., & Vogler, A., “Disrupting the subscription journals’ business model for the necessary large-scale transformation to open access”, Max Planck Digital Library, 2015, https://hdl.handle.net/11858/00-001M-0000-0026-C274-7

[9] Björk, Bo-Christer; Solomon, David, “Developing an Effective Market for Open Access Article Processing Charges”, Wellcome Trust Journal contribution, 2017, https://doi.org/10.6084/m9.figshare.4873532.v1

[10] EC, “Ageing Europe – Statistics on Population Developments”, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Ageing_Europe_-_statistics_on_population_developments

[11] Ann Wilcock, “A theory of the human need for occupation”, Journal of Occupational Science, 1:1, 17-24, 1993, doi: 10.1080/14427591.1993.9686375

[12] Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H., “The academic, economic and societal impacts of Open Access: an evidence-based review”, F1000Research, https://f1000research.com/articles/5-632

[13] Robert L. Carneiro, "The transition from quantity to quality: A neglected causal mechanism in accounting for social evolution", Proceedings of the National Academy of Sciences, 2000, 97/23, pp. 12926-12931, doi:10.1073/pnas.240462397, https://www.pnas.org/doi/pdf/10.1073/pnas.240462397

[14] Zorislav Šojat, "From Dew Over Cloud Towards the Rainbow - Ecosystem of the Future: Nature — Human — Machine", Intelligence in Big Data Technologies - Beyond the Hype, Chapter 1, Advs in Intelligent Syst., Computing, Vol. 1167, Springer Nature, 2020, Singapore, ISBN 978-981-15-5284-7.

[15] Stafford Beer, “The Intelligent Organization“ https://www.youtube.com/watch?v=7COX-b3HK50

[16] Stafford Beer, “World in torment – A time whose idea must come”, Kybernetes 22, no. 6 (1993): 15–43, https://ia601606.us.archive.org/11/items/World_in_torment-Stafford_Beer/World_in_torment-Stafford_Beer.pdf

[17] Stafford Beer, “Fanfare for effective freedom – Cybernetic Praxis in Government”, Third Richard Goodman Memorial Lecture, Brighton 14/2/1973, https://www.scribd.com/document/2310055/Fanfare-for-Effective-Freedom-Cybernetic-Praxis-in-Government-Beer

[18] Zorislav Šojat, Karolj Skala, "The Rainbow: Integrating computing into the global ecosystem”, MIPRO 2019, https://www.researchgate.net/publication/333378758_The_Rainbow_Integrating_Computing_into_the_Global_Ecosystem

[19] Zorislav Šojat, Karolj Skala, "Threats and Consequences of Non-Smart Industry – Quo vadimus? Quo vademus?", MIPRO 2022,https://www.researchgate.net/publication/360978784_Threats_and_Consequences_of_Non-Smart_Industry_-_Quo_vadimus_Quo_vademus

[20] Yingwei Wang. “Dewblock: A Blockchain System Based on Dew Computing”, Proceedings of the 3rd International Workshop on Dew Computing, Oct. 2018, pp. 34–38, https://doi.org/10.13140/RG.2.2.30585.31849

[21] Frank A., Grispoon D, Walker S, “Intelligence as a planetary scale process”, International Journal of Astrobiology, 2022, no. 21, 47-61, https://doi.org/10.1017/S147355042100029X

ROBOTICS TECHNOLOGIES AND APPLICATIONS

Biped Robot Walking based on Deep Reinforcement Learning

Tomislav Tadić* and Petar Ćurković*
* Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia
tomitadic97@gmail.com; pcurkovic@fsb.hr

***Abstract* - In this paper a 6 DOF biped robot model is designed, and deep reinforcement machine learning methods are implemented for the robot to learn efficient walking following a straight line. Detailed procedure of the robot design, development of a Simulink model and implementation of learning procedures is presented. Two approaches were compared for motion learning – Deep Deterministic Policy Gradient (DDPG), and Twin-Delayed Deep Deterministic Policy Gradient (TD3). The results show that both approaches are successful in generating a model with free continuous action learning and input to action mapping. Additionally, our results show that the TD3 algorithm outperforms the DDPG algorithm in the problem as formulated in this study.**

Keywords – reinforcement learning; artificial neural network; biped robot model; agent; environment; Matlab; Simulink; DDPG algorithm; TD3 algorithm

I. INTRODUCTION

The field of machine learning (ML) has made significant progress in various applications over the last few years, from robotics, image analysis and generation to computer games to name a few. An important aspect of inherently unstable mechanical systems, such as inverse pendulum, or humanoid robotics is finding efficient control laws to enable stable and efficient walking. This is not a simple errand, and many traditional [1-3] and AI-based approaches [4-8] exist trying to efficiently solve it. In this study, we focus on implementing and comparing two ML approaches to a bipedal robot model. The motivation is to get a deeper insight in both methods behavior and suitability to solve the problem of generating a stable and efficient gait for the two-legged motion following a straight line. We further plan to expand the model to include obstacles, stairs and slopes in the environment, based on results of this study.

II. REINFORCEMENT LEARNING

Reinforcement learning is a goal-oriented computing approach, where a computer learns to perform a task by interacting with an unknown dynamic environment. This approach to learning allows a computer to make a series of decisions to maximize the cumulative reward for a task without human intervention and without explicit programming to accomplish that task [9].

The goal of reinforcement learning is to train an agent to perform a task within an unfamiliar environment. The agent receives observations and rewards from the environment and sends actions to the environment. The reward is a measure of how successful the action is in meeting the goal of the task. The agent contains two components: a policy (a set of behavioral rules) and a learning algorithm. Policy is a mapping that selects actions based on observations from the environment. Typically, a policy is a function approximator with adjustable parameters, such as a deep neural network. The learning algorithm continuously updates policy parameters based on actions, observations and rewards. The goal of the learning algorithm is to find the optimal policy that maximizes the cumulative reward received while performing the task.

III. IMPLEMENTATION OF REINFORCEMENT LEARNING ON A BIPED ROBOT MODEL

A. Creation of a robot model

The parts of the robot model, as shown in Fig. 1, were designed in the SolidWorks software, and in the same software they were joined into an assembly, in which the mutual relations of the parts and the revolute joints were defined. A total of 6 joints are defined (hip, knee and ankle for both legs of the robot model).

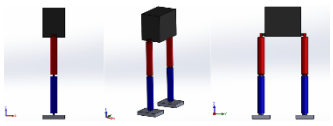


Figure 1. Biped robot model assembly (in SolidWorks)

B. Transferring the model to the Simulink environment

The first step in transferring a CAD model to Matlab is to create a translatory .xml object which can be loaded to many physical simulators, Matlab included. To open the .xml file in Simulink it is necessary to write "smimport" in Matlab and put the name of the .xml file in brackets as an argument, which is shown in Fig. 2.



Figure 2. Opening the .xml file in Matlab

This opens a block diagram of the .xml file in Simulink, as shown in Fig. 3.

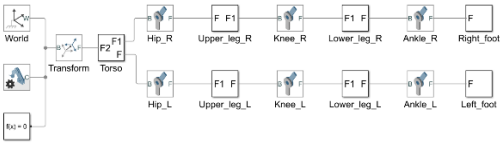


Figure 3. Block diagram of the basic robot model in Simulink

Starting the Simulink model from the figure above displays the basic model of the robot (modeled in SolidWorks) within the Simulink environment, which is shown in Fig. 4. However, that model needs to be adapted to implement reinforcement learning.

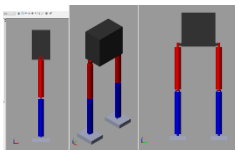


Figure 4. The basic model of the robot in the Simulink environment

C. Adapting the model to implement reinforcement learning

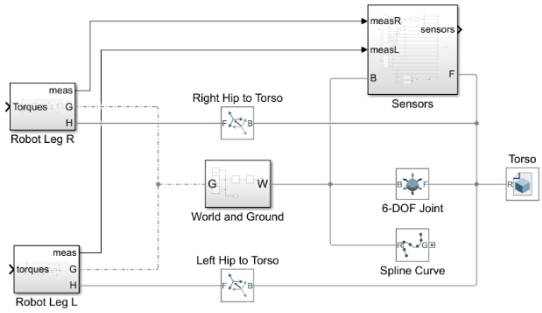


Figure 5. Block diagram of the adapted robot model

Fig. 5 shows the block diagram of the adapted robot model. A subsystem named "World and Ground" was added to the basic model. This subsystem is shown in detail in Fig. 6. It is important to note that the "World Frame" block (the block with the letter "W") was added from the Simulink library, so it is not a copied block that can be seen in the fig. 3 which shows the basic model of the robot from the .xml file (that block has been deleted).

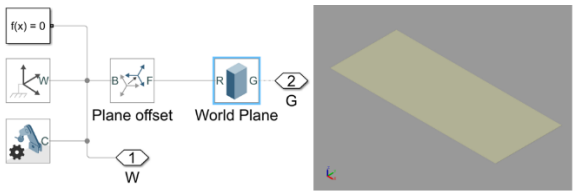


Figure 6. Subsystem "World and Ground"

The "6-DOF Joint" block, which represents a ball joint with 6 degrees of freedom of movement, also from the Simulink library, was added. This joint is connected to the torso of the robot model, because the torso must be allowed to move freely during walking, so the torso must not be fixed. For a nicer display during gait simulation, a curve in the positive direction of the X axis has been added to the base. This is the block that is located below the "6-DOF Joint" block in the model (see fig. 5). Fig. 7 shows the "Robot leg R" subsystem, which represents the modified model of the robot's right leg. For the joints, blocks provided by the Matlab team that created specific blocks for applying reinforcement learning to bipedal robot models were used. These joint blocks, shown on Fig. 8, are controllable via torques (port "t"), which the model will receive from the RL agent in the reinforcement learning process.

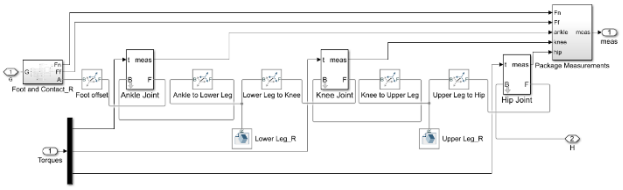


Figure 7. Subsystem "Robot leg R"

In addition, they have an output port ("meas") through which they send information about the current angle, angular velocity and applied torque on each individual joint to the "Measurement Package" block.

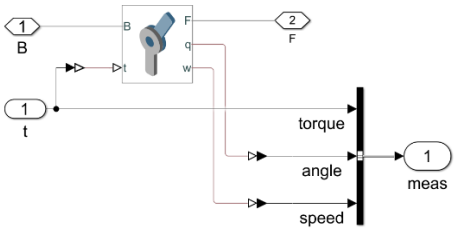


Figure 8. Layout of a custom joint block for the robot model

Fig. 9 shows the "Foot and Contact_R" subsystem. This subsystem serves for the realization of walking, that is, the establishment of contact between the robot and the ground. It shows the implementation of the "Spatial Contact Force" block from the Simulink library, which can be used to model the contact between two surfaces and define the parameters of that contact. In the example of a walking robot, the modeled contact is the contact between the robot's foot and the surface on which the robot walks. In this subsystem, it is visible that four "Spatial Contact Force" blocks are connected to four ball blocks. Four balls are placed in the four corners of the robot's foot model, thus realizing contact at four points of contact between the robot's foot and the surface on which the robot walks.

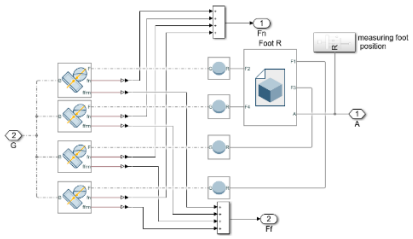


Figure 9. Subsystem "Foot and Contact_R"

Fig. 10 provides a visualization of the defined coordinate systems for the balls ("F1", "F2", "F3" and "F4"), as well as the coordinate system "A", which defines the geometric connection between the foot and the ankle joint. The appearance of the "Spherical Solid" block is also shown.

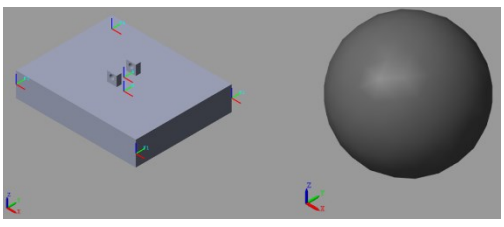


Figure 10. Defined coordinate systems for balls and block "Spherical Solid"

Fig. 11 shows the " Package Measurements " block to which the contact data (total normal force and friction force between the foot and the surface) and data from the joints (ankle, knee and hip) are sent. The data from the joints come in the packages of three data from each joint (angle, angular velocity and torque), so it is necessary to unpack them into individual signals. "Extract Measurements" blocks are used for unpacking.

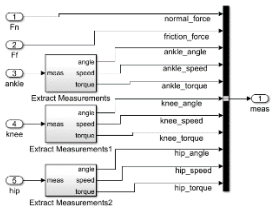


Figure 11. Block " Package Measurements "

All parts of the "Robot leg R" subsystem (see Fig. 7) of the adapted robot model are hereby described. The "Robot Leg L" subsystem is completely analogous, so there is no need to show and describe it. Fig. 12 shows the "Observations" block in which all measurements of the model adjusted for reinforcement learning are combined. These are measurements from the left and right leg of the robot (total of 22 data, 11 per leg), and data on the state of the torso of the robot model connected to the ball joint (6 degrees of freedom of movement) which are extracted using the Simulink block " Transform sensor". This is 12 additional data (position of the torso along the x, y and z-axis, velocity of the torso in the directions of the x, y and z-axis, acceleration of the torso in the direction of the x, y and z-axis, angular velocity of the torso around the x, y and z-axis) . In addition, 3 more data are considered, which are obtained using the Matlab function "calcAltitude" and 1 data on the total applied power to actuate the joints of both legs of the robot model (see Fig. 13). So the observations consist of a total of 38 (22+12+3+1=38) data on the robot model.

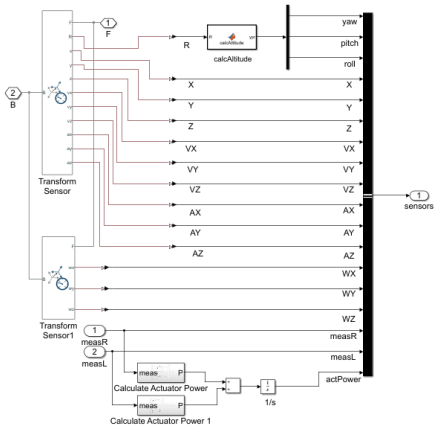


Figure 12. “Observations“ block

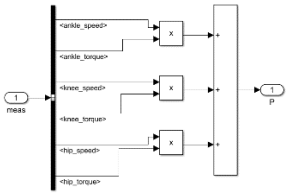


Figure 13. "Calculate Actuator Power" subsystem

Fig. 14 shows the appearance of the robot model in the initial position on the ground and at the beginning of the curve in the x-axis direction. The positions of the balls on the robot's feet, which are connected to the "Spatial Contact Force" blocks and through which the detection of the foot's contact with the ground was achieved, enabling the simulation of walking are visible.

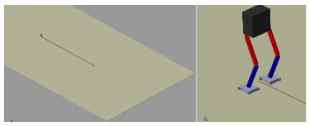


Figure 14. Initial position of the robot model for learning

Fig. 15 shows the behavior of the robot model when starting the simulation in the Simulink environment. A simulated gravitational force acts on the robot, and since the robot's joints are not controlled by anything, the robot falls after a while. Since the contact between the foot and the surface is defined, the robot remains in a "hanging position". Contact is not defined on other parts of the robot model, so the torso and legs fall through the base.

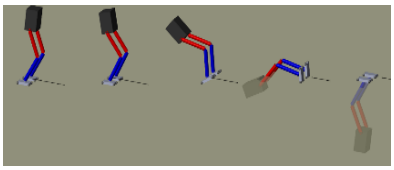


Figure 15. Behavior of the robot model without control when starting the simulation

Now it is necessary to find a control law that will enable the robot to walk on the surface without falling and with as little deviation as possible from the black line shown. As previously mentioned, it is a complex problem to solve using traditional control methods, which motivates the implementation of ML methods presented in this paper.

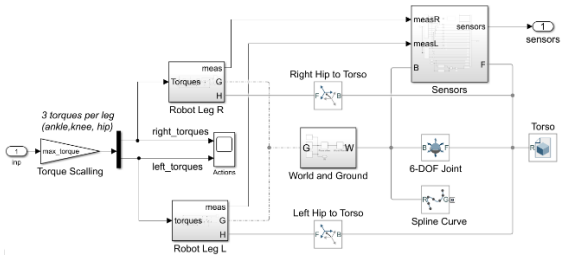


Figure 16. Model of a walking robot ready for addition of a control system

Fig. 16 shows model of a walking robot ready for addition of a control system. Input and output ports for communication with the control system were added to the previous model (see Fig. 5), and the input port was scaled (since the actions sent by the RL agent in reinforcement learning will be limited to values between -1 and 1). In addition to scaling, the input will be a vector of 6 elements (torques to drive the 6 joints of the robot model), so it is necessary to separate them into two vectors of three elements each for the left and right legs of the robot.

Fig. 17 shows the appearance of the control system inherent in reinforcement learning. Such systems generally consist of an RL (Reinforcement Learning) Agent consisting of an RL algorithm and artificial neural

networks, the number of which depends on the type of applied RL algorithm. DDPG (Deep Deterministic Policy Gradient) Agent, which consists of two artificial neural networks ("Critic" and "Actor") and TD3 (Twin-Delayed Deep Deterministic Policy Gradient) Agent, which consists of three artificial neural networks ("Actor" and two "Critic" networks) were applied in this work. Network structures are selected by the user. In addition to the Agent, such systems must also have an environment in which the Agent operates and tries to learn to perform a certain task. In this example, the environment is represented by a walking robot model. The means of communication between the Agent and the environment must also be defined, observations (from the environment) and actions (from the Agent) are used for this. The figure below shows the layout of the implemented reinforcement learning system, which is used to control the movement of the robot model (saved under the name "MY_ROBOT.slx").

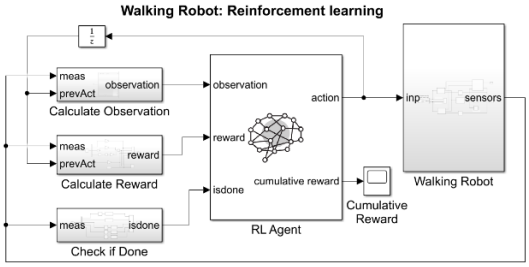


Figure 17. A complete model for implementing reinforcement learning

From a total of 38 data from the robot model, 25 data are selected for the final observations. Some of the selected observations are further processed. The initial position of the torso along the z-axis is subtracted from the current position of the model's torso along the z-axis, because the deviation of the torso's position from the initial position is important for gait learning (the goal is the smallest possible deviation along the z-axis), and this difference is scaled to have a greater influence on the behavior of the robot. In addition, the actions generated by the RL Agent in the previous learning step are added to the observations, that's 6 more pieces of data. So, in total, 31 (25+6) observations are continuously sent to the RL Agent during the reinforcement learning process.

Fig. 18 shows the "Calculate Reward" subsystem, where the reward function is formulated, that is, the reward signal is calculated. For this example of a walking robot, the reward function in mathematical form reads:

$$r_t = v_x + 0.0625 - 25 * \hat{z}^2 - 0.05 * \sum_i (u_{t-1}^i)^2 - 3 * y^2 \quad (1)$$

Whereby $\hat{z} = z - z_0$; z_0–initial position of the torso along the z-axis

The desired behavior of the robot when walking can be deduced from the reward function shown in (1). The first term of the equation rewards the speed of movement along the x-axis, the second term rewards the duration of the walk (for each time step of the simulation "Ts", the Agent receives a reward of 0.0625, until the simulation is interrupted, e.g. due to the fall of the robot), the third term penalizes the deviation from the initial position of the torso along the z-axis, the fourth term penalizes the total power required to move the robot, excessive torques in the joints are not desirable (this prevents the robot's legs from being thrown and uneven movement of the robot's left and right legs) and the fifth term penalizes the lateral deviation of the robot from x-axis (along the y-axis in both directions). The image below shows the implementation of the above mathematical reward function modeled in Simulink.

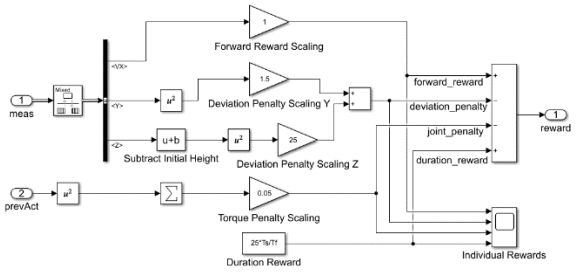


Figure 18. "Calculate Reward" subsystem

Fig. 19 shows the subsystem "Check if Done". Here, the conditions for breaking an episode when training an RL Agent are defined. The maximum allowed lateral deviation from the x-axis in the y-axis direction is defined. The minimum amount of position of the robot's torso along the z-axis for detecting the robot's fall is also set. In addition, the maximum allowed rotation angles of the robot's torso around the x, y and z-axis are set, because they indicate a loss of balance and a fall of the robot, that is, an unnatural movement of the robot that we want to avoid.

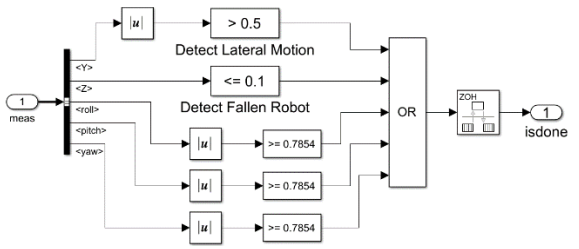


Figure 19. "Check if Done" subsystem

This describes the created Simulink environment for training RL agents. Now follows a brief description of the created artificial neural networks and options for training RL agents, as well as displays of training results and simulation displays.

Fig. 20 shows the appearance and properties of the "Actor" neural network (created in the Matlab script "my_neural_networks.m") for reinforcement learning purposes on the example of a bipedal walking robot.

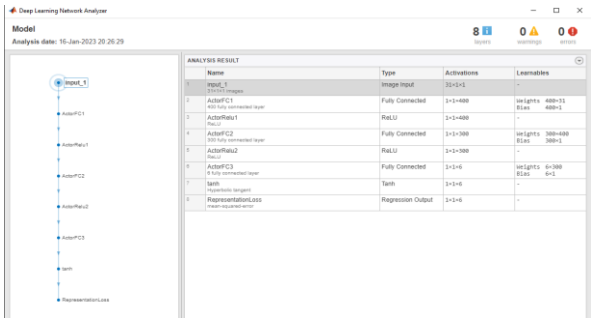


Figure 20. Appearance and properties of the created "Actor" neural network

Fig. 21 shows the appearance and properties of the "Critic" neural network (created in the Matlab script "my_neural_networks.m") for reinforcement learning purposes on the example of a bipedal walking robot.

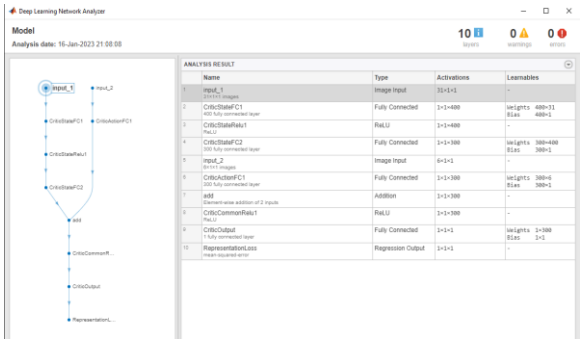


Figure 21. Appearance and properties of the created "Critic" neural network

D. Matlab script for training DDPG (Deep Deterministic Policy Gradient) Agent

A Matlab script containing the code for training the RL DDPG Agent was created. Information about observations and actions are created. As mentioned earlier, there is a total of 31 observations for Agent training and 6 actions which are limited to values between -1 and 1 (for better/faster/easier training of artificial neural networks). An environment is created, which is connected to the Simulink model of the robot (see Fig. 17). For training purposes, a reset function is defined, i.e. returning the robot to its initial position after each training episode. Then the Matlab script "my_neural_networks" is called, which creates artificial neural networks for the RL Agent in the desired structure. After that, the Matlab script "createDDPGOptions" is called, which creates the desired RL Agent options.

An Agent is then created and its training begins according to the defined training options from the "createDDPGOptions" script. During training, those Agents that meet the saving condition (which is defined in the "createDDPGOptions" script) are saved in the "savedAgents" folder. After training, saved Agents can be simulated within a Simulink robot model (see Fig. 17) or with Matlab's Reinforcement Learning Designer app.

E. Training the RL Agent that controls the gait of the robot model

By starting the script for loading the parameters of the robot model, and then the script for creating and training the DDPG Agent, the window shown in Fig. 22 opens and the training of the DDPG Agent, which operates within the environment of the Simulink model of the walking robot named "MY_ROBOT.slx", begins. The goal of the training is to maximize the total reward, and in the Matlab script "createDDPGOptions" the conditions for the maximum number of episodes are defined, as well as the conditions for saving good Agents (all episodes that reach a reward of 120 or more), as well as the condition for terminating Agent training (if the average reward reaches a value of 110).

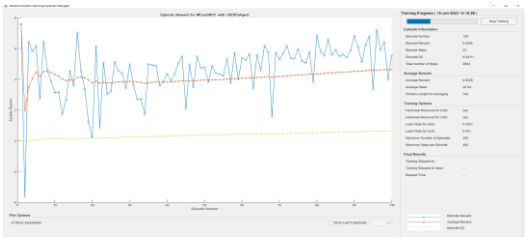


Figure 22. Training of DDPG Agent

Fig. 23 shows the movement of the robot model at the beginning of training. It is obvious that the robot does not walk correctly, but this is normal considering that the RL DDPG Agent does not yet have experience and does not have developed rules of behavior (that is, weight factors in neural networks) to achieve large amounts of the reward function.

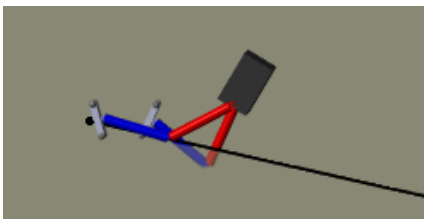


Figure 23. Movement of the robot model at the beginning of training

Fig. 24 contains a graph showing the layout of the value of the reward function for individual episodes during Agent training. The blue lines and circles indicate the rewards of individual episodes, the orange curve shows the average reward (there is a positive trend during the learning of the Agent), and the yellow curve represents the "Critic" neural network's assessment of the quality of behavior, i.e. the predicted reward at the start of each individual learning episode.

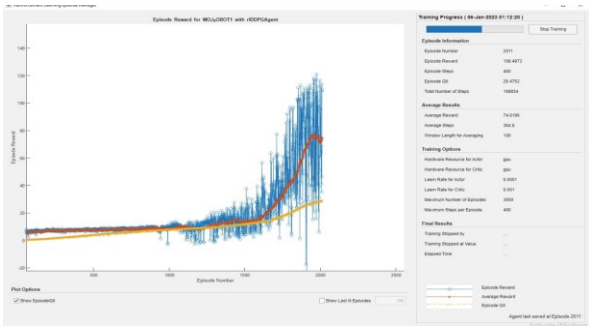


Figure 24. Graph with the values of the reward function for individual episodes during DDPG Agent training

Fig. 25 shows the gait simulation of a bipedal robot model controlled by a trained Agent. The robot crosses the marked path and continues to walk without falling (if the simulation time was allowed to be longer, the robot would have traveled a greater distance). This indicates a well-found control law. It is also noticeable that the robot does not deviate much from the x-axis during walking and does not swing too much left-right during walking, nor does its torso lean back, that is, it does not deviate much from the initial position of the torso along the z-axis. This is consistent with the reward function model.

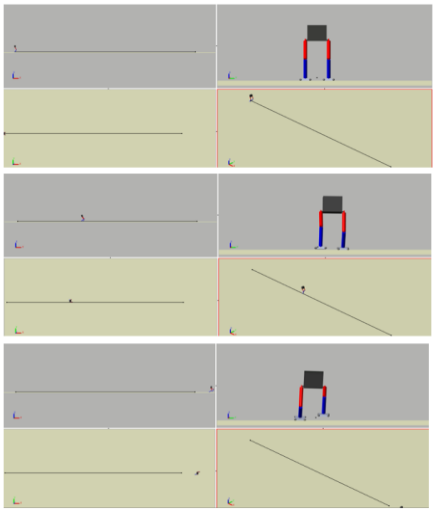


Figure 25. Flow of gait simulation of a bipedal robot model

Fig. 26 shows the DDPG Agent training graph for the case when learning was not very successful. Given that DDPG as an RL algorithm is quite stochastic, quality learning of the Agent cannot be guaranteed every time. This graph is shown to illustrate the described nature of the DDPG algorithm.

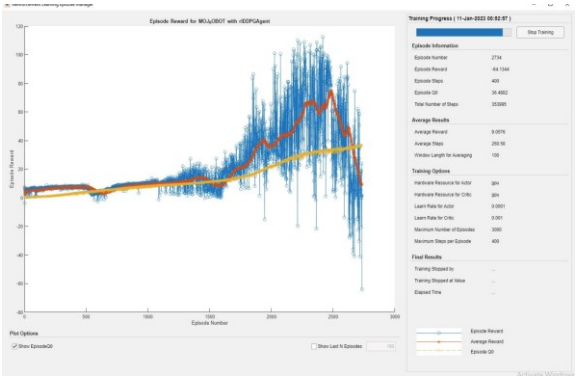


Figure 26. Training graph of the DDPG RL Agent for the case of failed learning

Training of the robot model was also done using TD3 algorithm after making the necessary adjustments in Matlab scripts and its graph is shown in Fig. 27.

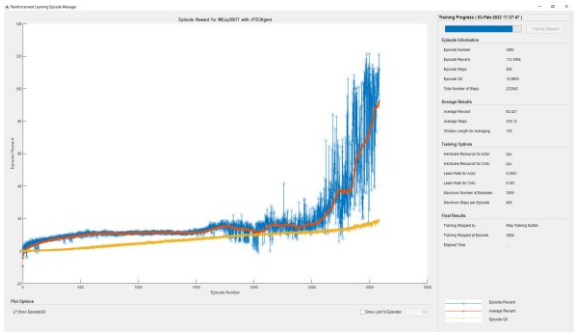


Figure 27. Training graph of a TD3 RL Agent

IV. Conclusion

If we compare the reinforcement learning graphs for the DDPG and TD3 RL algorithms, we can conclude that the DDPG algorithm is more aggressive, which is reflected in the fact that it has larger oscillations in episodic rewards in earlier learning episodes. With the DDPG algorithm, reward amounts of around 100 were achieved already around the 1800th learning episode, which is much faster compared to the TD3 algorithm, which only achieves reward amounts of around 20 around the 1800th episode. Reward amounts of around 120 were achieved with the DDPG algorithm around 1900 episodes, and with the TD3 algorithm only around the 2900th learning episode. On the other hand, the obvious advantage of the TD3 algorithm is that the average reward per episode has a continuous and stable growth, which promises better control laws proportional to the learning time. The same cannot be said for the DDPG algorithm, because even though it reaches good control laws faster, it does not guarantee a continuous growth of the average reward, that is, with a longer learning time, there may be a decline in the quality of the control laws. This flaw of the DDPG algorithm is very well shown by the graph in Fig. 26. Therefore, the conclusion is that the TD3 algorithm has an advantage in the application of reinforcement learning for finding control laws for the gait of a bipedal humanoid robot. However, it would not be impossible to find quality solutions using the DDPG algorithm as well, since it allowes greater exploration.

References

[1] Reher, J. and Ames, A.D., 2021. Dynamic walking: Toward agile and efficient bipedal robots. *Annual Review of Control, Robotics, and Autonomous Systems*, *4*, pp.535-572.

[2] Gong, Y., Hartley, R., Da, X., Hereid, A., Harib, O., Huang, J.K. and Grizzle, J., 2019, July. Feedback control of a cassie bipedal robot: Walking, standing, and riding a segway. In *2019 American Control Conference (ACC)* (pp. 4559-4566). IEEE.

[3] Kim, I.S., Han, Y.J. and Hong, Y.D., 2019. Stability control for dynamic walking of bipedal robot with real-time capture point trajectory optimization. *Journal of Intelligent & Robotic Systems*, *96*, pp.345-361.

[4] Z. Li et al., "Reinforcement Learning for Robust Parameterized Locomotion Control of Bipedal Robots," 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 2021, pp. 2811-2817, doi: 10.1109/ICRA48506.2021.9560769.

[5] Rudin, N., Kolvenbach, H., Tsounis, V. and Hutter, M., 2021. Cat-like jumping and landing of legged robots in low gravity using deep reinforcement learning. *IEEE Transactions on Robotics*, *38*(1), pp.317-328.

[6] Siekmann, J., Green, K., Warila, J., Fern, A. and Hurst, J., 2021. Blind bipedal stair traversal via sim-to-real reinforcement learning. *arXiv preprint arXiv:2105.08328*.

[7] Li, T., Geyer, H., Atkeson, C.G. and Rai, A., 2019, May. Using deep reinforcement learning to learn high-level policies on the atrias biped. In *2019 International Conference on Robotics and Automation (ICRA)* (pp. 263-269). IEEE.

[8] T. Tiong, I. Saad, K. T. K. Teo and H. b. Lago, "Deep Reinforcement Learning with Robust Deep Deterministic Policy Gradient," *2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, Kuala Lumpur, Malaysia, 2020, pp. 1-5, doi: 10.1109/ICECIE50279.2020.9309539.

[9] Russell, S.J., 2010. *Artificial intelligence a modern approach*. Pearson Education, Inc.

Human Intention Recognition in Collaborative Environments using RGB-D Camera

Jelena Marić*,[1], Luka Petrović[1], Ivan Marković[1]

[1] University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
*maric298@gmail.com

***Abstract*—The increasing presence of robots and other autonomous systems in everyday environments, such as households, workplaces, schools and hospitals, requires their safe cohabitation and collaboration with humans. Accurate and reliable human intention recognition that fosters safe and efficient human-robot collaboration is thus a crucial component of these autonomous systems, especially if attainable from relatively cheap and widely available sensors such as RGB-D cameras. In this paper, we propose a hand-crafted model of human intention when reaching for one of multiple objects present on the table in front of a person. By coupling our hand-crafted model with a method for human skeleton tracking from RGB-D images, we devise a computationally efficient human intention recognition method suitable for collaborative pick-and-place scenarios. We experimentally verify our method in real-world scenarios of a person reaching for one of two, three and four objects placed on a table. We complement the paper with an open-source implementation of our human intention recognition method.**

***Keywords*—intention recognition, rgb-d, collaborative environments, action prediction, human-robot collaboration**

I. Introduction

As robots become more advanced and integrated into human-centric environments such as households, offices, schools and hospitals, there is a growing need for them to effectively collaborate and interact with humans, which presents new challenges in terms of system efficiency and human safety. Unlike robots, whose behavior can be fully controlled, human behavior is inherently unpredictable. Consider, for example, a person setting the dining table. While the task is well-defined, the order of subtask execution will clearly differ between people, especially if there are unforeseen changes in the environment. If the same task is carried out by a robot in collaboration with a person, it is of paramount importance for the robot to be aware of uncertainties and nuances of human behavior in order to ensure human safety and comfort. Human intention recognition thus presents one of the biggest challenges in collaborative robotics and has been an active area of research in recent years.

Intention recognition in the early stage of human actions can facilitate a variety of tasks in different scenarios, such as helping people overcome physical impediments [1], ensuring safety of human participants in traffic situations [2]–[4], enabling robots to purposefully partake in the manufacturing assembly process [5], or helping distinguish between healthy and unhealthy movement patterns in clinical settings [6]. The need for efficient human intention recognition is accentuated in collaborative environments where humans and robots simultaneously try to achieve a given task, especially in object-picking where a person and robot arm may often be in relative proximity. For such scenarios, a multimodal approach to intention recognition problem was presented in [7], where intention prediction is based on fusion of skeletal tracking measurements of a person's hand and pupil position measurements. Besides an RGB-D sensor used for obtaining skeletal data, additional piece of wearable equipment was needed for pupil tracking, reducing practicality of the proposed system.

Many researches have yielded successful results in intention recognition tasks using learning approaches [8], [9], where neural networks are often used as time series forecasters based on spatio-temporal movement patterns. Using kinematic data that includes motion and task-specific forces, a neural network was trained in [10] to approximate the function that represents dynamics of human movement and then predicts future human motions. Therein, intention refers to how the human hand will move next. A recurrent neural network (RNN) was trained in [11] to provide a prediction of future human movements given its history. Additionally, the RNN was then used for identifying the most probable goal of the predicted trajectory. In [12], the authors presented two methods for intention recognition that relied on RNNs, where the first method used person's hand trajectory as input, while the second method was trained on RGB-D videos. While both approaches showcased good performance when tested in an industrial setting, the high price and complex setup of motion tracking systems makes RGB-D sensor-based method more convenient and effective to use. A convolutional neural network (CNN) was proposed in [13] to recognize intention in each of the N captured images by learning subtle patterns in spatial features and accumulating those individual predictions over a time interval to produce the final prediction. Although spatial information in still images could be sufficient for action recognition, it might not be enough to distinguish subtle variations in motions which are a key aspect of intention recognition. To incorporate temporal aspect of motion, the authors in [14] proposed a CNN in a two-stream architecture which uses optical flow along with RGB data in an intention prediction task. A long short-term memory (LSTM) network combined with a CNN was used in [15] to predict reaching intention from gaze cues. In [16], [17], the

This research has been supported by the European Regional Development Fund under the grant KK.01.2.1.02.0119 (A-UNIT).

authors proposed using an ensemble of LSTMs for action prediction in collaborative environments, showcasing a degree of generalization to changes in the environment. While all of the mentioned methods achieve good performance in human intention recognition in collaborative environments, they require substantial amount of training data and hardware resources for the training process itself.

In this paper, we propose a lightweight human intention recognition method that uses RGB-D camera input. By relying on relatively cheap and widely available RGB-D sensors, as well as focusing on the ease of implementation, the proposed method is suitable for fast implementation and deployment on a variety of autonomous systems. Specifically, we craft a simple model of human intention when reaching for one of multiple objects present on the table in front of a person. Input data is processed in few simple steps using a limited amount of previously computed and stored values which makes the method computationally efficient. The proposed method considers distances between possible goal objects and key points of the tracked human skeleton, which are obtained from an existing method for RGB-D human skeleton tracking. The proposed method is suitable for collaborative pick-and-place tasks, and we experimentally verify its performance in real-world scenarios of a person reaching for one of two, three and four objects placed on a table. We complement the paper with an open-source implementation of our human intention recognition method[1].

II. Human intention recognition method

Throughout this paper, we consider a collaborative application where a human and a robot manipulate objects placed at known locations at the same table. In such scenarios, a person will typically move towards an intended goal, rather than in a random manner. Therefore, we need a human intention recognition method to determine a person's intended goal in order for the robot to choose its tasks appropriately, diminishing the risk of violating human safety and comfort. In this section we propose an easy-to-implement, computationally-efficient human intention recognition method that is suitable for the aforementioned scenario. We present each component of our method, including the definition of our weighted distance (Sec. II-A) and intention probability (Sec. II-B) functions, followed by the description of the utilized sliding window technique (Sec. II-C). Finally, our algorithm is summarized in Sec. II-D.

A. Weighted distance function

Consider a scenario where a person is grasping an object present on the table with their right hand. When a person begins their movement from a standing or sitting anatomical position, they start to reach with their hand towards the object. As the reaching motion continues, the distance between the target object and the person's hand becomes smaller and smaller. When the object is reached, that distance becomes zero. Moreover, the distance between target object and some of the other key points of the human skeleton, such as the right arm elbow or torso, becomes smaller with time.

We use this observation as the foundation of our intention recognition method. If there are multiple objects on the table, the distance between human skeleton keypoints and the preferred object at a certain time instant before grasping becomes smaller than the distance from every other object. Therefore, to be able to detect that the person is approaching the location of a certain goal g, we propose a general weighted distance function

$$d(g) = \sum_{j=1}^{J} w_j d_{gj} \tag{1}$$

where d_{gj} denotes the three-dimensional Euclidean distance between a goal location g and a given human skeleton joint j (out of J tracked skeleton joints), while w_j represents the weighting factor indicating relative importance of the given joint j, with $w_j \geq 0$, $\sum_{j=1}^{J} w_j = 1$.

While the proposed distance function is general in a sense that it can be used with any sensor that can provide skeleton detection, in this paper we employed a particular implementation of skeleton tracking from RGB-D sensor data. For data acquisition we used the widely available Xbox One Kinect sensor, that was initially intended for videogaming purposes, but quickly gained popularity in research applications due to its depth sensing ability and financial affordability. Since Kinect is a natural interaction device which can be accessed by open source APIs provided by OpenNI (Open Natural Interaction) framework, we used PrimeSense NiTE middleware which includes modules for OpenNI providing gesture recognition and skeleton tracking.

The mentioned skeleton tracking module provides position information for 15 different skeleton key points, namely hand, elbow and shoulder points for both arms, hip, knee and foot points for both legs, as well as head, neck and torso points. While this implies that $J = 15$, in our implementation of (1) we focus on the right hand, right elbow and torso information. We handcraft weights w_j in a way that right hand position information has the most influence on the total weighted distance. This corresponds to the aforementioned property of the distance between hand position and the desired object becoming zero when the object is grasped. The elbow and torso position information serve as a correction component in cases when the tracked human hand position provided by the used OpenNI framework is particularly inaccurate. We also assume that goal locations are invariable and known in advance, although they could be determined by using an open-source 3D object detection framework, e.g., [18].

B. Intention probability

When the weighted distance function in (1) is calculated for each possible goal location, we are in some cases able to simply infer the intended goal. For example, if we have three goals and one of them has comparatively much smaller weighted distance value than the other two possible goals, we may conclude that it is the intended

[1] https://github.com/maricjelena/intention-prediction

goal. Taking the argument of the minimum of the weighted distance function will provide reliable intention recognition in the aforementioned case that often happens when the person is already very close to the goal location. However, if the weighted distance is reasonably high from each possible goal, for example in the standing anatomical position, we might be unable to infer or be very uncertain about the intended goal. Taking the argument of the minimum in such situations may lead to clearly incorrect human intention estimates.

For the aforementioned reasons, we want to have a notion of uncertainty about our intention estimate and thus we utilize the *softmax* function to provide a probabilistic interpretation of the calculated weighted distances for each goal. A probability of each potential goal g_i (from a set of G possible goals) is given as

$$p(g = g_i) = \frac{e^{-\beta d(g_i)}}{\sum_{n=1}^{G} e^{-\beta d(g_n)}}, \tag{2}$$

where $d(g_i)$ is the weighted distance function given in (1), while β is a parameter for tuning the behavior of the softmax function, dependent on the specific requirements of a given task. With larger values of β, the function given in (2) has higher sensitivity. If we choose relatively large value of β, the probability of the predicted goal will always be close to 1, providing overconfident estimates. On the other hand, if chosen β is too small, it will output values close to 0.5 for the predicted goal, providing no confidence in the estimate. For the particular object reaching problem considered in this paper, we empirically concluded that $\beta = 40$ provides good estimates.

C. Sliding window

The only source of information we rely upon for intention recognition is skeleton tracking from RGB-D images, which is prone to providing noisy detections, e.g., due to self-occlusion. Unreliable input can significantly degrade estimation performance of the proposed intention estimation method. To mitigate the negative impact of wrong detections, we incorporate the sliding window technique in the estimation process to produce the final prediction formulated as the weighted moving average

$$\widetilde{p}_i = \frac{\sum_{f=1}^{F} w_f p_{i,f}}{\sum_{f=1}^{F} w_f} \tag{3}$$

where $p_{i,f}$ is the prediction probability of the goal g_i in the frame f, w_f is the weighting factor in the given frame f, and F is the number of frames considered in the weighting group. Parameters w_f and F (i.e., sliding window size) are manually tuned to work well for the observed motion.

Generally, applying the moving average with time-series data smooths out short-term fluctuations. In the context of this work, noticeable and sudden changes in tracked position do not occur naturally, because it is not physically possible for people to significantly change position of any joint in short period of time such as 0.03 s. However, this kind of deviation is not rare in skeleton detection. The purpose of the sliding window technique is to smooth out detection errors by predicting the intention using weighted contributions of the last F predictions, including the current one. Earlier predictions contribute less than the newer ones to better represent the time connection of the tracked position changes during the motion. For the particular problem considered in this paper, we empirically chose $F = 7$ which provided good balance between stability and fast prediction times in our experiments.

D. Algorithm summary

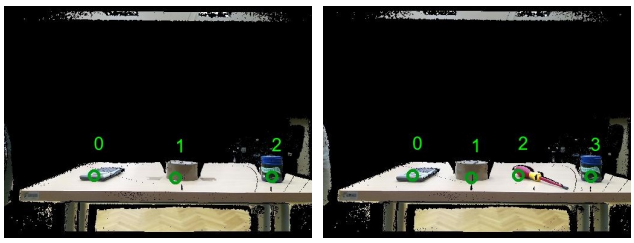
Given the skeletal pose of a human in a frame f, we first compute the weighted distance function given in (1) for each goal location g_i based on the current measurement of the considered joint positions. We then calculate the probabilities for each goal g_i following (2). Afterwards, the sliding window is recursively applied using (3) to get the final probability distribution at the current time instant. If there are not enough previous frames when estimating the intention in the current frame f, i.e., $f < F$, then the predictor takes f frames into account with the adjusted weight of each frame. The object location with the highest associated probability is considered to be the estimated intended goal.

III. Experimental results

In this section we demonstrate the performance of our method on data we collected in our laboratory. We start by explaining our dataset collection process and the obtained data in Sec. III-A. Then we present the results of the conducted performance analysis for scenarios with differing levels of prediction difficulties determined by the number of objects and their spatial arrangement. We first utilize the proposed intention prediction method in a two-object scenario (Sec. III-B) to verify the applicability of the proposed method in a fairly simple intention discrimination task. Experimentation area is then expanded by adding more objects (Sec. III-C), since the intention prediction method can be of great use in a real-life application based on recognizing the intended goal among many others. This also gives us an opportunity to induce possible problems that rarely occur in the basic two-object scenario and to show how they are handled with the proposed method. Lastly, in Sec. III-D we observe the effect of the sliding window technique on the stability of the proposed method when detection errors occur.

A. Dataset collection

We recorded 3D coordinates of a person's skeleton position while the person was reaching for objects placed on a table at least 9 cm apart. The number of those objects $K \in \{2, 3, 4\}$ and their locations are known for each scenario in advance and are depicted in Fig. 1. Each recorded sequence includes data acquired during multiple reaching movements. The duration of each reaching motion is in range of $1.5 - 2$ s, starting from the neutral body posture and ending in object reaching, after which the person returns to the neutral position or continues to reach for an another object directly from the location of the previously reached object. Sequences are labeled with timestamps of moments when an object is reached.



(a) (b)

Fig. 1: Scene with: (a) 3 and (b) 4 objects

B. Toy problem: Two object scenario

The method was first tested in the basic case of two objects being located with maximum distance apart (Fig. 2). As shown in Fig. 3, intention is correctly predicted for each reaching motion with a high probability of the intended object throughout the whole sequence. This is a somewhat simple decision because of the arrangement of objects. They are positioned such that approaching one object increases the distance to the other one which means that individual direction of probability change can point directly to the estimate of the method.

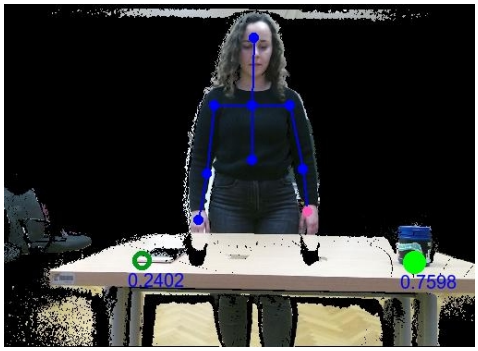


Fig. 2: Scene with two objects located with a maximum distance

C. Multi-object scenario

In this section we examine the prediction performance in scenarios featuring two, three and four possible goal objects. A comparison of prediction performance for three different combinations of tracked human skeleton joints being used in probability computation is given in Table I. For the right hand and right elbow combination, we used $w_j = 0.8$ for the hand and $w_j = 0.2$ for the elbow. For the right hand, right elbow and torso combination, we used $w_j = 0.8$ for the hand, and $w_j = 0.1$ for both the elbow and the torso. The table of prediction results shows average time left between the time instant of intention recognition and a manually determined time instant of reaching the goal location. This time is calculated by averaging results for chosen reaching motions in the observed sequence. Sequence selection criteria are explained in the following paragraph.

A given time instant is considered to be the prediction time instant if: i) the probability $\widetilde{p}_i$ of the goal g_i is greater than 0.5 in scenarios containing 2 objects and greater than 0.4 in scenarios containing 3 or 4 objects

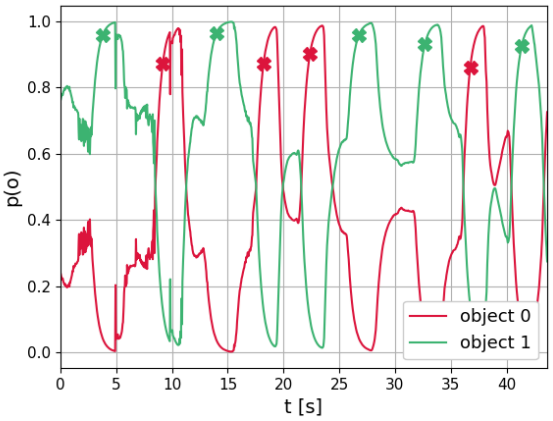


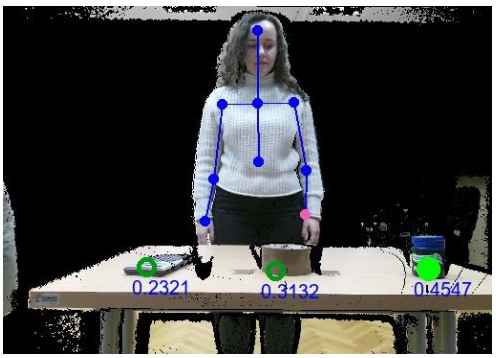
Fig. 3: Probabilities of objects from sequence shown Fig. 2 being reached. Note: moments when the goal objects are reached are marked with 'x' coloured with the ground truth object's color.

TABLE I: Average time left between the moment when intention is recognised and the reaching moment. Note: higher value represents better performance.

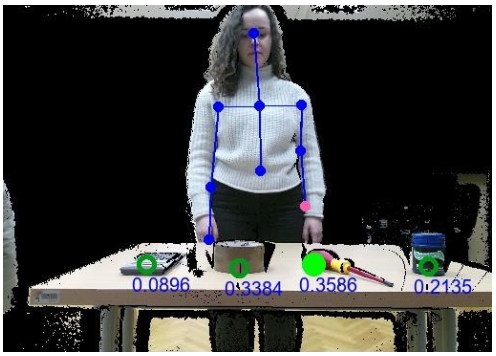
Reference skeleton joints	Number of objects		
	2	3	4
right hand	0.7559	0.6472	0.4789
right hand + right elbow	0.7219	0.6029	0.4152
right hand + right elbow + torso	0.7570	0.6279	0.4727

(Fig. 4); and ii) probability $\widetilde{p}_i$ of the goal g_i being reached is constantly greater than other objects' probabilities until (at least) the moment of reaching. There are cases when the initial position of a person suggests that a certain object is being reached, i.e., the initial probability of an object is higher than the defined threshold (Fig. 4a). Reaching that particular object is ignored in the analysis because the results do not reflect methods true prediction ability. The unevenness of initial probabilities could be alleviated with predefined location of a person such that the body position is truly neutral. Reaching motions which continue from the previous goal location were also not taken into account here due to significant differences in duration in comparison to trajectories starting from a neutral position which made them unsuitable for direct comparison. Note that the probability threshold of 0.4 was empirically determined as a suitable value for scenarios with three and four objects due to task complexity increase with the increase in the number of objects. An example of introducing complexity in an estimation task is a situation when approaching intended object results in comparable distance decrease for other objects as well.

The results from Table I show that the proposed method can provide reliable human intention recognition up to $0.75\,\mathrm{s}$ before reaching the desired object. Expectedly, the increased number of objects leads to decreased performance in average time left to correct intention recognition. These results also indicate that different joint combinations lead to similar results in the general case, implying that the intention could be determined by tracking only the hand joint, which is true for the chosen sequences. However, it is clear that the prediction performance declines when the detection does not correspond with the true position



(a)



(b)

Fig. 4: Scenarios containing (a) 3 and (b) 4 goal locations marked with initial probabilities

TABLE II: Average time left between the moment intention is recognised and the reaching moment in a scenario with 3 objects where input data suffers from detection errors.

Reference skeleton joints	Δt
right hand	0.5830
right hand + right elbow	0.5327
right hand + right elbow + torso	0.5940

of a person, as is often the case for hand joint. Taking the elbow and torso measurements into account can enhance output stability when hand detection significantly changes due to unreliable tracking or possible occlusions. To support this claim, we chose another sequence, depicted in Fig. 5, to show how the suggested sets of joints affect the final prediction. The average time left before correct intention recognition for the pertaining sequence is shown in Table II and suggests that the utilization of different tracked body joints may improve performance when input data suffers from detection errors.

D. Sliding window impact

We conducted additional performance analysis to show the importance of using sliding window. For that purpose, a sequence containing detection errors in multiple frames within one sliding window was chosen. Our analysis showed that, when using the sliding window, intention prediction method has the ability to resist negative influence of few false detection inputs, as showcased in Table III. The visual representation of prediction changes before and after the application of the sliding window technique is shown in Fig. 6. If skeleton detection was consistently

TABLE III: Probabilities when reaching for object $i = 1$ in 8 consecutive frames from Fig. 5 depending on the use of sliding window. Number of possible goal locations is 3, i.e., $i \in \{0, 1, 2\}$ seen from left to right. Predictions which differ from the ground truth are marked red.

time instant t	no sliding window		sliding window	
	i	$max(p_i)$	i	$max(\widetilde{p}_i)$
t_k	1	0.8431	1	0.5184
t_{k+1}	1	0.9188	1	0.5528
t_{k+2}	0	0.7632	1	0.4371
t_{k+3}	1	0.9508	1	0.5752
t_{k+4}	1	0.9526	1	0.5999
t_{k+5}	1	0.7210	1	0.5734
t_{k+6}	0	0.8137	1	0.4790
t_{k+7}	0	0.7510	1	0.4763

bad, then the prediction would have also underperformed since we rely solely on these measurements. The overall performance of the proposed method depends greatly on the provided input which can be unreliable since a single Kinect is used for obtaining the data. Introducing one or more additional RGB-D cameras would offer better scene coverage, hence improving the fused measurement accuracy.

IV. Conclusion

In this paper, we proposed a simple skeletal data-based model for recognizing intention in human actions when reaching for one of multiple objects present on the table in front of a person. By coupling our hand-crafted model with a method for human skeleton tracking from RGB-D images, we devised an intention recognition method suitable for collaborative pick-and-place scenarios. We experimentally verified the performance of the proposed method in real-world scenarios of a person reaching for one of two, three and four objects placed on a table. The proposed method can also be adapted for usage in different scenarios by picking a subset of skeleton joints with weights adjusted accordingly to the nature of the task. The obtained results support the decision of considering multiple joints in the distance function along with the sliding window technique for smoothing the prediction curve, both ensuring the stability of prediction method to some extent. If objects were placed in a straight line between the human and the farthest object, ambiguity of movement path would lead to method becoming greedy and declaring the location of the nearest object as the intended one; therefore, the proposed method is not suitable for every arrangement of objects.

In future work, it would be interesting to implement a robot controller for a human-robot collaborative assembly task to evaluate the performance of presented method in an real-life industrial environment.

References

[1] C. A. Chin, A. B. Barreto, J. G. Cremades, and M. Adjouadi, "Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities." *Journal of rehabilitation research and development*, vol. 45 1, pp. 161–74, 2008.

[2] H. Berndt, J. Emmert, and K. Dietmayer, "Continuous driver intention recognition with hidden markov models," in *2008 11th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 2008, pp. 1189–1194.

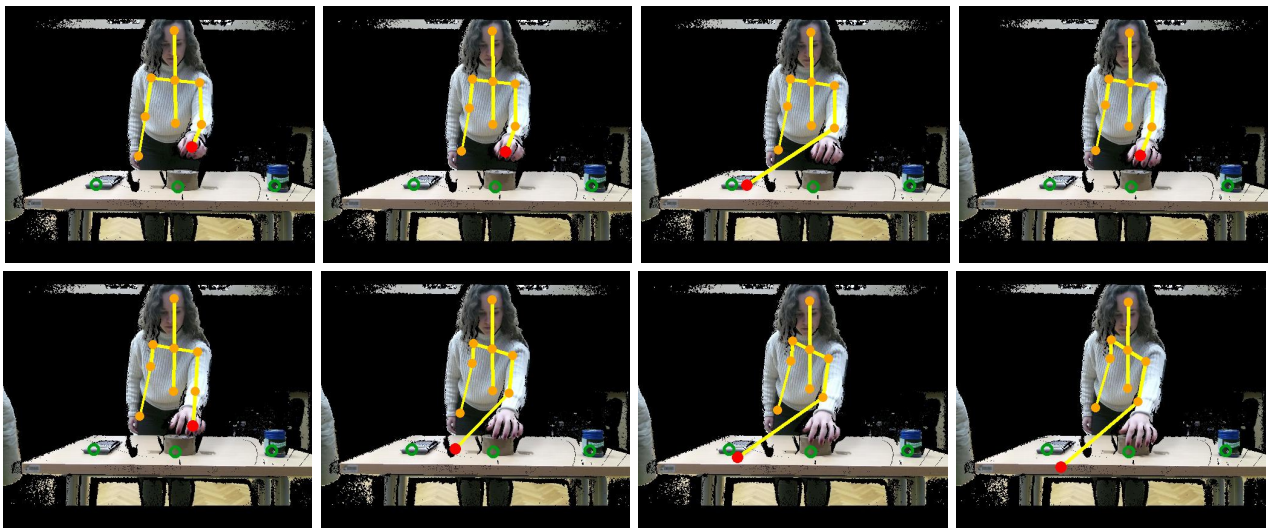
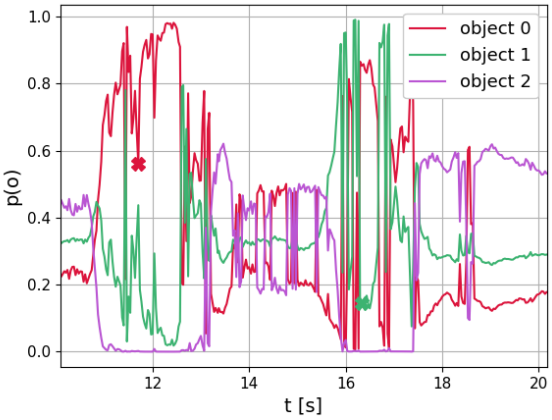
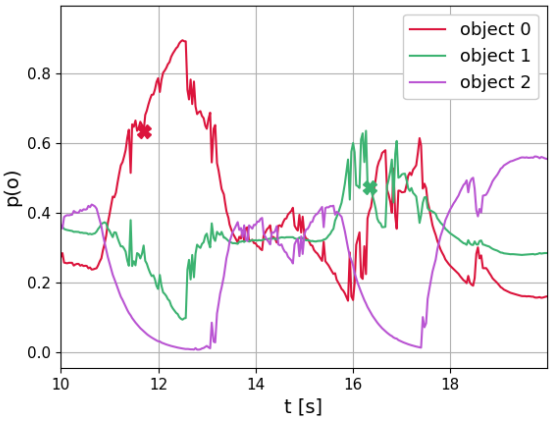


Fig. 5: Skeleton detection in consecutive frames



(a)



(b)

Fig. 6: Probabilities of reaching objects (a) without and (b) with sliding window.

[3] Z. Fang and A. López, "Intention recognition of pedestrians and cyclists by 2d pose estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, pp. 1–11, 10 2019.

[4] F. Camara, N. Bellotto, S. Cosar, F. Weber, D. Nathanael, M. Althoff, J. Wu, J. Ruenz, A. Dietrich, G. Markkula *et al.*, "Pedestrian models for autonomous driving part ii: high-level models of human behavior," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 9, pp. 5453–5472, 2020.

[5] Z. Zhang, G. Peng, W. Wang, Y. Chen, Y. Jia, and S. Liu, "Prediction-based human-robot collaboration in assembly tasks using a learning from demonstration model," *Sensors*, vol. 22, no. 11, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/11/4279

[6] F. Ragni, L. Archetti, A. Roby-Brami, C. Amici, and L. Saint-Bauzel, "Intention prediction and human health condition detection in reaching tasks with machine learning techniques," *Sensors*, vol. 21, no. 16, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/16/5253

[7] D. Trombetta, G. S. Rotithor, I. Salehi, and A. P. Dani, "Human intention estimation using fusion of pupil and hand motion," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 9535–9540, 2020, 21st IFAC World Congress. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S240589632033113X

[8] P. Schydlo, M. Rakovic, L. Jamone, and J. Santos-Victor, "Anticipation in human-robot cooperation: A recurrent neural network approach for multiple action sequences prediction," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 5909–5914.

[9] P. Kratzer, N. B. Midlagajni, M. Toussaint, and J. Mainprice, "Anticipating human intention for full-body motion prediction in object grasping and placing tasks," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2020, pp. 1157–1163.

[10] E. C. Townsend, E. A. Mielke, D. Wingate, and M. D. Killpack, "Estimating human intent for physical human-robot co-manipulation," *arXiv preprint arXiv:1705.10851*, 2017.

[11] D. Nicolis, A. M. Zanchettin, and P. Rocco, "Human intention estimation based on neural networks for enhanced collaboration with robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1326–1333.

[12] M. Mavsar, M. Deniša, B. Nemec, and A. Ude, "Intention recognition with recurrent neural networks for dynamic human-robot collaboration," in *2021 20th International Conference on Advanced Robotics (ICAR)*. IEEE, 2021, pp. 208–215.

[13] L. Zhang, S. Li, X. Diao, and O. Ma, "An application of convolutional neural networks on human intention prediction," *International Journal of Artificial Intelligence Applications*, vol. 10, pp. 1–11, 09 2019.

[14] S. Li, L. Zhang, and X. Diao, "Deep-learning-based human intention prediction using rgb images and optical flow," *Journal of Intelligent & Robotic Systems*, vol. 97, pp. 95–107, 2020.

[15] P. Festor, A. Shafti, A. Harston, M. Li, P. Orlov, and A. A. Faisal, "Midas: Deep learning human action intention prediction from natural eye movement patterns," *ArXiv*, vol. abs/2201.09135, 2022.

[16] T. Petković, L. Petrović, I. Marković, and I. Petrović, "Ensemble of lstms and feature selection for human action prediction," in *Intelligent Autonomous Systems 16: Proceedings of the 16th International Conference IAS-16*. Springer, 2022, pp. 429–441.

[17] ——, "Human action prediction in collaborative environments based on shared-weight lstms with feature dimensionality reduction," *Applied Soft Computing*, vol. 126, p. 109245, 2022.

[18] T. Yin, X. Zhou, and P. Krahenbuhl, "Center-based 3d object detection and tracking," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 11 784–11 793.

ROS Framework for Distributed Control of Networks of Dynamical Systems

M. Rossi*, A. Jokić†
* sees.ai, United Kingdom
† Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia
mrossi@sees.ai, andrej.jokic@fsb.hr

***Abstract*—In this paper we present a software package developed to accommodate flexible and efficient modelling and simulation of networks of dynamical systems with distributed control structures. The package is implemented within the Robot Operating System (ROS), which is becoming a standard software implementation tool in robotics, mechatronics and wider. The developed package is characterized with the following features: implementation of distributed controllers with both off-line predefined control laws in closed form and with on-line iterative solving of an optimisation problem (distributed model predictive control); modular structure which allows the user to easily modify the subsystems in the network, as well as the number of subsystems; allows for straightforward implementation of controllers onto real physical systems.**

***Keywords*—control systems, dynamical networks, robot operating system, distributed control**

I. INTRODUCTION

With the recent advances in communication and information technologies, and the development of new generations of sensors and actuators, a whole new set of systems started to emerge, which were unfeasible only years ago [1]. Some examples of such systems are: the so-called smart structures, composed of large amounts of sensors and actuators mounted on structural elements with the purpose of vibration dumping [2] or fluid flow control [3]; adaptive optics [4]; smart electrical power grids [5]; platoons of vehicles on automated motorways [6]; or large groups of ground mobile robots, unmanned aerial vehicles, or autonomous underwater vehicles, that collaborate towards a common goal [7], [8], [9]. The common characteristic of those systems is that they are composed of a relatively large number of spatially distributed dynamical subsystems, which interact with each other trough physical interconnections and/or communication links. Such systems are called networks of dynamical systems, or, shorter, dynamical networks.

The main focus of today's research in this area is the development of methods for control algorithm synthesis [1]. In most cases the challenge is not the design of subsystems, which can be very reliable when operating individually, but instead the design of local control laws, possibly in combination with distributed coordination schemes, for achieving common objectives at the network level. At the moment there is still no well understood, mature and widely applicable theory that offers scalable, robust, and reliable solutions to real-life network control problems. The research is currently still scattered, as problems of this nature are being explored also in biology, economics, sociology, game industry, etc. [10].

This paper contributes to the field by introducing a software framework that can be used for development, simulation, and real-life implementation of control algorithm for networks of dynamical systems. The framework allows modelling networks in the Python language, while the framework itself runs within the Robot Operating System (ROS).

The framework, called Dinsdale, is available as free software, under the General Public Licence, at `https://github.com/mross-22/dinsdale`.

II. CLASSIFICATION OF DYNAMICAL NETWORKS

To have a better understanding of the complexity of the field of distributed and decentralised control, and therefore to recognise the potential of the framework presented in this paper, it is useful to be familiar with the different types of networks present in control systems today.

A. Control structures

In the past decade, it has been acknowledged that fully centralised control structures are not capable to cope with the complexity of large spatially distributed systems. A centralised control structure is one in which one central control unit collects all the measurements and sends commands to all actuators. Such structures can guarantee globally optimal results of the controlled network, and their design is a mature field with a very well developed underlying theory. It has been proved that many practical problems with this structure can be formulated in terms of convex optimisation problems, and therefore efficiently solved [11]. However, some of their main limitations are that they are not scalable, not robust to failures, and are often practically impossible to implement.

The opposite of centralised structures are decentralised ones. In this case every subsystem is controlled by a local controller, and there is no collaboration between local controllers. Their main advantage is that no long distance communication is required, which makes them theoretically ideal for practical implementation. They also have some significant disadvantages, like the possibility to offer only suboptimal solutions which could possibly be

far from the global optimum, resulting in very inefficient behaviour. Using this structure the control synthesis is a non-convex problem, and there are no constructive algorithms for solving it [11].

The structure which has been recognised as often the most suitable for control of such large systems is the distributed structure. In such structure, each subsystem is controlled by a local controller which, apart from operating with locally available measurements, cooperates with a usually small set of neighbouring controllers. The controller communication network topology is in most cases required to be the same as the plant interaction network topology, but there could be exceptions.

B. Subsystem connection typology

The subsystems of a dynamical network can be physically coupled or decoupled. The former means that they directly affect each other's dynamics. Such systems range from inverted pendulums interconnected with springs, to complex electrical power grids, or the Internet.

If the systems are not physically coupled, it means that their interaction is through information exchange or measurements. They are connected by common goals or constraints, rather than physical links. Some examples are multirobot systems, vehicles platooning on motorways, and sensor networks.

C. Subsystem connection topology

Each subsystem has a certain set of neighbours, with which it interacts. This set can be either constant or change over time. In the former case the network topology is said to be static, while in the latter dynamic.

An example of subsystems that interact in a dynamic network topology are mobile robots which have a limited interaction range, due to limitations of sensing or communication equipment. This means that the set of each robot's neighbours is a function of their relative positions, which change over time.

D. Subsystem typology

The subsystems connected in a network can be all of the same type, or of different kinds. If the network is composed of identical systems, it is called homogeneous. Examples of such systems can be sensor networks or swarms of robots, given that all agents are of the same type. An interesting subset of such networks are spatially invariant networked systems, in which not only the subsystems are all equal, but also the dynamics of the system does not change moving along any spatial axis. An example are large segmented telescopes [12].

Networks composed of more than one type of subsystems are said to be heterogeneous. This is the case in the majority of applications, like electrical power grids, communication networks, or teams of different autonomous vehicles.

E. Control laws

The last important classification that is important to describe in this chapter is the difference of how the control laws of subsystems are being computed. The first possibility is that the controllers are predefined, which means that the control laws are computed offline, before operation. The second possibility is online solving of optimisation problems. In this case, each controller is computing its control output by solving an optimisation problem for every sampling period. This requires the controllers to iteratively exchange information between them inside one sampling period.

III. PROBLEM DEFINITION

Control of dynamical networks is a rapidly developing area of research, where most of the effort is put towards the development of a unified underlying theory. However, a universal software framework for development, simulation, testing, and practical implementation of such control algorithms is still missing. For this reason the transition from a general simulation to implementation on real systems requires big efforts and time investments. For example, it can be the case that people working on the theoretical development lack the programming skills required to implement complex distributed systems on specific hardware. Furthermore, even when having excellent programming skills and knowledge of embedded systems, this transition often requires a significant amount of time, having to reformulate the control algorithms in a form suitable for the particular system, and having to deal with communication protocols and limitations.

This has been the main motivation behind the work presented in this paper. The goal was to explore the possibilities of creating a software framework that will allow development of distributed and decentralised control laws for both simulated and real networks of dynamical systems, minimising or completely removing the transition from one to the other. The framework should be fully flexible to accommodate all the possible characteristics of dynamical networks.

A similar goal has already been achieved in robotics, by the Robot Operating System (ROS). The problem there was that a lot of robot control algorithms were being developed, but most of the time they were impossible to reuse across various systems and to connect with other existing software. Therefore, a lot of effort had to be put in the implementation process, often ending up rewriting existing software due to lack of portability. This is where ROS has contributed by offering a thin middleware which acts as a universal platform for software development, without being invasive (the software does not have to be designed specifically to work with ROS).

Since ROS already offers a standard systems to make different software communicate, and since there already exists a large amount of software packages for it, from hardware drivers to high level functionality, it has been

decided to use it as the base for developing the dynamical network control framework.

IV. ROBOT OPERATING SYSTEM

This is a very brief overview of the main concepts of ROS. A more detailed introduction, with focus on distributed system development, can be found in [13]. More in-dept information on ROS, as well as tutorials, can be found online, on the ROS website [14], and in books, like [15].

ROS is a framework primarily intended for development of robot software. It is a collection of tools, libraries, and conventions that aim to simplify the task of creating complex and robust robot behavior across a variety of robotic platforms [16]. ROS is free software, licensed under the permissive BSD licence. It was created, and, from 2007 to 2013, developed, and maintained mostly by Willow Garage, a research laboratory and technological incubator which produces hardware and free software for service robotics [17]. Since 2013 the project has been transfered to the Open Source Robotics Foundation (OSRF) [18], a non-profit organization born as a Willow Garage spin-off.

It is important to note that ROS is not a replacement for a computer operating system, but a middleware between the robot hardware and the control algorithms. From the robotics point of view however, it offers all the services and abstractions expected from an operating system, which explains its name. Some of the key services are hardware abstraction, low level device control, implementation of commonly used functions, process management, and package management. It also provides tools and libraries for obtaining, building, writing and running code across multiple computers [14]. At the moment, ROS runs on top of UNIX-like operating systems, with the best support on the Ubuntu GNU/Linux distribution and its derivatives.

Software in ROS is distributed in packages and stacks of packages. A package is simply a directory with a certain structure, which can contain executables, libraries, configuration files, or anything else.

At runtime, ROS manages the software inside what is called the ROS Computational Graph, which is basically a network of processes that can communicate with each other. The main concepts are: *Nodes, Master, Parameter Server, Messages, Topics, Services, Bags.*

The programs (processes) that are being executed are called nodes, while the edges of the graph are the interactions between them. The topology of the graph can be dynamical: the graph can be fully modified at runtime.

The main program (server) which allows the interaction between nodes, because it contains all the names and addresses, is called Master. The Parameter Server is a part of the Master which serves as a place for storing publicly available data in a centralised location. It is not very fast, so it is not used for runtime communication, but rather for storing some universal parameters.

There are two methods for communication between nodes: the asynchronous publishing of messages to topics, and the synchronous communication through services. A node can send a message by publishing it to a given topic. The topic is a name to identify the message. A node that is interested in a certain kind of data will subscribe to the appropriate topic. There is no limit to the number of publishers and subscribers for a topic. Services are used in cases when request/reply communication is required. A node can offer a service under a specific name, to which any other node can send a request and wait for the reply.

The last fundamental concept that is important to understand are bags. Bags are files used for storage and reproduction of logged data. They offer a convenient way to store all kind of data during execution, which is very valuable for development and testing. The data from a bag file can be reproduced in the order it was collected, which allows for example testing different algorithms on the same data set.

V. THE DINSDALE FRAMEWORK

This section offers an overview of the Dinsdale package for the Robot Operating System (ROS), a complete framework to solve the problems presented in this paper. It gives the possibility to implement control algorithms which can equally work on both simulated and real networks of systems, without the need to modify the code. Its flexibility also allows both the controllers and the plants to be interconnected in any desired way, allowing implementation of all control structures from sec:control.

An additional characteristic is that networks of dynamical systems (controllers and simulated plants) are fully contained inside Python packages, completely independent from both Dinsdale and ROS. This allows separate development and testing, and encourages reuse of system models for any other purpose. The core of the Dinsdale framework itself is a ROS wrapper around those Python packages, which takes care of the execution, coordination, communication, and data logging, and which gives a standard interface for integration with other ROS software.

Python's popularity within the scientific community is increasing very rapidly, often at the expense of proprietary scientific computing software, making it the obvious choice for the end user side (system modelling). Python is, however, not only being used for implementation of dynamical networks, but the entire framework has been written in it. Its beauty and flexibility enabled the development of a powerful core functionality, while simultaneously keeping it relatively simple for users to understand, modify or extend.

ROS has been chosen as the underlying framework because it is rapidly becomingh the *de facto* standard in the robotics community for implementing control structures. This offers the possibility to integrate control algorithm developed in Dinsdale with other ROS packages, e.g. the ones for low-level control, localisation and mapping, or various virtual reality visualisation software. It is also

possible to use standard ROS tools (e.g. `rostopic`, `rosbag`, or `rqt`) to inspect the dynamical network, or to store and reproduce data.

The Dinsdale package is free software, with all of its components released under the GNU General Public Licence (GPL) version 3, as published by the Free Software Foundation. Dinsdale is hosted on GitHub, at `https://github.com/mross-22/dinsdale`.

A. Package structure

```
dinsdale
├─bags
├─msg
│  └─Floats.msg
├─scripts
│  ├─node_controller
│  ├─node_plant
│  ├─node_simulation_time
│  ├─node_topology_plants
│  ├─result_analysis
│  └─setup_simulation
├─src
│  └─dinsdale_system
│     ├─input_parameters
│     │  ├─A_controllers.txt
│     │  ├─A_plants.txt
│     │  ├─controllers_iterative.txt
│     │  ├─plants_topology.txt
│     │  ├─system_types.txt
│     │  ├─T.txt
│     │  └─x0.txt
│     ├─tools
│     │  ├─__init__.py
│     │  ├─read_matrix.py
│     │  └─result_analysis.py
│     ├─__init__.py
│     ├─controller.py
│     ├─plant.py
│     └─plants_topology.py
├─CMakeLists.txt
├─LICENSE
├─package.xml
├─README
└─setup.py
```

Fig. 1: Dinsdale directory tree

The Dinsdale package is a directory containing all the source code and data files required to run dynamical networks inside a ROS computational graph. The minimal structure of Dinsdale is illustrated in Fig. 1. All the source files can be divided in three categories:

- System description ⟶ `./src`
- Logged data ⟶ `./bags`
- ROS wrapper ⟶ everything else

The Python package containing the entire definition of the dynamical network that is being simulated is `./src/dinsdale_system`. Inside it, the user defines the dynamics, control laws, and topology of the network. `./src` can contain more than one system, but only the one in the `./src/dinsdale_system` package is executed. This allows to store multiple systems, and choose the one to use by simply renaming its directory.

Each time a simulation is set up, a directory containing the current date and time in its name is created inside `./bags`. In that subdirectory all the data logged during the simulation will be stored, inside ROS `.bag` files. The user does not necessarily need to access or manipulate those files, as Dinsdale has a script which allows data analysis directly in Python.

Fig. 2 gives a somewhat simplified overview of the Dinsdale architecture, which is the topic of the rest of this chapter.

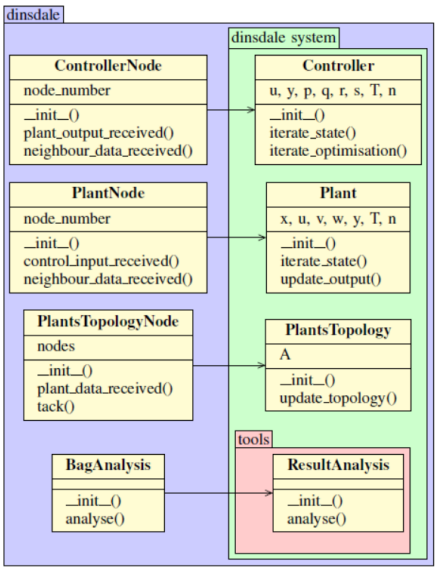


Fig. 2: Dinsdale directory tree

B. System definition

The dynamical network is defined in a Python package called `dinsdale_system`, located inside the `src` directory of the main Dinsdale package (see Fig. 1).

1) Plants: Each plant in the network is defined by a `Plant` class. This class is located inside `plant.py`, and its structure is shown in Fig. 2. It is necessary to define as many classes, each in its own file, as the number of different types of plants in the network.

Table I explains what the attributes of the class are. All, except for `n` and `T`, are of the `numpy.matrix` type. They need always to have the shape of column vectors, but their length is not limited.

As shown in Fig. 2, the `Plant` class has three methods:

- `__init__(n, x0, T)` – The initialisation of a `Plant` instance. It initialises all the attributes, and executes the initialisation code set by the user. This method is called only once, at the beginning of the execution.
- `iterate_state()` – The user here sets the equations for the update of the plant's states `x` and output

TABLE I: Plant attributes

Plant	
x	plant states
u	input from controller
y	output for controller
w	input from other plants
v	output for other plants
n	ordinal number of the node
T	sample time

for other plants `v`. This method is executed when new inputs from the controller are received.

- `update_output()` – This method is executed when the new set of input from neighbouring plants `w` is received, and its purpose is to contain the equation for updating the output for the controller `y`.

2) Controllers: Each controller in the network is defined by a `Controller` class. Following the same logic applied to plants, the class for a controller is inside a `controller.py` file. In case of multiple different types of controllers, there will be multiple files, called `controller.py`, `controller_1.py`, `controller_2.py`, etc.

TABLE II: Controller attributes

Controller	
y	input from plant
u	output for plant
q	input from other controllers
p	output for other controllers
s	iterative input from other controllers
r	iterative output for other controllers
finished	iterative communication finished
n	ordinal number of the node
T	sample time

Fig. II explains what the attributes of the `Controller` class are. Except for `n`, `T`, and `finished`, the rest are of the `numpy.matrix` type. They need always to have the shape of column vectors, but their length is not limited.

As shown in Fig. 2, the `Controller` class has three methods:

- `__init__(n, T)` – The initialisation of a `Controller` instance. It initialises all the attributes, and executes the initialisation code set by the user. This method is called only once, at the beginning of the execution.
- `iterate_state()` – The user here sets the equations for the update of the controller's outputs, both for the plant and for neighbouring controllers (`u` and `p`). This method is executed when new data from the plant (`y`) is received.
- `iterate_optimisation()` – This method is executed after `iterate_state()` if the controllers communicate iteratively within one simulation step. This method is usually used when the control law is computed by collaboratively solving an optimisation problem on-line. If this is the case, the vectors used for iterative communication are `r` and `s`. In each simulation step this method is called until `finished` is set to `True`. In the next iteration it will automatically be reset to `False`.

3) Plants interaction topology: The topology of the plants interaction network can be static or dynamic. When it is static, nothing has to be done. In case it is dynamic, the user has to determine the law according to which the topology is changing. This is done in the `PlantsTopology` class (shown in Fig. 2), inside the `plants_topology.py` file (note that there can be only one such file for a network).

The class `PlantsTopology` initialises:

- `A` – The adjacency matrix of the plants interaction network (a square $(n \times n)$ matrix, where n is the number of plants).
- `nodes` – The number of plants.
- `w` – The output from each plant (a list of n elements, each being a one dimensional `numpy.array`).
- `communication` – A list of neighbours of each plant (a list of n lists). It gets populated during the initilaisation by looping through the adjacency matrix. Once the neighbours of each plant are found, in each simulation step the plant will receive their outputs.

The method `update_topology()` is empty by default, and this is where the user can define the laws according to which the topology is changing, in case it is dynamic. The laws could use the data sent by plants, or be functions of time. Another possible application of the `update_topology()` method, apart from having a dynamic topology, is multiplying the values exchanged by plants with some weights, which can also change over time. Whatever its usage may be, the objective in this method is to update the list `communication` with lists of neighbours of each plant in every time step.

4) Parameters: Inside `./input_parameters` there are seven items to be set to fully define the dynamical network:

- `A_controllers.txt` – The adjacency matrix of the controller communication network. The file can be empty if the controllers do not communicate.
- `A_plants.txt` – The adjacency matrix of the plant communication network. Even if the plants do not interact, it has to be of size $(n \times n)$, where n is the number of plants, but filled with zeros.
- `controllers_iterative.txt` – Determines whether the controllers communicate iteratively inside one time step (1 if they do, 0 or empty otherwise).
- `plants_topology.txt` – 0 if the plants topology is static, 1 if dynamic.
- `system_types.txt` – Describes the type of controller and plant for each subsystem. It is a matrix of size $(n \times 2)$, where n is the number of subsystems. The first column corresponds to controllers, the second to plants.
- `T.txt` – A (3×1) vector, the elements of which correspond to the sampling period (used for discretisation of the system dynamics), the real duration of a simulation time step, and the final time of the simulation, all given in seconds.
- `x0.txt` – Contains the initial conditions for each plant. It is an $(n \times m)$ matrix, where n is the number of plants and m the number of states of a plant (in case of heterogeneous systems, the plant with the largest dimension of the state space determines this number).

5) Tools: The additional package `./tools` contains useful tools which are not part of the dynamical network. Two modules can be find inside it:

- `read_matrix.py` – This is the function all other modules use for reading matrices from files. It is placed here to make it available also for the user, in case he or she wants to load some additional matrices.
- `result_analysis.py` – This module is where all data from a simulation is made available to the user for analysis of any kind.

The `result_analysis.py` file contains the `ResultAnalysis` class, shown in Fig. 2. In its initialisation method, the class initialises a dictionary called `data`. This dictionary gets filled by a core Dinsdale class with data stored during a simulation. The keys of the dictionary correspond to names of controller and plant attributes. The value of a key is a list of a list of all values of the specific attribute of every node in the network. In the method `analyse()` the user is free to manipulate this data in any way, e.g. plotting it.

C. Runtime

The `dinsdale_system` package is a pure Python package, which contains the description of a dynamic network. To generate as many instances of its classes as required by the network specification, initialise their values, transform them in ROS nodes, set up the communication networks, and coordinate the execution, there is a wrapper around it. This wrapper is the core Dinsdale functionality. Fig. 2 shows how every class in `dinsdale_system` is being use by Dinsdale.

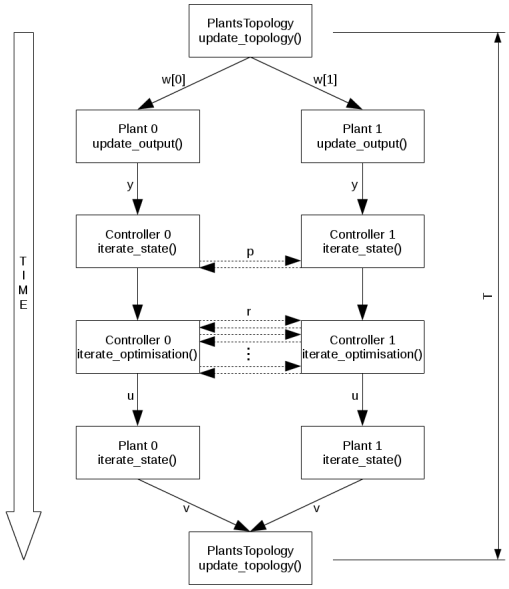


Fig. 3: Execution flow

At runtime, Dinsdale sets up as many ROS nodes as defined by the user, and connects them accordingly in a ROS computational graph. After all nodes are set up in a ROS graph and the simulation time is started, Dinsdale starts simulating the network using the classes defined in `dinsdale_system`. Fig. 3 shows the order in which the methods are being called inside one time step, when simulating a network of two systems. The dashed arrows represent optional communication between controllers (single and iterative). When the controllers are not communicating iteratively, the method `Controller.iterate_optimisation()` is not called at all. The letters near the arrows represent the name of the attribute that is being sent.

D. Replacing simulated plants with real systems

One of the main motivations for the development of the Dinsdale package was to make the transition from simulation to implementation as simple as possible. This was kept in mind during the entire design process, which can be seen for example in the structure of controller nodes, which have all their input and output topics set within their namespaces. This makes them agnostic about who is publishing their inputs and who is using their outputs. Not only this gives the possibility to replace the simulated plants with real systems, but also to use Dinsdale together with the big variety of existing ROS packages, e.g. for simultaneous localisation and mapping, visualisation, image processing, etc.

As ROS is gaining popularity, more and more hardware drivers are available as ROS packages. At the same time, small single-board ARM computers are rapidly being adopted as cheap and efficient control units for a huge variety of systems. Since those computers are able to run GNU/Linux distributions, and therefore ROS, these facts makes it very convenient to use Dinsdale generated controllers for control of real networks of systems.

VI. CONCLUSION AND FUTURE DEVELOPMENTS

This paper presented the motivation and explored the possibilities of creating a universal framework for development of networks of dynamical systems. It introduced a new framework, which fully satisfies the requirements, derived by the needs of today's research community. With the testing conducted so far, it has been shown that the Dinsdale framework offers a solid and reliable infrastructure for research and development of dynamical networks. Although the benefits from using such a framework are theoretically clear, they still have to be proven in practice. Feedback from other research groups would be immensely valuable for determining the direction of the future development of this software.

Generally speaking, the need for such framework exists, and considering its foundations are the increasingly popular ROS and Python, its adoption could be very convenient. It will certainly remain available and continue being maintained and developed, at least in the foreseeable future.

Dinsdale is hosted on GitHub, at `https://github.com/mross-22/dinsdale`.

REFERENCES

[1] R. Murray, K. Astrom, S. Boyd, R. Brockett, and G. Stein, "Future directions in control in an information-rich world," *Control Systems, IEEE*, vol. 23, no. 2, pp. 20–33, Apr 2003.

[2] A. Preumont, *Vibration Control of Active Structures: An Introduction*, ser. Solid Mechanics and Its Applications. Springer, 2011. [Online]. Available: http://books.google.hr/books?id=MUQUQyB4bEUC

[3] P. Koumoutsakos and I. Mezic, *Control of Fluid Flow*, ser. Lecture Notes in Control and Information Sciences. Springer, 2006. [Online]. Available: http://books.google.hr/books?id=7T2lrA-YVCwC

[4] R. Duffner, *The Adaptive Optics Revolution: A History*. University of New Mexico Press, 2009. [Online]. Available: http://books.google.hr/books?id=6nJ9PQAACAAJ

[5] J. Ekanayake, N. Jenkins, K. Liyanage, J. Wu, and A. Yokoyama, *Smart Grid: Technology and Applications*. Wiley, 2012. [Online]. Available: http://books.google.hr/books?id=AmpSxODiF7sC

[6] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968. [Online]. Available: http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal\&id=SJCODC000006000001000131000001\&idtype=cvips\&gifs=yes

[7] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, pp. 401–420, 2006.

[8] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, ser. Princeton Series in Applied Mathematics. Princeton University Press, 2010.

[9] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, `http://coordinationbook.info`.

[10] G. Antonelli, "Interconnected dynamic systems: An overview on distributed control," *Control Systems, IEEE*, vol. 33, no. 1, pp. 76–88, 2013.

[11] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 274–286, Feb 2006.

[12] S. Jiang, P. Voulgaris, L. Holloway, and L. Thompson, "Distributed control of large segmented telescopes," in *American Control Conference, 2006*, June 2006, pp. 6 pp.–.

[13] M. Rossi, "ROS u distribuiranom upravljanju dinamičkim sustavima," University of Zagreb, Tech. Rep., 2014.

[14] ROS, "Documentation," `http://wiki.ros.org/`.

[15] J. M. O'Kane, *A Gentle Introduction to ROS*. Independently published, Oct. 2013, available at `http://www.cse.sc.edu/ jokane/agitr/`.

[16] ROS, "About ROS," `http://www.ros.org/about-ros/`.

[17] Willow Garage, "About Us," `http://www.willowgarage.com/pages/about-us`.

[18] OSRF, "ROS @ OSRF," `http://osrfoundation.org/blog/ros-at-osrf.html`.

BIOMEDICAL ENGINEERING

Heart Rate Variability Monitoring with Savvy ECG Sensor During Dental Surgery

I. Tomasic[1], A. Prkic[2], A. Lesin[3,4], D. Kalibovic Govorko[3,5], N. Tomasic[6], and I. Medvedec Mikic[3,7]
[1] Division of Intelligent Future Technologies/Mälardalen University, Västerås, Sweden
[2] Private dental practice, Zagreb, Croatia
[3] Department of Maxillofacial Surgery/University Hospital of Split, Split, Croatia
[4] Department of Oral Surgery/School of Medicine, University of Split, Split, Croatia
[5] Department of Orthodontics/School of Medicine, University of Split, Split, Croatia
[6] Karolinska University Hospital, Stockholm, Sweden
[7] Department of Endodontics and Restorative Dentistry/School of Medicine, University of Split, Split, Croatia
Email: ivan.tomasic@mdu.se

***Abstract* - Surgical procedures are accompanied with anxiety and stress for both patients and surgeons. This paper investigates the utility of various heart rate variability parameters, estimated from Savvy patch ECG recordings, obtained on patients during surgical extraction of impacted lower third molars under local anesthesia, to estimate psychological state of the patient, as well as pain levels, in real time. This could be particularly beneficial for patients with chronic conditions, such as cardiovascular diseases, that can deteriorate under psychophysical stress. Statistical analysis showed significant correlation between several heart rate variability parameters and changes in autonomous nervous system functionality. Additionally, there was a significant correlation between several heart rate variability parameters on one side, and patient reported anxiety, pain levels, and the procedure difficulty, on the other.**

Keywords - Patch ECG; Savvy; heart rate; heart rate variability; anxiety; stress; dental surgery; pain; local anestesia.

I. INTRODUCTION

Because of anxiety and stress, patients experience hemodynamic changes during and before surgical procedures [1]–[4]. Anxiety related to surgery, as well as anxiety caused by the local anesthetic injection; increase cardiac oxygen demand, which increases the risk for arrhythmic events [5], [6]. Oral surgical procedures are therefore riskier for hypertensive, cardiac, cerebrovascular, and senior patients. Healthy individuals cope with stress better, but sudden and asymptomatic heart rhythm disorders may occur during oral surgical procedures even for them [7]. Some disorders, such as hypertension or coronary vascular disease (CVD), might manifest only during psychophysical stress [8]. Monitoring of vital signs is therefore recognized as beneficial, and even necessary, in some dental procedures [9].

Pulse oximetry has been accepted as a reliable device, for monitoring oxygen saturation (SaO_2) and heart rate, in dental procedures [10]. Recently however, novel wearable ECG devices have emerged, among which patch ECG monitors are the most prominent [11]. In addition to the heart rate, they provide electrocardiogram (ECG) and respiratory rate. In this study we used Savvy patch ECG monitor which is CE certified as a medical device [12].

In our previous studies, we investigated heart rate (HR) and ECG recordings during surgical extraction of impacted lower third molars under local anesthesia, considering patients' gender and dental anxiety levels [13], [14]. This study calculates all the prominent heart rate variability (HRV) parameters and correlates them with reported anxiety and pain levels in patients, as well as with the procedure complexity.

A. HRV relation to anxiety and pain

HRV estimates are measures of balance between sympathetic (SNS) and parasympathetic (PNS) autonomous nervous systems. Increased SNS, or decreased PNS activity, causes an increase in HR, and vice-versa. Furthermore, the response time of PNS is shorter than the SNS (PNS has faster reaction). Consequently, SNS activity is associated with low-frequency measures of HRV, whereas PNS is associated with medium and high frequency measures of HRV [15], although some authors consider low-frequency to be influenced by both systems.

Stress, anxiety, and pain are in general associated with an increase in sympathetic and decrease in parasympathetic output [15]–[18].

Two previous studies that evaluated patient [19] and surgeon [20] stress during surgery, have reported increased stress as measured by frequency characteristics of the HR signal.

II. METHODS

A. Data

Thirty (15 males and 15 females) healthy, normotensive, medication-free patients with ages ranging from 19 to 43 years, scheduled for surgical extraction of lower third molars, were enrolled in the study. The data was collected between Nov 2019 and March 2020 in the Department of Oral and Maxillofacial Surgery, University Hospital of Split, Croatia. The Bioethics Committee of the University Hospital of Split approved the study. The participants signed written informed consent. All procedures were performed in the same dental office, by the same oral surgeon during morning hours.

B. Study protocol

Before the procedure, while in the waiting room, patients were asked to fill Corah's Dental Anxiety Scale (DAS) [21]. It consists of four multiple-choice questions scoring from 4 to 20. At the end of the procedure, the visual analog scale (VAS) from 0 to 10 was handed to patients to report intraoperative pain.

ECG was recorded using the Savvy patch ECG monitor connected to a Smartphone through low-power Bluetooth. The sensor was placed with positive electrode on the sternum approximately 5 cm above the xiphoid, whereas the negative electrode was placed as much above and to the left as the connecting wire allowed, without causing tension on the electrode-skin contact.

All surgical procedures were performed in aseptic conditions using a buccal approach and rising mucoperiosteal flap with the vertical incision in the lower second molar area. The osteotomy site was irrigated with saline, and after bone removal, extraction was performed using either forceps or elevators, separating crown and roots if needed. The flap was repositioned and sutured using silk 3-0 stitches (Johnson & Johnson Medical Ltd Simpson Parkway, Krikton Campus, Livingston, United Kingdom).

The procedure's difficulty was assessed by the modified Parant scale (PARANT) [22]. Since all patients required osteotomy, grade II was the minimal grade. The procedure duration time was measured from the time of incision until the placement of the last suture. Each patient received the inferior alveolar nerve block using the same local anesthetic 4% articaine with a vasoconstrictor (adrenalin 1:100000) in 1.8 ml dose (Ubistesin 40 mg/ml+0.005 mg/ml injection solution, 3M ESPE). No more than 5.4 ml (3 carpules) of anesthetic solution was used in each case.

By using the MobECG app that accompanies Savvy, various time moments were marked during the procedure, enabling the total time to be split in the following intervals: 1-Waiting room; 2-Anestesiation; 3-Procedure; 4-After procedure. Patients were aware of each stage in the procedure.

C. HR and HRV estimates used

HRV parameters were calculated in MATLAB from the raw ECGs by using PhysioNet Cardiovascular Signal Toolbox [23], [24].

Measures of normal R-R intervals (NNs) rate used:
- NNmean (ms) - mean value of NN intervals
- NNskew - skewness of NN intervals
- NNkurt - kurtosis of NN intervals.

Measures or NN intervals variability used:
- NNiqr (ms) - interquartile range of NN intervals
- SDNN (ms) - standard deviation of NN intervals
- RMSSD (ms) - square root of the mean of the sum of the squares of differences between adjacent NN intervals
- pnn50 (%) - number of pairs of adjacent NN intervals differing by more than $50\,ms$, divided by the total number of all NN intervals.

Since RMSSD and pnn50 measure faster changes in NNs, they are considered as estimates of parasympathetic activity, whereas NNiqr and SDNN are measures of sympathetic activity.

Frequency domain measures (using Lomb periodogram method):
- ulf (ms^2) - power in the ultra-low frequency range (< 0.003 Hz)
- vlf (ms^2) - power in very low frequency range ($0.003 <= \text{vlf} < 0.04$ Hz)
- lf (ms^2) - power in low frequency range ($0.04\text{Hz} <= \text{lf} < 0.15$ Hz)
- hf (ms^2) - power in high frequency range ($0.15 <= \text{hf} < 0.4$ Hz)
- lfhf - ratio lf/hf]
- ttlpwr (ms^2) - total spectral power.

The hf component is seen as indicators of parasympathetic activity, whereas the lf component represents the activity of the sympathetic nervous system, although some studies suggest that it reflects both the sympathetic and parasympathetic nervous system. The lfhf-ratio is used as indicator of the balance between the sympathetic and parasympathetic nervous system; higher values of lfhf indicate a predominance of the SNS, whereas lower values indicate a predominance of the PNS. ttlpwr is the total variance in the heart rate [25].

Other HRV measures:
- ac (ms) - acceleration capacity
- dc (ms) - deceleration capacity

dc of heart rate is a measure of cardiac parasympathetic modulation as it captures the lengthening of NN interval within 2-4 successive beats. ac of heart rate captures the shortening of NN interval.

Poincaré plot (PP) features:
- SD1 (ms) - standard deviation of projection of the PP on the line perpendicular to the line of identity ($y = -x$)
- SD2 (ms) - standard deviation of the projection of the PP on the line of identity ($y = x$)
- SD1SD2 - SD1/SD2 ratio.

SD1 is related to the short-term variability, whereas SD2 describes the longer-term variability [15]. The ratio SD1/SD2 describes the relationship between the two.

Entropy measures: SampEn - Sample entropy and ApEn – Approximate entropy, are measures of the regularity and complexity of time series. ApEn decreases as the NN variation decreases. The only difference between dSampEn and ApEn is a small computational difference.

D. Statistical analysis

Linear mixed effect models were fitted between each HRV measure, as the output variable, and interval number, as input variable. All the models were corrected for gender and age, since there is a well-established relation between HRV and these two variables [15], [26]. Additionally, the influence of DAS, VAS and PARANT were estimated as slopes in each time interval. All the models included

random intercept per patient to model dependency between measurements on the same patient.

In the continuation we used asterisk (*) to mark p values lower than 0.05. Influences that are not presented were not found to be statistically significant. All graphs show mean and 95% confidence intervals.

The first time-interval (waiting room) was set as the baseline for HRV analysis, so the other intervals are contrasted to the first.

III. RESULTS

Fig. 1 presents HR properties, whereas Fig. 2 presents time domain HRV measures. Figures 3 and 4 present frequency domain HRV properties whereas other, Poincaré and entropy measures are presented in figures 5, 6, and 7, respectively.

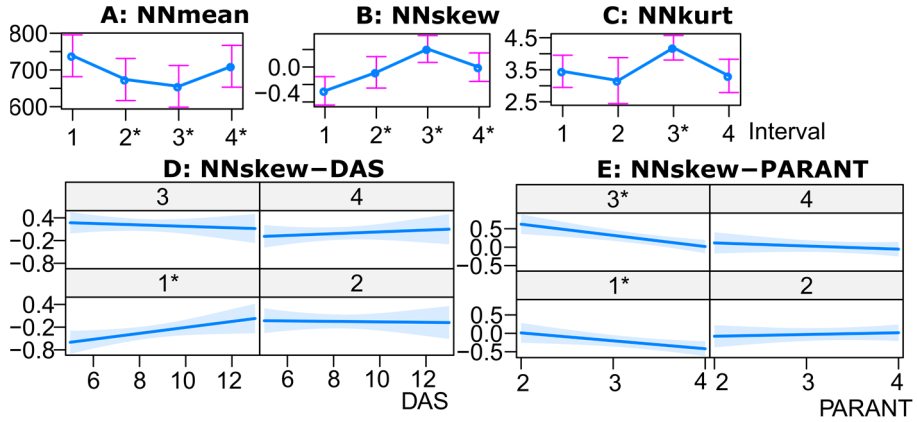


Figure 1. Estimates for NN central tendency (mean) and departure from normality (skewness and kurtosis) (D – skewness as a function of DAS per interval; E – skewness as a function of PARANT per interval)

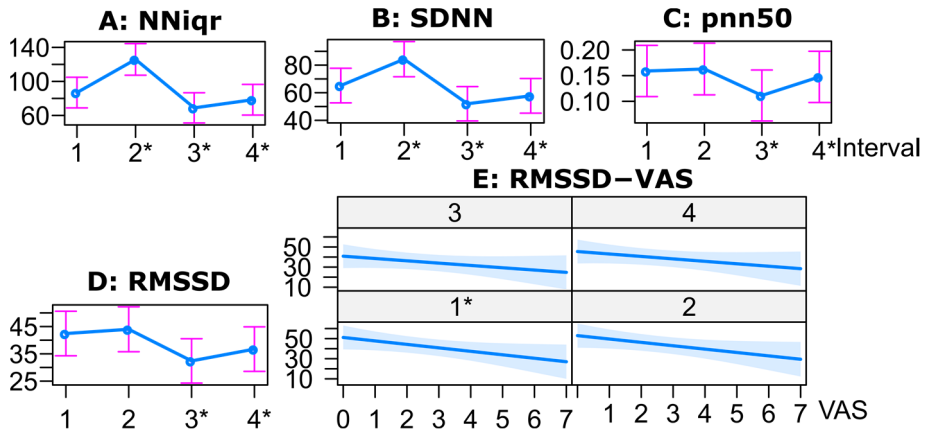


Figure 2. Estimates of NN intervals variability (D- RMSSD as a function of VAS per interval)

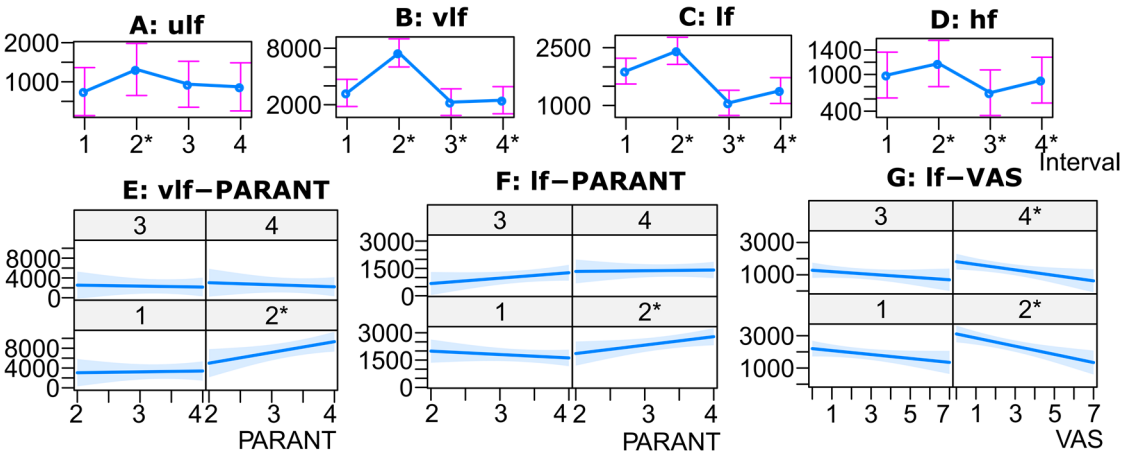


Figure 3. Frequency domain measures (E - vlf as a function of PARANT per interval; F – lf as a function of PARANT per interval; G – lf as a function of VAS per interval)

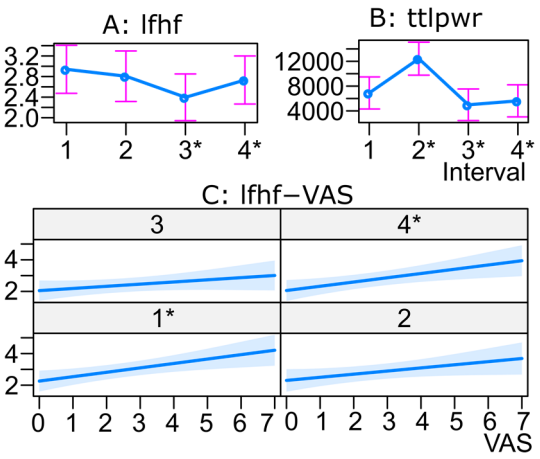


Figure 4. Low-to-high-frequency ratio and total power (C – lfhf as a function of VAS per interval)

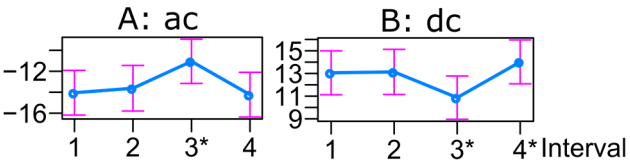


Figure 5. Other HRV measures.

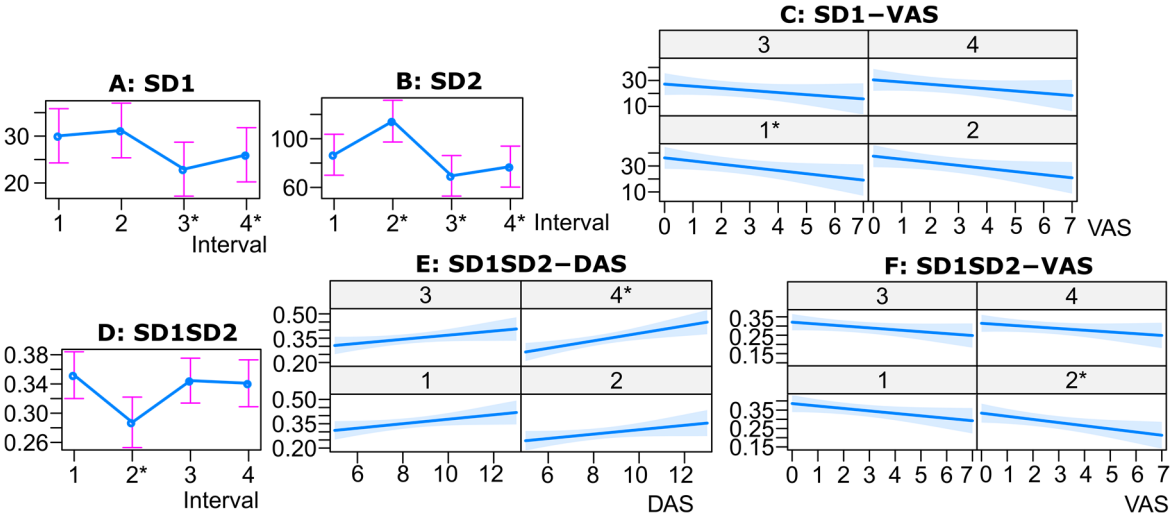


Figure 6. Poincaré plot features (E – SD1SD2 ratio as a function of DAS for each interval; F – SD1SD2 ratio as a function of VAS for each interval)

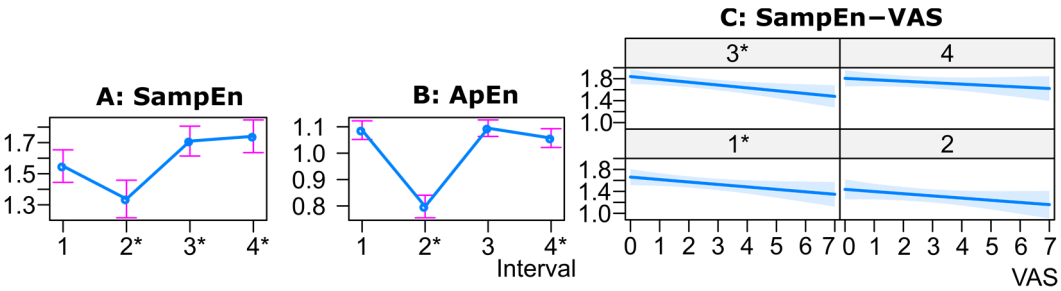


Figure 7 Entropy measures (C – SampEn as a function of VAS per interval)

IV. Discussion

Panel A in Fig. 1 shows that patients are excited because of the procedure since NNmean is less than 850 ms for all time intervals (average value can be considered to be 926 *ms* [26]). Mean and skewness of NN intervals were significantly different for all intervals compared to the baseline. Kurtosis however was significantly different only during the procedure. There is a significant correlation between skewness and DAS in the waiting room. Skewness decreases with PARANT during procedure but also in the waiting room indicating that patients might have been informed about the severity of the foregoing procedure.

In Fig. 2 NN variability (panels A, B) increases during anesthesia, indicating increased sympathetic activity caused by pain and/or fear [18]. Variability decreases significantly during procedure (for all estimates), indicating lower parasympathetic activity (panels C, D). Panel E shows that even though RMSSD drops with VAS for all intervals, the drop was significant only during baseline, which is interesting because there is no pain in the waiting room. This is however in accordance with previous research showing that lower rest parasympathetic activity means lower resilience to stress [25].

lf is significantly increased in the anesthesia segment, and significantly decreased during the procedure and after the procedure (Fig. 3). This indicates higher sympathetic activity during anesthesia and calming during procedure. This is in accordance with previous findings in fig. 2, and with panel D, which shows that parasympathetic activity follows sympathetic activity for homeostasis. The significant increase in sympathetic activity (lf) in interval 2 is probably not caused by pain, but by stress. This is because pain is accompanied by a decrease in parasympathetic activity [18]. Panel E and F show significant increase in power of low frequencies with PARANT during anesthesia, which is possibly because of more intensive anesthetizing required for more complicated procedures. Panel G shows that lf in interval 2 can be used as a predictor of perceived pain. Interestingly, sympathetic activity is significantly lower for patients who reported pain that is more intensive.

In Fig. 4 we see that lfhf ratio drops in time from the baseline value of about 2.9 which is slightly above the average value of 2.8 [26]. The drop becomes significant during the procedure, indicating imbalance in the nervous systems caused by the surgery. In other words, PNS becomes significantly more prevalent compared to SNS, which we already observed in the previous figure. This is in accordance with the previous study, which also showed a drop in lfhf ratio during the procedure [19]. Panel C shows an increase in lfhf ratio with higher reported pain. This is expected since sympathetic activity is increased with pain but considering we have seen decrease of PSN with VAS in the first time-interval on Fig. 2 panel E, the decrease of the lfhf ratio was probably because of a drop in PNS activity. It is interesting that the increase was significant also in the waiting room, indicating that anxiety before the procedure causes higher perception of pain. The anxiety before surgery (in the waiting room) might have been increased by the details of surgery that patients have received [27], which in turn could have increased the perception of pain.

Fig. 5. shows significantly decreased ac and dc during procedure.

Both SD1 and SD2 drop significantly during procedure, whereas the ratio SD1SD2 is significantly lower during anesthesia (Fig. 6). Panel C and F show that the SD1 and SD1SD2 could be used as indicators of VAS, whereas SD1 is significant already in the waiting room. We can see on panel E that SD1SD2 increases with DAS (significantly in interval 4).

Since HRV increases in segment 2 (Fig. 2), one might expect that entropy also increases, but surprisingly Fig. 7 shows a significant decrease in entropy. Panel C shows a significant correlation between SampEn and VAS, already in the waiting room.

V. CONCLUSION

Several HRV parameters significantly change during dental surgery, indicating changes in the balance between SNS and PNS, caused by anxiety and pain during the procedure. In particular, NNiqr, SDNN, and SD2 were significantly different in all time intervals with respect to the waiting room.

We confirmed the significance of the lfhf parameter for monitoring patients during surgery. Moreover, lfhf has potential for estimating postoperatively reported VAS, already in the waiting room. RMSSD, SD1, SD1SD2 as well as SampEn, also significantly correlated with VAS. On the other hand, NNskew and SD1SD2 have shown a significant correlation with DAS, which makes them DAS estimates. Even though vlf and lf significantly correlated with PARANT, it is only the NNskew that correlated significantly with PARANT during the procedure.

The estimates of HRV were calculated offline, but they could be integrated in the mobile application MobECG, that receives data in real time from the Savvy ECG monitor. This will enable real time assesment of stress, pain, and anxiety during dental procedures. Additionally, patients' responses to VAS and DAS could be predicted already in the waiting room.

REFERENCES

[1] H. S. Brand, "Cardiovascular responses in patients and dentists during dental treatment," *Int Dent J*, vol. 49, no. 1, 1999, doi: 10.1111/j.1875-595X.1999.tb00509.x.

[2] F. L. Liau *et al.*, "Cardiovascular influence of dental anxiety during local anesthesia for tooth extraction," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 105, no. 1, pp. 16–26, Jan. 2008, doi: 10.1016/j.tripleo.2007.03.015.

[3] A. Sharma, R. Pant, S. Priyadarshi, N. Agarwal, S. Tripathi, and M. Chaudhary, "Cardiovascular Changes Due to Dental Anxiety During Local Anesthesia Injection for Extraction," *J Maxillofac Oral Surg*, vol. 18, no. 1, 2019, doi: 10.1007/s12663-018-1085-4.

[4] R. Mestre Aspa, I. Carrera Grano, L. Berini Aytes, and C. Gay Escoda, "Pulsioxymetry monitorization during lower third molar extraction. A comparative study of three local anesthetics with epinephrine 1:100,000.," *Med Oral*, vol. 6, no. 3, 2001.

[5] B. Chernow *et al.*, "Hormonal Responses to Graded Surgical Stress," *Arch Intern Med*, vol. 147, no. 7, 1987, doi: 10.1001/archinte.1987.00370070087013.

[6] "Analgesia and the Metabolic Response to Surgery," *The Lancet*, vol. 325, no. 8436, pp. 1018–1019, May 1985, doi: 10.1016/S0140-6736(85)91618-6.

[7] R. E. Davenport, R. J. Porcelli, V. J. Iacono, C. F. Bonura, G. I. Mallis, and P. N. Baer, "Effects of Anesthetics Containing Epinephrine on Catecholamine Levels During Periodontal Surgery," *J Periodontol*, vol. 61, no. 9, 1990, doi: 10.1902/jop.1990.61.9.553.

[8] C. M. Hill, P. Mostafa, A. G. Stuart, D. W. Thomas, and R. v. Walker, "ECG variations in patients pre- and post-local anaesthesia and analgesia," *Br Dent J*, vol. 207, no. 12, 2009, doi: 10.1038/sj.bdj.2009.978.

[9] Jt. Lambrecht, A. Filippi, and J. Arrigoni, "Cardiovascular monitoring and its consequences in oral surgery," *Ann Maxillofac Surg*, vol. 1, no. 2, 2011, doi: 10.4103/2231-0746.92766.

[10] M. P. Shindova, A. Belcheva-Krivorova, and Z. Taralov, "Pulse oximetry in paediatric dentistry," *Folia Med (Plovdiv)*, vol. 64, no. 2, pp. 202–206, 2022, doi: 10.3897/folmed.64.e61136.

[11] I. Tomasic, N. Tomasic, R. Trobec, M. Krpan, and T. Kelava, "Continuous remote monitoring of COPD patients—justification and explanation of the requirements and a survey of the available technologies," *Med Biol Eng Comput*, vol. 56, no. 4, pp. 547–569, 2018, doi: 10.1007/s11517-018-1798-z.

[12] A. Rashkovska, M. Depolli, I. Tomašić, V. Avbelj, and R. Trobec, "Medical-grade ECG sensor for Long-Term Monitoring," *Sensors*, vol. 20, no. 6, pp. 1–17, 2020, doi: https://doi.org/10.3390/s20061695.

[13] A. Prkic, I. Tomasic, A. Lesin, T. Becic, D. K. Govorko, and I. M. Mikic, "Using patch ECG device during lower third molar surgery," *International Journal of Reliable and Quality E-Healthcare*, vol. 10, no. 4, 2021, doi: 10.4018/IJRQEH.2021100102.

[14] A. Lesin *et al.*, "Wireless ECG device for heart rhythm monitoring during dental extractions-a pilot study," in *2021 6th International Conference on Smart and Sustainable Technologies, SpliTech 2021*, 2021. doi: 10.23919/SpliTech52315.2021.9566379.

[15] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Med Biol Eng Comput*, vol. 44, no. 12, pp. 1031–1051, Dec. 2006, doi: 10.1007/s11517-006-0119-0.

[16] S. García-Gonzálvez, D. López-Plaza, and O. Abellán-Aynés, "Influence of Competition on Anxiety and Heart Rate Variability in Young Tennis Players," *Healthcare*, vol. 10, no. 2237, pp. 1–9, 2022, doi: 0.3390/healthcare10112237.

[17] Y. Cheng, M. Su, C. Liu, Y. Huang, and W. Huang, "Heart rate variability in patients with anxiety disorders: A systematic review and meta-analysis," *Psychiatry Clin Neurosci*, vol. 76, no. 7, pp. 292–302, Jul. 2022, doi: 10.1111/pcn.13356.

[18] J. Koenig, M. N. Jarczok, R. J. Ellis, T. K. Hillecke, and J. F. Thayer, "Heart rate variability and experimentally induced pain in healthy adults : A systematic review," vol. 18, pp. 301–314, 2014, doi: 10.1002/j.1532-2149.2013.00379.x.

[19] S. H. Le, K. Tonami, S. Umemori, L. T.-B. Nguyen, L. T.-Q. Ngo, and S. Mataki, "The potential of heart rate variability for exploring dental anxiety in mandibular third molar surgery," *Int J Oral Maxillofac Surg*, vol. 47, no. 6, pp. 809–815, Jun. 2018, doi: 10.1016/j.ijom.2018.01.019.

[20] K. I. Jones, F. Amawi, A. Bhalla, O. Peacock, J. P. Williams, and J. N. Lund, "Assessing surgeon stress when operating using heart rate variability and the State Trait Anxiety Inventory: will surgery be the death of us?," *Colorectal Disease*, vol. 17, no. 4, pp. 335–341, Apr. 2015, doi: 10.1111/codi.12844.

[21] N. L. Corah, "Development of a Dental Anxiety Scale," *J Dent Res*, vol. 48, no. 4, 1969, doi: 10.1177/00220345690480041801.

[22] S. Pathak, S. Vashisth, S. Mishra, S. P. Singh, and S. Sharma, "Grading of extraction and its relationship with post-operative pain and trismus, along with proposed grading for trismus," *Journal of Clinical and Diagnostic Research*, vol. 8, no. 6, Jun. 2014, doi: 10.7860/JCDR/2014/8560.4470.

[23] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, pp. E215-20, 2000, doi: 10.1161/01.CIR.101.23.e215.

[24] A. N. Vest *et al.*, "An open source benchmarked toolbox for cardiovascular waveform and interval analysis," *Physiol Meas*, vol. 39, no. 10, 2018, doi: 10.1088/1361-6579/aae021.

[25] M. Meeus *et al.*, "Heart rate variability in patients with fibromyalgia and patients with chronic fatigue syndrome: A systematic review," *Semin Arthritis Rheum*, vol. 43, no. 2, pp. 279–287, Oct. 2013, doi: 10.1016/j.semarthrit.2013.03.004.

[26] D. Nunan, G. R. H. Sandercock, and D. A. Brodie, "A Quantitative Systematic Review of Normal Values for Short-Term Heart Rate Variability in Healthy Adults," *Pacing and Clinical Electrophysiology*, vol. 33, no. 11, pp. 1407–1417, Nov. 2010, doi: 10.1111/j.1540-8159.2010.02841.x.

[27] S. K. S. Ng, A. W. L. Chau, and W. K. Leung, "The effect of pre-operative information in relieving anxiety in oral surgery patients," *Community Dent Oral Epidemiol*, vol. 32, no. 3, 2004, doi: 10.1111/j.1600-0528.2004.00161.x.

Optimizing Heartbeat Classification using Bézier Interpolation

M. Gusev*, N. Petrovski*, L. Tonkovikj*
* Sts Cyril and Methodius University in Skopje, Faculty of Computer Science and Engineering, North Macedonia
E-mail: marjan.gushev@finki.ukim.mk, {nikola.petrovski1, lucija.tonkovikj}@students.finki.ukim.mk

***Abstract*—Analog to digital conversion of electrocardiograms depends on the sampling frequency influencing the determination of a proper heartbeat location and precision of further digital processing. We set a research question to find the optimal number of interpolation points to reduce the mistakes in the similarity check of heartbeats and classify ventricular beats. In addition, another research question aims at finding the optimal number of interpolation points applying Bézier interpolation to reveal the optimal performance/cost ratio. Final research question is to find the sampling frequency that will reveal optimal performance in classification of ventricular beats. The experiments evaluate all neighbouring pairs of heartbeats from the standard benchmark MIT-BIH arrhythmia dataset resampled to a 125 Hz sampling frequency. The results show that even one more interpolation point, which corresponds to a sampling frequency of 250 Hz, will increase the performance versus the original 360 Hz sampling frequency. At the same time, the optimal is interpolation with additional five or seven points corresponding to 750 Hz, and 1000 Hz respectively. We found that a threshold value of 34 reveals the optimal performance to conclude a change between ventricular heartbeats and others, even in a 10-bit precision of the analog-digital conversion. The processing time and performance/cost-benefit analysis show that one interpolation point is the most beneficial.**

***Keywords*—Interpolation, ECG, upsampling**

I. Introduction

Detection of ventricular beats and arrhythmia disorders motivates researchers to develop new effective and efficient algorithms. The best available tool in their effort is to analyze the measured electrical heart activity represented by an electrocardiogram (ECG). Most of the approaches for automated detection are based on pattern matching and signal processing algorithms, or developed machine learning (ML) models. However, all these approaches depend on the quality to represent the ECG in the digital form, especially considering the requirement for efficient transfer and storage of data. On one side, efficiency requires a smaller amount of data to be transferred and processed, and on the other side, detection precision requires more data. The analog-to-digital conversion (ADC) unit is responsible to create samples with a specific size determined by the bit resolution. Samples are created on a predefined time interval according to the sampling frequency.

Each heartbeat can be determined by the R-peak in the ECG (the extreme in the QRS complex representing a heartbeat in the ECG), and the sampling frequency affects its real-time stamp (location). Most research analyzed the impact of the sampling frequency on the detection of the R peaks, compression techniques for ECG signals, or calculation of Heart Rate Variability (HRV) focusing on reconstructing the missing data within ECG sampling. In our recent paper [1], we address the problem of classifying a heartbeat type by performing a similarity check on two heartbeats. After correcting the proper time location with Bézier interpolation [2], we perform the similarity check of neighbouring heartbeats and evaluate the performance versus different threshold values to distinguish between a class of ventricular (V) and class (NonV) of other heartbeat types. A ventricular beat is usually determined by a wider width of the QRS complex, meaning that the heart has not followed the normal heartbeat activation procedure. In this paper, we continue this research aiming at finding the optimal sampling frequency and perform efficient transfer from the end-user devices (sensor and smartphone or other IoT device) allowing effective algorithms. We address two research questions: to 1) find the threshold value that reveals optimal performance, and 2) find the number of points for applying Bézier interpolation which reveal the best performance/cost ratio. We conduct experiments with two methods for Bézier interpolation: Method I as a standard construction of a four control points Cubic curve [3]; and Method II as a modified construction using a centripetal Catmull-Rom curve [4], [5].

The benchmark ECG dataset is the MIT-BIH ECG benchmark with 30-minute ECG measurements of 44 patients (excluding four measurements for patients with paced rhythm out of 48 available in the dataset). The original measurements are sampled at a 360 Hz sampling frequency (assumed as default), and in our experiments, the measurements were downsampled to a 125 Hz sampling frequency, which generates 2.88 fewer data to be transferred. The test cases included the construction of Bézier curves by adding 1, 3, 5, 7, and 15 extra samples between a pair of two analyzed samples, which corresponds to a sampling frequency of 250 Hz, 500 Hz, 750 Hz, 1000 Hz, and 2000 Hz. Finally, each test case is evaluated according to achieved sensitivity, positive predictive value, and F1 score to detect a V beat, avoiding accuracy due to the imbalance of the V class.

II. Related work

This research can potentially reveal the lowest sampling frequency and bit resolution the ADC unit should use for a sufficient performance of detecting the features in the ECG signal. The conclusions in the analyzed literature

for QRS detection show that sampling frequencies below 100 Hz will result in decreased performance [7], with a large number of wrong QRS detections for sampling frequencies below 200 Hz [8], while [9] concludes negligible performance discrepancies for sampling frequencies above 125 Hz having 120 Hz as a cut-off for sufficient performance. Recommendations suggest sampling frequencies of 128 Hz [7], or between 250 - 500 Hz [8].

Analyzing the bit resolution, researchers have concluded that 8-bit resolution is unacceptable, satisfactory results are obtained over 10 bits, and 12 bits are recommended ADC value [10].

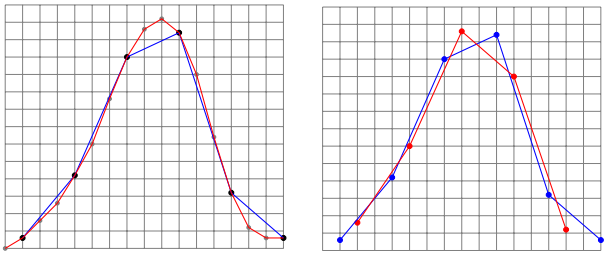
The impact of the sampling frequency on heart rate variability has been reported by several research papers [11]–[15]. Our earlier paper [1] contains a review of related work on the use of Bézier interpolation to improve ECG interpretation, the recovery of ECG missing parts with techniques tested on audio signals [16], or denoising techniques [17], developing a model of an ECG heartbeat [18], or analyzing ECG morphological features [19], [20], concluding that this method was not used to check the similarity of two heartbeats and reveal a tool to detect V versus NonV heartbeats.

III. Methods

Since the ECG is sampled at a particular sampling frequency, we do not know if the real R peak will be represented by the samples. Therefore, we aim at finding the real location of the R peak, and we need to search in the region between two points found to be extreme values. The main idea to apply this procedure is to transfer as little data as possible without losing precision performance and generate a digital presentation for a smooth curve and find a much better location of the time stamp.

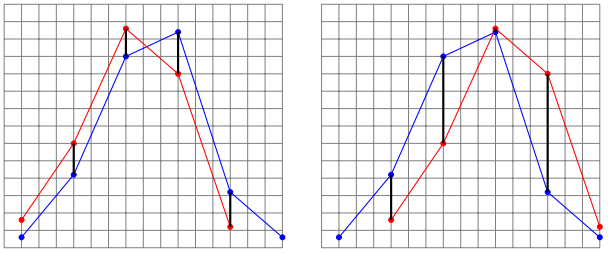
Experiments are performed on MIT BIH Arrhythmia ECG database (MITDB) [21] resampled to a sampling frequency of 125Hz and 10-bit resolution using the Mode-Median-Bucket downsampling method [6]. Each file of ECG samples is accompanied by an annotation file specifying the sample with a heartbeat. Heartbeats were annotated by two independent experts and a third one was engaged in the case of discrepancy. A heartbeat with wider QRS is classified as a V beat if it is annotated as a V (ventricular), E (escape), or F (fusion of ventricular and normal beat) type. All other beat types are classified as NonV class. The problem of detecting a wider QRS in the ECG is a complex task since the baseline fluctuates and detection of Q and S points may not be accurate, knowing that the ECG is usually measured on a real human where the muscle noise corrupts the ECG signal. This is why we apply the similarity check to find how much two heartbeats differ.

Fig. 1a) illustrates the difference between digital presentations sampled at different sampling frequencies (120 and 360 Hz), and Fig. 1b) shows differences in the digital presentations that depend on the location of the sample



a) different sampling freq. b) different start offsets

Fig. 1: QRS digital presentation: a) blue line sampled at 120 Hz and red line sampled at 360 Hz; b) both lines sampled at 120 Hz with different sample starts.



a) misaligned b) aligned to the R peak

Fig. 2: Similarity check of the same heartbeat sampled at 120 Hz with different sample starts - alignment according to the improper location of the R peak.

start. In both cases of the lower sampling frequency, the R peak can not be properly determined.

Similarity check between two heartbeats (blue and red lines in Fig. 2) is evaluated by a sum of Euclidean distances between corresponding samples in the digital representations (black lines in Fig. 2). In the case of a lower sampling frequency, the R peak will not be located on a proper location and a similarity check will generate large values even if the same heartbeat is checked with a different sampling start point.

In this paper, Bézier interpolation was applied to generate more samples to keep the ECG line smooth and determine the proper location of the R peak.

A. Method I

The first interpolation method follows the standard Cubic Bézier curve construction using four control points (P_1, P_2, P_3 and P_4, such that P_2 and P_3 do not have to be on the curve itself). Each segment S_i is defined by its starting point p_i and its ending point p_{i+1}. The segment S_i can be constructed using a Cubic Bézier curve, such that $P_1 = p_i$ and $P_4 = p_{i+1}$. The final graph is constructed by merging all of the segments together. However, the control points P_2 and P_3 remain to be found for each segment.

In our method, we use the algorithm [3] that starts with the Cubic Bézier curve $B(t)$ equation (1), where t

represents how far $B(t)$ is from p_i to p_{i+1}.

$$\begin{aligned} B(t) =&(1-t)^3 \cdot P_1 + 3t(1-t)^2 \cdot P_2 \\ &+ 3t^2(1-t) \cdot P_3 + t^3 \cdot P_4 \end{aligned} \tag{1}$$

Each segment S_i is constructed using a Cubic Bézier curve, such that $P_1 = p_i$, $P_2 = a_i$, $P_3 = b_i$ and $P_4 = p_{i+1}$, as represented in (2).

$$\begin{aligned} S_i(t) = (1-t)^3 \cdot p_i + 3t(1-t)^2 \cdot a_i \\ +3t^2(1-t) \cdot b_i + t^3 \cdot p_{i+1}, \quad 0 \leq i < 20 \end{aligned} \tag{2}$$

The transition between segment S_i and S_{i+1} is smooth around the point p_{i+1}, meaning that the first and second derivatives of S_i and S_{i+1} have to be equal in point p_{i+1}. (3) is obtained assuming that the segment S_i is in point p_{i+1} for $t = 1$, while the segment S_{i+1} is in the same point for $t = 0$.

$$\begin{aligned} S_i'(t=1) = S_{i+1}'(t=0), \quad i = 0, 1, \ldots, 19 \\ S_i''(t=1) = S_{i+1}''(t=0), \quad i = 0, 1, \ldots, 19 \end{aligned} \tag{3}$$

To find all missing control points (a_i and b_i) for all n segments, we need a system of $2n$ linear equations. However, only $2(n-1)$ can be obtained from (3). Therefore, the following (arbitrary) boundary conditions are imposed:

$$\begin{aligned} S_0''(t=0) = 0 \\ S_{n-1}''(t=1) = 0 \end{aligned} \tag{4}$$

(5) and (6) respectively present the first and second derivative of (2).

$$\begin{aligned} S'(t) =&3[-(1-t)^2 \cdot p_i + (1-3t)(1-t) \cdot a_i \\ &+ t(2-3t) \cdot b_i + t^2 \cdot p_{i+1}] \end{aligned} \tag{5}$$

$$\begin{aligned} S''(t) =&6[(1-t) \cdot p_i + (3t-2) \cdot a_i \\ &+ (1-3t) \cdot b_i + t \cdot p_{i+1}] \end{aligned} \tag{6}$$

(7) and (8) are constructed substituting (5) and (6) in (3) and (4).

$$\begin{aligned} a_{i+1} + b_i = 2 \cdot p_{i+1} \\ a_i + 2a_{i+1} = 2b_i + b_{i+1} \\ p_0 - 2a_0 + b_0 = 0 \\ a_{n-1} - 2b_{n-1} + p_n = 0 \end{aligned} \tag{7}$$

$$\begin{cases} a_{i+1} + b_i = 2 \cdot p_{i+1}, & i = 0, 1, \ldots, n-1 \\ a_i + 2a_{i+1} = 2b_i + b_{i+1}, & i = 0, 1, \ldots, n-1 \\ p_0 - 2a_0 + b_0 = 0 \\ a_{n-1} - 2bn - 1 + p_n = 0 \end{cases} \tag{8}$$

The matrix form of the system is given in (9) for easier computation.

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 4 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 4 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 4 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 4 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 2 & 7 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{17} \\ a_{18} \\ a_{19} \end{bmatrix} = \begin{bmatrix} P_0 + 2 \cdot P_1 \\ 2(2 \cdot P_1 + P_2) \\ 2(2 \cdot P_2 + P_3) \\ \vdots \\ 2(2 \cdot P_{17} + P_{18}) \\ 2(2 \cdot P_{18} + P_{19}) \\ 8 \cdot P_{19} + P_{20} \end{bmatrix} \tag{9}$$

All a_i control points are obtained by solving (9), and all missing b_i points can easily be calculated by (10).

$$\begin{cases} b_i + 2 \cdot p_{i+1} - a_{i+1}, & i = 0, 1, \ldots, 18 \\ b_{19} = \frac{a_{19} + p_{20}}{2} \end{cases} \tag{10}$$

B. Method II

The second interpolation method is a modified Cubic Bézier curve construction, which converts a centripetal Catmull-Rom curve [5] to a Cubic Bézier form. It uses 4 consecutive points on the ECG graph (P_0, P_1, P_2, P_3) to construct the curve between the points P_1 and P_2 by generating two additional control points T_1 and T_2.

To process a given heartbeat we start by extracting 23 points $(p_i, i \in [0, 22])$, which consist of: the peak (p_{11}), 10 points before the peak $(p_i, i \in [1, 10])$, 10 points after the peak $(p_i, i \in [12, 21])$ and 2 points before and after the heartbeat (p_0 and p_{22}). Similarly to the first method, the heartbeat is divided into 20 segments $(S_i, i \in [1, 20])$, where S_i is the segment from point p_i to point p_{i+1}. The two additional points (p_0 and p_{22}), are only used to assist in the generation of segments S_1 and S_{20}.

Interpolation of a given segment S_i starts off by selecting the 4 consecutive points (P_0, P_1, P_2, P_3) needed for the generation of the curve running from p_i to p_{i+1}. The points used for the generation of the segment S_i are assigned in the following way: $P_0 = p_{i-1}$, $P_1 = p_i$, $P_2 = p_{i+1}$ and $P_3 = p_{i+2}$. Following the assignment of the points, we must generate the required control points T_1 and T_2. In this paper, we use a fascinating proposal for the creation of these points given by Cem Yuksel, Scott Schaefer and John Keyser in [5] and additionally implemented by [4], based on (11), where the four points (P_0, P_1, P_2, P_3) are the previously assigned points and the values d_1, d_2, d_3 are described such that $d_i = |P_i - P_{i+1}|^\alpha = \sqrt{|P_i - P_{i+1}|}$,

and $\alpha = \frac{1}{2}$ due to the Catmull-Rom curve being centripetal.

$$T_1 = \frac{d_1^2 P_2 - d_2^2 P_0 + (2d_1^2 + 3d_1 d_2 + d_2^2) P_1}{3d_1(d_1 + d_2)}$$
$$T_2 = \frac{d_3^2 P_1 - d_2^2 P_3 + (2d_3^2 + 3d_3 d_2 + d_2^2) P_2}{3d_3(d_3 + d_2)} \quad (11)$$

Following the generation of the control points T_1 and T_2, we will use the points $\{P_1, T_1, T_2, P_2\}$ in a Cubic Bézier curve equation, denoted by $B(t)$ (for $t \in [0, 1]$), given in (12) to generate the interpolated points for a given segment S_i.

$$B(t) = (1-t)^3 \cdot P_1 + 3t(1-t)^2 \cdot T_1 + 3t^2(1-t) \cdot T_2 + t^3 \cdot P_2 \quad (12)$$

C. Experiment testbed

A threshold value for the calculation of the similarity check is used to detect whether a beat belongs to the same class as the previous one. We test each pair of successive neighbouring heartbeats in MITDB and calculate the similarity coefficient as a sum of Euclidean distances between corresponding 21 samples around the R peak.

The following experiments are specified according to the applied method:

- *R360* - raw original ECG benchmark data sampled at 360 Hz,
- *R125* - data resampled at a sampling frequency of 125 Hz, and
- *B1,B3,B5,B7,B15* - data sampled at 125 Hz adding 1, 3, 5, 7, and 15 points applying Bézier interpolation.

Both methods of Bézier interpolation are tested.

D. Evaluation method

The MITDB ECG benchmark contains 100733 heartbeats with a large imbalance of the V beats versus NonV class. We evaluate relevant statistical measures: sensitivity (SEN) as a ratio of correctly detected out of all V beats, positive predictive value (PPV) as a portion of correctly detected out of all V detections, specificity (SPC) as a ratio between correctly detected NonV beats out of all NonV beats, and negative predictive value (NPV) as a portion of correctly detected NonV beats out of all NonV detections. However, due to the class imbalance, we focus on F1 score as a relevant performance measure.

An improvement factor (expressed in %) is calculated as a ratio between the achieved performance of applied Bézier interpolation method and the benchmark with data sampled at 360 Hz and resampled at 125 Hz. The performance/cost ratio aims at finding the experiment that revealed the highest performance versus cost. Also, we calculate the gain as a ratio between the increased performance (improvement factor) and increased cost, both expressed in %. If the GAIN is more than 1, then the experiment results with higher performance versus the cost.

TABLE I: Average execution times for Bézier interpolation to process 21 samples around R peak.

Exp.	Method I	Method II
B1	1.41 ms	0.25 ms
B3	2.23 ms	0.32 ms
B5	2.92 ms	0.39 ms
B7	3.66 ms	0.46 ms
B15	6.62 ms	0.75 ms

TABLE II: Similarity check performance (in %) for optimal threshold values Thr applying Bézier curves.

Method I	Thr	SEN	SPC	PPV	NPV	F1
360	-	89.21	96.87	82.24	98.22	85.58
125	-	87.29	96.70	81.15	97.91	84.11
B1	34.0	89.50	96.87	82.30	98.27	85.75
B3	34.0	89.74	96.95	82.71	98.31	86.08
B5	34.0	89.35	97.06	83.19	98.25	86.16
B7	33.0	88.84	97.17	83.61	98.17	86.15
B15	33.0	89.38	97.06	83.19	98.25	86.18
Method II	Thr	SEN	SPC	PPV	NPV	F1
360	-	89.21	96.86	82.22	98.22	85.58
125	-	87.29	96.70	81.13	97.91	84.10
B1	34.0	89.45	96.87	82.31	98.26	85.73
B3	34.0	89.56	96.96	82.75	98.28	86.02
B5	34.5	89.17	97.08	83.22	98.22	86.09
B7	34.5	89.20	97.07	83.18	98.22	86.09
B15	34.5	89.22	97.07	83.22	98.23	86.12

IV. Evaluation of Results

We conducted seven experiments of calculating the similarity check of successive neighbouring heartbeats. Bézier interpolation was applied in the last five experiments with both Method I and Method II.

Table I presents the average execution times for interpolating 1, 3, 5, 7 and 15 points for each detected heartbeat in the analyzed dataset. Optimal threshold values and corresponding performance results are presented in Table II.

In our earlier paper we concluded that both methods revealed almost equal performance (Table II) although Method II was much faster than Method I (Table I). In this paper, we focus on optimization and determine the threshold value which differentiates the similarity check coefficient if the beat has changed the class was found to be approximately 34 with slight differences for different experiments.

Fig. 3 presents details on optimal threshold values and the corresponding performance measures. Adding one point for Bézier interpolation (experiment B1), the threshold value of 34 yields the maximum F1 score of 85.7%. The results for different values of interpolated points per section are very similar. When interpolating three, five or seven points per section (Experiments B3 and B5 and B7), the maximum value for the F1 score was slightly increased to 86%, with the ideal threshold being 34 in both cases.

Additionally, Experiment B1 Method I results (Fig. 4) present the performance measures SEN, SPC, PPV, NPV

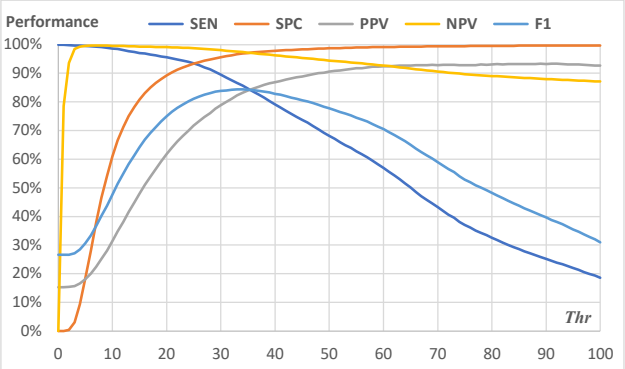


Fig. 3: Performance (in %) versus threshold values for Bézier interpolation (experiment B1 Method I)

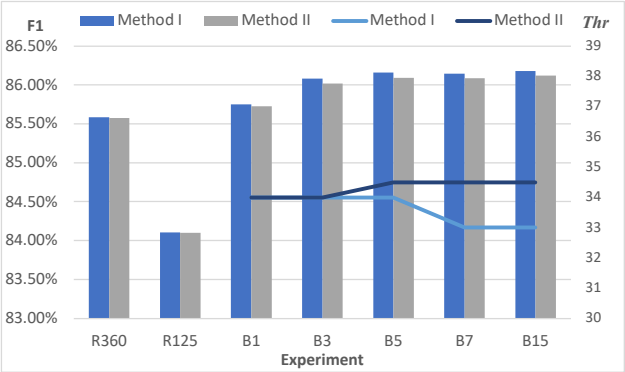


Fig. 4: Comparing optimal threshold values and corresponding F1 Scores of both Bézier interpolation methods

and F1. The lower threshold is the higher values of SEN and lower values of PPV, and the opposite, as threshold values get higher, SEN decreases and PPV increases. The optimal value is obtained for $Thr = 34$. The same trends and behaviour are observed for performance measure values in other experiments with both methods of Bézier interpolation.

Table III presents the improvement factor in % comparing the achieved F1 score versus the original benchmark (with a sampling frequency of 360 Hz), and also versus data obtained with downsampling to 2.88 times lower frequency (125 Hz). As expected, downsampling showed a decrease in performance by a factor of -1.73. However, we found that even adding one point with Bézier interpolation (experiment B1) achieves a slightly better performance (0.2%). This is a very interesting result, meaning that instead of transferring almost three times more data, we apply rather simple calculations and obtain better performance, which is a good lead toward ECG compression techniques.

Table IV presents the performance/cost ratio, revealing the optimal to be B1 experiment (adding one point) with method II (we get the highest performance for the smallest cost).

TABLE III: Improvement factor (in %) for similarity check with optimal threshold value when applying Bézier interpolation (Methods I and II)

	IF versus 360		**IF versus 125**	
Exp.	Method I	Method II	Method I	Method II
360	/	/	1.76	1.76
125	-1.73	-1.73	/	/
B1	0.20	0.18	1.96	1.94
B3	0.58	0.52	2.35	2.29
B5	0.67	0.60	2.44	2.37
B7	0.66	0.60	2.43	2.36
B15	0.69	0.63	2.46	2.40

TABLE IV: Performance/cost ratio for similarity check with optimal threshold value when applying Bézier interpolation (Methods I and II)

Exp.	Method I	Method II
B1	0.608	3.429
B3	0.386	2.688
B5	0.295	2.208
B7	0.235	1.871
B15	0.130	1.148

Adding more points with Bézier interpolation (experiments B3, B5, B7 or B15) achieves better performance, and the performance trend saturates after three added points. Table V presents the increased F1 score values (expressed in %) and also increased cost (measured as an average response time for calculations applying the corresponding Bézier interpolation method).

We observe that GAIN of 0.64 is a very small value to increase the number of interpolated points from one to three, and this value is even smaller for a larger number of points. This suggests that adding even one point with Bézier interpolation will increase the performance, with the highest performance/cost ratio, and the higher performance achieved with more points is small and is not cost-effective.

The recommendations of the American Heart Association' [22] suggest that the minimum sampling frequency for a digital ECG recording should be 500 Hz.

However, high sampling rates imply increased network throughput and processing requirements demanding computers with larger memories and faster processors. The results from the presented experiments show that a lower sampling rate would retain adequate information for ECG representation without introducing significant errors in the results.

Pizzuti et al. [11] conclude that a sampling frequency of 500 Hz gives redundant information and recommend 250 Hz without loss of statistical accuracy for ECG measurements. Kwon et al. [15] show that down-sampling to 500 or 250 Hz resulted in excellent concordance, while down-sampling to 100 Hz produced acceptable results for time-domain HRV analysis, but not for frequency-domain analysis, and down-sampling to 50 Hz proved to be unacceptable. The overall conclusion that a 250-Hz

TABLE V: GAIN versus B1 for similarity check with optimal threshold value when applying Bézier interpolation (Methods I and II)

	Method I			Method II		
Exp.	inc F1	inc cost	GAIN	inc F1	inc cost	GAIN
B3	20.06	58.16	0.35	17.87	28.00	0.64
B5	24.74	107.09	0.23	22.34	56.00	0.40
B7	23.99	159.57	0.15	21.88	84.00	0.26
B15	25.93	369.50	0.07	23.92	200.00	0.12

sampling frequency is acceptable for HRV analysis is also confirmed by our approach.

Pizzuti et al. [11] show a significant decrease in differences among the amplitude measurements between a sampling rate of 500 Hz and 125 Hz. down to 65% for the detection of the S or 83% for the detection of R in the QRS wave.

V. Conclusion

We have conducted experiments to determine the optimal threshold value for similarity check and detection (classification) of ventricular beats versus normal beats when applying Bézier interpolation. The results confirmed that a threshold value of 34 is the optimal value. Our experiments confirmed that the best performance/cost ratio is obtained by adding one more interpolation point. An ECG digital representation sampled at 125 Hz reveals sufficient information, and adding one more point (upsampling to 250 Hz) with Bézier interpolation reveals optimal performance.

Downsampling from 360 Hz to 125 Hz has decreased the similarity check performance by 1.73%. Interestingly adding one sample by Bézier interpolation, on downsampled data from 360 Hz to 125 Hz improves the similarity check by 0.2%, that is, it finds R peak more precise than the original 360 Hz sampled data. We can conclude it is good to realize sampling at a lower frequency (such as 125 Hz), transfer a smaller amount of data, add only one point to obtain a smooth curve with Bézier interpolation and then process data.

Although the highest improvement is revealed for adding 15 samples (corresponding to a sampling frequency of 2000 Hz), we observe a saturated performance increase after adding three more samples, which can be concluded that after sampling the ECG with a sampling frequency of more than 500 Hz will not be beneficial.

This research can potentially ensure that 125 Hz is the lowest sampling frequency and the 10-bit bit resolution of the ADC unit should provide sufficient performance for detecting the features in the ECG signal. The added value is running Bézier interpolation and adding one more point for upsampling that will reveal increased performance in detecting specific features in the ECG, such as classification of ventricular beats.

References

[1] M. Gusev, L. Tonkovikj, and N. Petrovski, "Effect of bézier interpolation on similarity of heartbeats in electrocardiograms," in *2022 30th Telecommunications Forum (TELFOR)*, pp. 1–4, IEEE, 2022.

[2] M. E. Mortenson, *Mathematics for computer graphics applications*. Industrial Press Inc., 1999.

[3] O. Aflak, "Bézier interpolation," 2020. Last accessed online on 15.09.2022 at: https://towardsdatascience.com/b%C3%A9zier-interpolation-8033e9a262c2.

[4] A. Joshi, "Constructing a cubic bezier that passes through four points," 2019. Last accessed online on 23.09.2022 at: https://apoorvaj.io/cubic-bezier-through-four-points/.

[5] C. Yuksel, S. Schaefer, and J. Keyser, "Parameterization and applications of catmull–rom curves," *Computer-Aided Design*, vol. 43, no. 7, pp. 747–755, 2011.

[6] S. Steinarsson, *Downsampling time series for visual representation*. PhD thesis, 2013.

[7] G. G. Berntson, K. S. Quigley, J. F. Jang, and S. T. Boysen, "An approach to artifact identification: Application to heart period data," *Psychophysiology*, vol. 27, no. 5, pp. 586–598, 1990.

[8] M. Malik, "Task force of the european society of cardiology and the north american society of pacing and electrophysiology. heart rate variability. standards of measurement, physiological interpretation, and clinical use," *Eur Heart J.*, vol. 17, pp. 354–381, 1996.

[9] R. J. Ellis, B. Zhu, J. Koenig, J. F. Thayer, and Y. Wang, "A careful look at ECG sampling frequency and r-peak interpolation on short-term measures of heart rate variability," *Physiological measurement*, vol. 36, no. 9, p. 1827, 2015.

[10] E. Ajdaraga and M. Gusev, "Analysis of sampling frequency and resolution in ecg signals," in *2017 25th Telecommunication Forum (TELFOR)*, pp. 1–4, IEEE, 2017.

[11] G. Pizzuti, S. Cifaldi, and G. Nolfe, "Digital sampling rate and ecg analysis," *Journal of biomedical engineering*, vol. 7, no. 3, pp. 247–250, 1985.

[12] L. Hejjel and E. Roth, "What is the adequate sampling interval of the ecg signal for heart rate variability analysis in the time domain?," *Physiological measurement*, vol. 25, no. 6, p. 1405, 2004.

[13] M. Gusev, S. Tudjarski, and A. Anagelevska, "Sampling rate impact on heart rate variability," in *2022 30th Telecommunications Forum (TELFOR)*, pp. 1–4, IEEE, 2022.

[14] S. Mahdiani, V. Jeyhani, M. Peltokangas, and A. Vehkaoja, "Is 50 hz high enough ecg sampling frequency for accurate hrv analysis?," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5948–5951, IEEE, 2015.

[15] O. Kwon, J. Jeong, H. B. Kim, I. H. Kwon, S. Y. Park, J. E. Kim, and Y. Choi, "Electrocardiogram sampling frequency range acceptable for heart rate variability analysis," *Healthcare informatics research*, vol. 24, no. 3, pp. 198–206, 2018.

[16] A. Prieto-Guerrero, C. Mailhes, and F. Castanié, "Lost sample recovering of ecg signals in e-health applications," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 31–34, IEEE, 2007.

[17] S. M. M. Islam, M. S. Farid, and M. A. Kiber, "Denoising ecg signal using adaptive filter algorithms and cubic spline interpolation for regaining missing data points of ecg in telecardiology system," *International Journal of Enhanced Research in Science, Technology and Engineering*, vol. 4, no. 12, 2015.

[18] R. Soontornvorn, H. Fujioka, V. Chutchavong, and K. Janchitrapongvej, "Modeling ecg waveform using optimal smoothing bézier-bernstein curves," in *TENCON 2017-2017 IEEE Region 10 Conference*, pp. 1235–1238, IEEE, 2017.

[19] E. K. Roonizi and R. Sameni, "Morphological modeling of cardiac signals based on signal decomposition," *Computers in biology and medicine*, vol. 43, no. 10, pp. 1453–1461, 2013.

[20] R. Chaturvedi and Y. Yadav, "Analysis of ecg signal by polynomial approximation," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 2, no. 5, pp. 1029–1033, 2014.

[21] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[22] H. Pipberger, G. Baule, A. Berson, S. Briller, D. Geselowitz, L. Horan, and O. Schmitt, "Recommendations for standardization of instruments in electrocardiography and vectorcardiography," *IEEE Trans. Biomed. Eng.*, no. 1, pp. 60–68, 1967.

Fetal Monitoring: Multi-Channel Fetal ECG Denoising Based on Artificial Intelligence Approach

D.D. Taralunga*
*University Politehnica from Bucharest, Bucharest, Romania
dragos.taralunga@upb.ro

***Abstract* - Continuous electronic fetal monitoring using cardiotocography (CTG) represents the standard of evaluating the health status of the fetus and the risk of the pregnancy, in developed countries. However, the CTG has many limitations: high false positive rates, cannot be used for long term monitoring, poor sensitivity, it offers just the fetal heart rate and its variability etc. In this context, the fetal electrocardiogram (fECG) signal is used to obtain additional diagnostic information. On the other hand, the standard in clinical practice for obtaining the fECG is invasive, can pose a risk for both mother and fetus, can only be used during birth (very limited time window). An alternative is the abdominal fECG, that is recorded using a matrix of electrodes placed on the maternal abdomen. This approach is noninvasive and can be used for long term monitoring. The main drawback is the small signal to noise ratio for the abdominal fECG. Thus, the challenge is to isolate the fECG signal from other types of noise that are recorded by the abdominal electrodes: the maternal electrocardiogram (mECG), the electromyogram (EMG), the electrohysterogram (EHG), power line interference (PLI) etc. In this paper the author proposes an algorithm based on artificial neural network approach to extract the fECG signal waveform from abdominal recorded signals (ADS). The performance evaluation of proposed approach is realized on a database with simulated abdominal signals. A comparison is introduced, with other approaches described in literature for fECG denoising from abdominal signals.**

Keywords–fetal monitoring; artificial intelligence, fetal electrocardiogram, abdominal recorded signals

I. INTRODUCTION

Electronic fetal monitoring (EFM) is a technique used to track the fetal heart rate (FHR) and uterine contractions during labor and delivery. The goal of EFM is to ensure the well-being of the fetus and identify any signs of distress that may require intervention to prevent harm to either the fetus or the mother. EFM has been widely used in obstetrics for several decades and is considered the standard of care for many high-risk pregnancies.

There are two main types of EFM: external and internal. External EFM uses ultrasound and a tocodynamometer placed on the mother's abdomen to monitor the fetus heart rhythm and the contractions of the uterus. This is the most commonly used method and is non-invasive, making it a popular choice for low-risk pregnancies. Internal EFM, on the other hand, uses an electrode attached to the fetal scalp to measure the direct fetal electrocardiogram (fECG) signal that offers a more accurate FHR. This method is more invasive, as it requires the insertion of a fetal scalp electrode into the uterus, but it provides more accurate readings of the FHR and is usually reserved for high-risk pregnancies.

The main drawback of external EFM technique are: it offers an average determined FHR, it cannot be used for long term monitoring, there is a high intra- and inter-observer variability that sometimes leads to unnecessary C-section interventions. On the other hand, the internal EFM is invasive, can pose a risk for both mother and fetus and it is restricted to be used only in the short window of delivering (after the rupture of the membrane).

An alternative technique that can tackle all the drawbacks of the current EFM approaches is to record the fECG signal via electrodes placed on the maternal abdomen. Hence, it is non-invasive, it offers not only the FHR but also the beat-to-beat morphological variations of the fECG signal, *i.e.*, more diagnostic information, and it allows for continuous monitoring of the fetal health status during labor, which is important for detecting any changes in fetal well-being.

However, the main drawback is the small signal to noise ratio (SNR) for the abdominal fECG. Thus, the challenge is to isolate the fECG signal from other types of noise that are recorded by the abdominal electrodes: the maternal electrocardiogram (mECG), the electromyogram (EMG), the electrohysterogram (EHG), power line interference (PLI) etc.

There are many different approaches described in literature that try to extract the abdominal fECG: adaptive filtering [1], [2], wavelet analysis [3]–[5], blind source separation [6]–[8], empirical mode decomposition [9], [10], etc.

With the development of artificial intelligence techniques, there are few researchers that considered machine and deep learning approaches for denoising the abdominal fECG. In [11] the authors propose a encoder – decoder framework based on deep convolutional network for removing different types of noise from abdominal recorded signals (ADS). Fotiadou *et al* introduce a multi-channel deep convolutional encoder – decoder network that is able to learn how to obtain the best combination of

input channels in order to offer at the output, a clean multi-channel fECG signal [12].

In this paper, a multi-channel blind source separation method for extracting the fECG signal from abdominal recordings is introduced, which is based on deep-learning stacked-long short-term memory (LSTM) network. The model is evaluated on simulated abdominal signals.

II. Materials and Methods

A. Problem Definition

Each channel from a multi-channel abdominal recorded signals represents a mixture of different types of signals: fECG, mECG, EHG, EMG, PLI etc. Let $\mathbf{x}(t)$ be the set of multi-channel abdominal signals, then $\mathbf{x}(t) = \{\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots \mathbf{x}_m(t)\}^T$, where m is the number of abdominal recorded channels (the number of electrodes placed on the maternal abdomen). Let $\mathbf{s}(t)$ be the independent signal sources, then $\mathbf{s}(t) = \{\mathbf{s}_1(t), \mathbf{s}_2(t), \ldots \mathbf{s}_p(t)\}^T$, where p is the number of independent source signals (e.g. fECG, mECG, PLI, EMG etc). The recorded mixture of signal sources on different locations on the maternal abdomen can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where $\mathbf{A}$ is the mixing matrix, $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{m \times p}$. Hence, an abdominal channel recorded by the ith electrode is $\mathbf{x}_1(t) = a_{i,1}\mathbf{s}_1(t) + a_{i,2}\mathbf{s}_2(t) + \ldots a_{i,p}\mathbf{s}_p(t)$.The aim is to obtain the estimated sources, $\mathbf{s}^*(t) = \{\mathbf{s}^*_1(t), \mathbf{s}^*_2(t), \ldots \mathbf{s}^*_p(t)\}^T$,by separating the set of abdominal signals, through the determination of an, so called, unmixing matrix, $\mathbf{B} = [\mathbf{b}_{i,j}]^{p \times m}$:

$$\mathbf{s}^*(t) = \mathbf{B}\mathbf{x}(t) \quad (2)$$

If $m>p$ then the system of equations is overdetermined and a linear approach can be used for determining the unmixing matrix. Usually, other types of noise can be easily removed from the abdominal signals in the preprocessing step, thus one can assume that the abdominal signal is a mixture of mECG and fECG. In this case, at least two abdominal channels are required.

In this paper an artificial intelligence approach, based on stacked – LSTM network, is used to estimate the source signals. Hence, the multi-dimensional mapping of the mixing matrix in the neural network is obtained by training the weight coefficients, as opposed to the classic blind source separation method, which generates the mixing matrix by iterative calculation.

B. Proposed Method

Figure 1 depicts the components that compose the proposed Stacked-LSTM network structure: an encoding component, a separation component, and a decoding component. The mixed signal's features are represented in the encoding phase. To determine the source signal mask, the separation component is trained using a stacked LSTM block, and the source signal waveform is restored using the decoding component.

The encoding feature representation of the abdominal signals as input is extracted using 1-D convolution layer.

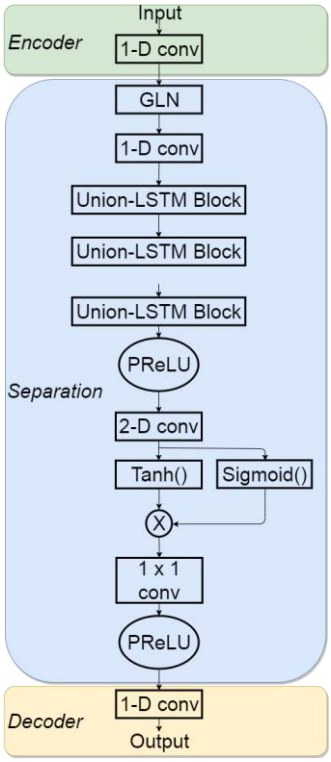


Figure 1. The proposed separation structure based on the Stacked-LSTM network

Linear encoding is used to encode mixed signals for further processing. The abdominal signal is processed using 512 convolution kernels to create a multi-dimensional coding feature that serves as input for the separation network.

Next, the source signal mask is generated using the separation component of the proposed Stacked-LSTM network. The separation phase involves determining the mask for each source signal, which is achieved through training the network. The mask obtained for each source signal corresponds to the mixing matrix A in (1). The separation component has the following steps:

- Group normalization is performed, i.e. the channels are grouped and the variance and mean are computed for each group. The 1-D convolution layer is used to extract the features.
- The extracted feature sequence is used as input for the stack Union-LSTM that consists of six LSTM blocks.
- Next, the parametric rectified linear unit (PReLU) activation function is used in order to prevent the he vanishing gradient problem. The Tanh() and Sigmoid() activation function are used for obtaining the multi-dimensional characteristics of the mixed source signal and, respectively the mask information of the two source signals.
- The time domain mask obtained in the previous step is multiplied with the encoding feature representation of the mixed signal in order to obtain the feature encoding of the source signals.

The final 1-D deconvolutional layer is used to decode the source signals and obtain a time-domain waveform.

The performance index used to evaluate the separation process is the scale-invariant source-to-noise ratio (SI-SNR) (a higher SI-SNR determines a lower separation error). Normalization is used prior to the calculation of the SI-SNR, to maintain the same scale.

$$SI\text{-}SNR = -10\log\frac{\|s_{target}\|^2}{\|s^* - s_{target}\|^2} \quad (3)$$

$$s_{target} = \frac{\langle s^*, s\rangle s}{\|s\|^2} \quad (4)$$

Equation (3) represents the loss function used during the training of the network with the gradient descent method.

C. Dataset

The mECG and fECG are generated using an electrocardiogram generator described in [13], ECGSYN. It creates a synthetic ECG signal with customizable parameters such as, number of beats, sampling frequency, mean heart rate, waveform morphology, low frequency (LF) / high frequency (HF) ratio, standard deviation of the RR interval. By utilizing a model that involves three linked ordinary differential equations, ECGSYN can generate an ECG signal that mimics many features of a human ECG, such as respiratory sinus arrhythmia, beat-to-beat variations in morphology and timing, R-peak amplitude modulation, and QT dependence on heart rate [13].

The model creates a path in a 3D state-space with coordinates (*x, y, z*). The ECG's quasi-periodic nature is represented by the trajectory's motion around a limit cycle with a radius of one in the (*x, y*) plane. Each complete revolution of this circle represents one heartbeat or one RR interval. The trajectory's movement in the *z* direction mimics the interbeat variability in the ECG signal. The P, Q, R, S, and T waves of the ECG, are determined by events linked to negative and positive attractors / repellers in the *z* direction [13].

Three differential equations are used to describe the dynamical motion:

$$\dot{x} = \alpha x - \omega y$$
$$\dot{y} = \alpha y - \omega x$$
$$\dot{z} = -\sum_{i\in\{P,Q,R,S,T\}} a_i \Delta\theta_i exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \quad (5)$$

where $\alpha = 1 - \sqrt{x^2 + y^2}$, $\Delta\theta_i = (\theta - \theta_i)\ mod\ 2\pi$, $\theta = atan2(y, x)$ and ω is the angular velocity of the trajectory around the limit cycle.

With the help of the signal generator described, the mECG and fECG signals are obtained with specific characteristics for adult and fetal heart, respectively. Thus, the RR intervals in the fECG signal are chosen to be as least double of the ones in the mECG signals, since usually, the normal FHR is much higher (almost double) than the maternal HR. The two signals are sampled at 1kHz, with a length of 20000 samples and are mixed together with different SNRs (5 - 20 dB with the step of 2.5 dB). The mixed dataset has 140.000 samples.

III. Results

The flowchart that describes the approach used in the paper is depicted in Figure 1

An example of a fECG and mECG signals generated with the ECGSYN generator are presented in Figure 3 and Figure 4, respectively. The mixture between the two signals, that form the synthetic abdominal signal, can be observed in Figure 5. The parameter configuration of the proposed method is as following: encoder filter number - 512, frame length – 16, kernel size – 3, 1D block channel number – 512, 1D block number - 8, repeat number – 3. Five-fold cross-validation was used in all experiments.

The performance of the model introduced in this paper is compared with the performance of the Independent Component Analysis (ICA) and Non-negative Matrix Factorization (NFM) as blind source separation (BSS) reference methods. The TABLE 1 presents the results obtained with the stacked-LSTM model and the ICA and NFM methods.

The experimental results show that the Stacked-LSTM method has the best performance, with a loss value of – 20.03 dB in average and it is significantly better than the traditional BSS methods.

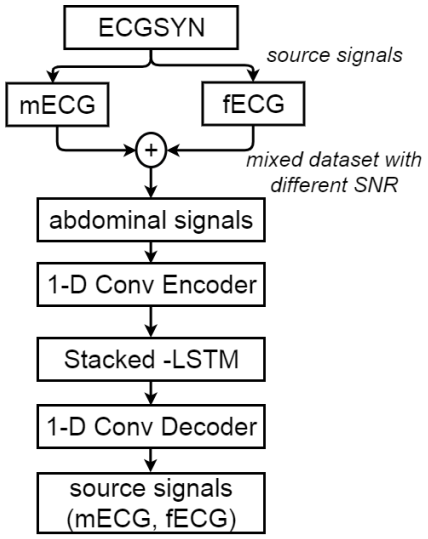


Figure 2 The flowchart of the proposed approach

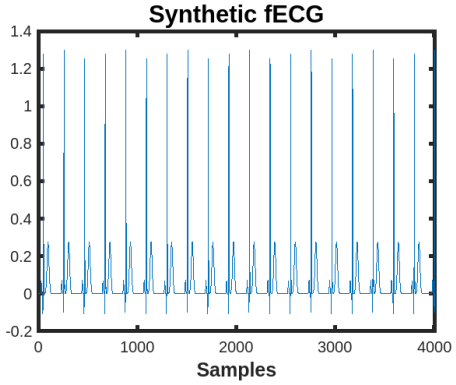


Figure 3. Synthetic fECG generated with ECGSYN

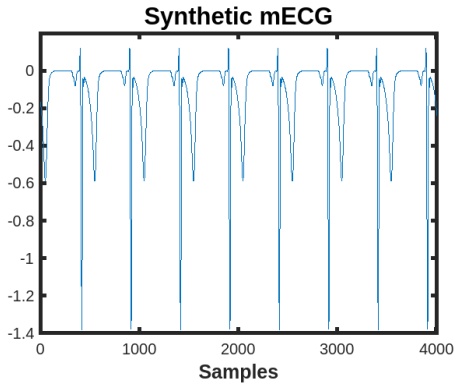


Figure 4. Synthetic mECG generated with ECGSYN

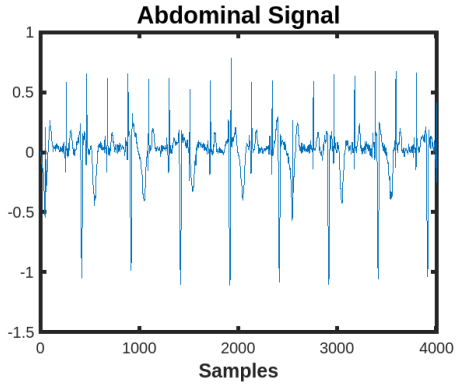


Figure 5. The synthetic abdominal signal as a mixture of the synthetic fECG and mECG signals

TABLE 1 LOSS OF THE COMPARED METHODS (dB)

Mixture	Stacked-LSTM	ICA	NFM
ADS 5 dB	-29.4	1.07	4.22
ADS 7.5 dB	-25.61	1.56	4.68
ADS 10 dB	-22.92	2.92	5.14
ADS 12.5 dB	-20.37	3.21	5.72
ADS 15 dB	-18.21	3.83	6.23
ADS 17.5 dB	-13.53	4.60	7.67
ADS 20 dB	−10.19	5.07	9.95

The proposed method is a BSS approach and in consequences is compared with classical BSS methods. However, when compared at a glance with approaches that use CNN for ADS denoising, the results obtained are comparable and, in some cases, exceed the ones obtained in [11], [12]. Nevertheless, one should interpret this comparison with extreme care because it is not a fair comparison due to the different databases that were used in the studies.

IV. CONCLUSIONS

The paper introduces a novel Stacked-LSTM-based deep learning model that can address the issue of multi-channel blind source separation in the case of abdominal recorded signals. The proposed model can effectively separate and restore the mixed signals, the mECG and the fECG, into their respective source signals by learning relevant information. The experiments demonstrate that the proposed model has excellent performance. One of the limitation of the introduced approach is that the model needs to be validated also on real abdominal signals.

ACKNOWLEDGMENT

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P1-1.1-TE-2021-0393, within PNCDI III

REFERENCES

[1] A. S. Eltrass, "Novel cascade filter design of improved sparse low-rank matrix estimation and kernel adaptive filtering for ECG denoising and artifacts cancellation," *Biomed Signal Process Control*, vol. 77, p. 103750, Aug. 2022, doi: 10.1016/j.bspc.2022.103750.

[2] E. Fotiadou, J. O. E. H. van Laar, S. G. Oei, and R. Vullings, "Enhancement of low-quality fetal electrocardiogram based on time-sequenced adaptive filtering," *Med Biol Eng Comput*, vol. 56, no. 12, pp. 2313–2323, Dec. 2018, doi: 10.1007/s11517-018-1862-8.

[3] M. Farahi *et al.*, "Beat-to-beat fetal heart rate analysis using portable medical device and wavelet transformation technique," *Heliyon*, vol. 8, no. 12, p. e12655, Dec. 2022, doi: 10.1016/J.HELIYON.2022.E12655.

[4] A. Jaba Deva Krupa, S. Dhanalakshmi, N. L. Sanjana, N. Manivannan, R. Kumar, and S. Tripathy, "Fetal heart rate estimation using fractional Fourier transform and wavelet analysis," *Biocybern Biomed Eng*, vol. 41, no. 4, pp. 1533–1547, Oct. 2021, doi: 10.1016/J.BBE.2021.09.006.

[5] Y. S. Alshebly and M. Nafea, "Isolation of Fetal ECG Signals from Abdominal ECG Using Wavelet Analysis," *IRBM*, vol. 41, no. 5, pp. 252–260, Oct. 2020, doi: 10.1016/j.irbm.2019.12.002.

[6] A. Jiménez-González and N. Castañeda-Villa, "Blind extraction of fetal and maternal components from the abdominal electrocardiogram: An ICA implementation for low-dimensional recordings," *Biomed Signal Process Control*, vol. 58, p. 101836, Apr. 2020, doi: 10.1016/J.BSPC.2019.101836.

[7] S. Ziani, A. Jbari, L. Bellarbi, and Y. Farhaoui, "Blind Maternal-Fetal ECG Separation Based on the Time-Scale Image TSI and SVD – ICA Methods," *Procedia Comput Sci*, vol. 134, pp. 322–327, Jan. 2018, doi: 10.1016/J.PROCS.2018.07.179.

[8] D. A. Ramli, Y. H. Shiong, and N. Hassan, "Blind Source Separation (BSS) of Mixed Maternal and Fetal Electrocardiogram (ECG) Signal: A comparative Study," *Procedia Comput Sci*, vol. 176, pp. 582–591, Jan. 2020, doi: 10.1016/J.PROCS.2020.08.060.

[9] P. Gupta, K. K. Sharma, and S. D. Joshi, "Fetal heart rate extraction from abdominal electrocardiograms through multivariate empirical mode decomposition," *Comput Biol Med*, vol. 68, pp. 121–136, Jan. 2016, doi: 10.1016/J.COMPBIOMED.2015.11.007.

[10] D. D. Taralunga and G. Mihaela Neagu, "An Ensemble Empirical Mode Decomposition Based Method for Fetal Phonocardiogram Enhancement," 2019, pp. 387–391. doi: 10.1007/978-981-10-9038-7_73.

[11] W. Zhong, X. Guo, and G. Wang, "Non-invasive Fetal Electrocardiography Denoising Using Deep Convolutional Encoder-Decoder Networks," 2020, pp. 1–10. doi: 10.1007/978-981-32-9682-4_1.

[12] E. Fotiadou and R. Vullings, "Multi-Channel Fetal ECG Denoising With Deep Convolutional Neural Networks," *Front Pediatr*, vol. 8, Aug. 2020, doi: 10.3389/fped.2020.00508.

[13] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans Biomed Eng*, vol. 50, no. 3, pp. 289–294, Mar. 2003, doi: 10.1109/TBME.2003.808805.

ECG Compression Based on Successive Differences

M. Gusev*+, M. Jovanov*, A. Shekerov*, G. Temelkov+
* Sts Cyril and Methodius University in Skopje, Faculty of Computer Science and Engineering
+ Innovation Dooel, Skopje, North Macedonia
E-mail: {marjan.gushev, mile.jovanov}@finki.ukim.mk, andrej.shekerov@students.finki.ukim.mk,
goran.temelkov@innovation.com.mk

***Abstract*—This paper aims to reduce the amount of physical space the electrocardiogram data occupies and needs to be transferred, designing an efficient algorithm that can run on hardware commonly found in a medical device setting. The algorithm needs to be compatible with real-time data transfer and be fast and lossless. We introduce an algorithm inspired by pulse code modulation in telecommunication theory based on successive differences between neighboring electrocardiogram samples. The research hypothesis addressed in this paper finds if the new algorithm can reduce the data size by a meaningful amount and efficiently. Further on we tackle the research questions to find the level of distortion, compression, and processing speed of compression and decompression, along with the analysis of benefits and disadvantages. We reduced the storage space or bits to be transferred up to 46% reaching a compression ratio of 2.17 for an electrocardiogram benchmark dataset of up to 100 Mbits.**

***Keywords*—ECG, Sampling, Analog to digital conversion, ADC, Compression, Consecutive differences**

I. Introduction

Advances in wearable technologies enable small medical devices to measure an electrocardiogram (ECG) in real time. The convenient sampling frequency to provide sufficient and relevant data for medical experts and automated Artificial Intelligence (AI) algorithms are in the range between 125 and 500 Hz, although some professional devices use sampling frequencies up to 1000 Hz, while the bit resolution is in the range between 10 and 16 bits.

The problems associated with systems for data transfer and real-time processing are classified in the Big Data domain, since data comes in high velocity generating large volumes. A simple calculation shows that such a system produces from 10,800,000 samples or 108 Mbits per day up to 86,400,000 samples or 1.38 Gbits per day. Considering that date/time stamp maybe accompanied to ECG data, the amount of data to be transferred and processed is from 20 MB up to 180 MB per day.

CardioHPC [1] is a project to realize an experiment for High Performance Computing (HPC) data processing center for thousands of real-time ECG streams. Such a system requires huge storage capacity and any savings with lossless compression are beneficial and highly appreciated. In addition, this method allows transferring of a smaller number of bits of the end-user device (ECG sensor) to the nearby smartphone or edge/dew device, which consumes less energy and saves the battery life. The related application [2] is a software solution [3] to detect arrhythmia or other dangerous heart conditions.

In this paper, we present an algorithm to compress ECG data based on pulse code modulation concepts from telecommunication theory encoding the differences between successive ECG samples. The idea is to build an algorithm that will efficiently store only successive differences in a specific way without losing data.

The purpose of this algorithm is to help with the transfer, storage, and real-time practical application of ECG data since the amount of storage space that needs to be allocated for even a day's worth of data can be overwhelming. With our algorithm, the storage space required to store this data can be reduced by more than half in some cases, and by 46% on average.

We realize an experimental method to find if the newly developed algorithm can efficiently reduce the data size by a meaningful amount. In addition, we aim at finding the level of distortion, compression, and processing speed of compression and decompression, along with the analysis of benefits and disadvantages, specifying optimization as the research question.

Related work is presented in Section II and methods in Section III with explanation of the algorithm the experiment and evaluation methodology. Implementation issues, benefits and disadvantages, along with a comparison to other tools are discussed in Section V. Finally, conclusions are presented in Section VI.

II. Related work

The need for compressing ECG data has been around for a long time. The earliest methods make use of linear interpolation and Huffman encoding, and some more modern methods apply approaches such as neural networks, wavelets, and compatibility with portable devices.

A. Overview

One of the earliest papers that tackled ECG data compression are reported when tape-less ECG recorders were starting to become feasible. Moody et al. [4] introduce compression by reducing the amount of redundant information which arise from a lack of statistical independence between symbols. Another relevant topic discussed is regarding "lossless" compression, where the accuracy of

the compression depends on the clinical interpretation; a change in the data can range from insignificant to extremely significant, depending on the context.

Singh et al. [5] review a few ECG data compression techniques, including direct, transform, parameter extraction, 2D transformation, etc. The authors introduce several evaluation parameters (compression ratio, percent mean square, difference, etc.) to compare and evaluate the different approaches. A key observation form their study is that even though there are a lot of ECG compression techniques, not so many are utilized in real monitoring systems and telemedicine, mainly due to fear of recovery from the distortions.

Another comprehensive review [6] presents various ECG compression implementations, using similar evaluation parameters. In addition, the authors analyze some effective algorithms (such as the wavelet transform method) and conclude that, due to their complexity, they haven't been implemented in real-life scenarios.

An overview of ECG compression methods suitable for specific scenarios is analyzed [7] from different evaluation perspectives. They conclude that threshold-based algorithms require low computational cost, transform-based methods are good for scenarios with a lot of noise due to their insensitiveness, etc.

A methodological review of different ECG data compression techniques [8], is reported with implementation of real-world data to evaluate the analyzed methods by several evaluation metrics. An interesting conclusion is that although simple direct methods achieve a smaller compression ratio, due to their simplicity and ease of implementation are most suitable for Internet of Things (IoT) health-care applications.

ECG compression techniques for wireless ECG devices have been analyzed in [9] with a focus on two direct data compression methods: delta coding and Huffman coding, as well as their variations. The experiments were checked on a smaller proprietary dataset and selected MITDB records achieving a compression ratio of 1.6. A combination of a second-order delta encoding and run-length encoding (RLE) based data compression is proposed [10] for ECG and PPG compression after quantizing using optimal quantization level to achieve a lesser number of representation bits.

Variable-length code was applied in developing an electronic circuit [11] on determined four regions of an ECG (initial region, QRS region, flat region and unrest region) adaptively predicting (detecting) unique waveform features, followed by a variable length code.

These methods differ from our approach which uses delta coding and variable-length code on successive differences.

B. Specialized techniques

An overview of a few known compression methods [12] states the need and benefits of standardization in terms of ECG compression evaluation methods. Due to a lack of any standardized way to make comparisons between different algorithms and manufacturer equipment, the clinical engineering community have no way of direct comparison of equipment and services.

There are several algorithms that utilize wavelets to compress the data [13]–[16] and belong to the category of transform methods, which are generally considered to be very effective. Among a variety of wavelet types, notably scalar wavelets and multiwavelets are generally more optimal.

Specialized techniques are often utilized when designing a new compression algorithm, and 2D transformation [17] has been used to compress ECG data in an efficient manner. Additionally, it attempts to highlight the current and future issues regarding the development of suitable ECG data compression techniques.

A more modern take on the problem - a compression algorithm that is based both on wavelets and on neural networks, and manages to achieve a very respectable compression ratio using these techniques [18]. This algorithm uses self-learning to gradually improve it's accuracy and reliability, and manages to achieve results that exceed other wavelet-based methods.

One can use direct methods [19] with the goal of finding an algorithm that is most suitable for telemedicine. The algorithms presented in this paper need to be transferred over a network reliably and efficiently, so they had additional requirements of security and the goal of a higher network transfer rate.

Jacobi polynomials have also been used in an ECG compression algorithm [20] to achieve interesting results compared to wavelet compression methods. The results of this algorithm could be greatly improved, since the procedure shown here is subject to rounding errors that pile up and eventually give modified results.

C. Special purpose

Development of ECG compression algorithms for portable devices [21] and Bluetooth are two specialized scenarios that require additional attention. These algorithms are generally suitable for portable devices and embedded systems with limited computing power. Some of the more interesting pre-processing methods listed here include filtering, down-sampling and peak-detection. An analysis [22] is given of some existing methods to find ones which are suitable to be used in an remote environment that uses Bluetooth to transfer the data. Important properties of the algorithms are latency, error-tolerance and power-consumption. Due to the limited processing and transmission power of small devices, these algorithms should not be very complex.

Lossless ECG data compression algorithms [23] are preferable to lossy methods due to the importance of the data that is being transmitted; medical regulatory board worldwide advocate for lossless compression. Since a

change in the original data may have grave consequences on the well-being of patients, ECG compression algorithms should aim to be lossless, or to at least have an acceptable amount of distortion.

An evaluation of a few known ECG compression methods [24], based on a few parameters. The authors mention the need for a standard way for comparing compression algorithms, and give a brief overview and recommendation on how to measure the performance on new methods.

III. Methods

Further on, we present the new compression algorithm and describe the experimental and evaluation methodology.

A. Algorithm

The ECG data stream contains a series of integer samples. Each ECG sample is compared to the preceding sample to determine the successive difference. The successive difference represents only an increase or decrease of an ECG sample from the previous sample. The sampling frequency determines the magnitude of the differences, and it determines the code for compression. The idea is that the code will be chosen to represent the average magnitude of the differences (or to encompass the most of cases).

The new ECG compression based on successive differences takes advantage of the fact that successive ECG values are usually similar, so it is better to store the difference between successive values instead of the values themselves. Still an open question arises to detect the efficient algorithm and the optimal number of bits to allocate for each value. Determination of the optimal number of bits to store the difference is not the only issue, we need to find a method to deal when the difference is bigger than the space we have allocated for it.

To find the optimal number of bits, we conduct an experiment by specifying test cases with different number of allocated bits and observe the achieved compression ratio. The testing domain is in the range from 2 bits to 7 bits.

This concept is implementation of the differential pulse-code modulation which encodes the differences between successive samples into n-bit data streams used in telecommunications or in analog-to-digital conversion, where the code modulation with 1-bit data stream is also known as the Delta-sigma $\Delta\Sigma$ modulation. This or other forms of pulse code modulations simply reject cases when successive difference is larger than the selected code, while in our algorithm we implement a solution to cope with these differences and do not lose any data in the compression.

The code to be selected is supposed to contain the successive difference and also the sign considering that the difference may be positive or negative. The minimal code to be used is 2 bits, one is the sign and the other for the difference. In a general case, the code of n bits will be able to store successive differences in the range from $-2^{(n-1)}$ up to $+2^{(n-1)}$.

The method to decide how to deal with successive differences is based on the no data loss concept, so we develop an algorithm that uses a special code to write a normal ECG sample instead of a successive difference.

The algorithm starts with a predefined zero value (the mid point in the selected bit-resolution). Then, a repetitive procedure analyzes the ECG values one by one in a sequence, to calculate the difference between the values of the current sample and the previous sample. If the calculated difference is small enough to fit into the space determined by the code, then it is stored with the sign and absolute value. The cases when the calculated difference is out of the range determined with a magnitude of $(n-1)$ bits, then a special code is stored, usually interpreted as negative zero (-0), and then store the real ECG sample instead of the successive difference.

B. Experiment

In our experiment, we use 48 ECG records from the MITDB ECG benchmark dataset. This collection of 48 ECG records totals 46.8 MB for 650000 samples per record each stored as a 12-bit sample. We have excluded the four records for patients with paced beats, so our dataset contains only 44 ECG records. Note that although the original declaration of the MITDB is 12 bit analog-to-digital conversion, still data are presented in 11-bit resolution.

Besides conducting the first *Experiment A* on the original dataset that uses the original sampling frequency of 360 Hz and bit resolution of 11 bits we have resampled the ECG records from 360 to 125 Hz, and rescaled from 11 to 10 bits and conduct an *Experiment B* on a second smaller dataset with a total of 13.5 MB. This resampling and rescaling achieves a compression ratio of 3.456. Although this method loses specific data features, still it is proven [25], [26] to be efficient in detecting the main heartbeat features and classifying different arrhythmia. Note that a very small number of values in the MITDB were found out of a 10-bit range, for specific signal inytervals of a pysically active person, with a lot of noise and artefacts, and these values were simply truncated.

The starting point in experiment A is the mid value of 1023 (midpoint of 11-bit resolution) and for the Experiment B the starting point is 511 (midpoint of 10-bit resolution). Experiment A stores original data of 11 bits after specifying a special code (-0) when the successive difference is larger then the selected code (number of bits), and Experiment B stores only 10 bits.

Test cases cover codes (number of bits) from 2 up to 7 bits. The testing environment that the algorithm is executed and tested is a system running with g++ version 12.1.0, Intel® Core™ i7-9750H processor, 16GB of RAM.

Additionally, to make sure that the compression is correct, the algorithm does a self-check to decompress the

data and compare it to the initial input. This is also used to measure decoding time.

C. Evaluation methodology

Additionally, this algorithm is lossless, meaning that the data should remain identical to the original after decompression. Due to this there is no need to measure the distortion. Considering the nature of lossless algorithms, the evaluation metrics [5] that compare different compression methods based on the magnitude of errors, such as Percent Mean Square Difference (PRD); Percent Root Mean Square Difference Normalized (PRDN); Quality Score (QS); Root Mean Square Error (RMS), etc. are not applicable. Our compression algorithm does not generate an error due to compression, these metrics are PRD=0, PRDN=0, RMS=0 and $QS \rightarrow \infty$.

A valid evaluation measure is the *compression ratio*, calculated as a ratio between the uncompressed size and compressed size. In our experiments, we will measure the the sizes of the benchmark dataset and compare to the size of the compressed data. In addition, we will express the reciprocal value of the compression ratio to reveal the capacity of the reduced size.

Additionally, we measure the time the algorithm takes in milliseconds for each test case and for each number of bits. This time includes reading the data, compressing the data, writing the data to a file, and then verifying the compression by reading the compressed data from a file and comparing it to the initial values we had for the test case. In practice the verification could be excluded to save on I/O and computing time, but here it is included to make sure that the algorithm gives us correct values. To make a real comparison this test is also executed with uncompressed data (original dataset).

IV. Results

Table I and Table II contain details on the number of analyzed bits as a sum and average per record, results of achieved capacity in percentage of the number of bits in the uncompressed dataset, and the compression rate (total, minimum, maximum, average and standard deviation) for selected 44 records from MITDB without records with paced beats. Table I refers to Experiment A for data sampled on the original 360 Hz sampling frequency and 11-b resolution, while Table II to Experiment B for data resampled to a 125 Hz sampling frequency and rescaled to 10-bits.

The best compression rate is achieved for 4 bits regardless the sampling frequency (applied for both experiments), reaching a value of 2.17 and 1.87 for Experiments A and B correspondingly.

Processing time (measured in milliseconds) is presented in Table III. The fastest compression is again achieved by using 4 bits, and the average time this per case is 123.85ms, whereas without compression the data is processed with an average of 108.96ms per case.

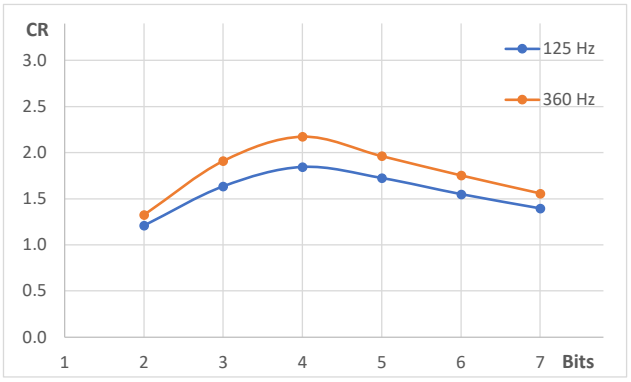


Fig. 1: Measured compression ratio in the experiments using data sampled at 360 Hz and 125 Hz

V. Discussion

Even though the algorithm works similarly regardless of the number of bits to allocate the space for successive differences, the compression achieved depends on this number. We have found that using 4 bits for this compression is the optimal number of bits in most cases. There are a few exceptions where 3 bits or 5 bits are better by a small margin, and even a case where using 6 bits is optimal, but the 4-bit compression has the lowest overall average and lowest total sum across all our tests. Additionally, the compression always performs better than the non-compressed data, except for 2-bit compression, which only falls behind on 4% of cases. The compression does take extra time to compute, but on average the fastest compression (depending on bits we use) is only 13.71 milliseconds behind the non-compressed data.

Many other methods of ECG compression use bit manipulation as part of their algorithm to help reduce the size of the data. Here we have provided a simple yet effective way to reduce the data in a lossless way. The data generated from this method could then be further compressed by some of the other methods instead of the raw ECG data.

The benefits of using this algorithm over others is the fact that our algorithm has no distortion; the original signal can always be recovered, and quite efficiently too as seen from the results section. The time it takes the algorithm to compress and decompress the data only slightly exceeds the time it takes for the raw data to be processed. Additionally, the algorithm is very simple and doesn't require much processing power to run, being able to be developed in a low-level language such as C, C++ or other similar languages.

Figure 1 presents the dependence of the achieved compression ratio versus different code (number of selected bits) in our algorithm.

Analysing the original 360 Hz sampling frequency (Experiment A dataset) and the sampling frequency of 125 Hz for resampled data (Experiment B dataset) we confirm the theoretical conclusion that the higher the sampling

TABLE I: Achieved compression rates for 44 MITDB ECG records in Experiment A with data sampled on 360 Hz with a 11-bit resolution.

	Total bits			**Compression Rate (CR)**				
Method	Sum	Avg	Capacity	Total	MIN	MAX	AVG	STD
No C.	314600000	7150000						
2 Bits	237733276	5403029	75.57%	1.32	1.07	1.70	1.34	0.15
3 Bits	164798403	3745418	52.38%	1.91	1.31	2.56	1.96	0.33
4 Bits	144762464	3290056	46.01%	2.17	1.80	2.46	2.19	0.17
5 Bits	160336044	3644001	50.97%	1.96	1.76	2.14	1.97	0.08
6 Bits	179489937	4079316	57.05%	1.75	1.60	1.83	1.75	0.05
7 Bits	202024306	4591461	64.22%	1.56	1.50	1.57	1.56	0.02

TABLE II: Achieved compression rates for 44 MITDB ECG records in Experiment B with data resampled on 125 Hz with a 10-bit resolution.

	Total bits			**Compression Rate (CR)**				
Method	Sum	Avg	Capacity	Total	MIN	MAX	AVG	STD
No C.	99305360	2256940						
2 Bits	81843372	1860077	82.42%	1.21	0.97	1.55	1.23	0.13
3 Bits	60081008	1365477	60.50%	1.65	1.06	2.26	1.70	0.29
4 Bits	53088234	1206551	53.46%	1.87	1.31	2.14	1.89	0.20
5 Bits	57243710	1300993	57.64%	1.73	1.56	1.87	1.74	0.08
6 Bits	64025726	1455130	64.47%	1.55	1.43	1.66	1.55	0.05
7 Bits	71216762	1618563	71.71%	1.39	1.30	1.43	1.39	0.03

TABLE III: Processing time for compression of 44 MITDB ECG records in Experiment B with data sampled on 125 Hz (measured in milliseconds)

Name	Sum	Avg
No comp.	5230	108.96
2 Bits	6878	143.29
3 Bits	6428	133.92
4 Bits	5945	123.85
5 Bits	6093	126.94
6 Bits	6115	127.40
7 Bits	6380	132.92

frequency is, the smaller differences between neighbours and the higher the compression rate is achieved.

The standard deviation calculated on different records shows small deviations, not just in absolute, but also in relative value, in the range from 2.0% (for 7 bits) - 17.6% (for 3 bits) for Experiment B (dataset sampled on 125 Hz), and from 1.2% (for 7 bits) - 17.1% (for 3 bits) for Experiment A (dataset sampled on 360 Hz). The best compression achieved for a 4-bit code results in 10.5% and 7.8% relative standard deviation correspondingly for Experiments B and A.

Note that several published approaches, including simple successive difference and Huffman coding [9] or their variations, or a combination of delta encoding and RLE [10] are different from our approach. Merdjanovska et al. [9] use successive differences followed by Huffman coding as a very popular lossless coding allocating specific codes to repeating data values. The authors do not expose the applied code table and conclude that a combination of Huffman and delta coding is worth exploring. In this paper, we use a variant of variable-length coding based on allocating a fixed code for the most frequent values (for example, successive differences of 1, 0, and -1) and if the successive difference is outside the values covered with the code then the successive difference is stored. This is quite different from Huffman code, which builds a binary tree for all the values. In our case, we do not build a binary or other tree and use an n-bit code to store both positive and negative numbers and the code associated to a negative zero is the special code to represent the successive difference.

Banerjee et al. [10] use successive differences of second order (difference of differences) instead of our approach using only first order differences. The authors apply quantization to different numbers of bits and in our approach we use conversion from 11 to 10 bits for the MITDB example. Repetition of zeros (a sequence of several zeros) is eliminated using the RLE, and in. our approach, we use a completely different approach to variable-length code. The proposed combination of second-order delta encoding is limited by the space allocated to represent the maximum absolute value (MAV) which was estimated to be 6 bits, out of which one bit is for a sign, meaning that the quantization was done to a smaller number of bits, generating representation errors.

These approaches generate a modified (distorted) signal within the compression algorithm and thus produce $PRD > 0$. Therefore, a direct comparison of our algorithm to other algorithms is not possible since our algorithm does not distort the signal (PRD=0). We use successive differences since their magnitude is much smaller than the ECG value, and the histogram of their representation follows a normal Gaussian distribution, such that our code that maps successive differences to a variable number of bits, allows compression with zero error (lossless data compression) and still be read back symbol by symbol.

Some of the following features might be considered as disadvantages:

- Misinterpretation of a code might generate a nonrecoverable error, solvable by an integration of time

stamps within each new data chunk.
- Variable bandwidth is expected with such a code to achieve higher compression against fixed bandwidth.
- The worst case scenario of sending noise with maximal alternating successive differences (not an ECG signal) will result in $CR \leq 1$.

VI. Conclusion

There are many ways in which ECG data can be compressed, and depending on the needs of the situation different algorithms have different strengths and weaknesses. We introduced a new compression algorithm for ECG measurements, based on pulse code modulation techniques exploited in telecommunication theory, focusing on successive differences between neighbouring ECG samples. In addition to the conventional approach in the pulse code modulation methods, we implement a new code that will save the original value of the successive difference instead of cutting the value to fit to the pulse code value. This approach allows a suitable implementation and flexibility without loosing in quality or distorting the data values that represent the signal.

Besides requiring a simple implementation our algorithm is effective and efficient (CR=1.6 and PRD=0) suitable in any scenario where the ECG data needs to be transferred or stored without any data loss. This is especially important to save energy for end-user devices and sensors which need to save energy and prolong battery life. Our future work addresses the dependence of the achieved compression rate with a wide range of different sampling frequencies, evaluating not only the compression achieved with our algorithm but also with reducing the sampling frequency.

Acknowledgement

The experiment "CardioHPC - Improving DL-based Arrhythmia Classification Algorithm and Simulation of Real-Time Heart Monitoring of Thousands of Patients" has received funding from the European High-Performance Computing Joint Undertaking (JU) through the FF4EuroHPC project under grant agreement No 951745. The JU receives support from the European Union's Horizon 2020 research and innovation program and Germany, Italy, Slovenia, France, and Spain.

References

[1] Innovation Dooel, "CardioHPC project "Real-time heart monitoring of thousands of patients"," online, 2023, as seen on 01.02.2023. [Online]. Available: http://cardiohpc.innovation.com.mk/

[2] M. Gusev, A. Stojmenski, and A. Guseva, "ECGalert: A heart attack alerting system," in *ICT Innovations 2017: Data-Driven Innovation. 9th International Conference, ICT Innovations 2017, Skopje, Macedonia, September 18-23, 2017, Proceedings 9*. Springer, 2017, pp. 27–36.

[3] E. Domazet and M. Gusev, "Improving the qrs detection for one-channel ECG sensor," *Technology and Health Care*, vol. 27, no. 6, pp. 623–642, 2019.

[4] G. B. Moody, K. Soroushian, and R. G. Mark, "ECG data compression for tapeless ambulatory monitors," *Computers in Cardiology*, vol. 14, pp. 467–470, 1987.

[5] B. Singh, A. Kaur, and J. Singh, "A review of ECG data compression techniques," *International journal of computer applications*, vol. 116, no. 11, 2015.

[6] S. B. Kale and D. H. Gawali, "Review of ECG compression techniques and implementations," in *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*. IEEE, 2016, pp. 623–627.

[7] S. O. Rajankar and S. N. Talbar, "An electrocardiogram signal compression techniques: a comprehensive review," *Analog Integrated Circuits and Signal Processing*, vol. 98, no. 1, pp. 59–74, 2019.

[8] C. K. Jha and M. H. Kolekar, "Electrocardiogram data compression techniques for cardiac healthcare systems: A methodological review," *IRBM*, 2021.

[9] E. Merdjanovska, M. Mohorcic, M. Depolli, A. Rashkovska, and T. Javornik, "Data compression for wireless ecg devices." in *BIOSIGNALS*, 2022, pp. 15–21.

[10] S. Banerjee and G. K. Singh, "A new real-time lossless data compression algorithm for ecg and ppg signals," *Biomedical Signal Processing and Control*, vol. 79, p. 104127, 2023.

[11] K. Li, Y. Pan, F. Chen, K.-T. Cheng, and R. Huan, "Real-time lossless ecg compression for low-power wearable medical devices based on adaptive region prediction," *Electronics Letters*, vol. 50, no. 25, pp. 1904–1906, 2014.

[12] S. M. Jalaleddine, C. G. Hutchens, R. D. Strattan, and W. A. Coberly, "ECG data compression techniques-a unified approach," *IEEE transactions on Biomedical Engineering*, vol. 37, no. 4, pp. 329–343, 1990.

[13] R. Besar, C. Eswaran, S. Sahib, and R. Simpson, "On the choice of the wavelets for ECG data compression," in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*, vol. 6. IEEE, 2000, pp. 3614–3617.

[14] M. S. Manikandan and S. Dandapat, "Wavelet-based electrocardiogram signal compression methods and their performances: A prospective review," *Biomedical Signal Processing and Control*, vol. 14, pp. 73–107, 2014.

[15] J. Chen, S. Itoh, and T. Hashimoto, "ECG data compression by using wavelet transform," *IEICE TRANSACTIONS on Information and Systems*, vol. 76, no. 12, pp. 1454–1461, 1993.

[16] N. V. Thakor, Y.-c. Sun, H. Rix, and P. Caminal, "Multiwave: a wavelet-based ECG data compression algorithm," *IEICE TRANSACTIONS on Information and Systems*, vol. 76, no. 12, pp. 1462–1469, 1993.

[17] R. Kumar, A. Kumar, and G. K. Singh, "Electrocardiogram signal compression based on 2d-transforms: A research overview," *Journal of Medical Imaging and Health Informatics*, vol. 6, no. 2, pp. 285–296, 2016.

[18] B. Zhang, J. Zhao, X. Chen, and J. Wu, "Ecg data compression using a neural network model based on multi-objective optimization," *PloS one*, vol. 12, no. 10, p. e0182500, 2017.

[19] V. Kumar, S. C. Saxena, and V. Giri, "Direct data compression of ECG signal for telemedicine," *International Journal of Systems Science*, vol. 37, no. 1, pp. 45–63, 2006.

[20] D. Tchiotsop, D. Wolf, V. Louis-Dorr, and R. Husson, "ECG data compression using jacobi polynomials," *IEEE engineering in medicine and biology magazine*, vol. 1, p. 1863, 2007.

[21] D. C. Pandhe and H. Patil, "Review and enhancement of ecg data compression and reconstruction method for portable devices," in *2015 International Conference on Energy Systems and Applications*. IEEE, 2015, pp. 490–495.

[22] B. Yu, L. Yang, and C.-C. Chong, "ECG monitoring over bluetooth: data compression and transmission," in *2010 IEEE Wireless Communication and Networking Conference*. IEEE, 2010, pp. 1–5.

[23] A. Tiwari and T. H. Falk, "Lossless electrocardiogram signal compression: A review of existing methods," *Biomedical Signal Processing and Control*, vol. 51, pp. 338–346, 2019.

[24] A. Němcová, R. Smíšek, L. Maršánová, L. Smital, and M. Vítek, "A comparative analysis of methods for evaluation of ECG signal quality after compression," *BioMed research international*, vol. 2018, 2018.

[25] E. Ajdaraga and M. Gusev, "Analysis of sampling frequency and resolution in ECG signals," in *2017 25th Telecommunication Forum (TELFOR)*. IEEE, 2017, pp. 1–4.

[26] M. Gusev and E. Domazet, "Optimizing the impact of resampling on QRS detection," in *ICT Innovations 2018. Engineering and Life Sciences: 10th International Conference, ICT Innovations 2018, Ohrid, Macedonia, September 17–19, 2018, Proceedings 10*. Springer, 2018, pp. 107–119.

Smart Glasses for Gait Analysis of Parkinson's Disease Patients

Ivana Kiprijanovska*, Filip Panchevski*, Simon Stankoski*, Martin Gjoreski**, James Archer*, John Broulidakis*, Ifigeneia Mavridou*, Bradley Hayes***, Tiago Guerreiro****, Charles Nduka*, and Hristijan Gjoreski *, *****

* Emteq Ltd., Brighton BN1 9SB, UK
** Faculty of Informatics, Università della Svizzera italiana, Lugano, Switzerland
*** Canterbury District Health Board, Christchurch Hospital, Christchurch, New Zealand
**** LASIGE, Faculdade de Ciencias, Universidade de Lisboa, Lisboa, Portugal
***** Ss. Cyril and Methodius University in Skopje, Skopje, N. Macedonia
ivana.kiprijanovska@emteqlabs.com

***Abstract* - Parkinson's disease (PD) is one of the most common neurodegenerative disorders of the central nervous system, which predominantly affects patients' motor functions, movement, and stability. Monitoring movement in patients with PD is crucial for inferring motor state fluctuations throughout daily life activities, which aids in disease progression analysis and assessing how patients respond to medications over time. In this preliminary study, we examine the possibility of using smart glasses equipped with Inertial Measurement Unit (IMU) sensors for providing objective information on the motor state of PD patients. Data was collected from seven patients with PD with varying levels of symptom severity, who performed a total of 35 trails of the Timed-Up-and-Go (TUG) test while wearing the glasses. We present an IMU-based step-detection algorithm with a potential for continuous monitoring of patients' gait. Furthermore, the analyses reveal patient-specific clusters in the sensor data, suggesting the possibility of developing personalized models for monitoring symptom progression. The findings suggest that smart glasses have the potential for unobtrusive and continuous screening of PD patients' gait, enhancing the medical assessment and treatment.**

Keywords - Parkinson's disease, smart glasses, inertial measurement unit, gait analysis

I. INTRODUCTION

Parkinson's disease (PD) is one of the most common progressive neurological disorders. In the past 25 years, the prevalence of Parkinson's disease has doubled globally. Global estimates in 2019 showed over 8.5 million individuals living with PD. PD predominantly impairs patients' motor abilities but is also associated with a wide variety of non-motor complications, including cognitive impairment, mental health and sleep disorders, sensory disturbances, and other behavioral problems. The severity and frequency of the symptoms usually increase as the disease progresses over time, impacting the patient's mental health and self-esteem, and significantly worsening their quality of life. Gait impairments are considered a good indicator of the progression of the disease, and can provide valuable insight into a patient's overall health, cognitive function, fall risk, and likelihood of institutionalization [1], [2]. Therefore, having efficient and reliable tool for gait monitoring and analysis is crucial for early diagnosis and tracking the progression of PD.

Fortunately, mobile health and behavior monitoring technology has advanced significantly with the availability of small, wearable, low-cost sensors combined with advanced signal processing, machine learning, and information extraction methods [3]. In the PD field, advanced analytics applied to speech (e.g., audio data), gait (e.g., accelerometer data), handwriting (e.g., touchscreen data), and face movements (e.g., video data), have offered great potential for disease monitoring and management [4], [5]. Adopting wearable and smart devices for PD has enabled an understanding of patients' symptoms outside the clinic (e.g., while performing everyday activities). The mobile-health tools also enable clinicians to continuously monitor the response of people with Parkinson's disease to various treatments, opening the opportunity to adapt medications efficiently [6]. Furthermore, recent meta-reviews demonstrate that people with Parkinson's disease want to use digital technologies [6], [7], [8]. Examples of perceived advantages include access to specialists, convenience, and time savings.

However, the meta-reviews also demonstrate that existing digital solutions for PD do not live up to the expectations [9]. One important challenge is acceptance, i.e., in future digital tools for improved care of PD patients should be much more user-friendly than average digital tools, given the motor and cognitive challenges PD patients may face. Another important challenge is personalization. Personalization is essential for PD given the symptoms' unpredictability and high inter-person variability [6], [8]. A third challenge is validation, i.e., the development of PD digital tools must be performed in collaboration with expert medical teams. The digital tools must also be validated against established clinical measures [10], [5].

This study is a step towards addressing the three challenges of the existing digital tools, acceptance, personalization, and validation. We aim to maximize user acceptance by providing a digital solution based on an object that is omnipresent in the life of PD patients, i.e., eyeglasses. The eyeglasses should also enable user-specific data, provided that they are worn daily by the PD patients. The user-specific data provided by the glasses could then

be used for personalized solutions. Finally, aiming towards clinical validation, this study provides initial analysis of data collected in a clinic with actual PD patients while performing a clinically validated protocol, Timed-Up-and-Go (TUG) [11].

II. Background and Related Work

The Timed-Up-and-Go (TUG) test is a clinical tool commonly used for measuring motor dysfunction in PD. The subject stands up from sitting, walks three meters, turns, walks back to the chair, and sits back down. The examiner measures, in seconds, the time taken to complete the task [11]. Total time taken has been shown to correlate with disease severity [12], risk of falls over the next year [13], [14], and response to dopaminergic therapies [15]. The TUG test is limited in that it measures the patient at a single point and is not truly representative of all activities of daily living (ADL). In addition, dopaminergic medications have "on" phases (peak effect) and "off" phases (trough effect), which makes the comparison of TUG testing over time challenging as patients may be in a different part of this cycle [16].

Many studies have looked at wearable devices which can track movements of PD patients without the need for a direct observer. Multiple studies have validated wearable devices during TUG testing and found reliability in measuring the total time taken [17], predicting fall risk, and gait and mobility parameters [18]. Wearable sensors have also been shown to delineate the time taken for specific movements such as the sit-to-stand or stand-to-sit transition [19], gait speed, and turning speed [20]. A study by Weiss et al. [21] measured data from a sensor worn on the lower back for three days and found a significantly improved ability to predict fall risk compared to commonly used clinical scoring measures. Wearable sensors have also been used to measure Freezing Of Gait (FOG) symptoms, i.e., when the patient is unable to initiate movement, over longer time periods [22], [23]. This is important as there may be specific triggers for this, such as walking through a narrow space, which cannot be replicated during a TUG test. Therefore, wearable devices can be used to enhance PD monitoring and aid medication decisions by giving clinicians information about the patient in a wider variety of settings and over a longer time period.

This study aims to assess the feasibility of using smart glasses equipped with an Inertial Measurement Unit (IMU) sensor for monitoring gait in patients with PD as an objective measurement of their motor state. From wearable sensing perspective, the head is a promising position for IMU sensor placement that should provide less noisy sensor data compared to other sensor placements (e.g., wrist/smartwatch). This is because our hands can be involved in a variety of micro and macro tasks (from running to typing on a keyboard). On the other hand, the human body has a natural mechanism to keep the head steady in space when the body is moving in order to maintain a stable visual field [24].

III. Dataset

A total of seven PD patients (four females and three males, with a mean age of 77 ± 7.3 years, range 68–87 years) with different levels of disease symptom severity were recruited. All participants provided written informed consent before participating in the study. For data collection, we used the OCOsense™ smart glasses (Figure 1), developed by Emteq Labs. The smart glasses are equipped with: (i) seven OCO™ sensors that measure skin movement in three dimensions (location indicated by the green rectangles in Figure 1), and (ii) a 9-axis IMU sensor (accelerometer, gyroscope, and magnetometer), including an altimeter (located in the right arm of the glasses frame, indicated by the purple rectangle in Figure 1). The orientation of the inertial sensor in the glasses can be determined by aligning the black dot on the zoomed inertial sensor in Figure 1 and the white dot below the purple rectangle on the glasses frame. This results in the x-axis being vertically oriented (opposite of the gravitational force), the y-axis being aligned with the glasses arm (pointing towards the ear), and the z-axis being the horizontal axis (pointing to the head).

The experimental protocol was based on the Timed-Up-and-Go (TUG) test: the participants sat on a chair, stood up, walked straight for three meters at their normal speed, turned 180° around an obstacle, walked straight back to the chair, did another 180° turn, and finally sat down on the chair. They completed five repetitions of the TUG test, resulting in a total of 35 walking trials in the dataset. All participants were able to perform all trials independently, without an assistive device. The sessions were also recorded with a video camera. The study was reviewed and approved by the NHS research ethics committee, (ref: 18/WM/0205) and took place at Queen Victoria Hospital in East Grinstead, England.

All participants completed two questionnaires: the Freezing of Gait (FOG) Questionnaire and the Parkinson's Disease Questionnaire (PDQ-8). The former includes six questions and is focused on FOG severity and gait impairments over the last week, while the latter includes eight questions about mobility, activities of daily living, emotional well-being, stigma, social support, cognition, communication, and bodily discomfort, and is used to quantify the quality of life among PD patients. Each question was scored between zero and four. A higher score on the FOG questionnaire corresponds to more severe FOG episodes. A higher score for the PDQ-8 questionnaire signifies a poorer quality of life and a more severe form of the disease.

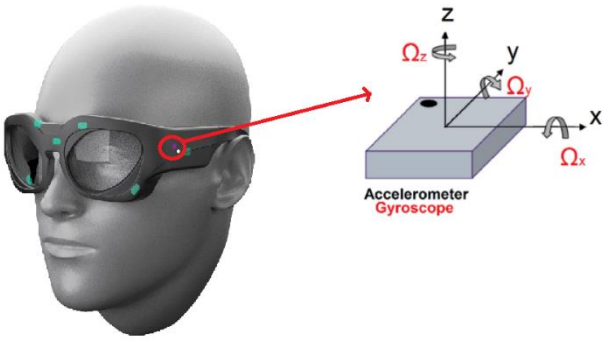


Figure 1. OCOsense™ glasses and the sensors' location. The green rectangles represent the OCO™ sensors, and the purple rectangle represents the 9-axis inertial sensor.

TABLE 1. DEMOGRAPHICS DATA AND QUESTIONNAIRES' SCORES.

ID	Age (Years)	Sex	FOG Score (%)	PDQ-8 Score (%)
1	70	Male	20.8	46.9
2	73	Female	4.2	40.6
3	83	Female	87.5	75.0
4	68	Female	66.7	37.5
5	73	Male	25.0	31.3
6	86	Male	66.7	65.6
7	87	Female	33.3	34.4

The final scores from the questionnaires, alongside the participant demographics, are presented in Table 1. The questionnaire's score for each participant was divided by the total possible score (for each questionnaire separately), and the final scores in the table are given as a percentage out of 100.

IV. SENSOR DATA ANALYSIS

A. *Sensor data preprocessing*

The sensor data was collected using Emteq's OCOsense™ smart-glasses that provided 3-axis accelerometer, 3-axis gyroscope, and 3-axis magnetometer data, all sampled at 50 Hz. To reduce the impact of high-frequency artifacts and preserve the gait information within the signals, a 3rd order Butterworth bandpass filter was applied to the sensor data. It allowed only frequencies within the range of 0.1 to 5 Hz to pass through, effectively filtering out frequencies irrelevant to gait analysis and ensuring that the signals are not impacted by noise. Figure 2 presents an example of a raw acceleration signal recorded during a walking activity, and the same signal after the pre-processing.

B. *Step detection – heathy participants*

We implemented a step detection method that involves analyzing the periodic differences present in the pre-processed accelerometer magnitude data. The peak detection is based on the first order difference of the signal.

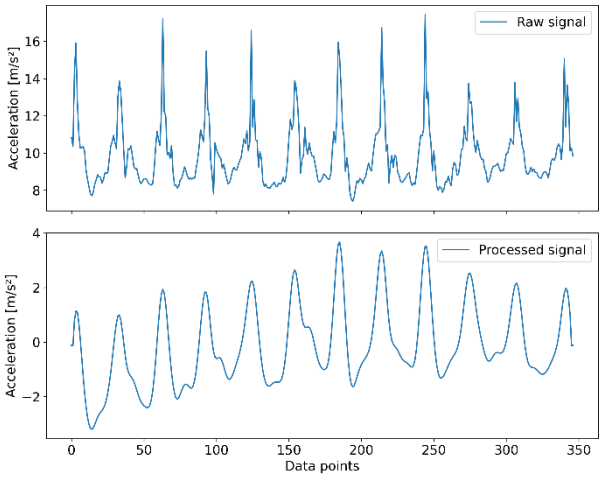


Figure 2. An example of a raw and processed acceleration signal (x-axis).

The algorithm has two parameters, a minimum amplitude threshold above which a data point can be considered as a peak, and a minimum time distance required between two peaks. The detected peaks in the acceleration signal correspond to the moment of highest impact during each step. Similarly, if we invert the acceleration signal and detect the peaks in the inverted signal, we are also able to detect the start and end of each step, or the heel strike and toe off. An example for the step detection algorithm is presented in Figure 3. The figure displays the change in acceleration (represented by the x-axes and measured in meter per second squared). We consider the x-axis to be the most informative for the walking activity because of its orientation (see Figure 1).

Before applying the step-detection method on the data from the PD patients, we performed in-lab testing with data from four healthy participants (with a mean age of 23.3 ± 1.64) following a simple testing scenario. The participants wore the Emteq's OCOsense™ glasses and completed two sessions, in each performing twenty steps at their normal walking pace. Thus, the speed of walking varied between the participants. We evaluated the method using Mean Absolute Percentage Error (MAPE) (1):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{A_i - D_i}{A_i} \right| \quad (1)$$

where n is the number of participants; A_i is the actual number of steps that the *i-th* participant performed; and D_i is the number of detected steps based on the step detection method. This dataset was also used to fine-tune the parameters of the peak detector. The final MAPE score achieved by the method was of 1.3%. This means that error was below one step per participant (on average).

C. *Step detection – PD participants*

The step-detection method was then applied to analyze the number of steps performed by the Parkinson's patients during the five trials of the TUG test. The number of detected steps in each trial for each participant is reported in Table 2. The average number of steps across all participants was 15.9 with a standard deviation of 5.45. The average MAPE in step detection across the participants was 8.1%, which represents an average error of 1.3 steps (out of 16).

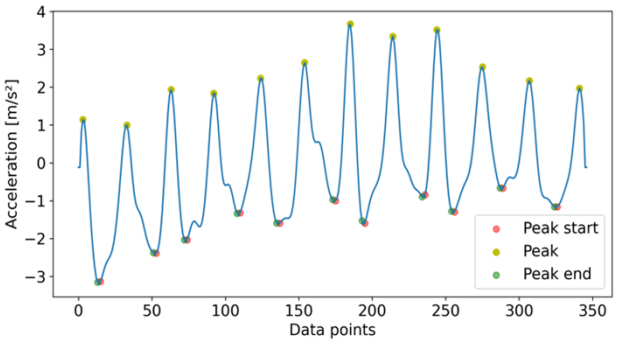


Figure 3. An example of a processed acceleration signal (x-axis) and the detected steps.

TABLE 2. NUMBER OF DETECTED STEPS FOR EACH PARTICIPANT IN FIVE TUG SESSIONS.

ID	S1	S2	S3	S4	S5	FOG	PDQ-8
1	16	16	16	14	14	20.8	46.9
2	11	12	11	10	12	4.2	40.6
3	33	25	24	27	24	87.5	75.0
4	13	13	14	15	14	66.7	37.5
5	11	11	11	11	11	25.0	31.3
6	24	19	19	19	20	66.7	65.6
7	16	13	13	13	13	33.3	34.4

Patients with PD experience changes in their gait, including a decrease in walking speed and reduced step length. This is demonstrated by the number of detected steps required by the patients to complete the TUG test. The results show that participant three, who has the most severe PD symptoms according to the scores of the questionnaires (FOG Score = 87.5, PDQ Score = 75.0), shows notably increased number of steps for completion of the TUG test (26.6 ± 3.1), compared to the rest of the participants. On the other hand, participant five, who is ranked 5th on the FOG Score scale and 7th on the PDQ-8 scale (indicating that this participant is the one with the least severe form of the disease across the dataset), had the lowest number of detected steps, on average (11.0 ± 0.0).

To further determine the strength of the relationship between the questionnaires' scores and the number of detected steps performed during the TUG test, we performed linear regression analysis and obtained the correlation coefficients. These results are depicted in Figure 4.

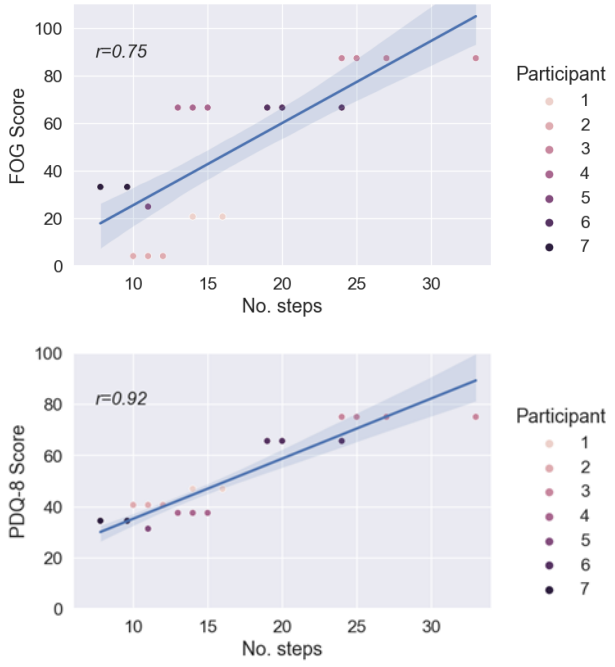


Figure 4. Linear Regression Analysis of the FOG and PDQ-8 Questionnaires Scores versus the number of detected steps performed during the TUG test.

A positive correlation was observed in both tested cases. In general, the results show that the number of detected steps required for completing the TUG test are stronger correlated with the PDQ-8 Score (r=0.92) that the FOG Score (r=0.75). Despite the lower performance of the step detection method on the PD patient dataset compared to the healthy participants dataset, the results still demonstrate a strong correlation between the detected steps and the scores of the PD-specific questionnaires (FOG and PDQ-8). This correlation provides evidence for the potential utility of the step detection method as a tool for monitoring the progression of PD symptoms.

D. Personalized Gait analysis

Next, we explored whether there are intra- and inter-person characteristics in the acceleration signals. The sensor data contains five walking signatures per participants – given that each PD participant performed five repetitions of the TUG test. The resulting walking signatures for each participant are presented in Figure 5. In the figure, each of the horizontal subplots corresponds to data from a single participant, which are further divided into five sessions marked by the vertical gray lines. The x-axis represents the traveled distance (5 x 300 cm), and the y-axis represents the acceleration as measured through the IMU sensor. The walking signature is generated by interpolating the scattered dots, each represent–ting a single step, as detected by the algorithm. The steps (dots) are plotted on the graph based on the distance travelled during each step as determined by assuming a constant walking speed (on the x-axis), and the magnitude of acceleration which they produce (on the y-axis).

We used the dynamic time warping (DTW) algorithm to search for similarities among and within people (intra- and inter-person). The algorithm evaluates the similarity of two time-sequences by using the Euclidian formula to calculate distance. The intra-person distance for a participant was calculated as the average DTW distance between all pairs of walking signatures generated by that participant (one participant has five walking signatures). Similarly, the inter-person distance for a pair of participants was calculated as the average DTW distance between all pairs of walking signatures generated by the two participants (each of them has five signatures), where one item in the pair always belongs to the first participant and the second item belongs to the second participant. The results showed clear examples of contrasting and similar walking signatures. Participants at the opposing ends of the PDQ-8 spectrum (participant three and participant five) had a notable difference in gait, with the lowest similarity value recorded compared to the other pairs of participants. Similarly, the participants with comparable PDQ-8 scores (e.g., participant seven and participant four) had the highest similarity score in the inter-person analysis. In contrast, the intra-person comparison yielded interesting results regarding the participants with high FOG scores. Participants three, four and six had the highest similarity scores. This could be due to their premeditated walking style; a technique used to manage and prevent FOG episodes. A larger dataset is needed to obtain a definitive answer.

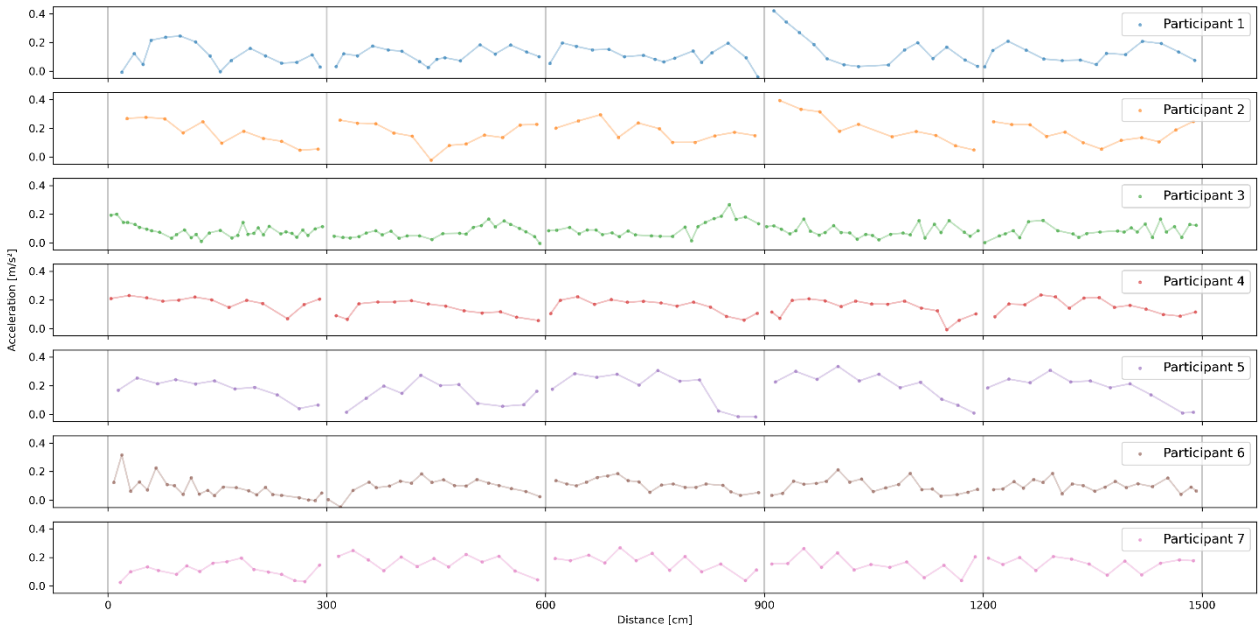


Figure 5. Participant's walking signature, throughout the five sessions. Each horizontal plot represents a single participant, with the grey line dividing the plot into the five separate sessions.

With a view towards real-time machine-learning analysis, we performed additional experiments with segmented data. More specifically, the data was first filtered and then segmented using a sliding window of two seconds and a 50% overlap. We extracted nine general-purpose statistical features from the accelerometer and gyroscope sensor signals, resulting in a total of 72 features. The features included the mean, standard deviation, minimum, maximum, kurtosis, skewness, value range, root mean square, and interquartile range. We conducted the initial analysis in this study with a small number of simple features that are computationally inexpensive in an effort to provide real-time processing on the glasses in the future. The extracted features were then used the with the t-distributed stochastic neighbour embedding (t-SNE) method [25] for visualization, and to determine whether the sensor data recorded from the glasses could reveal additional information about the patients. The method was applied to the feature vectors extracted from the walking segments from the TUG test trials. The result of the t-SNE analysis is shown in Figure 6. The figure shows several distinct clusters of data, highlighting visible differences in walking styles among PD patients.

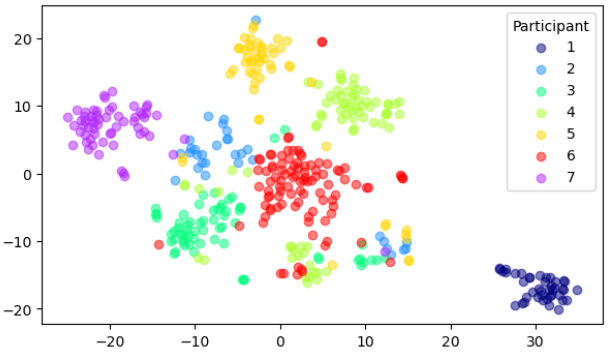


Figure 6. t-SNE visualization of the walking data from the TUG test.

The observed variations in the walking styles could potentially be related to the severity of the PD. While further research with a larger sample size and more diverse patient population is needed to draw definitive conclusions about the uniqueness of each patient's walking style, these findings suggest that the glasses-based data is sufficiently sensitive to be used for the development of personalized models for patient-specific monitoring of symptom progression. Such personalized monitoring could provide valuable information for the management of PD, enabling more targeted interventions based on individual patient needs.

V. CONCLUSION

In this paper, a preliminary analysis was presented with regards to the ability of Emteq's novel OCOsense™ smart glasses equipped with IMU sensors to provide objective information on the motor state in patients with Parkinson's disease. We used data from seven Parkinson's disease patients with varying levels of the disease's symptoms severity who were performing the Timed-Up-and-Go (TUG) test while wearing the glasses.

Although only a small group of PD patients were examined in the experiment, the analysis suggests that IMU-equipped smart glasses have the potential to provide information about patients' gait and can be used to assess the severity level of Parkinson's disease as measured by two standardized questionnaires. They can therefore be considered as a screening tool that will continuously monitor Parkinson's disease patients' gait and motor activity. Moreover, the analysis of the walking data has shown visible differences walking styles among PD patients, which may also depend on the severity of the disease. This indicates that the glasses-based sensor data is sensitive enough to develop personalized models for patient-specific monitoring of symptom progression.

Some limitations of the presented work are the small number of PD patients in the study and the narrow focus of gait analysis only during the TUG test. While the preliminary results are promising, it is crucial to validate our findings on a larger scale and analyze gait during different activities of daily living. To address these limitations, we are currently undertaking the systematic organization of a large-scale data collection, which aims to compile data from numerous individuals diagnosed with Parkinson's disease. This data collection will also include different positions of the IMU sensors that will enable comparison with our glasses-based approach. Additionally, we plan to add other daily-life activities, including different household and core self-care activities, which will allow us to expand our analysis of gait patterns during day-to-day activities, but also other PD related symptoms that occur during daily living, such as body tremors. Moreover, we intend to conduct a longitudinal study to determine whether data acquired from the glasses could be used to track changes in the disease symptoms over time, which will be beneficial for inferring early disease progression or deterioration.

ACKNOWLEDGMENT

This project was partially supported by the WideHealth project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 952279.

REFERENCES

[1] Morris R, Hickey A, Del Din S, Godfrey A, Lord S, Rochester L. A model of free-living gait: A factor analysis in Parkinson's disease. Gait Posture. 2017 Feb;52:68-71. doi: 10.1016/j.gaitpost.2016.11.024. Epub 2016 Nov 16. PMID: 27883986.

[2] Bouça-Machado R, Jalles C, Guerreiro D, Pona-Ferreira F, Branco D, Guerreiro T, Matias R, Ferreira JJ. Gait Kinematic Parameters in Parkinson's Disease: A Systematic Review. J Parkinsons Dis. 2020;10(3):843-853. doi: 10.3233/JPD-201969. PMID: 32417796; PMCID: PMC7458503.

[3] Gjoreski, M. (2020). A Fusion of Classical and Deep Machine Learning for Mobile Health and Behavior Monitoring with Wearable Sensors: Doctoral Dissertation, Jožef Stefan International Postgraduate School, Ljubljana, Slovenia.

[4] Canning, C. G., Allen, N. E., Nackaerts, E., Paul, S. S., Nieuwboer, A., & Gilat, M. (2020). Virtual reality in research and rehabilitation of gait and balance in Parkinson disease. Nature Reviews Neurology, 16(8), 409-425.

[5] Fröhlich, H., Bontridder, N., Petrovska-Delacréta, D., Glaab, E., Kluge, F., El Yacoubi, M., ... & Klucken, J. (2022). Leveraging the Potential of Digital Technology for Better Individualized Treatment of Parkinson's Disease. Frontiers in neurology, 13.

[6] Chandrabhatla, A. S., Pomeraniec, I. J., & Ksendzovsky, A. (2022). Co-evolution of machine learning and digital technologies to improve monitoring of Parkinson's disease motor symptoms. NPJ digital medicine, 5(1), 1-18.

[7] Spear, K. L., Auinger, P., Simone, R., Dorsey, E., & Francis, J. (2019). Patient views on telemedicine for Parkinson disease. Journal of Parkinson's Disease, 9(2), 401-404.

[8] Riggare, S., Stamford, J., & Hägglund, M. (2021). A long way to go: patient perspectives on digital health for Parkinson's disease. Journal of Parkinson's Disease, 11(s1), S5-S10.

[9] Espay, A. J., Bonato, P., Nahab, F. B., Maetzler, W., Dean, J. M., Klucken, J., ... & Movement Disorders Society Task Force on Technology. (2016). Technology in Parkinson's disease: challenges and opportunities. Movement Disorders, 31(9).

[10] Severson, K. A., Chahine, L. M., Smolensky, L. A., Dhuliawala, M., Frasier, M., Ng, K., ... & Hu, J. (2021). Discovery of Parkinson's disease states and disease progression modelling: a longitudinal data study using machine learning. The Lancet Digital Health, 3(9), e555-e564.

[11] Podsiadlo, D. and Richardson, S., 1991. The Timed "Up & Go": A Test of Basic Functional Mobility for Frail Elderly Persons. Journal of the American Geriatrics Society, 39(2), pp.142-148.

[12] Matinolli, M., Korpelainen, J., Korpelainen, R., Sotaniemi, K., Matinolli, V. and Myllylä, V., 2009. Mobility and balance in Parkinson's disease: a population-based study. European Journal of Neurology, 16(1), pp.105-111.

[13] Nocera, J., Stegemöller, E., Malaty, I., Okun, M., Marsiske, M. and Hass, C., 2013. Using the Timed Up & Go Test in a Clinical Setting to Predict Falling in Parkinson's Disease. Archives of Physical Medicine and Rehabilitation, 94(7), pp.1300-1305.

[14] Okumiya, K., Matsubayashi, K., Nakamura, T., Fujisawa, M., Osaki, Y., Doi, Y. and Ozawa, T., 1998. THE TIMED "UP & GO" TEST IS A USEFUL PREDICTOR OF FALLS IN COMMUNITY-DWELLING OLDER PEOPLE. Journal of the American Geriatrics Society, 46(7), pp.928-929.

[15] Morris, S., Morris, M. and Iansek, R., 2001. Reliability of Measurements Obtained With the Timed "Up & Go" Test in People With Parkinson Disease. Physical Therapy, 81(2), pp.810-818.

[16] Armstrong, M. and Okun, M., 2020. Diagnosis and Treatment of Parkinson Disease. JAMA, 323(6), p.548.

[17] Kleiner, A., Pacifici, I., Vagnini, A., Camerota, F., Celletti, C., Stocchi, F., De Pandis, M. and Galli, M., 2018. Timed Up and Go evaluation with wearable devices: Validation in Parkinson's disease. Journal of Bodywork and Movement Therapies, 22(2), pp.390-395.

[18] Greene, B., Premoli, I., McManus, K., McGrath, D. and Caulfield, B., 2021. Predicting Fall Counts Using Wearable Sensors: A Novel Digital Biomarker for Parkinson's Disease. Sensors, 22(1), p.54.

[19] Postural transitions detection and characterization in healthy and patient populations using a single waist sensor

[20] Mariani, B., Jiménez, M., Vingerhoets, F. and Aminian, K., 2013. On-Shoe Wearable Sensors for Gait and Turning Assessment of Patients With Parkinson's Disease. IEEE Transactions on Biomedical Engineering, 60(1), pp.155-158.

[21] Weiss, A., Herman, T., Giladi, N. and Hausdorff, J., 2014. Objective Assessment of Fall Risk in Parkinson's Disease Using a Body-Fixed Sensor Worn for 3 Days. PLoS ONE, 9(5), p.e96675.

[22] Pardoel, Kofman, Nantel and Lemaire, 2019. Wearable-Sensor-based Detection and Prediction of Freezing of Gait in Parkinson's Disease: A Review. Sensors, 19(23), p.5141.

[23] Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J., Giladi, N. and Troster, G., 2010. Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptom. IEEE Transactions on Information Technology in Biomedicine, 14(2), pp.436-446.

[24] Jiménez Ortega, Laura, Katrin Stoppa, Onur Güntürkün, and Nikolaus F. Troje. "Vision during head bobbing: are pigeons capable of shape discrimination during the thrust phase?." Experimental brain research 199 (2009): 313-321.

[25] Van der Maaten, L., & Hinton, G., 2008. Visualizing data using t-SNE. Journal of machine learning research, 9(11).

Attention-based U-net: Joint Segmentation of Layers and Fluids from Retinal OCT Images

M. Melinščak
Zagreb University of Applied Sciences, Zagreb, Croatia
martina.melinscak@tvz.hr

Abstract **– Since its introduction in 2015, U-net has become state-of-the-art neural network architecture for biomedical image segmentation. Although many modifications have been proposed, few novel concepts were introduced. Recently, some significant breakthroughs have been achieved by introducing attention or, more specifically, Transformers. Many attempts to incorporate self-attention mechanisms into solving computer vision tasks resulted in Vision Transformer (ViT). As ViT has some downsides compared to convolutional neural networks (CNNs), neural networks which merge advantages from both concepts prevail, especially in small data regimes we often face in medicine. U-net architecture still outperforms ViT models as their high accuracy relies on massive data. This paper investigates how attention added in U-net architecture affects results. We evaluate the outcomes on a publicly available dataset which consists of 1136 retinal optical coherence tomography (OCT) images from 24 patients suffering from neovascular age-related macular degeneration (nAMD). Also, we compare results to previously published results, and it could be noted that the Attention-based U-net model achieves higher Dice scores by a significant margin. The code is publicly available.**

Keywords – Vision Transformer; attention; convolutional neural networks; U-net; automatic segmentation; retinal optical coherence tomography images

I. INTRODUCTION

Optical coherence tomography (OCT) allows non-invasive 3D imaging where high-resolution images provide information about the retinal structure. Consequently, OCT imaging found an essential role in ophthalmology. Morphological features that can be seen and measured from OCT tomograms, such as the thickness of individual retinal layers, shapes, spatial distributions, and optical properties of various lesions and blood vessels, can serve as identifiers in the diagnosis of retinal diseases [1]. Some of the most often diseases diagnosed by OCT imaging are diabetic retinopathy (DR), age-related macular degeneration (AMD), retinal vein occlusion (RVO), central serous retinopathy (CSR), and glaucoma.

There is increased demand for automatic segmentation due to a lack of ophthalmologists and an increased prevalence of patients suffering from retinal diseases (primarily because of population ageing and the diabetes epidemic). However, detailed inspection of 3D volumes is time-consuming, prone to errors, and mostly impossible within a typical clinical setup. Also, there is no consensus between different ophthalmological opinions. On the other hand, tremendous advances in computer vision have been made, but we still lack a reliable model for automatic segmentation. Deep learning models are still too brittle and unable to generalize well on unseen data, which makes automatic segmentation an open and vital research field.

Since introducing AlexNet architecture, which achieves breakthrough results on an ImageNet benchmark [2], convolutional neural networks (CNNs) have become state-of-the-art architectures for computer vision tasks. Architectures that followed (e.g., VGG [3], GoogleNet [4], ResNet [5], Inception v3 [6], MobileNet [7], and EfficientNet [8]) have had more layers and parameters. However, there were minor enhancements – convolutions were still the predominant paradigm in computer vision. Mostly, concepts introduced in the following years were various regularization techniques and overcoming problems like overfitting due to an enormous number of parameters.

In a highly cited paper, "Attention is all you need," Vaswani et al. introduced the Transformers. Primarily Transformers were used for natural language processing (NLP), and mostly they gained popularity due to the latest large language models (LLM) applications. Next, Dosovitskiy et al. introduced a Vision transformer (ViT) [9]. Unlike convolutional networks, Transformers lack inductive bias like locality and translational equivariance. Furthermore, Transformers work with image patches, and similarly, like in multi-layer perceptron (MLP), all patches attend to all others, making the ViT model enormously computationally costly. Finally, the authors highlighted the remaining challenges, such as solving segmentation and detection tasks with Transformers. Although Transformers became highly popular, they were appropriate for tasks with accessible large amounts of data and immense computing power. Soon after, there were many proposals on how to merge the advantages of self-attention architectures and convolutional neural networks.

In the case of medical image segmentation, introducing U-net architecture [10] made a significant breakthrough. Furthermore, encoder-decoder architecture with skip connections enables concurrent capturing of localization and context. Many modifications to the U-net architecture were proposed [11]–[14]. Nevertheless, U-net-based architectures remain predominant in biomedical image segmentation, mainly due to the small data regime we often face in biomedical fields.

This paper is mainly inspired by Attention U-net [15]. We made all modifications to make it applicable to the

OCT dataset we use for evaluation. Also, we made a comparison to standard U-net architecture and state-of-the-art architectures like U-net++ architecture (a nested U-net architecture for medical image segmentation) [16] and U-net-like architecture [17] to see how adding attention affects results.

In summary, the main contributions of this work are as follows:

- Proposing an attention-based model for joint segmenting layers and fluids in OCT retinal images.
- Evaluating the model on the publicly available AROI dataset of 1136 B-scans from patients with neovascular AMD.
- Comparing obtained outcomes with previous results on the benchmark dataset and improving segmentation results by a significant margin.

The rest of this paper we organize as follows. First, section 2 mentions relevant related work using attention in ophthalmology. Then, in section 3, we describe the database we use for evaluation purposes and give a detailed description of the model we use. Next, in section 4, we provide results and discuss them in the following section; in the last section, we conclude by presenting the limitations of our work and possible future research directions.

II. Related work

Wang et al. [18] proposed a Multi-scale Transformer Global Attention Network (MsTGANet) for drusen segmentation on the Kermany dataset [19]. The Kermany dataset is the largest and most popular publicly available dataset for classifying OCT images into four classes: choroidal neovascularization (CNV), diabetic macular edema (DME), drusen, and images from healthy persons. Jiang et al. [20] used a Vision Transformer to classify images on the publicly available DUKE dataset [21]. As a result, images are classified into three classes: patients suffering from AMD, DME, and healthy persons. Cao et al. [22] proposed a regression method based on Transformer that performs segmentation of seven layers in OCT images from the publicly available DUKE DME dataset [23].

Playout et al. [24] presented Vision Transformer for image classification: the paper mainly focuses on classification and mechanisms for generating interpretable predictions via attribution maps. They also perform the segmentation of fluids in images from the AROI dataset. It is hard to compare results with other papers that used images from the AROI dataset as they only segment fluids and do not distinguish different types of fluids. Nevertheless, they show that segmentation with Vision Transformer outperforms classical CNN models.

A few papers [25]–[29] combine attention with CNNs. However, it is out of the scope of this paper to give an extensive survey of all published papers. Instead, we would like to refer the reader to the review study from Khan et al. [30].

III. Materials and methods

A. AROI dataset

To evaluate the proposed model and compare results with CNNs without attention, we use the publicly available AROI dataset [31]. Macular SD-OCT volumes were recorded with the Zeiss Cirrus HD OCT 4000 device: each OCT volume comprises 128 B-scans with a 1024 x 512 pixels resolution. There are annotated 1136 OCT B-scans from 24 patients suffering from late neovascular AMD. An ophthalmologist did all annotations. There are annotations for four boundaries between the layers: internal limiting membrane – ILM, retinal pigment epithelium (RPE), the boundary between the inner plexiform layer and inner nuclear layer (IPL/INL), and Bruch's membrane (BM). Also, fluids are annotated: pigment epithelial detachment (PED), subretinal fluid (SRF), and intraretinal fluid (IRF). Images are prepared for semantic segmentation with eight classes (Figure 1).

Data collection adhered to the tenets of the Declaration of Helsinki and the standards of Good Scientific Practice of Sestre milosrdnice University Hospital Center (Zagreb, Croatia). All patients signed informed consent, and the data were anonymized. The Ethics Committee of the Sestre milosrdnice University Hospital Center approved the presented study.

We opt for this dataset as it is publicly available, and annotations are provided for both layers and fluids. Also, results for human variability are given. Furthermore, images of patients suffering from neovascular AMD (in some cases simultaneously from geographic atrophy) are highly challenging for segmentation due to severe pathological changes. Commercial software for automatic segmentation associated with OCT devices usually works well for healthy persons but with significant errors in pathological cases.

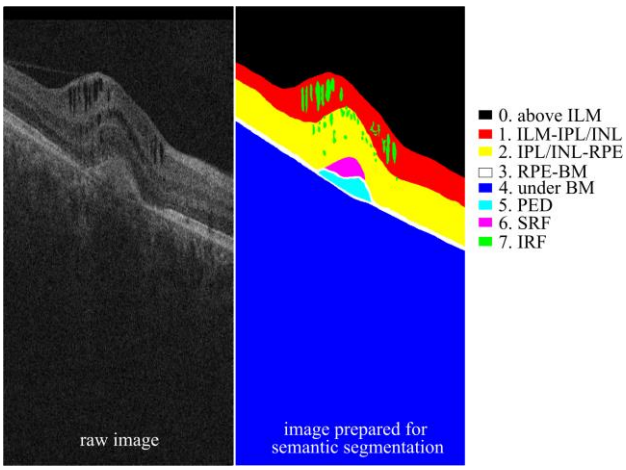


Figure 1. Example of an image from the AROI dataset [31].

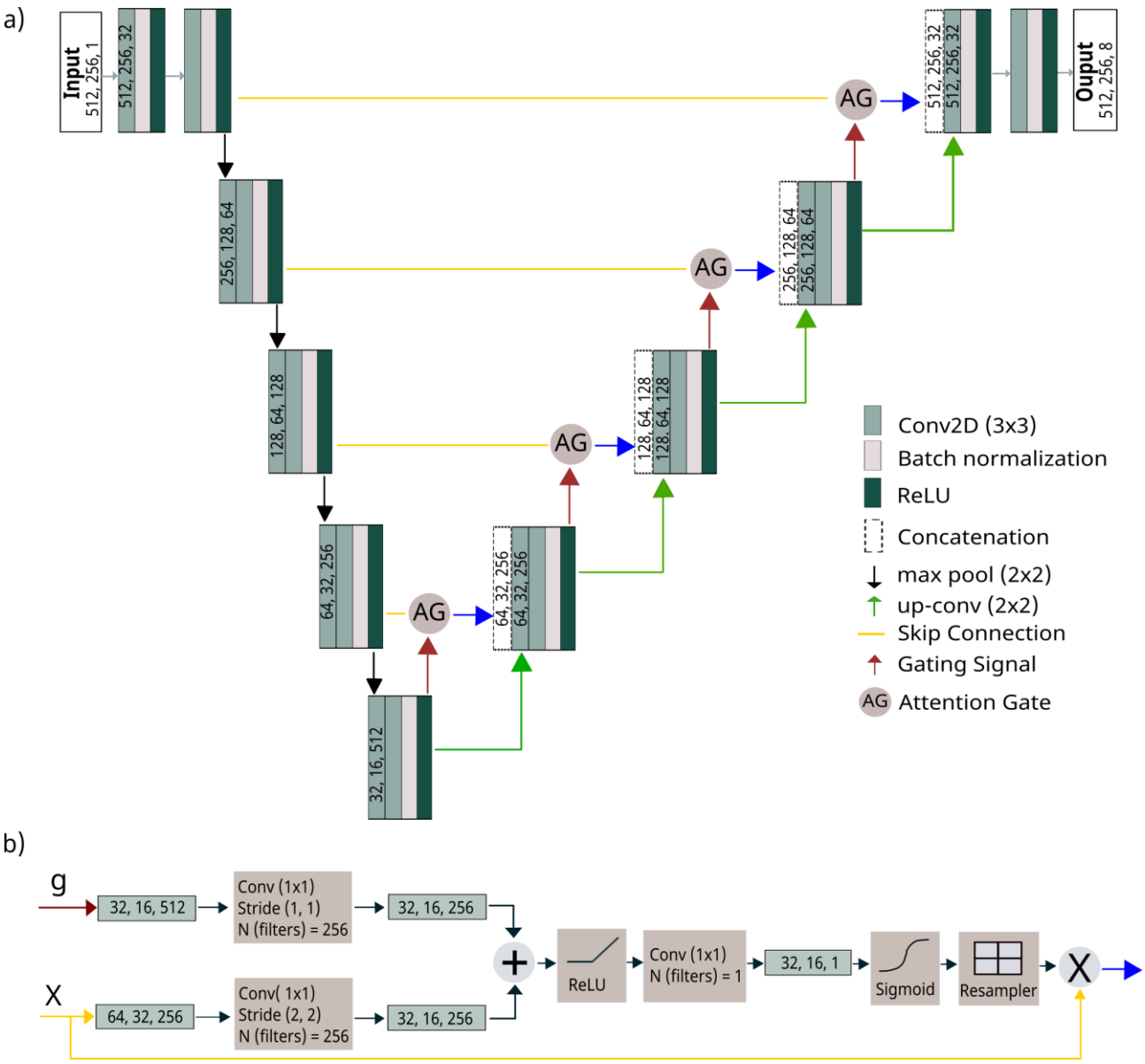


Figure 2. a) Proposed network architecture: unlike in standard U-net architecture, there is an Attention Gate instead of a direct skip connection. b) Attention Gate in which signal X from skip connection and a gating signal G from the next lowest layer of the network is merged.

B. *Proposed network architecture*

A paper from Oktay et al. [15] primarily inspires the architecture we use, in which they introduce the attention mechanism for pancreas segmentation. Standard U-net consists of an encoder, decoder, and skip connections. Each block in the encoder consists of convolutions, activation function (ReLU), and Batch normalization [32]. In addition, there is max-pooling between two blocks, which reduces the image size by half. Similarly, in the decoder, there is upsampling, which doubles the size of the image. Finally, skip connections combine information from the encoder with more spatial information and the decoder with more locally relevant features. The downside of standard U-net architecture is that there are plenty of filters from initial layers without much information.

Attention added to U-net architecture is a way to highlight relevant activations during the training, and in that way, it should lead to better generalization. There are two types of attention: hard and soft attention. Hard attention usually highlights relevant features by cropping. Therefore, it is not differentiable and cannot be used with backpropagation. Contrary, with soft attention, the relevant parts of images get larger weights and vice versa. It is differentiable, and therefore it can be used with backpropagation. With training, attention is increasingly focused on regions of interest (ROI).

We use soft attention to keep all advantages of backpropagation, or more precisely, soft attention is added to skip connections to restrain activations from irrelevant regions. Inputs to the Attention Gate are signals G and X, as shown in Figure 2. G is the gating signal. It comes from the deeper part of the network and has better feature representations. On the other hand, X is a signal from a skip connection. As it comes from the previous layer, it has better spatial information than G. Attention Gate combines X and G. We use stride equals two for X to reduce the image size to be able to sum both signals. By

summing them, aligned weights get larger and unaligned weights are suppressed. After the ReLU activation function, weights range from zero to infinity, so they should be constrained with the sigmoid function, which brings them back between zero and one. Then, upsampling is done to get the original size of X. Finally, X and weights are multiplied element-wise; in that way, vector X is rescaled based on relevance.

C. Training the model

We have used the same parameters for training as the previously published paper [33] to enable better comparison: the categorical cross-entropy loss was used to train the model, the batch size was set to 4, and the AdaBound optimizer [34] was used. Images were resized to 512x256 pixels (i.e., half of the original size). Early stopping was used to prevent overfitting. Also, data augmentation was applied (horizontal flipping and rotating in a small range of angles [-8º, 8º]). As the previously published paper recommended, we used K-fold cross-validation with K equals 6.

The model was trained on Google Colab [35] with a GPU. The model was implemented in Python, using the Keras library with the TensorFlow backend.

IV. Results

We use a Dice score as an evaluation metric as it is the most common evaluation metric for semantic segmentation. In Table 1 are reported Dice scores (mean and standard deviation) for each class and the inter- and intra-observer errors. Also, the prediction errors from published results [33] are reported for the standard U-net model, U-net-like model, and U-net++ model. U-net-like model [17] has encoder-decoder architecture and is enhanced so that each block in the encoder and decoder has residual blocks inspired by ResNet [5], but it lacks skip connections. U-net++ architecture (a nested U-net architecture for medical image segmentation) [16] is inspired by DenseNet architecture [36]. Therefore, dense blocks and convolution layers exist between the encoder and decoder instead of direct skip connections.

It could be noted that the Dice scores for Attention-based U-net are higher in the case of all three fluids. Also, the Dice scores are higher for the area between the IPL/INL (inner plexiform layer and inner nuclear layer) and RPE (retinal pigment epithelium) and the area between RPE and BM (Bruch's membrane). Significantly, segmentations of IRF and SRF are diagnostically most important, and Dice scores are much improved for those classes. The Dice scores for segmenting IRF achieved by the standard U-net model and the Attention-based U-net model are 0.480 (±0.252) and 0.563 (±0.146), respectively. Similarly, the Dice scores for segmenting SRF achieved by the standard U-net model and the Attention-based U-net model are 0.513 (±0.294) and 0.600 (±0.276), respectively.

Figure 3 shows images in case of no severe pathological changes, in case of moderate pathological changes, and a case of extreme pathological changes. In the first case, it could be observed that an Attention-based U-net is the only model which detects SRF. Other models cannot determine SRF when it occupies a small region. In the second case, it could be noticed that Attention-based U-net segmentation prediction is better than predictions from all other models. Better segmentation prediction is even more visible in the third case of extreme pathological changes (all three fluids are present).

V. Discussion

From the overall results, it could be observed that Dice scores are still lower for any model predictions than for inter-observer variability. However, it is questionable how appropriate this comparison is as ophthalmologists never perform manual segmentation as part of clinical practice. Also, there is no consensus among the ophthalmologist on which segmentation accuracy we need. Probably it depends on the segmentation purpose, whether it is just detecting some pathological changes or a case where fluid volume is guidance for therapy.

Also, it could be seen that higher Dice scores are accomplished with standard U-net for highly represented classes and are background from a diagnostical point of view (e.g., the area under BM and the area above ILM). Conversely, lower Dice scores achieved with Attention-based U-net for these classes result from paying more attention to regions of higher interest (e.g., IRF, SRF, PED, RPE-BM, IPL/INL-RPE). As mentioned in previous work [33], a considerable class imbalance makes segmentation extremely challenging, and adding an attention mechanism has improved model performance.

TABLE I. The Dice score (mean and standard deviation) in the inter-observer case, the intra-observer case, for the standard U-net model, the U-net-like model, the U-net++ model, and the Attention-based U-net model.

	Above ILM	**ILM - IPL/INL**	**IPL/INL - RPE**	**RPE - BM**	**Under BM**	**PED**	**SRF**	**IRF**
Inter-observer [33]	0.982 (0.072)	0.950 (0.111)	0.948 (0.112)	0.699 (0.129)	0.989 (0.114)	0.860 (0.301)	0.876 (0.366)	0.735 (0.280)
Intra-observer [33]	0.998 (0.003)	0.973 (0.008)	0.970 (0.117)	0.778 (0.092)	0.998 (0.001)	0.912 (0.242)	0.924 (0.331)	0.844 (0.140)
Standard U-net [33]	**0.995** (0.011)	**0.950** (0.028)	0.923 (0.083)	0.669 (0.129)	**0.988** (0.016)	0.638 (0.173)	0.513 (0.294)	0.480 (0.252)
U-net-like [33]	0.995 (0.004)	0.899 (0.040)	0.890 (0.066)	0.476 (0.132)	0.988 (0.014)	0.533 (0.139)	0.365 (0.291)	0.040 (0.061)
U-net++ [33]	0.992 (0.011)	0.944 (0.032)	0.924 (0.064)	0.641 (0.133)	0.986 (0.017)	0.622 (0.159)	0.465 (0.297)	0.432 (0.265)
Attention-based U-net	0.991 (0.021)	0.945 (0.032)	**0.932** (0.051)	**0.682** (0.126)	0.985 (0.021)	**0.674** (0.147)	**0.600** (0.276)	**0.563** (0.146)

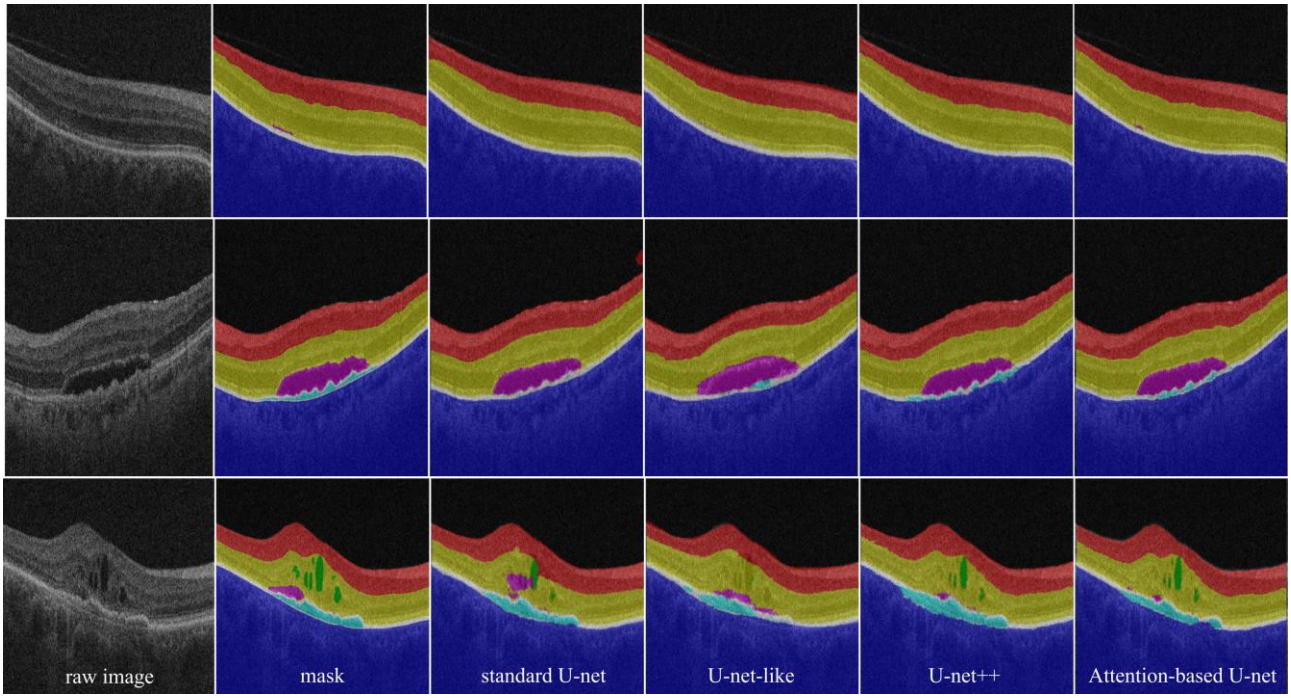


Figure 3. From left to right: raw image, mask (ground truth), predictions from the standard U-net model, U-net-like model, U-net++ model, and Attention-based U-net model. First raw: a case of minor pathological changes (only SRF is present). Second raw: a case of moderate pathological changes (SRF and PED are present). Third raw: a case of severe pathological changes (all fluids are present). All images are cropped to ROI.

By careful inspection of segmentation predictions, it could be spotted that some of the errors present in models without attention remain and probably could not be solved with attention mechanisms. It could be surmised that dominant errors are consequences of a lack of understanding of the retinal structure and pathological traits and hence could not be solved only based on attention mechanisms.

VI. Conclusion

This paper proposes a novel model for the automatic segmentation of biomarkers in retinal OCT images, which combines the attention mechanism with a convolutional neural network. By evaluating results on the publicly available dataset, it could be inferred that the attention mechanisms lead to improved segmentation predictions. Notably, errors due to class imbalance are decreased and lead to better fluid segmentation, which is of prime importance.

Although images in the AROI dataset are highly challenging for segmenting due to extreme pathological deviations in retinal structure, it remains to examine how attention mechanisms improve generalization on unseen images. The lack of extrapolation to unseen data is a significant deficiency of deep learning. The medical field is even more critical due to the small data regime.

This research shows that despite significant progress in the field, we still lack some fundamental concepts to overcome shortcomings. Moreover, even combining the most advanced methods, segmentation predictions are still deficient in some cases of severe pathological alterations. Nevertheless, in our opinion, much progress can be made by active learning and by introducing OCT devices with accompanying automatic segmentation models in clinical practice to help ophthalmologists partially and simultaneously enhance model performance via active learning.

Data and code availability

The AROI dataset is publicly available [31].

Code is available on GitHub:
https://github.com/mmelinscak/OCT-images-segmentation

References

[1] A. Mishra, A. Wong, K. Bizheva, and D. A. Clausi, "Intra-retinal layer segmentation in optical coherence tomography images," *Opt. Express*, vol. 17, no. 26, pp. 23719–23728, Dec. 2009, doi: 10.1364/OE.17.023719.

[2] "ImageNet." http://www.image-net.org/ (accessed Sep. 14, 2020).

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Apr. 2015, Accessed: Jul. 07, 2021. [Online]. Available: http://arxiv.org/abs/1409.1556

[4] C. Szegedy *et al.*, "Going Deeper With Convolutions," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9. Accessed: Aug. 23, 2022. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, Accessed: Sep. 14, 2020. [Online]. Available: https://arxiv.org/abs/1512.03385v1

[6] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.

[7] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,"

arXiv:1704.04861 [cs], Apr. 2017, Accessed: Jul. 07, 2021. [Online]. Available: http://arxiv.org/abs/1704.04861

[8] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning*, May 2019, pp. 6105–6114. Accessed: Aug. 23, 2022. [Online]. Available: https://proceedings.mlr.press/v97/tan19a.html

[9] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." arXiv, Jun. 03, 2021. doi: 10.48550/arXiv.2010.11929.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]*, May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[11] J. Wang *et al.*, "S-D Net: Joint Segmentation and Diagnosis Revealing the Diagnostic Significance of Using Entire RNFL Thickness in Glaucoma," p. 10.

[12] Z. Gu *et al.*, "CE-Net: Context Encoder Network for 2D Medical Image Segmentation," *IEEE Trans. Med. Imaging*, vol. 38, no. 10, pp. 2281–2292, Oct. 2019, doi: 10.1109/TMI.2019.2903562.

[13] J. I. Orlando *et al.*, "U2-Net: A Bayesian U-Net model with epistemic uncertainty feedback for photoreceptor layer segmentation in pathological OCT scans," *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 1441–1445, Apr. 2019, doi: 10.1109/ISBI.2019.8759581.

[14] W. Liu, Y. Sun, and Q. Ji, "MDAN-UNet: Multi-Scale and Dual Attention Enhanced Nested U-Net Architecture for Segmentation of Optical Coherence Tomography Images," *Algorithms*, vol. 13, no. 3, p. 60, Mar. 2020, doi: 10.3390/a13030060.

[15] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas." arXiv, May 20, 2018. doi: 10.48550/arXiv.1804.03999.

[16] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *arXiv:1807.10165 [cs, eess, stat]*, Jul. 2018, Accessed: Mar. 30, 2020. [Online]. Available: http://arxiv.org/abs/1807.10165

[17] K. Team, "Keras documentation: Image segmentation with a U-Net-like architecture." https://keras.io/examples/vision/oxford_pets_image_segmentation/ (accessed Sep. 14, 2020).

[18] M. Wang *et al.*, "MsTGANet: Automatic Drusen Segmentation From Retinal OCT Images," *IEEE Trans. Med. Imaging*, vol. 41, no. 2, pp. 394–406, Feb. 2022, doi: 10.1109/TMI.2021.3112716.

[19] D. Kermany, K. Zhang, and M. Goldbaum, "Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images," vol. 3, Jun. 2018, doi: 10.17632/rscbjbr9sj.3.

[20] Z. Jiang *et al.*, "Computer-aided diagnosis of retinopathy based on vision transformer," *J. Innov. Opt. Health Sci.*, vol. 15, no. 02, p. 2250009, Mar. 2022, doi: 10.1142/S1793545822500092.

[21] P. P. Srinivasan *et al.*, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomed Opt Express*, vol. 5, no. 10, pp. 3568–3577, Sep. 2014, doi: 10.1364/BOE.5.003568.

[22] G. Cao, S. Zhang, H. Mao, Y. Wu, D. Wang, and C. Dai, "A single-step regression method based on transformer for retinal layer segmentation," *Phys. Med. Biol.*, vol. 67, no. 14, p. 145008, Jul. 2022, doi: 10.1088/1361-6560/ac799a.

[23] S. J. Chiu, M. J. Allingham, P. S. Mettu, S. W. Cousins, J. A. Izatt, and S. Farsiu, "Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema," *Biomed Opt Express*, vol. 6, no. 4, pp. 1172–1194, Mar. 2015, doi: 10.1364/BOE.6.001172.

[24] C. Playout, R. Duval, M. C. Boucher, and F. Cheriet, "Focused Attention in Transformers for interpretable classification of retinal images," *Medical Image Analysis*, vol. 82, p. 102608, Nov. 2022, doi: 10.1016/j.media.2022.102608.

[25] X. Liu, S. Wang, Y. Zhang, D. Liu, and W. Hu, "Automatic fluid segmentation in retinal optical coherence tomography images using attention based deep learning," *Neurocomputing*, vol. 452, pp. 576–591, Sep. 2021, doi: 10.1016/j.neucom.2020.07.143.

[26] G. Lazaridis, M. Xu, S. S. Afgeh, and D. Garway-Heath, "Bio-Inspired Attentive Segmentation of Retinal OCT imaging," p. 10.

[27] X. Mao *et al.*, "Deep Learning with Skip Connection Attention for Choroid Layer Segmentation in OCT Images," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Jul. 2020, pp. 1641–1645. doi: 10.1109/EMBC44109.2020.9175631.

[28] X. Xi, X. Meng, Z. Qin, X. Nie, Y. Yin, and X. Chen, "IA-net: informative attention convolutional neural network for choroidal neovascularization segmentation in OCT images," *Biomed. Opt. Express*, vol. 11, no. 11, p. 6122, Nov. 2020, doi: 10.1364/BOE.400816.

[29] Z. Fu *et al.*, "MPG-Net: Multi-Prediction Guided Network for Segmentation of Retinal Layers in OCT Images," *arXiv:2009.13634 [cs, eess]*, Sep. 2020, Accessed: Apr. 06, 2022. [Online]. Available: http://arxiv.org/abs/2009.13634

[30] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in Vision: A Survey," *ACM Comput. Surv.*, vol. 54, no. 10s, p. 200:1-200:41, Rujan 2022, doi: 10.1145/3505244.

[31] M. Melinščak, M. Radmilović, Z. Vatavuk, and S. Lončarić, "AROI: Annotated Retinal OCT Images database," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Sep. 2021, p. accepted for publication.

[32] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," p. 11.

[33] M. Melinščak, M. Radmilović, Z. Vatavuk, and S. Lončarić, "Annotated retinal optical coherence tomography images (AROI) database for joint retinal layer and fluid segmentation," *Automatika*, vol. 62, no. 3–4, pp. 375–385, Oct. 2021, doi: 10.1080/00051144.2021.1973298.

[34] L. Luo, Y. Xiong, Y. Liu, and X. Sun, "Adaptive Gradient Methods with Dynamic Bound of Learning Rate," *arXiv:1902.09843 [cs, stat]*, Feb. 2019, Accessed: Feb. 04, 2021. [Online]. Available: http://arxiv.org/abs/1902.09843

[35] "Google Colaboratory." https://colab.research.google.com/notebooks/intro.ipynb (accessed May 17, 2020).

[36] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *arXiv:1608.06993 [cs]*, Jan. 2018, Accessed: Jul. 19, 2020. [Online]. Available: http://arxiv.org/abs/1608.06993

Recognizing Activities of Daily Living using Multi-sensor Smart Glasses

Simon Stankoski*, Borjan Sazdov*, John Broulidakis*, Ivana Kiprijanovska*, Bojan Sofronievski*, Sophia Cox*, Martin Gjoreski**, James Archer*, Charles Nduka*, and Hristijan Gjoreski *, ***
* Emteq Ltd., Brighton BN1 9SB, UK
** Faculty of Informatics, Università della Svizzera Italiana, Lugano, Switzerland
*** Ss. Cyril and Methodius University in Skopje, Skopje, N. Macedonia
simon.stankoski@emteqlabs.com

***Abstract* - Continuous and automatic monitoring of an individual's physical activity using wearable devices provides valuable insights into their daily habits and patterns. This information can be used to promote healthier lifestyles, prevent chronic diseases, and improve overall well-being. Smart glasses are an emerging technology that can be worn comfortably and continuously. Their wearable nature and hands-free operation make them well suited for long-term monitoring of physical activity and other real-world applications. To this end, we investigated the ability of the multi-sensor OCOsense™ smart glasses to recognize everyday activities. We evaluated three end-to-end deep learning architectures that showed promising results when working with IMU (accelerometer, gyroscope, and magnetometer) data in the past. The data used in the experiments was collected from 18 participants who performed pre-defined activities while wearing the glasses. The best architecture achieved an F1 score of 0.81, demonstrating its ability to effectively recognize activities, with the most problematic categories being standing vs. sitting.**

Keywords - human activity recognition, smart glasses, IMU

I. INTRODUCTION

Smart sensing with wearable devices has enabled numerous applications in ubiquitous computing. One of the tasks that has emerged as an important pre-requisite for other applications is Human Activity Recognition (HAR). HAR is a broad field of study that identifies a person's movements or actions based on sensor data. It covers a broad range of activities, from typical indoor activities, such as walking, talking, standing, and sitting, to more focused activities, such as those performed in a kitchen or factory floor [1]. HAR can also be used for tracking transportation modes [2] and stress levels [3], or as a part of disease severity detection methods for various diseases such as Parkinson's disease and depression monitoring [4]. In healthcare, HAR is used in identifying symptoms of movement disorders or neurological diseases, aiding in diagnosis or symptom management [5].

Historically, collecting sensor data for HAR was a challenging and expensive process, which required custom hardware. This made it difficult to collect and analyze data on a large scale and limited the potential applications of the technology. However, with the widespread availability of smartphones and other wearable devices for fitness and health monitoring, this scenario has changed significantly. With the recent development of technology, sensor data can be easily obtained, and therefore a larger amount of HAR-based studies can be made.

Body movements are among the main measurable components of behavior and play a significant role in the study of HAR. The use of motion sensor devices, such as accelerometers and gyroscopes, has made it possible for scientists to measure human behavior outside of the laboratory environment. These devices supply researchers with objective and quantitative data based on the three spatial dimensions, making it possible to analyze human movements in detail.

Over the years, a number of wearable devices and sensors have been proposed for automatic detection of everyday activities. Early studies in this field explored the use of sensors placed on different parts of the body, so they can enhance detection accuracy. Over time, the selection of sensors was refined, taking into account two key factors, the capability of the sensors to precisely detect daily activities, and the practicality that encompasses user comfort and acceptability. As a result, the most commonly used devices for HAR are smartphones, wristbands, and smartwatches. However, the recent advancements in technology have enabled the development of compact, lightweight, and stylish smart glasses. Therefore, the researchers in this field started exploring the possibility to detect everyday activities using data coming from sensors embedded in smart glasses [6][7]. This device is suitable for HAR because it can provide an immersive and hands-free experience, enabling the user to perform physical activity without any distractions.

In this study, we investigated the ability of the multi-sensor OCOsense™ smart glasses to recognize everyday activities. For this purpose, we evaluated three end-to-end deep learning architectures that showed promising results when working with sensor data in the past. Additionally, we explored the inter-subject variability for the recorded activities. Finally, we explored how we can include temporal information and detect transitions between activities.

II. RELATED WORK

The HAR domain has been thoroughly explored in the past using body-worn sensors. Most of the multi-sensor HAR approaches involve machine learning (ML) algorithms (e.g., Random Forest (RF), Support-vector

Machines (SVM), and k-Nearest Neighbors (KNN)) to build models from features extracted from each modality of the body-worn sensor. More recently, researchers started exploring novel approaches based on deep learning (DL) for HAR [8][9].

In the field of HAR, sensors are usually placed on the wrists, ankles, hips, waist, or torso of the user. Approaches using head-mounted devices are rather scarce. Loh et al. [10] used a head-worn accelerometer, barometer, and GPS sensors with an SVM for fitness activity classification. Additionally, Zhang et al. [6] and Farooq et al. [11] proposed the use of head-mounted sensors to detect eating and chewing events. Head-AR [12] is a method based on weighted ensemble learning used for HAR from sensors mounted on a VR device.

Despite the limited work in the field, the use of smart glasses for human activity recognition is a promising approach with a unique set of challenges. To this end, Faye et al. [13] published a dataset that is collected with commercially available glasses, smartwatch, and smartphone. The smart glasses provide data from an embedded IMU. This dataset has, however, some noticeable drawbacks. First, only one user participated in the experiment. Moreover, there is no well-defined set of activities or well-defined protocol, which makes it difficult to evaluate or extend.

Another use of Electrooculography (EOG) J!NS MEME glasses have been demonstrated by Ishimaru et al. [14]. The study provides a signal level assessment of MEME glasses and shows the ability to distinguish 4 activities (typing, reading, eating, talking) with an accuracy of 70% for 6 second windows and up to 100% for a 1-minute majority decision.

Meyer et. al [15][16] conducted two studies that propose context-aware human activity recognition (HAR) system for smart glasses by combining eye movement features from laser feedback interferometry (LFI) sensors and head movement features from an IMU. The method presented in his studies was DL based approach using convolutional neural networks (CNNs).

More recently, UCA-EHAR was introduced, and it consists of 20 subjects performing 8 different activities, while wearing Ellcie Healthy (EH) smart glasses equipped with a gyroscope, accelerometer, and barometer, was proposed by Novac et al. [7]. RNN was used for classification of different activities and analysis of the power consumption during live inference on the smart glasses' microcontroller was performed.

Another effort to develop a system for HAR using smart glasses (Google Glass Explorer Edition XE 22) has been made by Wang et al. [17]. The authors compare the classification performance of a Support Vector Machine (SVM) with data collected either from a smartphone or smart glasses for four activities (biking, jogging, movie watching, and video gaming).

III. DATA

A total of 18 young adult research volunteers (9 males and 9 females) aged between 18 – 37 (23 ± 5) participated in the data collection process, which included individuals engaging in activities of daily living (ADL) in a realistic home environment. There were no strict criteria that needed to be met for individuals to participate in the study. Instead, all individuals who were able to comfortably wear the one-size-fits-all glasses were considered eligible to participate. This approach ensured that the sample size was not limited by factors such as age, gender, or physical ability. Each participant was equipped with emteq's OCOsense™ smart glasses (Figure *1*), which are fitted with: (i) a 9-axis inertial measurement (IMU) sensor (accelerometer, gyroscope, and magnetometer), (ii) pressure sensor; and (iii) optomyography sensors that measure skin movement. For this study, we used only the data provided by the IMU and pressure sensor. During the data collection procedure, the sensors were continuously sampled at a fixed rate of 50Hz.

The data from the smart glasses was streamed in real-time via Bluetooth to an application running on an iPad. The researcher conducting the data collection also performed real-time labeling. To facilitate this, the app featured a screen with buttons representing pre-determined activities. The researcher selected the appropriate button based on the activity being performed by the participant. Additionally, the dataset consists of synchronized video data from five external cameras placed around the apartment, which ensures labeling the activities as accurately as possible.

Prior to the start of the data collection procedure, each participant was instructed on how to wear the glasses properly, and we ensured the fit and comfort. The data collection procedure involved three predefined scenarios that covered the activities of interest.

Each participant was instructed about the scenarios, however, there was no limitation on how to perform a specific activity. The predefined scenarios resulted in three categories of activities:

- Posture activities – included activities related to the posture of the participant, such as sitting, standing, walking, lying and transitions between sitting and standing.
- Personal hygiene activities – included activities that are part of an individual's daily hygiene routine, namely, washing hands and brushing teeth.
- Dietary intake activities – included activities that are related to nutrient intake, such as eating and drinking.

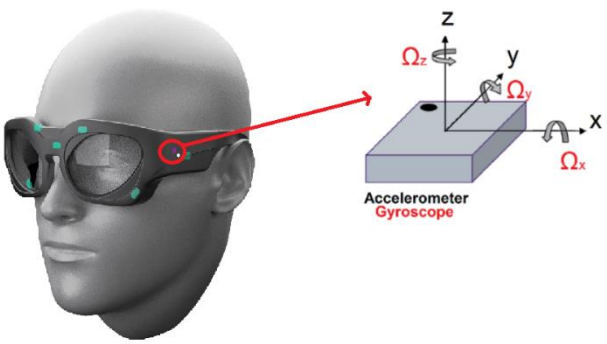


Figure 1. OCOsense™ smart glasses. The green colored rectangles represent the OCO™ sensors. The 9-axis IMU is represented by the purple rectangle.

Each participant was recorded for 2 hours, resulting in a dataset with 36 hours of labelled data. For the purposes of this study, we analyzed the following activities: sitting, standing, lying, walking, eating, drinking and hygiene, which includes washing hands and brushing teeth. The distribution of the data is shown in Figure 2.

IV. METHOD

Before passing the data as an input to the selected DL architectures we filtered the signals using a fifth-order median filter. Once the data was filtered, we generated magnitude for the accelerometer and gyroscope sensors. The next step in the process was to select the optimal window size for windowing the data. We tested various window sizes, and experimentally we determined that window size of 4 seconds contains enough temporal information from where the models can learn the patterns between the different activities. As a result, the signals were segmented using a window size of 4 seconds with a 2 second slide between consecutive windows.

Once the signals were prepared, we used them as an input to three end-to-end DL architectures. The chosen DL architectures are based on well-known architectures used in the past for various problems, such as computer vision and time-series related topics. In this study, the selected architectures were adjusted to work with one dimensional signals.

- Spectro-Temporal Residual Neural Network (STResNet). For each sensor utilizes channel-specific residual blocks and spectrograms to extract temporal and frequency information. This architecture already proved to be successful in previous studies for various applications, such as HAR from smartphone sensors [18], for heart failure detection from heartbeat sounds [19], for monitoring driver distractions from physiological and visual signals [20].
- Inception Network – combines multiple blocks of channel-specific inception modules, which should increase the performance of vanilla CNNs. This performance of this architecture was evaluated in a previous study aimed at detection of eating habits based on smartwatch data [21].
- Feature Level Fusion Network – combines channel-specific CNNs and bidirectional LSTM networks to extract features from the raw sensor data and further aggregates the extracted features by using feature level fusion. This network was successfully implemented in detecting gait abnormalities using data from wrist-worn inertial sensor [22].

For all three architectures the cross-entropy loss was used as the objective function for training. Furthermore, all the models were trained using an Adam optimizer with a dynamic learning rate. The initial learning rate was set to 1e-3 and decreased by a factor of 0.1 every fifty epochs. To avoid overfitting to the training data, we used early stopping callback for all architectures. The callback was set to monitor the F1-macro score of the validation set. If the F1-macro score was not improved over ten epochs the training was stopped. The main reason for monitoring the F1-macro score is because it provides unweighted mean of the F1-scores calculated for each class separately. This way can have a general sense of how well the trained models perform over all classes. For all models NVIDIA Titan Xp GPU was used to accelerate the training process.

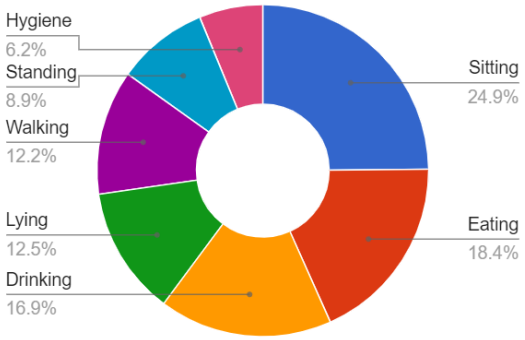


Figure 2. Distribution of the seven activities.

V. EXPERIMENTAL SETUP AND RESULTS

To estimate the performance of the previously described DL architectures, we used a 6-fold cross validation (CV) approach. The created folds were participant-based, which means that the data from one participant only appeared in a single fold. The training procedure is repeated 6 times, each time using a different fold as the testing set. By using subject-based k-fold cross-validation, the inter-subject variability of the data is taken into account and helps to reduce the risk of overfitting. To assess the performance of the trained models, we used the F1-macro score and accuracy. The reported results in this study are obtained using predictions for all folds, across all tasks and they are shown as a mean value with the standard deviation between folds.

A. *Performance Evaluation*

Table 1 shows the performance of each architecture described in Section IV, including F1-macro score, model size, and the number of parameters. Based on the obtained results, we can see that all architectures show promising results.

As shown, STResNet performed the best with an F1- macro score of 0.81. Additionally, the standard deviation of the results obtained with the STResNet architecture is quite small, which suggests that this DL approach can learn general characteristics and perform well on data from unseen participants. Also, obtaining stable results over all participants proves our hypothesis that IMU data from head-mounted device is less prone to movement noise, hence a model based on such data can perform well on new participants. The other two architectures show similar results, however, it can be seen that the Inception architecture is a bit unstable based on the standard deviation between the folds. Additionally, if we compare the F1-macro score of the Inception and Feature Fusion architectures to the STResNet architecture and take into account the size of the models, we can say that the difference in performance is acceptable. To gain better understanding of each architecture performance the confusion matrices are shown in Figure *3*. This analysis provides us with an understanding of the model's strengths and weaknesses, which can be useful for optimization.

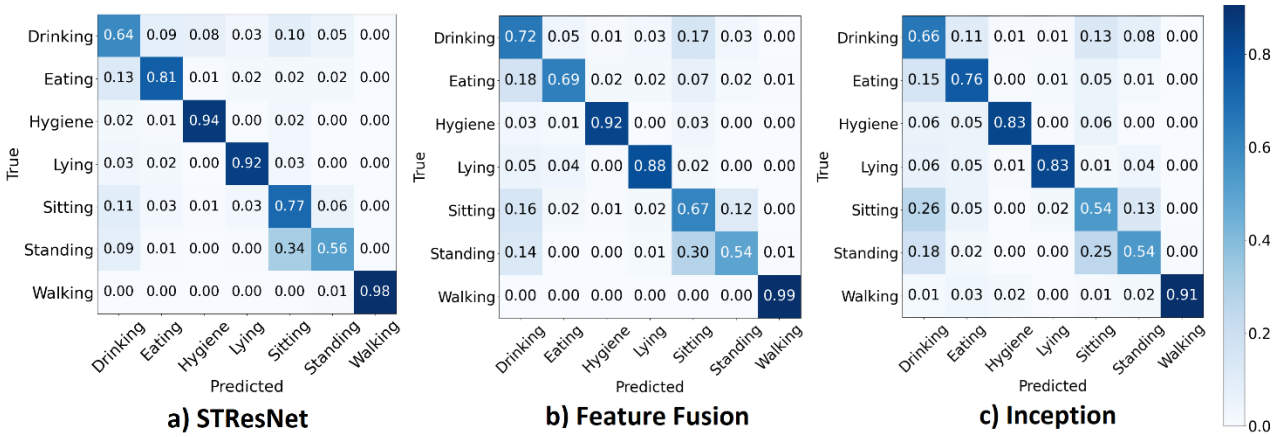


Figure 3. Confusion matrices for each DL HAR model.

TABLE 1: PERFORMANCE EVALUATION OF EACH DL ARCHITECTURE

	F1-Macro	Model size (MB)	Network Parameters
STResNet	0.81±0.025	516	44,044,203
Inception	0.74±0.07	121	7,988,775
Feature Fusion	0.76±0.04	143	9,648,395

For the STResNet architecture, it is evident that the model is able to accurately detect activities such as hygiene, lying, and walking. However, the model faced some challenges when differentiating between standing and sitting. Additionally, in some cases, the model misclassifies eating and drinking activities. The drinking activity sometimes is predicted as sitting and less often as standing.

The Feature Fusion and Inception architectures also exhibit similar limitations as STResNet. Specifically, the Feature Fusion network had a higher misclassification rate for the eating activity. Meanwhile, the Inception architecture had the lowest F1-Macro score for drinking due to a higher rate of misclassifying drinking as sitting. In comparison with STResNet, the misclassification of drinking and eating activities with sitting and standing is noticeably higher.

One of the key issues for all three architectures is the differentiation between standing and sitting. This was in a way expected, given that these two activities resemble each other when looking through the prism of an IMU in smart glasses, i.e., the orientation of the IMU is similar during the two activities. Mixing drinking and eating activities with sitting and standing also arises from this issue. During the drinking activity, the subjects were in a seated position. Additionally, the eating activity was also primarily performed while sitting. This is the cause for misclassifying these activities. Nonetheless, we can see that the STResNet architecture in some situations managed to differentiate between standing and sitting. The reasoning why we expect the models to learn the difference is analyzed in more details in the subsection Exploring Standing vs. Sitting.

B. Sampling Rate Analysis

The OCOsense™ smart glasses provide sensor data sampled at a fixed rate of 50 Hz. Based on studies done in this field [18], it was established that this sampling frequency is sufficient for most activity recognition tasks. However, high sampling frequency and processing large amount of data can significantly reduce the device's battery life. To try to mitigate these issues we decided to further evaluate the best performing architecture (STResNet) with down sampled sensor data to 25Hz and 10 Hz.

The results for this experiment are shown in Table 2. It can be clearly seen that the best performance was achieved with 50 Hz data. However, if we compare the results with those achieved when the data was down sampled to 25 Hz, it can be noticed that the performance was decreased minimally. Given the results in similar fields, it's not surprising that 25 Hz is sufficient for recognizing human motion. Additionally, the training time for the model was reduced by almost one third. The reduced training time of the model with 25 Hz data was also an important consideration, as it allows multiple experiments to be conducted and do analysis faster.

The experiment where the data was down sampled to 10Hz resulted in more significant loss of 0.06 F1-macro score compared to the results where 50 Hz data was used. Nonetheless, these results still show that the model is capable of learning how to differentiate the activities in most of the cases. Similarly as before, by further reducing the sampling frequency to 10 Hz, the training time of the models was reduced to only one third of the original training time when the data was sampled with 50 Hz.

C. Sensor-specific Experiments

To understand the usefulness of each sensor modality, we investigated the performance of the STResNet architecture by using each modality individually as well as their combinations. The list of sensor modalities tested in this experiment includes data from accelerometer (A), gyroscope (G), magnetometer (M), and pressure (P) sensor. As a result, we had to repeat the training procedure 15 times. Based on the results from the previous section we decided to reduce the training time and worked with 25 Hz data.

TABLE 2: SAMPLING RATES ANALYSIS - RESULTS

Sampling rate (HZ)	F1-Macro	Elapsed time (hh:mm:ss)
50	0.81±0.025	01:23:50
25	0.79±0.025	0:58:49
10	0.75±0.025	0:27:01

The obtained results are shown in Figure 4. The comparison of the results using data from a single sensor modality is colored with blue bars and shows that the accelerometer and the gyroscope are the most informative. Additionally, we can see that when using only data from the pressure sensor we cannot differentiate between activities.

Furthermore, based on the combinations of data from two sensor modalities, which are colored with yellow, we can see that in certain situations we get improvement over models based on a single sensor modality. The best performing model is based on a combination of data from gyroscope and magnetometer. Next, we have the combination of accelerometer and gyroscope, which also outperforms all single modality results. The remaining combinations did not introduce any improvement over the best performing single sensor modality.

The results where combinations of three sensor modalities are combined are shown with green bars on the plot. Here we can see that in all cases the models based on three sensor modalities outperform all previously discussed combinations. It is interesting to notice that by combining accelerometer, gyroscope, and pressure data the results were improved by 0.07 compared to only accelerometer and gyroscope data combination.

Finally, the red colored bar shown in Figure 4 represents the results when data from all four sensor modalities is combined. Here, it can be clearly seen that model based on all sensor modalities outperforms all previously discussed combinations. Also, we can see that our idea to include pressure data in the model results in improvement, compared to only working with accelerometer, gyroscope, and magnetometer.

D. Exploring Standing vs Sitting

The confusion matrix presented in Figure *3* showed that the STResNet model wrongly predicted classes standing and sitting. To further investigate the separability of these activities we included transitions from sitting to standing and vice versa and analyzed the raw signal acquired from the pressure sensor.

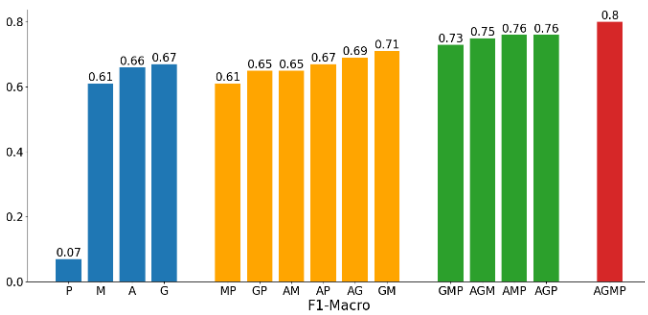


Figure 4. Sensor-specific experiments. A-Accelerometer, G-Gyroscope, M-Magnetometer, P-Pressure sensor.

The result from the analysis is visualized in Figure 5. The different activities are color-coded. The figure shows that the pressure of the subject changes during the four different activities. One interesting observation is that during the transitions "sitting down" and "standing up" the pressure increases and decreases accordingly. Being able to recognize these transitions accurately using data from the pressure sensor, may be useful in distinguishing between standing and sitting activities. Thus, a possible solution for the standing vs. sitting problem would be to have a temporal model that tracks the transitions ("sitting down", "standing up"). Once we know the last transition of the user, we can more easily infer whether the user is sitting or standing. For example, if the last transition was "sitting down," then "sitting" is a more probable next activity than "standing."

VI. CONCLUSION

In this study, we analyzed data from the novel OCOsense™ multi-sensor smart glasses and their ability to recognize everyday activities in combination with deep learning models. For this purpose, we collected a HAR dataset that includes common activities of daily living.

We explored three end-to-end deep learning models for HAR, including STResNet, Inception, Feature Fusion models. The best performing model, STResNet, achieved an overall f1-macro score of 0.81 and an accuracy of 0.80. Notably, the model achieved these results without the need for extensive feature engineering in an end-to-end manner, highlighting the power of deep learning in activity recognition.

Additionally, we conducted an experiment where we tested different sampling frequencies. The performance decreased slightly when the data was down-sampled to 25 Hz. A bit larger reduction was obtained when testing with 10 Hz, however, the results were still acceptable.

The analysis where we evaluated the contribution of individual sensor modalities and their combinations showed that the gyroscope is the most informative when working with a single modality. Additionally, this experiment showed that combination of all sensor modalities resulted in the best performance.

Although the presented results in this study are promising, there are still open challenges that we plan to improve in the near future. For instance, the challenge of accurately distinguishing between sitting and standing activities remained present in all experiments. The reason is that the orientation of the head (i.e., the glasses and the IMU) is the same during these two activities. However, we showed that by explicitly analyzing transitions using pressure data we might be able easily distinguish these classes. Additionally, future work for this study can include expanding the dataset to include a more diverse population and a wider range of activities. Furthermore, testing the model in real-world scenarios and evaluating its performance in various environments can provide valuable insights into its practical applicability.

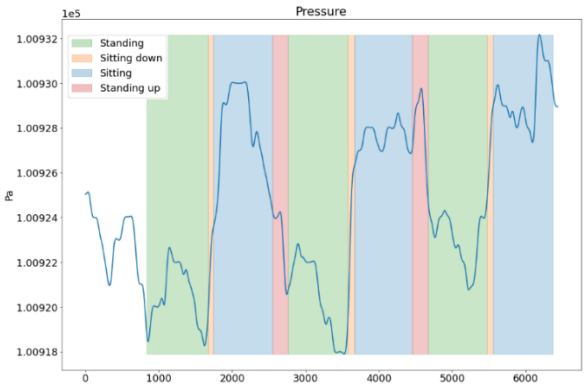


Figure 5. Pressure sensor signal visualization during sitting, standing sitting-to-standing and standing-to-sitting transition data.

Another important aspect is to investigate the generalization capability of the proposed model on other datasets.

ACKNOWLEDGEMENT

This work was supported by NIHR i4i FAST (reference no. 33180) under the project "Context-Aware Multimodal Pervasive Sensing (CAMPS) of behaviour in daily life", and by Innovate UK under the project "Mobile Observation of Depression (MOOD) platform for digital phenotyping" (Grant number 105207). Hristijan Gjoreski's work was partially funded by the WideHealth project (EU's Horizon 2020 research and innovation programme, grant agreement No. 952279).

REFERENCES

[1] Roitberg, A., Somani, N., Perzylo, A., Rickert, M., & Knoll, A. (2015, November). Multimodal human activity recognition for industrial manufacturing processes in robotic workcells. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (pp. 259-266).

[2] Kalabakov S, Stankoski S, Kiprijanovska I, Andova A, Reš?i? N, Janko V, Gjoreski M, Gams M, Luštrek M. What Actually Works for Activity Recognition in Scenarios with Significant Domain Shift: Lessons Learned from the 2019 and 2020 Sussex-Huawei Challenges. Sensors. 2022; 22(10):3613.

[3] Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2016, September). Continuous stress detection using a wrist device: in laboratory and real life. In proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct (pp. 1185-1193).

[4] Albert, M. V., Toledo, S., Shapiro, M., & Kording, K. (2012). Using mobile phones for activity recognition in Parkinson's patients. Frontiers in neurology, 3, 158.

[5] Kiprijanovska I, Gjoreski H, Gams M. Detection of Gait Abnormalities for Fall Risk Assessment Using Wrist-Worn Inertial Sensors and Deep Learning. Sensors. 2020; 20(18):5373

[6] Rui Zhang and Oliver Amft. Monitoring chewing and eating in free-living using smart eyeglasses. IEEE journal of biomedical and health informatics, 22(1):23-32, 2017.

[7] Novac, P.-E.; Pegatoquet, A.; Miramond, B.; Caquineau, C. UCA-EHAR: A Dataset for Human Activity Recognition with Embedded AI on Smart Glasses. Appl. Sci. 2022, 12, 3849. https://doi.org/10.3390/app12083849

[8] Wenchao Jiang and Zhaozheng Yin. 2015. Human Activity Recognition Using Wearable Sensors by Deep Convolutional Neural Networks. In Proceedings of the 23rd ACM international conference on Multimedia (MM '15). Association for Computing Machinery, New York, NY, USA, 1307–1310. https://doi.org/10.1145/2733373.2806333

[9] Martin Gjoreski, Stefan Kalabakov, Mitja Luštrek, Matjaž Gams, and Hristijan Gjoreski. 2019. Cross-dataset deep transfer learning for activity recognition. In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, New York, NY, USA, 714–718. https://doi.org/10.1145/3341162.3344865

[10] Darrell Loh, Tien J Lee, Shaghayegh Zihajehzadeh, Reynald Hoskinson, and Edward J Park. Fitness activity classification by using multiclass support vector machines on head-worn sensors. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 502-505. IEEE, 2015.

[11] Muhammad Farooq and Edward Sazonov. Accelerometer-based detection of food intake in free-living individuals. IEEE sensors journal, 18(9):3752-3758, 2018.

[12] Hristijan Gjoreski, Ivana Kiprijanovska, Simon Stankoski, Stefan Kalabakov, John Broulidakis, Charles Nduka, and Martin Gjoreski. "Head-ar: Human activity recognition with head-mounted imu using weighted ensemble learning." In Activity and Behavior Computing, pp. 153-167. Springer, Singapore, 2021.

[13] Faye, S.; Louveton, N.; Jafarnejad, S.; Kryvchenko, R.; Engel, T. An Open Dataset for Human Activity Analysis using Smart Devices. 2017. Available online: https://www.kaggle.com/datasets/sasanj/human-activity-smart-devices (accessed on 22 September 2021).

[14] Shoya Ishimaru, Kai Kunze, Yuji Uema, Koichi Kise, Masahiko Inami, and Katsuma Tanaka. Smarter eyewear: using commercial eog glasses for activity recognition. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pages 239-242, 2014.

[15] Johannes Meyer, Adrian Frank, Thomas Schlebusch, and Enkeljeda Kasneci. 2022. A CNN-based Human Activity Recognition System Combining a Laser Feedback Interferometry Eye Movement Sensor and an IMU for Context-aware Smart Glasses. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 5, 4, Article 172 (Dec 2021), 24 pages. https://doi.org/10.1145/3494998

[16] Johannes Meyer, Adrian Frank, Thomas Schlebusch, and Enkelejda Kasneci. 2022. U-HAR: A Convolutional Approach to Human Activity Recognition Combining Head and Eye Movements for Context-Aware Smart Glasses. Proc. ACM Hum.-Comput. Interact. 6, ETRA, Article 143 (May 2022), 19 pages. https://doi.org/10.1145/3530884

[17] Ho, J.; Wang, C.M. User-Centric and Real-Time Activity Recognition Using Smart Glasses. In Proceedings of the 11th International Conference on Green, Pervasive, and Cloud Computing, Xi'an, China, 6-8 May 2016; pp. 196-210. https://doi.org/10.1007/978-3-319-39077-2_13

[18] Martin Gjoreski, Vito Janko, Gašper Slapničar, Miha Mlakar, Nina Reščič, Jani Bizjak, Vid Drobnič, Matej Marinko, Nejc Mlakar, Mitja Luštrek, Matjaž Gams, Classical and deep learning methods for recognizing human activities and modes of transportation with smartphone sensors, Information Fusion,Volume 62,2020,Pages 47-62, https://doi.org/10.1016/j.inffus.2020.04.004

[19] M. Gjoreski, A. Gradisek, B. Budna, M. Gams and G. Poglajen, "Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds", IEEE Access, vol. 8, pp. 20313-20324, 2020.

[20] M. Gjoreski, M. Ž. Gams, M. Luštrek, P. Genc, J. -U. Garbas and T. Hassan, "Machine Learning and End-to-End Deep Learning for Monitoring Driver Distractions From Physiological and Visual Signals" in IEEE Access, vol. 8, pp. 70590-70603, 2020, doi: 10.1109/ACCESS.2020.2986810.

[21] Stankoski, S.; Jordan, M.; Gjoreski, H.; Luštrek, M. Smartwatch-Based Eating Detection: Data Selection for Machine Learning from Imbalanced Data with Imperfect Labels. Sensors 2021, 21, 1902. https://doi.org/10.3390/s21051902

[22] Kiprijanovska, I.; Gjoreski, H.; Gams, M. Detection of Gait Abnormalities for Fall Risk Assessment Using Wrist-Worn Inertial Sensors and Deep Learning. Sensors 2020, 20, 5373. https://doi.org/10.3390/s20185373

Reconstruction of short genomic sequences with graph convolutional networks

Lovro Vrček*†, Xavier Bresson‡, Thomas Laurent§, Martin Schmitz†‡, Mile Šikić*†
* Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
† Genome Institue of Singapore, A*STAR
‡ National University of Singapore
§ Loyola Marymount University
lovro.vrcek@fer.hr, mile.sikic@fer.hr

***Abstract*—Genome reconstruction, without prior knowledge about the sequence we are reconstructing, is performed with tools called *de novo* genome assemblers. These tools rely on numerous heuristics and usually provide a fragmented reconstruction, even for sequences shorter than the entire genomes or chromosomes. One of the most common approaches to *de novo* assembly, called Overlap-Layout-Consensus, constructs a graph from short overlapping fragments, which heuristics then simplify and find a path through. In this work, we explore how graph neural networks (GNNs) can assist with this task, and show that the GNN-based Layout phase can reconstruct longer sequences than naive search algorithms or heuristics deployed in de novo assemblers, with no significant difference in compute time on sequences up to 10 Mbp in length.**

***Keywords*—genome assembly, graph neural networks**

I. INTRODUCTION

Fast and accurate *de novo* genome assembly is one of the most difficult problems in bioinformatics and it remains unsolved to this day. It focuses on reconstructing the original genomic sequence from a sample of shorter overlapping fragments, called reads, without any prior knowledge about the original sequence. The first major achievement of *de novo* genome assembly happened in the early 2000s when the Human Genome Project was finished, an effort that took over a decade and cost billions of dollars [1]. The reported results of the project were that 99% of the genome had been reconstructed with less than 400 gaps. Unfortunately, that was not entirely correct, as only the euchromatic portion of the genome was considered while the heterochromatin was left out. When the heterochromatin regions—which include centromeres, telomeres, and tandem gene arrays—are also taken into account, the final result is that more than 5% of the entire genome was either missing or incorrect.

Since then, the sequencing technologies have improved significantly in many ways, but most notably in terms of the lengths and accuracies of the reads they produce. At the forefront of the latest sequencing technologies are the HiFi reads developed by PacBio [2] and the ultra-long reads by Oxford Nanopore Technologies [3], both of which were crucial for the most recent breakthrough in the field of de novo genome assembly—full reconstruction of the entire human genome, with no regions left unsolved [4]. Nevertheless, this was enabled not only by the latest sequencing technologies, but also by a tremendous effort of numerous researchers and bioinformaticians who used various *de novo* assembly tools and manually inspected large genomic regions.

One of the more common approaches to *de novo* genome assembly, which was also the one used in the recent reconstruction of the human genome, is the Overlap-Layout-Consensus (OLC) paradigm. In the Overlap phase, the reads in the sample are mapped onto each other in an all-versus-all manner in order to find overlaps between them. All reads that are entirely contained in other reads are removed from further processing. From the rest of the overlapped reads, an assembly graph is built—a graph in which nodes represent reads and edges represent the suffix-prefix overlaps between the reads. In the Layout phase, the assembly graph is simplified in order to find a path through it that would reconstruct the original genome. Finally, in the Consensus phase, all the reads are aligned to the reference in order to clean the assembly sequence of errors that happened during the rest of the process.

In the ideal scenario, the Layout phase would be formulated as finding a Hamiltonian path over the assembly graph—visit every node in the graph exactly once. However, due to imperfect read qualities, sequencing artifacts, and repetitive genomic regions, finding a Hamiltonian path is usually not possible and an unknown number of nodes and edges has to be removed. Because of this, instead of finding a path through the graph directly, modern assemblers rely on heuristics to simplify the entire graph, by iteratively removing nodes and edges deemed unnecessary, such as removing transitive edges, trimming tips, and popping bubbles [5], [6]. Frequently, however, assembly graphs are highly complex, and even this is not a straightforward task, as some regions cannot be simplified by the current heuristics. One such region can be seen in Figure 1 on the left, where the blue nodes are those that should be traversed for the ideal assembly, while the orange nodes should be avoided. For lack of a better method, these complex regions are cut out of the graph, and we are left with multiple fragments of the original genome instead of the continuous sequence. To this day, the problem of fragmentation continues to plague all the existing assemblers.

The central part of reconstructing the full human genome was untangling these complex regions in the assembly graph instead of cutting them. This approach required an abundance of data, various tools, and a detailed manual inspection of certain regions and individual reads. Reproducing such efforts on new genomes seems infeasible, and thus fast and accurate de novo genome assembly remains elusive.

In this work, we propose a novel approach to de novo genome assembly, one based on deep learning and finding a path through the graph instead of relying on handcrafted heuristics to simplify the graph. We train a non-autoregressive model based on gated graph convolutional network (GatedGCN) introduced by [7] that takes an assembly graph and outputs a score for each edge. These scores can then be used to guide a search algorithm over the graph, producing a path that represents the reconstructed genome. Our idea is illustrated in Figure 1 on the right—a depiction of the graph traversal done during the training, where in each step we calculate the probabilities for traversing each neighbor of the current node. For this, we create a synthetic dataset of reads from the real human genomic data, and generate assembly graphs prior to any simplification steps in order to avoid any errors such simplifications might produce. We believe this approach could reduce the fragmentation of the reconstructed genomes, especially in the entangled regions.

Lately, graph neural networks (GNNs) have been applied to a variety of biological problems, ranging from drug design [9] and protein interactions [10], to predicting anticancer foods [11]. However, to the best of our knowledge, the only work based on geometric deep learning that addresses *de novo* genome assembly is [12], which only simulates the deterministic simplification algorithms in the Layout phase, instead of tackling the tangled regions themselves. Moreover, the graphs used in [12] were completely synthetic, not generated from the biological sequences. Here, we go a step further by starting with the complete human genome and focusing on one of the main problems in genome assembly.

The rest of the paper is composed as follows: In Section II we describe the dataset preparation, in Section III we describe the architecture of the used model, in Section IV we state the performed experiments, and in Section V we discuss the results. Finally, Section VI concludes this paper.

II. DATASET

Generating the dataset can roughly be separated into three tasks. First, we simulate the genomic data from the human genome. Second, we adapt an existing tool for *de novo* genome assembly, Raven [6], to build an assembly graph for each set of reads. Finally, we implement an algorithm that leverages the information stored during the read-simulation phase and finds a ground-truth path through an assembly graph. We use these paths as the supervision signal during the training phase.

A. Simulating reads

To simulate the reads, we start with the recently reconstructed human genome called CHM13 [4], which consists of one chromosome from each of 23 pairs of human chromosomes. In the assembly process, each chromosome would ideally be represented by a single component of the large disconnected graph of the entire human genome, but in a realistic case, several chromosomes could be connected into a single component. Starting from a simpler scenario, we isolated one chromosome and split it into shorter "mini-references", each 2 million base pairs (bp) long.

During the sequencing process of the real samples, all information about the ordering of the reads and their position on the genome is lost. This is the main reason why we are not using sequenced human data, but simulate our own by sampling the reads from the mini-references. Thus, we are able to store the positional information for each read, which allows us to construct the ground-truth path for training.

The reads are simulated by mimicking the sequencing process. The mini-reference is sampled in a manner that each base of the mini-reference is covered approximately 20 times, which represents the number of times genomes are copied and fragmented prior to sequencing (sequencing depth). The lengths of the reads are sampled from a normal distribution with a mean value of 20,000 and a standard deviation of 1,500. These values were determined empirically so that the simulated reads would resemble the real PacBio HiFi reads. In contrast to sequencers, we introduce no errors to the simulated data to facilitate. Considering that the average accuracies of the real PacBio reads are usually over 99% and that there are tools that perform error correction of reads [13], we believe that this gives a decent approximation of the real data.

B. Generating graphs

Once the reads are simulated, we can construct the graphs. For this purpose, we use an adapted version of Raven, an assembler that follows the OLC paradigm and can output the assembly graphs at different stages of the Layout phase [6]. Additionally, we required Raven to keep only the perfect overlaps between the reads. Therefore, unless two reads have a suffix and a prefix which are matching in all the bases, they will not be connected with an edge in the assembly graph.

At the end of the Overlap phase, the assembly graph is constructed. We output the generated graph at the start of the Layout phase, prior to any simplification algorithm applied, in order to avoid errors that can occur during the simplification steps. The end result of this entire process is 50 graphs, each containing around 3,500 nodes and 50,000 edges. Considering that the reads are perfect and the similarity score of each overlap is 1.0, we only rely on overlap lengths as edge features and use no node features (or rather, we specify the feature of each node to be 1). A possible approach would be to encode the

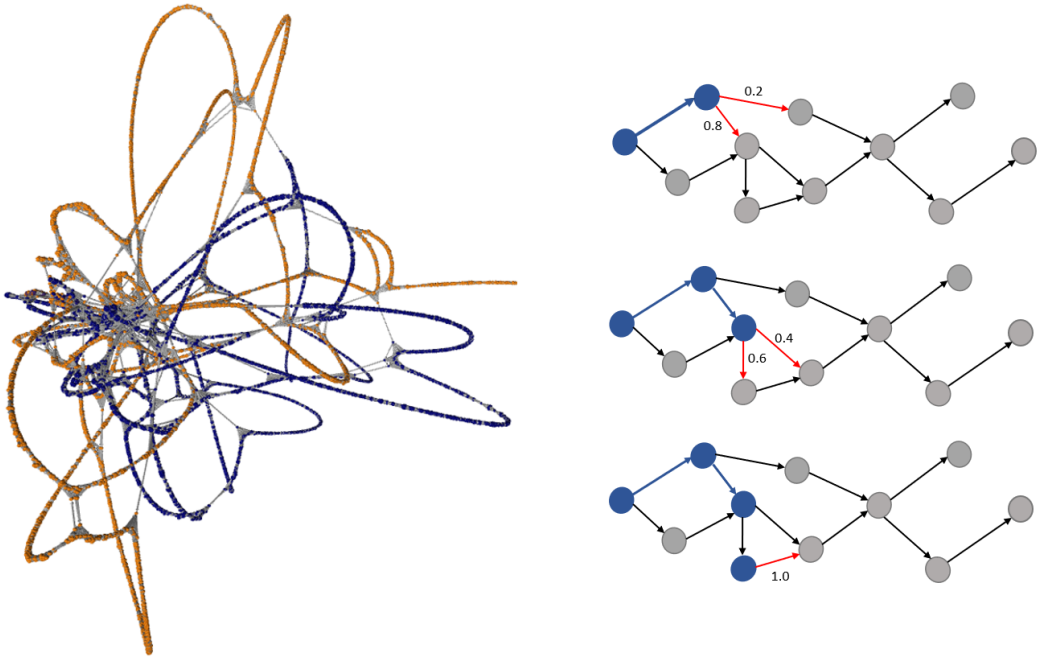


Fig. 1: **Left**: An assembly graph of one part of chromosome 11, containing the repetitive region. Blue nodes should be traversed, while the orange ones should be avoided. Figure generated with Graphia [8]. **Right**: Three steps of traversing the graph during the training where, in each step, the probabilities are computed for traversing all nodes neighboring the current node.

genomic sequences with a 1D-CNN and use the read-sequence encodings as node features and overlap-sequence encodings as additional edge features, but we leave that for future work. Therefore, in order to train on the generated graphs, the last thing needed is the supervision signal—a ground-truth path for each graph.

C. Ground-truth paths

A ground-truth path is a path in the assembly graph that produces the longest length of the reconstructed genome. For this, we utilize the positional information that was stored for each read during the sampling process and implement an algorithm resembling depth-first search, with an additional preference for successor nodes that are closer to the current node on the mini-reference. Neighbors that don't share a position on the mini-reference with the current node are avoided. Although the reads and the overlaps in the created graphs are both perfect, connections between the distant genomic regions can still exist. The main culprits for this are repetitive regions, regularly found in telomeres, centromeres, and highly duplicated ribosomal RNA genes.

In a way, this algorithm can be described as an exhaustive search with an oracle. Even though the oracle—the reads' positions on the mini-reference—guarantees that the reconstructed genome will be optimal, we noticed cases when the length of the reconstruction was less than that of the original mini-reference. There are two reasons for this. First, during the Overlap phase, all the reads are trimmed, which slightly reduces them in length. Since the ends of the mini-reference are covered only by the ends of the reads, trimming the reads necessarily leads to a loss of information. Second, and the more interesting case, happens due to an error during the Overlap phase resulting in a fragmented assembly graph. This occurs even though the sampled reads can cover the entire mini-reference (apart from a few bases on either end). We believe this happens due to a combination of repetitive regions and the discarding of reads completely contained inside the other reads. However, more analysis is needed to make a definite claim.

Notice that here we find just a single path through the graph, whereas in reality, multiple paths could lead to the optimal solution. This is particularly true in the case of diploid and polyploid genomes, but even in the case of haploid genomes, such as CHM13, there are transitive edges that do not influence the final assembly traversed. However, for the sake of an easier problem statement, we focus only on the best neighbor for each node.

III. MODEL ARCHITECTURE

The proposed model for obtaining the probability scores is non-autoregressive and can be split into three parts—encoder, processor, and decoder.

A. Encoder

A layer that transforms node features $x_i \in \mathbb{R}^{d_v}$ and edge features $z_{ij} \in \mathbb{R}^{d_e}$ into the d-dimensional node and edge representations. As stated in Section II, we use $x_i = 1.0$

and $z_{ij} = \text{len}(i \to j)$, thus making $d_v = d_e = 1$. Encoder for both node and edge features is a single linear layer:

$$h_i^0 = W_1 x_i + b_1 \in \mathbb{R}^d, \tag{1}$$

$$e_{ij}^0 = W_2 z_{ij} + b_2 \in \mathbb{R}^d, \tag{2}$$

where h_i^0 is the initial node representation of the node i, e_{ij}^0 is the initial representation of the edge $i \to j$, and $W_1 \in \mathbb{R}^{d \times d_v}, W_2 \in \mathbb{R}^{d \times d_e}$, $b_1, b_2 \in \mathbb{R}^d$ are learnable parameters.

B. Processor

The main part of the network consists of multiple GNN layers. For this task, we modify the GatedGCN [7] by including the edge representations and a dense attention map η_{ij} for the edge gates, as proposed in [14], [15].

The motivation for using this layer comes mainly from its performance on the Traveling Salesman Problem (TSP) [15]. Since the problem of finding the optimal walk on the assembly graph is similar to that of TSP, it makes sense to reuse the architecture. In addition, it was shown that the GatedGCN outperforms many other models in several tasks [16], which gives us further reason to use this architecture.

Let now node i be the node whose representation we want to update, and let its representation after a layer l be h_i^l. Also, let all the neighbors of node i be denoted with j. Then, in the layer $l+1$, the node and edge representations will be:

$$h_i^{l+1} = h_i^l + \text{ReLU}\left(\text{BN}\left(A^l h_i^l + \sum_{j \in N(i)} \eta_{ij}^{l+1} \odot B^l h_j^l\right)\right),$$

where all $A, B \in \mathbb{R}^{d \times d}$ are learnable parameters, ReLU stands for rectified linear unit, BN for batch normalization, and $\odot$ for Hadamard product. The edge gates are defined as:

$$\eta_{ij}^l = \frac{\sigma\left(e_{ij}^l\right)}{\sum_{j' \in N(i)} \sigma\left(e_{ij'}^l\right) + \epsilon} \in \mathbb{R}_+^d \tag{3}$$

where σ is the sigmoid function, and ϵ is a small value in order to avoid division by zero.

C. Decoder

A multi-layer perceptron (MLP) decodes the obtained representations into the probability scores. Probability score p_{ij} for traversing an edge $i \to j$ is computed from node representation of nodes i and j, as well as edge representations of the edge $i \to j$, all after the final GatedGCN layer L:

$$p_{ik} = \text{MLP}(h_i^L \parallel h_k^L \parallel e_{ij}^L), \tag{4}$$

where $(\cdot \parallel \cdot)$ is the concatenation operator.

IV. EXPERIMENTS

A. Training

Processing of a graph during one training epoch is done iteratively. In a single iteration, the graph is fed to the model providing us with scores for all the edges. We start the traversal from the starting node of the ground-truth path and take w steps, where w is a hyperparameter and we refer to it as the walk length. In each step, the best successor is predicted from their probability scores, while the correct next node is obtained from the ground-truth path. The cross-entropy loss is computed over the current node's successors, followed by teacher forcing where we choose the successor given by the ground-truth as the next node. This is repeated for w steps, after which the computed losses are averaged, the backpropagation is performed, and the next iteration starts, continuing from the last visited node. This is repeated until the end of the ground-truth walk. It is important to notice that, while computing the loss, the number of candidates in each step can vary, as the number of successors also varies.

The training was performed on a dataset consisting of 50 graphs with a 30/10/10 train/validation/test split. Each graph had about 3,500 nodes and 50,000 edges, all of them generated from 2 Mbp mini-references coming from chromosome 11. We used Adam optimizer [17], with the initial learning rate of 10^{-4}. We also decay the learning rate multiplying by with 0.9 in case there was no improvement in the validation loss for 5 epochs. The evaluation metric used during training to keep track of the learning process was the accuracy of predicting the best next neighbor, which we also refer to as the local reconstruction metric. The entire training was done on a single Nvidia A100 GPU.

Predicting the best successor is the only metric used during the training, since it is easy to calculate the corresponding loss and accuracy. In a way, the better the model follows the ground-truth, the better it should reconstruct the entire sequence. However, there are some situations where choosing the incorrect node makes no difference (e.g., transitive edges such as $i \to k$, when there also exist $i \to j$ and $j \to k$), as well as situations where choosing the correct node is crucial (e.g., dead-end nodes). Therefore, even though used for training, this is not the best metric for genome assembly in general. Still, we expect a certain transferability between the local and the global task, and this approach simplifies the training significantly. In addition, we also perform teacher forcing during the validation and testing—even though the model makes a mistake, we will put it back on the right path for the sake of easier evaluation of the training process.

B. Inference

At inference, the predictions are not performed iteratively, but all in one go—we feed the graph to the model and find a path with the greedy algorithm following the probability scores. The greedy algorithm runs until it

reaches a node without outgoing edges or all of the node's successors have already been visited. Since there is no good option for choosing the starting node, we run the greedy search from all the nodes that have in-degree zero and choose the longest walk. At inference, we cannot evaluate the performance by the accuracy of choosing the best next neighbor, since there is no ground-truth. Moreover, *de novo* genome assemblers usually work in a completely different way, so this would make the comparison against them impossible. Therefore, we evaluate our model on two other tasks—length of the reconstruction and the execution time:

a) Reconstructed sequence length.: We measure the reconstructed length as the percentage of the original mini-reference length from which the reads were sampled. This metric is robust to errors while choosing transitive edges, but harshly punishes choosing dead-end nodes and wrong paths in general. Ideally, the length ratio would be 1, but it is expected that it will be slightly less due to the way the reads are simulated and processed prior to the Layout phase (e.g., trimming of reads mentioned in Section II-C).

b) Execution time.: One of the main pitfalls for many assemblers is their execution time. We evaluate the time needed for the model to process the graph and the search algorithm to find the best path through it. All the experiments related to execution time were performed on a single Intel Xeon E5-2698 v4 CPU.

We benchmark our model on the mentioned tasks against three other approaches. First, the naive approach where the scores are calculated as normalized overlap lengths. This approach comes down to running a greedy algorithm over the overlap lengths, and thus we refer to it as the greedy in the next section. The second approach is the one used for obtaining the ground-truth paths as explained in Section II-C, the exhaustive search with oracle. Finally, we also compare the developed model against an existing assembler Raven, which was used to generate the assembly graphs in Section II-B.

V. RESULTS

The results reported here were obtained by the best-performing model, which had 4 GatedGCN layers, latent dimension 32, and the walk length was 10. The MLP classifier consisted of a single layer. During training this model achieved a 99.90% accuracy on the test set and upon deeper inspection, we noticed that the errors mostly consisted of traversing transitive edges.

Table I shows the evaluation results of our model (GNN) benchmarked against greedy approach, Raven, and exhaustive search with oracle (ES*). For the reported lengths, first the percentage of the reconstruction was calculated for each graph, after which all the percentages were averaged. With the aim to test how our model scales to larger graphs, and thus larger genomes, we trained it on graphs generated from 2 Mbp mini-references, and evaluate on graphs generated from 2, 5, and 10 Mbp mini-references. The number of nodes and edges scaled roughly linearly with the length of the mini-reference. For the evaluation on graphs generated from 2 Mbp mini-reference, we reuse the graphs from the test set, coming from chromosome 11. Additionally, we cut chromosome 10 into 5 Mbp mini-references and chromosome 12 into 10 Mbp mini-references. There was no particular reason for choosing the chromosomes 10, 11, and 12. All three chromosomes are of similar complexity, so the extracted mini-references should also be similar, with length being the main difference between them.

a) GNN vs ES*.: The developed GNN-based method doesn't manage to reconstruct the genomes as well as the exhaustive search, but does not fall far behind either. The difference comes from the cases where our model got stuck in a dead-end node, similarly to what happened to the greedy algorithm.

b) GNN vs Greedy.: We notice that our model consistently outperforms the greedy algorithm in terms of reconstructed length. This means that the model managed to leverage the graph topology in order to avoid some pitfalls where the greedy algorithm got stuck, such as dead-end nodes. The greedy algorithm is slightly faster than our model, which comes as no surprise. Interestingly, the difference in execution times is relatively small, mainly because the neural network we used is also relatively small, and so are the graphs, so a single forward pass through it can be efficiently done on a single CPU.

c) GNN vs Raven.: We notice that Raven slightly outperforms the GNN model on the two larger datasets, but on the 2 Mbp dataset falls behind even the greedy approach. After a more thorough analysis, we noticed that Raven underperformed on only one graph, while on all the other graphs in that dataset, it was as good as the GNN. The graph on which Raven failed came from a highly repetitive region, and since Raven tries to simplify the graph by removing nodes and edges, it ends up cutting the graph into numerous fragments, the longest one being only around 15% of the original mini-reference. GNN managed to correctly find a path through that graph. This is a critical result, as the repetitive regions are the main reason why the assemblers fail to accurately reconstruct genomes. At the same time, Raven is faster than the GNN, which is partially due to more efficient implementation of the algorithms performed. However, with differences being only up to a few seconds, this is insignificant given that the bottleneck of the *de novo* assembly is usually the Overlap phase.

Overall, we show that the model performs well consistently over graphs of different sizes and even different chromosomes, both in terms of accuracy and speed. Yet, human chromosomes are up to 250 Mbp long, so a definite conclusion on how it would perform in such a setting cannot be made.

VI. CONCLUSION

In this work, we introduce a novel approach to solving *de novo* genome assembly based on graph neural networks and finding a path through the assembly graph. We created

TABLE I: Comparison of GNN, naive approach, Raven, and exhaustive search.

Method	2 Mbp		5 Mbp		10 Mbp	
	length [%]	time [s]	length [%]	time [s]	length [%]	time [s]
ES*	99.79 ± 0.16	2.5 ± 0.1	95.82 ± 14.67	7.5 ± 1.1	95.93 ± 13.94	23.0 ± 5.0
Greedy	93.50 ± 17.88	0.2 ± 0.0	90.21 ± 24.01	0.7 ± 0.1	82.11 ± 33.14	1.4 ± 0.6
Raven	90.85 ± 26.74	0.216 ± 0.002	93.68 ± 17.25	0.54 ± 0.01	95.92 ± 13.96	1.09 ± 0.01
GNN	99.20 ± 1.86	0.5 ± 0.1	93.51 ± 18.31	1.4 ± 0.3	95.73 ± 13.90	3.0 ± 0.4

a dataset of assembly graphs based on reads sampled from human genomic data on which the developed model was trained and evaluated. The model was also evaluated against a naive greedy approach, an exhaustive search using the positional information of reads, and an existing genome assembler Raven [6]. It was shown that it consistently outperforms the greedy approach in terms of reconstructed length, and is usually on-par with Raven while having no significant computational overhead. More interestingly, it was shown that the model outperformed Raven when given a highly complex graph from a repetitive region.

These results are particularly promising to solve challenging regions more accurately than the existing assemblers can. Future work will investigate the proposed technique on sequenced instead of generated data, with the ultimate aim of using it as a tool for untangling the assembly graphs to reduce fragmentation. We believe that combining path-finding techniques with deep learning will play a major role in improving *de novo* genome assembly.

ACKNOWLEDGMENT

Lovro Vrček has been supported by "Young Researchers" Career Development Program DOK-2018-01-3373, ARAP scholarship awarded by A*STAR, and core funding of Genome Institute of Singapore, A*STAR. Xavier Bresson has been supported by NRF Fellowship NRFF2017-10 and NUS-R-252-000-B97-133. Martin Schmitz has been supported by SINGA scholarship awarded by A*STAR. Mile Šikić has been supported in part by the European Union through the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS), by the Croatian Science Foundation under the project Single genome and metagenome assembly (IP-2018-01-5886), and by the core funding of Genome Institute of Singapore, A*STAR.

REFERENCES

[1] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh *et al.*, "Initial sequencing and analysis of the human genome," 2001.

[2] A. M. Wenger, P. Peluso, W. J. Rowell, P.-C. Chang, R. J. Hall, G. T. Concepcion, J. Ebler, A. Fungtammasan, A. Kolesnikov, N. D. Olson *et al.*, "Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome," *Nature biotechnology*, vol. 37, no. 10, pp. 1155–1162, 2019.

[3] M. Jain, S. Koren, K. H. Miga, J. Quick, A. C. Rand, T. A. Sasani, J. R. Tyson, A. D. Beggs, A. T. Dilthey, I. T. Fiddes *et al.*, "Nanopore sequencing and assembly of a human genome with ultra-long reads," *Nature biotechnology*, vol. 36, no. 4, pp. 338–345, 2018.

[4] S. Nurk, S. Koren, A. Rhie, M. Rautiainen, A. V. Bzikadze, A. Mikheenko, M. R. Vollger, N. Altemose, L. Uralsky, A. Gershman, S. Aganezov, S. J. Hoyt, M. Diekhans, G. A. Logsdon, M. Alonge, S. E. Antonarakis, M. Borchers, G. G. Bouffard, S. Y. Brooks, G. V. Caldas, H. Cheng, C.-S. Chin, W. Chow, L. G. de Lima, P. C. Dishuck, R. Durbin, T. Dvorkina, I. T. Fiddes, G. Formenti, R. S. Fulton, A. Fungtammasan, E. Garrison, P. G. Grady, T. A. Graves-Lindsay, I. M. Hall, N. F. Hansen, G. A. Hartley, M. Haukness, K. Howe, M. W. Hunkapiller, C. Jain, M. Jain, E. D. Jarvis, P. Kerpedjiev, M. Kirsche, M. Kolmogorov, J. Korlach, M. Kremitzki, H. Li, V. V. Maduro, T. Marschall, A. M. McCartney, J. McDaniel, D. E. Miller, J. C. Mullikin, E. W. Myers, N. D. Olson, B. Paten, P. Peluso, P. A. Pevzner, D. Porubsky, T. Potapova, E. I. Rogaev, J. A. Rosenfeld, S. L. Salzberg, V. A. Schneider, F. J. Sedlazeck, K. Shafin, C. J. Shew, A. Shumate, Y. Sims, A. F. A. Smit, D. C. Soto, I. Sović, J. M. Storer, A. Streets, B. A. Sullivan, F. Thibaud-Nissen, J. Torrance, J. Wagner, B. P. Walenz, A. Wenger, J. M. D. Wood, C. Xiao, S. M. Yan, A. C. Young, S. Zarate, U. Surti, R. C. McCoy, M. Y. Dennis, I. A. Alexandrov, J. L. Gerton, R. J. O'Neill, W. Timp, J. M. Zook, M. C. Schatz, E. E. Eichler, K. H. Miga, and A. M. Phillippy, "The complete sequence of a human genome," *bioRxiv*, 2021. [Online]. Available: https://www.biorxiv.org/content/early/2021/05/27/2021.05.26.445798

[5] H. Li, "Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences," *Bioinformatics*, vol. 32, no. 14, pp. 2103–2110, 2016.

[6] R. Vaser and M. Šikić, "Time-and memory-efficient genome assembly with raven," *Nature Computational Science*, vol. 1, no. 5, pp. 332–336, 2021.

[7] X. Bresson and T. Laurent, "Residual gated graph convnets," *arXiv preprint arXiv:1711.07553*, 2017.

[8] T. Freeman, S. Horsewell, A. Patir, J. Harling-Lee, T. Regan, B. B. Shih, J. Prendergast, D. A. Hume, and T. Angus, "Graphia: A platform for the graph-based visualisation and analysis of complex data," *bioRxiv*, 2020.

[9] J. M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N. M. Donghia, C. R. MacNair, S. French, L. A. Carfrae, Z. Bloom-Ackermann *et al.*, "A deep learning approach to antibiotic discovery," *Cell*, vol. 180, no. 4, pp. 688–702, 2020.

[10] P. Gainza, F. Sverrisson, F. Monti, E. Rodola, D. Boscaini, M. Bronstein, and B. Correia, "Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning," *Nature Methods*, vol. 17, no. 2, pp. 184–192, 2020.

[11] G. Gonzalez, S. Gong, I. Laponogov, M. Bronstein, and K. Veselkov, "Predicting anticancer hyperfoods with graph convolutional networks," *Human Genomics*, vol. 15, no. 1, pp. 1–12, 2021.

[12] L. Vrček, P. Veličković, and M. Šikić, "A step towards neural genome assembly," *arXiv preprint arXiv:2011.05013*, 2020.

[13] H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, and H. Li, "Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm," *Nature Methods*, vol. 18, no. 2, pp. 170–175, 2021.

[14] X. Bresson and T. Laurent, "A two-step graph convolutional decoder for molecule generation," *arXiv preprint arXiv:1906.03412*, 2019.

[15] C. K. Joshi, T. Laurent, and X. Bresson, "An efficient graph convolutional network technique for the travelling salesman problem," *arXiv preprint arXiv:1906.01227*, 2019.
[16] V. P. Dwivedi, C. K. Joshi, T. Laurent, Y. Bengio, and X. Bresson, "Benchmarking graph neural networks," *arXiv preprint arXiv:2003.00982*, 2020.
[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

Fortuna Detects Novel Splicing in Drosophila scRNASeq Data

B. Borozan*, L. Borozan*, D. Ševerdija*, D. Matijević*, S. Canzar†
* Department of mathematics, University of Osijek, Osijek, Croatia
† Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, USA
bborozan@mathos.hr lborozan@mathos.hr

***Abstract*—Recent developments in single-cell RNA sequencing techniques (scRNASeq) have made large quantities of sequenced data available across numerous species and tissues. Alternative splicing (AS) of pre-mRNA introns varies between tissues and even between cell-types and can be altered in disease. The study of novel AS, using standard RNASeq data, has been extensively studied for many years, while similar work on scRNASeq data has been scarce, despite its potential to offer a broader insight into cell-type specific processes. In this paper, we propose a novel pipeline that uses fortuna, a method that efficiently classifies and quantifies novel AS events, to process scRNASeq samples. Due to its short lifespan, high number of progeny, low maintenance cost, and intricate alternative splicing patterns similar in complexity to those of mammals, *Drosophila Melanogaster* (fruit fly) is a species of particular interest to researchers. Therefore, we experimentally evaluate our pipeline on real-world *Drosophila* single-cell data samples from the Fly Cell Atlas.**

***Keywords*—novel splicing events, scRNASeq, fortuna, drosophila, alignment**

I. Introduction

Multiple different mRNA molecules (transcripts) can be transcribed from the same genomic region [1] [2]. These overlapping transcripts, which contain codes for protein synthesis, can be the result of the mechanisms that regulate alternative splicing (AS) [3]. Known AS events are recorded in the transcriptome annotation, while novel and aberrant AS events can occur in disease [4] and are often unannotated. Deviations from the regular splicing process may have a significant impact on the organism, as discussed in [5]. Therefore, analysis of AS events is an important topic in the field of computational molecular biology. With the advent of single-cell RNA sequencing (scRNASeq) [6], an opportunity to study AS events from a different perspective has presented itself to researchers. As it is discussed in [7], the study of AS in the context of scRNASeq is challenging and the related research has been scarce, despite its clear benefits. The authors of [8] have created a comprehensive reference atlas comprising of nearly 500,000 cells from 24 different tissues and organs and have dedicated a part of their research to AS events. A similar reference atlas has been created for the fruit fly (*Drosophila Melanogaster*) [9], called the Fly Cell Atlas. It contains approximately 580,000 cells from 15 different tissues, but no analysis of AS events has been conducted. In data obtained by the means of traditional RNA sequencing (RNASeq), novel AS events can be indentified and quantified [10] after conducting the computationally challenging alignment process. Pseudoalignment methods such as [11] and [12], significantly outperform traditional alignment methods such as [13], [14], and [15] in terms of running time. Though, they are usually limited to pseudoaligning short reads to annotated transcripts and are unable to detect novel AS events. Only a single tool exists, to our best knowledge, that combines the speed of a pseudoaligner with an ability to detect novel splicing [16]. In the scRNASeq case, additional computation is required to account for differentiation between cells, correcting for in-vitro [17] or PCR [18] amplification and related sequencing errors.

In this paper we propose a novel pipeline which can be used to catalog novel AS events in scRNASeq data. It uses fortuna [16], a tool originally developed for novel AS event quantification in bulk RNASeq data, in conjunction with 10x Genomics Cellranger [19] to offset for the specifics of the scRNASeq analysis. Additional scripts are used to extract the information from fortuna and Cellranger output files and adjust the counts of novel AS events in line with corrections for scRNASeq. We test our pipeline on 12 real-world scRNASeq 10x samples from Fly Cell Atlas, obtained from ArrayExpress with the accession number E-MTAB-10519. The aim of our experiments is to produce a catalog of novel AS events which could be expanded to include the entire dataset or to process entirely different data. Apart from the catalog, we provide a running time and memory usage analysis. The code for the additional scripts can be found at fortuna fly atlas github page[1], while fortuna can be obtained at its official github page[2].

The structure of the paper is as follows. In the Methods section we will formulate the problem of building a novel AS catalog from scRNASeq data and present a novel pipeline which solves it. Then, in the Experiments section, we will present a catalog built from the real world data, analyze the novel splicing events found within and discuss the running time and memory usage of the tools that comprise our pipeline. Finally, in the Conclusion

[1] https://github.com/canzarlab/fortuna_flyatlas
[2] https://github.com/canzarlab/fortuna

section we will provide a discussion and a brief overview of our work.

II. Method

A. Problem formulation

There are a couple of challenges that need to be overcome in order to build a catalog of novel AS events from scRNASeq data. Reads coming from different cells are identified by their respective cell barcodes and each read is assigned a unique molecule identifier (UMI). During the sequencing process, both barcodes and UMIs may be incorrectly sequenced and, thus, have to be corrected. Furthermore, due to amplification, UMIs with the same corrected sequence have to be collapsed into a single UMI before further analysis can be conducted. Another issue that arises is the alignment of the reads to the transcriptome - a process in which we find the origin the read was likely sampled from. Each read that supports a novel AS event, i.e. splices over an unannotated intron, has to be recorded in the catalog. Finally, the catalog must only contain a list of novel AS events supported by the reads with corrected barcodes and UMIs.

B. Pipeline

Our proposed pipeline for classification and quantification of novel events in scRNASeq is comprised of two major steps:

1) sample processing, and
2) (pseudo-)alignment processing.

During the first step, we use fortuna to identify and quantify novel alternative splicing events, and Cellranger to perform barcode correction and UMI collapsing. During the second step, we adjust the counts of the events identified by fortuna according to Cellranger's corrections and output them into the final catalog. In the next two subsections we will explain each step in more detail.

C. Sample processing

The workflow of fortuna can be separated into three steps: indexing, pseudoalignment and postprocessing, out of which we will use only the first two. During the indexing step, fortuna supplements the annotation with a virtual transcriptome that enables the detection of novel AS events. An index constructed that way is used to pseudoalign an arbitrary number of samples and, at the same time, detect any potential novel splicing. We will use fortuna's pseudoalignments and novel AS catalog outputs (obtained by specifying *-bam* and *-alt* when running fortuna).

Cellranger is a software comprised of five pipelines intended for different steps of the single-cell analysis. We will use its *cellranger count* pipeline to correct cell barcodes and collapse UMIs. It stores known barcodes in a pre-defined whitelist and includes into the analysis all sampled barcodes that are a part of it. Each sampled barcode that has not been whitelisted, is included in the analysis if it is 1-Hamming-distance away from a whitlisted barcode, and additionally, if its probability of having originated from the whitelisted barcode is sufficiently high. Reads that contribute to UMI counts, i.e. are selected among other reads with the same UMI, are considered valid if their corresponding barcodes are valid and if they have sufficient mapping quality. Valid non-whitelisted barcodes have their UMIs merged with their respective whitelisted barcode. Cellranger outputs an alignment file with flags identifying the corrected values of UMIs and barcodes that we will use in the next step of our pipeline.

D. (Pseudo-)alignment processing

Cellranger's alignments are written in a binary file which can be converted to a plain text file using *samtools view* command [20]. On the other hand, similar binary file produced by fortuna has to be converted using *pseudo-to-genome-alignment* (p2g) tool provided as a part of our pipeline. This is a necessary step because pseudoalignments provided by fortuna are expressed in local fragment coordinates and have to be converted into standard genomic coordinates. The final part of our pipeline, *barcodeAS* tool, adjusts fortuna's catalog of novel AS events according to the information obtained from both pseudoalignment and alignment files by removing the contribution of the discarded reads (4th bit of the flag *xf:i:* in the alignment file).

III. Experiments

The experiments were conducted on dual 2.30 GHz Intel® Xeon® E5-2697 v4 processors, 320GB 2.40GHz DDR4 memory operating on Scientific Linux 7.5 (Nitrogen). C++ compiler used to compile fortuna and other scripts was GCC 4.8.5 20150623. GNU time command was used to record the running times and memory usage.

In the next subsection we will describe the dataset we processed and highlight the importance of correcting barcodes and UMIs. After that, we will present a novel AS catalog and will focus on the classification of novel AS events and their coverage over chromosomes in the dataset. Finally, we will discuss running times and memory usage of the tools we have used.

A. Experimental data

We obtained 12 Fly Cell Atlas testis samples sequenced using 10x Genomics technologies from ArrayExpress. Genome and annotation files were downloaded from FlyBase FTP, release number FB2019_06. The annotation was preprocessed for usage by Cellranger in accordance to [9], and for usage by fortuna using *processGTF* script

available at the fortuna github repository.

Each raw sample has between 125.6 and 183.1 million reads spread across between 6 and 7.7 million barcodes. The number of unique barcode and UMI combinations in the raw data ranges between 93.3 and 123.9 million. Cellranger *count* was run using default settings, 16 cores (*–localcores*), and 64GB memory (*–localmemory*). After the corrections made by Cellranger, the number of barcodes dropped to between 1.41 and 1.82 million across all samples, while the number of unique barcode and UMI combinations ranged between 61.47 and 81.62 million. The latter number coincided with the number of reads which we considered in our analysis.

The distribution of UMIs across barcodes in raw and corrected sample *S44* is depicted in Fig. 1. The number of barcodes with less than 120 UMIs steadily drops after we run Cellranger due to it discarding a lot of invalid ones. On the other hand, on the interval between 120 and 300 UMIs per barcode, the total number of barcodes increases. While correcting barcodes, Cellranger reassigns their UMIs to others, thus increasing the number of UMIs per barcode. But, the number of discarded raw barcodes is significantly larger than the number of those that have been merged, as can be seen in Fig. 1. There is little difference between raw and corrected barcodes with more than 300 UMIs, so we have omitted them from the figure. Similar conclusions can be made for all other samples on different intervals.

B. Novel alternative splicing catalog

We constructed a T^{as} fortuna index (*–index*) limiting the number of skipped exons (*-exs*) to 8 and the number of transcript fragments per gene (*-Mc*) to 25000. The rest of the settings were set to default. Alignment and novel AS event analysis were done using fortuna *–quant* using 4 processing threads and default settings. (Pseudo-)alignment files were processed using *samtools* and *p2g*, whose outputs were passed to *barcodeAS* to obtain the catalog.

Across all 12 samples, 23078 novel introns supported by a total of 2318688 reads were identified. We observed 5 different kinds of novel AS events induced by these introns, namely intron-in-exon (IE), alternative donors (AD), alternative acceptors (AA), alternative donor-acceptor pairs (AP) and exon skipping (ES). We have included them in our analysis only if their respective number of supporting reads was at least 2. The definitions of these novel AS events are consistent with those in [16]. The number of novel AS events and their read support is summarized in Table I. Note that most events are detected in multiple samples so the sum over the number of events in Table I does not yield the total number of unique events. Furthermore, some novel introns can be classified in multiple ways, and thus increase the count in multiple columns. Approximately 93.47% of all novel AS events we have detected have been exon skipping events, supported by about 81.26% of the reads. Alternative donors amounted to 4.68% and 6.28%, alternative acceptors to 3.93% and 6.01%, intron-in-exon events to 2.63% and 5.75%, and alternative donor-acceptor pairs to 1.07% and 0.69% of the total events and reads, respectively. Fig. 2 depicts the total amount of novel splicing events and their supporting reads across combined samples.

We analyzed events supported by at least 2 reads, and cells in which we have detected at least one novel AS event, for all combined samples. In 90.5% of the cells we

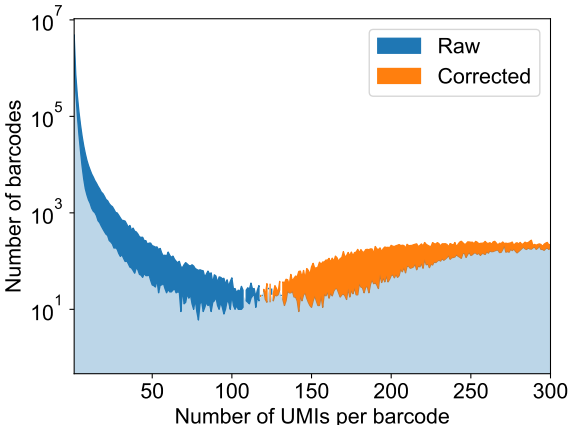


Fig. 1: Overlapped log-scaled histograms representing the number of barcodes having between 1 and 300 unique UMIs in raw (blue) and corrected (orange) sample *S44*. Overlap between bars has been colored light blue.

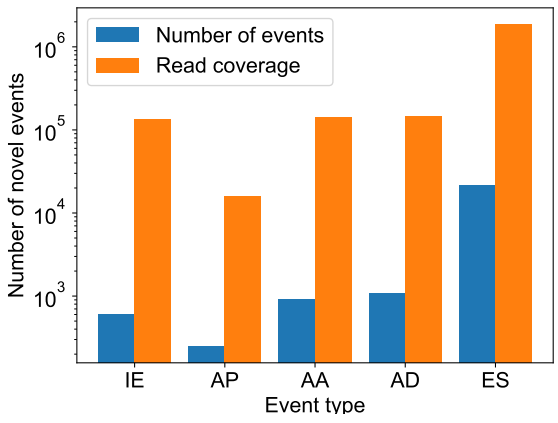


Fig. 2: Log-scaled amount of detected novel AS events separated by their type (intron-in-exon, alternative donor-acceptor pairs, alternative acceptors, alternative donors, exon skipping) and their supporting reads from twelve combined *Drosophila* samples.

Sample #	Number of events					Read coverage				
	IE	AP	AA	AD	ES	IE	AP	AA	AD	ES
S44	350	116	485	583	9577	10319	1333	11641	11935	151275
S48	391	121	509	628	10694	12861	1615	14192	14642	183180
S49	370	109	481	584	9934	11052	1382	10955	12541	161098
S56	337	98	454	573	9377	9665	1177	9655	10277	143079
S59	366	109	463	581	9739	10650	1269	11718	11617	150414
S60	389	111	478	590	9911	11825	1330	11353	12434	159158
S61	370	117	472	607	9820	10692	1272	11836	11970	152819
S65	364	99	473	564	9419	10861	1175	10465	11506	147321
S67	353	104	464	574	9514	10002	1214	10067	11002	143471
S68	357	112	489	583	9773	11039	1255	12209	12313	154495
S69	395	116	520	657	10707	13659	1586	14637	14817	183609
S70	365	114	479	587	9792	10807	1345	10703	10602	154297

TABLE I: Number of detected novel AS events separated by their type and their supporting reads in each of the twelve *Drosophila* samples.

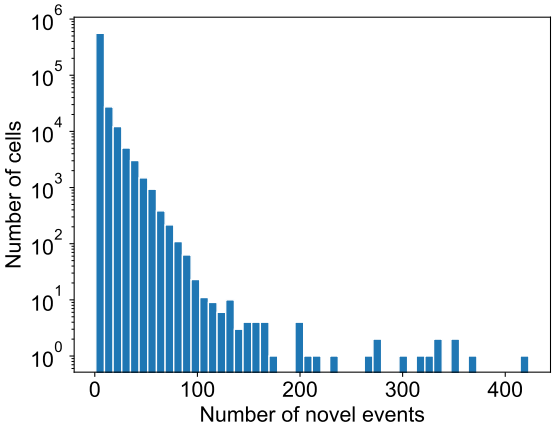


Fig. 3: A histogram of the log-scaled number of cells in relation to the number of novel AS events detected in them for the twelve combined *Drosophila* samples.

found at most 8 different novel AS events, and in 54.8% no more than 1 novel AS event. The highest number of events detected in a single cell was 423. We present the distribution of the number of novel AS events across the cells in Fig. 3.

In Fig. 4 we present the combined coverage of novel AS events (blue bars) across different chromosomes. X-axis represents genomic coordinates, while the left y-axis depicts the number of detected novel AS events. Additionally, the red line signifies the number of different cells in which we have detected the aforementioned novel AS events. Its corresponding values are located on the right y-axis. As can be seen in Fig. 4, there exist a few regions on each chromosome where certain novel AS events have been observed in a large amount of different cells, e.g. red line reaching a value over 60000 on chromosome 2R. This could indicate the existence of AS events that are common across multiple cells or tissues. Such cases could be investigated in the data comprised of multiple samples of different tissues and any suitable AS events could be considered for patching into the annotation. On the other hand, high coverage of novel AS events present in a low number of cells, visualized by tall blue bars and low red values, might point to processes affecting only specific cells that are worth investigating. Such examples can be visually detected in Fig. 4 on, e.g. chromosome 4.

We have investigated the outlier on chromosome 2R and found that a single novel AS event at genomic coordinates *2R:17478508-17485002* is supported by 135285 reads and is detected in 64511 different cells across all samples. It might suggest that the particular intron is a part of a common unannotated transcript. Investigation of this as well as other similar outliers through the genome might prove an interesting topic for further research.

C. Running time and memory usage

The experiments section will be concluded with a running time and memory usage analysis. There are five essential components of our pipeline - Cellranger, fortuna, samtools, *p2g*, and *barcodeAS*. Their memory usages and running times are presented in Table II. Cellranger and fortuna ran, on average, for 48m50s and 26m43s, respectively, while consuming 3.32 and 6.90 GB of memory. Increasing the amount of available threads would decrease the running time for both tools with some diminishing returns, likely due to disk-related limitations. Average samtools and *p2g* running times were 11m28s and 15m4s, while their memory consumtion was 1.29 and 310.39 MB. Finally *barcodeAS* ran for 24m56s while using 14.02 GB. That would mean that an average run of the entire pipeline would last slightly less than 130m and consume 14.02 GB of memory. Notice that *barcodeAS* consumed the most memory out of all listed tools. Reducing its memory usage could be an important part of our future work.

IV. Conclusion

We have presented a novel pipeline for building a catalog of novel alternative splicing events from scRNASeq data which uses fortuna and Cellranger in its core. Then, we have tested the pipeline using 12 real-world *Drosophila Melanogaster* testis samples from

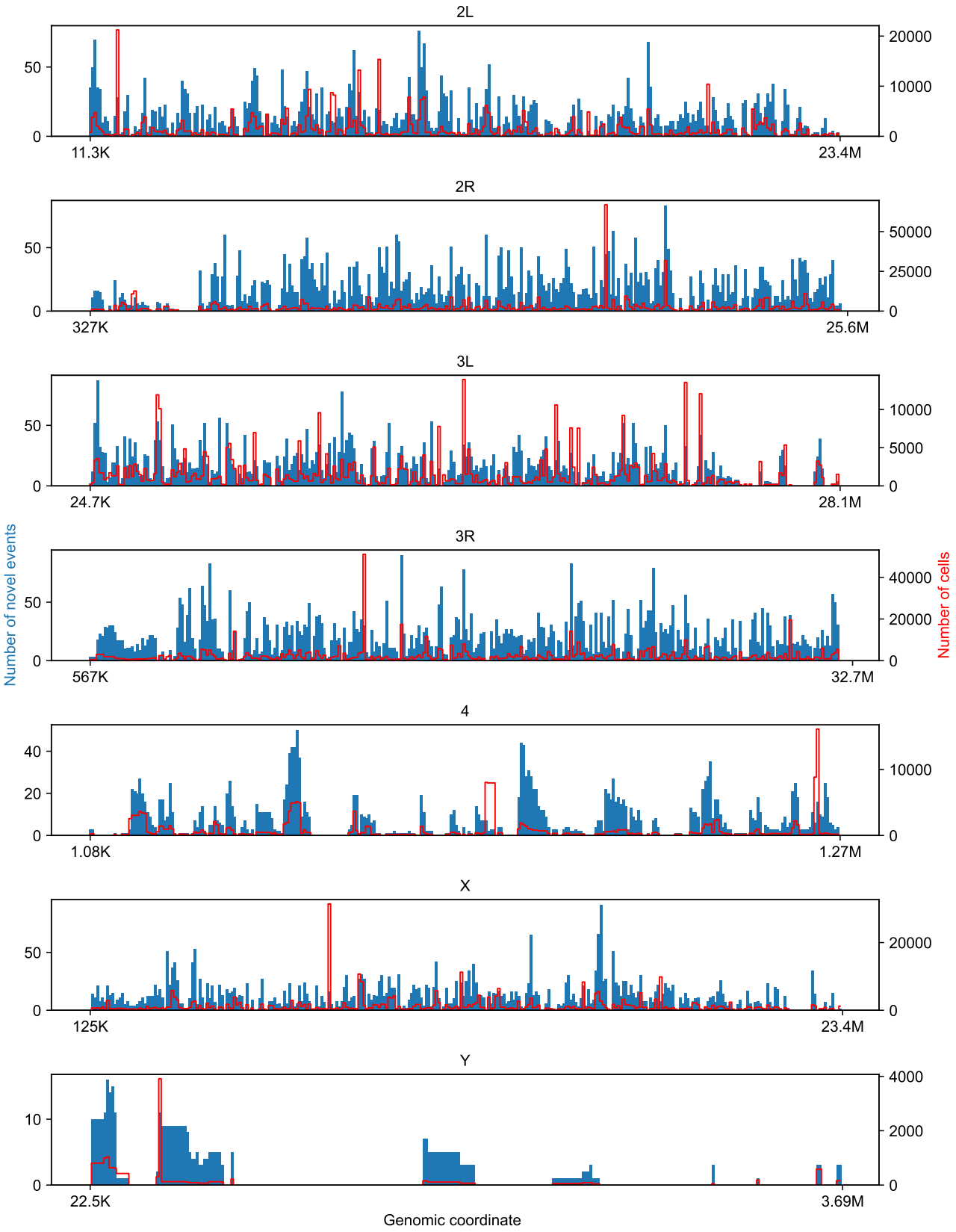


Fig. 4: The coverage of novel AS events across the chromosomes (blue) and the number of cells in which these events have been detected (red) for the twelve combined *Drosophila* samples.

Sample	Running time					Memory consumption (GB)			
#	fortuna	Cellranger	p2g	samtools	barcodeAS	fortuna	Cellranger	p2g	barcodeAS
S44	26m39s	50m54s	15m05s	8m47s	24m32s	6.88	3.26	0.30	13.85
S48	34m42s	60m08s	19m41s	12m04s	35m41s	6.90	3.50	0.30	16.79
S49	29m19s	55m39s	17m20s	10m03s	28m30s	6.90	3.32	0.30	14.76
S56	23m52s	44m18s	14m12s	11m46s	25m27s	6.88	3.17	0.30	13.03
S59	25m00s	44m33s	13m36s	14m24s	26m18s	6.90	3.29	0.30	13.37
S60	25m35s	47m25s	14m11s	9m20s	25m08s	6.90	3.36	0.30	13.98
S61	24m33s	49m28s	14m20s	15m10s	25m01s	6.90	3.34	0.30	13.38
S65	23m48s	42m08s	13m16s	8m26s	29m52s	6.90	3.16	0.30	12.94
S67	23m14s	41m25s	13m02s	13m47s	22m10s	6.90	3.27	0.30	12.65
S68	26m26s	45m20s	14m01s	9m26s	26m02s	6.90	3.20	0.30	13.57
S69	32m23s	56m01s	17m56s	11m47s	32m29s	6.91	3.50	0.30	16.16
S70	25m00s	48m40s	14m09s	12m37s	30m39s	6.90	3.40	0.30	13.71

TABLE II: Running times and memory usage of all tools in our pipeline recorded for each *Drosophila* sample. Samtools have been omitted from the memory consumption due to using less than 2 MB of memory.

Fly Cell Atlas. In our experiments, we have analyzed and discussed potential benefits of computing the coverage of novel AS events across the chromosomes, split novel AS events and their supporting reads by the type of AS, and briefly commented the running times and memory usage of the each part of our pipeline.

We acknowledge several new directions we could take in our future work. Working directly with binary alignment files without converting them into plain text, thus omitting samtools, could significantly speed up the process of building a catalog. Expanding the experiments to encompass the entire Fly Cell Atlas or even scRNASeq datasets of other species could result in potentially interesting and important biological findings. Finally, the inclusion of novel intron retention events could prove useful for downstream analysis.

REFERENCES

[1] E.T. Wang et al., "Alternative isoform regulation in human tissue transcriptomes," *Nature*, vol. 456, no. 7221, pp. 470–476, 2008.

[2] Q. Pan, O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe, "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing," *Nature Genetics*, vol. 40, no. 12, pp. 1413–1415, 2008.

[3] D. Brett, H. Pospisil, J. Valcárcel, J. Reich, and P. Bork, "Mechanisms of alternative pre-messenger rna splicing," *Nature Genetics*, vol. 30, no. 1, pp. 29–30, 2002.

[4] K. Jaganathan et al., "Predicting splicing from primary sequence with deep learning," *Cell*, vol. 176, no. 3, pp. 535–548, 2019.

[5] O. Kelemen et al., "Function of alternative splicing," *Gene*, vol. 514, no. 1, pp. 1–30, 2013.

[6] T. K.Olsen and N. Baryawno, "Introduction to single-cell rna sequencing," *Curr Protoc Mol Biol*, vol. 122, no. 1, 2018.

[7] A. Arzalluz-Luque and A. Conesa, "Single-cell rnaseq for the study of isoforms—how is that possible?" *Genome Biol*, vol. 19, no. 110, 2018.

[8] The Tabula Sapiens Consortium and S. R. Quake, "The tabula sapiens: a multiple organ single cell transcriptomic atlas of humans," *bioRxiv*, 2022. [Online]. Available: https://www.biorxiv.org/content/early/2022/03/04/2021.07.19.452956

[9] H. Li and J. Janssens et al., "Fly cell atlas: A single-nucleus transcriptomic atlas of the adult fruit fly." *Science*, vol. 375, no. 6584, 2022.

[10] A. Kahles, C. S. Ong, Y. Zhong, and G. Ratsch, "Spladder: identification, quantification and testing of alternative splicing events from rna-seq data," *Bioinformatics*, vol. 32, no. 12, pp. 1840–1847, 2016.

[11] N. L. Bray, H. Pimentel, P. Melsted, and L. Pachter, "Near-optimal probabilistic rna-seq quantification," *Nature Biotechnology*, vol. 34, pp. 525–527, 2016.

[12] R. Patro, G. Duggal, and M. Love et al., "Salmon provides fast and bias-aware quantification of transcript expression," *Nature Methods*, vol. 14, pp. 417–419, 2017.

[13] A. Dobin et al., "Star: ultrafast universal rna-seq aligner," *Bioinformatics*, vol. 29, no. 1, pp. 15–21, 2013.

[14] D. Kim, J. M. Paggi, and C. Park et al., "Graph-based genome alignment and genotyping with hisat2 and hisat-genotype," *Nature Biotechnology*, vol. 37, pp. 907–915, 2019.

[15] D. Kim, G. Pertea, C. Trapnell, R. K. H. Pimentel, and S. L. Salzberg, "Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions," *Genome Biology*, vol. 14, no. 4, p. R36, 2013.

[16] L. Borozan, F. Rojas Ringeling, S. Kao, E. Nikonova, P. Monteagudo-Mesas, D. Matijević, M. L. Spletter, and S. Canzar, "Counting pseudoalignments to novel splicing events," *article in revision*.

[17] J. Eberwine et al., "Analysis of gene expression in single live neurons," *Proc. Natl. Acad. Sci. USA*, vol. 89, pp. 3010–3014, 1992.

[18] G. Brady, M. Barbara, and N. N. Iscove, "Representative in vitro cdna amplification from individual hemopoietic cells and colonies," *Methods Mol. Cell Biol.*, vol. 2, pp. 17–25, 1990.

[19] G. X. Y. Zheng et al., "Massively parallel digital transcriptional profiling of single cells," *Nature Communications*, vol. 8, pp. 1–12, 2017.

[20] H. Li et al., "The sequence alignment/map (sam) format and samtools," *Bioinformatics*, vol. 25, pp. 2078–2079, 2009.

Using De Novo Metagenome Assembly for Improved Metagenomic Classification

Josipa Lipovac*, Krešimir Križanović*
* Faculty of Electrical Engineering and Computing, University of Zagreb
Unska 3, Zagreb, Croatia
josipa.lipovac@fer.hr
kresimir.krizanovic@fer.hr

***Abstract*—Metagenomics is a rapidly growing field that allows for studying complex microbial communities. One of the first steps in the metagenomic analysis is the classification of the organisms present in a sample. This is usually done by comparing sequencing reads to a database of known organisms. With the recent development of long-read sequencing technologies, such as PacBio and Oxford Nanopore Technologies (ONT), it is now possible to generate highly accurate assemblies of genomes from metagenomic samples. This is typically done using a combination of reference-based and de novo assembly approaches. Assembling the genomes from the metagenomic sample, prior to classification, could improve classification results and also aid in identifying new, previously unknown species. However, the evaluation of metagenome assemblies is a challenging task and it is important to assess the quality of the assemblies in order to ensure the accuracy of downstream analyses. In this paper, we provide a detailed overview of metagenomic classification, de novo metagenome assembly process, and evaluation of metagenome assembly, highlighting various tools and techniques currently available for each step. We also present initial results showing that metagenomic classification can benefit from a previously assembled metagenome.**

***Keywords*—metagenomics, metagenomic classification, metagenome assembly, metagenome assembly evaluation**

I. INTRODUCTION

The term microbiota refers to the microbial population present within the human body, including bacteria, viruses, archaea, protozoans, and fungi [1]. Over the course of genetic evolution, humans have formed symbiotic relationships with various microbes. To understand these relationships and the role of microbes in human health and disease, it is necessary to identify these microbes and determine their genome sequences [2].

Metagenomic sequencing is an alternative approach that allows for direct sequencing of a mixture of microbial DNA without the need for isolation, providing a more comprehensive understanding of microbial genomes with diverse characteristics. The field of metagenomics has seen significant growth in recent years due to increased interest in microbial communities and the development of techniques for analyzing their diversity and genetic potential [3].

Partial sequencing of microbiota DNA can provide information on the diversity of a given community, but more comprehensive insights into the genetic potential of the microbiome require the analysis of extended genomic regions or fully reconstructed genomes, which can be obtained through the use of metagenome assemblers [4]. Metagenomics involves sequencing and analysis of the genomic DNA of entire microbial communities in environmental samples, allowing researchers to better understand the makeup of these communities.

PacBio [5] and Oxford Nanopore Technologies (ONT) [6] are DNA sequencing methodologies that generate long reads with average lengths of 10-25 kilobase pairs (kbp) for Pacbio and 10-100 kbp for ONT. Recent advancements in ONT have enabled the production of ultra-long reads exceeding 1 megabase pair (Mbp). In metagenomic sequencing, long reads can be particularly useful as they provide more continuous DNA sequences and facilitate the identification of structural variations within a genome.

De novo metagenome assembly is the process of reconstructing genomic sequences of microorganisms from DNA sequencing reads without the use of reference genomes. This process involves the use of specialized software tools to assemble the reads into longer contiguous sequences, known as contigs. The complexity and computational intensity of this process may be heightened due to the potential for a high number of reads originating from diverse microorganisms, as well as the potential for significant differences between the reads.

One crucial aspect of the metagenomic analysis is metagenomic classification, which typically involves comparing the sequencing reads to databases of known genes and genomes to identify the microorganisms present in the sample. Once the microorganisms present in the sample have been identified, the genes present in the reads can be annotated by assigning functions to the genes based on their known or predicted roles. However, this approach may be problematic when the database is too large or when there is a new, unknown species in the sample. To address these issues, first, the metagenome can be assembled, and then metagenomic classification can be performed using the obtained contigs.

In order for contigs generated through metagenome assembly to be utilized for metagenomic classification, it is essential to accurately evaluate them. Many existing evaluation tools rely on the availability of a reference genome, which can complicate the evaluation process if the sample contains previously unknown species.

II. METAGENOMIC CLASSIFICATION

Metagenomic classification is the process of accurately identifying all species present in a sample and is a challenge because metagenomic sequencing generates genomic data from a mixture of species. During metagenomic classification, reads are assigned to taxonomic groups using various methods. This involves comparing the reads to databases of known reference genomes, which can be used to infer the taxonomic identity of the original organisms. Several classifiers have been developed to help with this process, but choosing the best method requires understanding the characteristics and limitations of each.

The first method for assigning taxonomic labels to unknown reads was using BLAST (basic local alignment and search tool) [7], a high-sensitivity DNA alignment tool. But, due to its high computational demands, it's infeasible for large metagenomic sequencing data.

Metagenomic classifiers can be roughly divided into two groups [8]:

1) **k-mer-based**: Kraken2 [9], Centrifuge [10], CLARK [11], CLARK-S [12], etc.
2) **mapping-based**: MetaMaps [13], MEGAN-LR [14], Minimap2 [15], MetaPhlAn3 [16], etc.

K-mer-based metagenomic classifiers use the presence and frequency of k-mers to classify DNA sequences to species or higher taxonomic groups. They compare the k-mer count of a query sequence to a reference database of labeled sequences and assign the query to the group with the highest similarity score. These classifiers are fast but often have lower accuracy (reporting many false positives) and can be affected by k-mer length and reference database quality. Kraken2 is a k-mer-based metagenomic classifier tool with fast classification and reduced memory usage. Its reference database is built from the NCBI RefSeq database [17] and offers high accuracy, especially at the genus level. However, it may be difficult to run on lower-end computers due to its high memory requirement. MiniKraken [18] is a smaller, less accurate version that requires less memory. Kraken2 also includes features like a spaced seed search scheme and compatibility with the Bracken algorithm [19] for estimating species-level sequence abundance. The Centrifuge uses Burrows-Wheeler Transform (BWT) and Ferragina-Manzini (FM) index for sequence storage and mapping. The transformed string from BWT enables efficient compression and string matching. FM index combines BWT and suffix array for efficient searching. CLARK and CLARK-S classify metagenomic reads by comparing them to a reduced k-mer database and classifying them as the target with the highest shared k-mers. CLARK-S has specific databases for bacteria, viruses, and fungi.

Mapping-based metagenomic classifiers use a sequence alignment algorithm to align the reads generated from metagenomic sequencing to a reference database. The basic idea is to identify the reference sequences that the reads in the query sample map to with the highest degree of similarity and use this information to classify the reads to the species or taxonomic group they belong to. These classifiers have the advantage of being able to classify reads with high accuracy, but they can be computationally intensive and may not be suitable for very large datasets. Minimap2 is a sequence alignment tool that aligns reads to a reference genome. In order for Minimap2 to be used as a metagenomic classifier, its output needs to be reformatted and specially interpreted, as the authors did in a recent study [8] where it achieved fairly successful results. When it encounters a sequence that is not in its database, Minimap2 adds it to the most similar organism, which helps increase its classification accuracy on the genus level. However, it is slower than k-mer-based tools. MetaMaps employs probabilistic scoring to approximate sample composition, while MEGAN-LR, built on MEGAN6, interprets long nucleotide sequences using a translation alignment method. MEGAN-LR assigns reads to taxa using the LCA (Lowest Common Ancestor) algorithm, along with other factors such as lcaCoveragePercent, minSupportPercent, and minPercentReadCover. In contrast, MetaPhlAn3 calculates the relative abundance of taxa by mapping reads to a database of unique, clade-specific marker genes and using coverage scores.

Metagenomic classifiers rely on a pre-computed database. A large and rapidly increasing size of such databases can make metagenomic analysis computationally demanding [20]. The most widely used reference databases are RefSeq complete genomes for microbial species, and the BLAST database [21]. Other commonly used databases include SILVA [22], which contains 16S ribosomal RNA sequences, and Genbank [23], which contains a larger number of genomes with lower quality control standards. Many metagenomics tools allow users to create their own reference database based on a specific set of sequences. While this process can be computationally demanding, especially for large databases, it gives users greater control over the analysis, particularly when studying rare, newly discovered, or highly diverse species. Using a uniform reference database is important when comparing results from different classifiers to prevent potential confounding effects from differences in default databases.

Benchmark papers show that most classification tools that obtain a low false positive rate tend to have a lower recall. Due to the inherent variability in the specific needs of metagenomic analyses, it is challenging to establish a universally superior classification tool. Longer reads tend to improve accuracy, but using only the longest reads can reduce accuracy due to varying read length distributions. Furthermore, the presence of novel species within a metagenomic sample has been found to have a significant impact on classification accuracy, yet this aspect has not been extensively investigated in previous benchmark studies [8] [24]. Currently, most metagenomic tools classify unknown species into a genus or report them as unclassified. A better solution would be having tools

that not only detect unknown species but also provide information on their closest related species. This can be achieved through metagenome assembly and contig classification. A large contig classified under a genus could indicate a previously unknown species in the sample.

III. Metagenome assembly

De novo genome assembly is the process of reconstructing a full genomic sequence from DNA reads without the use of a reference genome. It is more challenging than reference-based genome assembly, which uses a known genome as a starting point. De novo genome assembly is necessary when studying organisms for which no reference genome is available, and it is accomplished using computational algorithms that align the reads and reconstruct the full genomic sequence. Metagenome assembly is the process of reconstructing sequences from a metagenomic sample to obtain a representation of the genomes present in the sample. The full genetic potential of a microbial community can be determined by analyzing extended genomic regions or complete genomes, obtainable through modern sequencing technologies. However, genome assembly of metagenomic data is challenging due to factors such as large data volume, uneven representation of community members, and the presence of multiple strains and closely related microorganisms.

Tools for metagenome assembly can be divided into three groups:

1) **short-reads assemblers**: Omega [25], MetaVelvet [26], IDBA-UD [27], MEGAHIT [28], RayMeta [29], etc.
2) **long-reads assemblers**: Canu [30], metaFlye [31], HifiAsm-meta [32], etc.
3) **hybrid assemblers** metaSPAdes [33], DBG2OLC [34], OPERA-MS [35], etc.

Short-read metagenome assemblers typically have more difficulty reconstructing genomic sequences of the individual microorganisms present in a metagenomic sample compared to long-read metagenome assemblers. Short reads are typically 100-150 bp long and are very accurate (error rate is around 0.5%). On the other hand, long reads are several thousand bp long, so even though they are more error-prone, they provide more information about the genomic sequence, making it easier for the assembler to distinguish between the different genomes and correctly reconstruct the individual sequences and handle repetitive regions of the genome. Assembling a metagenome can be difficult, especially with short reads, but even an incomplete and fragmented assembly can help improve metagenomic classification, particularly in identifying new species. Omega uses hash tables to store the prefix and suffix sequences of each read and then uses these sequences to construct a bi-directed graph by linking the reads with their overlapping sequences. MetaVelvet creates a de Bruijn graph using Velvet and then divides it into subgraphs using coverage peaks of k-mers to separate different microbial genomes. IDBA-UD tries to trim the graph iteratively and merge bubbles using increasing k-mer sizes. MEGAHIT combines the process of selecting k-mer sizes with a succinct de Bruijn graph and has strong computational efficiency, while RayMeta generates the local depth distribution for each seed path in a de Bruijn graph.

Hybrid metagenome assembling merges the benefits of long reads (contiguous sequences) and short reads (error correction) to achieve more accurate and complete assemblies. A popular metagenome assembler is metaSPAdes, which combines various assembler strengths including the OLC (overlap-layout-consensus) approach, using long reads for assembly and short reads for error correction. DBG2OLC uses OLC and OPERA-MS combines long reads with low coverage and short reads to create a scaffold graph and group contigs into species-specific clusters using a Bayesian clustering algorithm that leverages read depth and long-read connections.

Long-read metagenome assemblers generate accurate and complete assemblies by producing contiguous sequences that can resolve complex regions of the genome. Canu is a long-read metagenome assembler that generates assemblies using the OLC approach. It error-corrects high-noise, high-coverage long reads, then uses "unitigging" to construct contigs from the graph representation of the genome/metagenome. MetaFlye generates genome assemblies from high-coverage long reads. It starts with constructing a de Bruijn graph from input reads, then identifies paths most likely to correspond to genomes. Error correction and polishing are done by aligning reads to the genome assemblies and correcting errors to improve the final genome quality. HifiAsm-meta uses the de Bruijn graph to identify probable paths of individual genomes from low-error PacBio-HiFi reads. Contigs are generated from these paths, which can then be used to reconstruct MAGs (Metagenome-Assembled Genomes) with the binning algorithm.

Most of the current assemblers do not represent complete microbial genomes with a single sequence so different binning algorithms are used. Metagenome binning algorithms group sequences into clusters. Many algorithms use features like TNFs (taxonomic novelty features), k-mer frequencies, and read depth to differentiate sequences and assign them to bins. MetaBAT2 [36] is a metagenome binning tool that uses TNFs and read depths to compute sequence similarities and partitions the graph of similarities into subgraphs using a modified label propagation algorithm (LPA).

IV. Evaluation of metagenome assembly

The evaluation of the quality of contigs assembled through metagenomic sequencing is a challenging problem. High-coverage regions of the contig, which are defined as regions that are overlapped by a large number of reads, are often associated with high-quality contigs. However, the development of a method for quantification

of these observations into a comprehensive quality score for the entire contig is a topic of ongoing research.

The majority of methods employed for the assessment of the quality of metagenomic assemblies depend on the availability of a set of annotated reference genomes for comparative analysis. One such commonly used reference-based method is metaQUAST [37], which aligns the contigs of the assembly to the reference genomes and subsequently computes various statistical measures such as: number of contigs, length of the largest contig, total length, N50, NG50, etc. MetaQUAST is a customized metagenomics version of the QUAST tool [38], which is one of the most widely used tools for genome assembly evaluation. Additionally, metaQUAST provides the capability to detect and visualize misassemblies and extract unrepresented genomic regions. However, it can be challenging to select a reference for novel MAGs from distantly related organisms, and the presence of similar organisms can lead to incorrect contig assignments. For the listed reason, the evaluation of the metagenomic assembly would be much more accurate and realistic if methods without reference requirements were used. An example of such a method is CheckM [39], which be utilized to evaluate the quality of metagenome assemblies by determining their completeness and contamination. The measures used by CheckM are based on the occurrence of specific genetic loci, and thus do not evaluate genome assembly at the level of individual contigs.

Repetitive genomic regions within the same genome or conserved sequences shared among different organisms can lead to assembly errors, including both inter- and intra-genome misassemblies. This is particularly likely to occur when multiple strains that are closely related are present in the same environment [40]. There are several existing reference-free methods for evaluating contig misassembly. ALE [41] provides nucleotide-level likelihood scores for assembled contigs, but not contig-level quality scores. SuRankCo [42] uses machine learning to provide contig quality scores based on length and coverage. VALET [43] detects misassemblies by combining multiple metrics extracted from the alignment of reads to contigs. DeepMAsED [44] is a deep learning-based tool for evaluating metagenomic assemblies by predicting the error rate using a CNN that analyzes k-mer content. The tool is trained on error-free reference assemblies and then used to predict the error rate of new assemblies by comparing k-mer content to the reference assemblies. metaMIC [45] is a tool for identifying and correcting misassemblies in metagenomic assemblies by localizing breakpoints. It uses features from both reads and assemblies, such as read coverage and k-mer consistency, to detect both intra- and inter-genome misassemblies. It can be adapted to work with assemblies from various assembler tools.

V. TEST EXPERIMENT

A. Experimental setup

To demonstrate the impact of metagenome assembly on metagenomic classification, we conducted two experiments, one based on species level and the other on strain level classification. First, we generated three metagenomic samples by simulating reads from known references using the Badread tool [46]. Two simulated datasets were used for experimental setup 1 and the third one was used for experimental setup 2. The simulation included ONT and PacBio long reads, incorporating the respective built-in qscore and error models (nanopore2020 and pacbio2016). To investigate differences between ONT and PacBio sequencing technologies, experimental setup 1 included simulated reads for both of them. Because the results showed no significant difference between the technologies, experimental setup 2 included only ONT simulated reads.

The presence of similar or low abundant organisms in the metagenomic sample may result in poor classification or assembly, however, we are using simulated reads with sufficient coverage. In contrast to actual metagenomic samples, our samples are characterized by a reduced number of organisms. The selection of a reduced number of organisms was done deliberately to facilitate the evaluation of the entire method. In the first sample, organisms were chosen based on their shared taxonomic levels, specifically order. The second metagenomic sample was simulated using organisms that are representative of those present in Zymo D6331 Gut Microbiome Standard metagenomic sample, where their relative abundance is well-established. This methodology ensures the preservation of authentic inter-organism associations present in the original sample. Simulated samples also contain errors, random reads as well as chimeric reads and thereby additionally reflect real samples. In addition, the second dataset contains five strains of E.coli, which also reflects the real situation, because metagenomic samples generally contain multiple strains of the same species, which makes strain level classification more challenging.

Once we simulated the reads, we performed metagenomic classification using the Kraken2 as a tool that provides the best balance between speed and accuracy. After the initial classification of raw reads, we assembled metagenomic samples using the metaFlye as a state-of-the-art metagenome assembly tool and additionally performed classification on the obtained contigs. Our main hypothesis was that assembling the metagenomic sample could enhance strain level classification and reduce the number of false positives.

To compare the classification outcomes, we took into account the abundance of individual organisms. Specifically, we estimated the approximate abundance of assembled contigs by considering their length and coverage. To calculate the approximate number of reads utilized for assembling a single contig, we applied the following

formula:

$$num_of_reads = ceil(\frac{contig_len * contig_cov}{mean_read_len}) \quad (1)$$

where $contig_len$ and $contig_cov$ are the length and coverage of the contig for which we estimate the number of reads, and $mean_read_len$ is the mean length of all reads in the dataset.

B. Results and discussion

Table I presents the number of false positive classifications on species and strain level for reads and contigs classification. From the results, it can be seen that with contigs classification the number of false positive species and strain identifications is significantly lower.

To reduce the number of incorrect classifications due to sequencing errors and other causes, a strain or species is considered identified if at least 50 reads are assigned to it. The used threshold (50 reads) was determined based on the study [8]. However, further investigation is necessary to establish the optimal threshold. Setting the threshold too high may result in true positive (TP) classifications reduction.

Table II shows calculated abundances for experimental setup 1. It can be seen that the classification at the species level is quite successful in both cases, but the abundance values of the contigs classification are still closer to the actual abundance values for all organisms. It can also be seen that there is no significant difference in the results for ONT and PacBio datasets. What additionally contributed to the success of the classification is the fact that all organisms in the sample are equally represented. However, even in this less complex scenario, it can be seen that the classification of reads reports a higher number of false positives and that the estimation of abundance is more accurate with the classification of contigs.

The calculated abundance values for experimental setup 2 are shown in Table III. The Table III does not show the results for 4 organisms: *Bifidobacterium adolescentis* (taxid: 367928), *Clostridioides difficile* (taxid: 1121308), *Escherichia coli* (taxid: 2605619), and *Roseburia hominis* (taxid: 585394). strain level classification for these organisms resulted in strains not being detected at all or being detected with abundance orders of magnitude lower than expected. Additionally, for 2 organisms (*Prevotella corporis* and *Veillonella rogosae*) the exact strain level was not known because the strain TAXID of those species is not present in the Kraken2 database. Abundance values for those two organisms were calculated at the species level.

In Table III, abundance values closer to the true abundances are presented in bold. The results show that there is a significant improvement in abundance values for contigs classification in 9 out of 14 cases. Additionally, in the case of non-bolded organisms with TAXIDs 83334 and 941322, the abundance values are quite different from the true values, but they are still better in the case of contigs classification.

TABLE I: **Number of FP classifications**

		reads classification	contigs classification
species level	experimental setup 1 ONT	14	2
	experimental setup 1 PacBio	15	1
	experimental setup 2 ONT	7	1
strain level	experimental setup 2 ONT	61	31

The results of experimental setup 2 show that, in general, strain level classification is significantly worse compared to species-level classification. Four of the strains were not detected by either method (reads and contigs), and in several cases reported abundance differs significantly from the true value. However, regardless of the fact that strain level classification in general requires further improvement, it is evident that the classification after assembling the metagenome gives better results in the context of the abundance of expected strains.

VI. Conclusion

The process of metagenomic analysis includes several steps, including metagenomic classification, de novo metagenome assembly, and the evaluation of metagenome assemblies. A variety of tools and techniques are available for each step, including long-read sequencing technologies such as PacBio and ONT, which have increased the accuracy of assemblies. It is challenging to identify a single tool that is superior to others, as most tools are well-suited to specific data or metagenomic analysis tasks.

One crucial aspect for future advancements in the field is the simulation of realistic metagenomic samples, which can greatly enhance the successful implementation of new tools for metagenomic classification, assembly, and assembly evaluation. The use of realistic datasets can provide a better understanding of the complexity of real-world samples and improve the performance of the new tools.

While single-genome assembly techniques have seen significant success in reconstructing a majority of genomes, metagenomic assemblers continue to face challenges in achieving high-quality reconstruction of the organisms present in a sample. It is necessary to develop appropriate methods for properly segregating reads according to the organisms present in the sample. Additionally, in order to evaluate the performance of metagenomic assemblers, quality assessment tools are needed, which particularly have difficulties when similar organisms are present in the sample, as it makes it challenging for these tools to distinguish between them.

In the context of this research, two simple experiments were conducted to test if there are indications of improvement in metagenomic classification with previous metagenome assembly. The datasets that we used in these experiments, although they contain errors as well as ran-

TABLE II: **Results for experimental setup 1.**

	ONT			PacBio		
species name - TAXID	true abundance	reads abundance	contigs abundance	true abundance	reads abundance	contigs abundance
Staphylococcus aureus - 1280	7,9%	8,7%	**7,9%**	8,0%	8,8%	**8,0%**
Staphylococcus epidermidis - 1282	7,0%	10,1%	**6,9%**	7,1%	10,6%	**7,0%**
Streptococcus mutans - 1309	5,7%	6,9%	**5,9%**	5,8%	7,6%	**5,7%**
Streptococcus agalactiae - 1311	5,3%	3,7%	**5,4%**	5,3%	3,6%	**5,3%**
Enterococcus faecalis - 1351	8,2%	10,4%	**8,4%**	8,2%	10,7%	**8,2%**
Bacillus cereus - 1396	15,0%	9,1%	**14,7%**	14,9%	9,3%	**14,9%**
Lactobacillus gasseri - 1596	6,1%	4,1%	**6,0%**	6,1%	3,4%	**6,1%**
Pseudomonas aeruginosa - 287	19,4%	20,1%	**19,2%**	19,3%	18,6%	**19,4%**
Acinetobacter baumannii - 470	10,5%	12,8%	**10,8%**	10,5%	12,8%	**10,6%**
Escherichia coli - 562	14,9%	14,0%	**14,7%**	14,8%	14,5%	**14,9%**

TABLE III: **Results for experimental setup 2.**

species name - strain TAXID	true abundance	reads abundance	contigs abundance
Akkermansia muciniphila - 349741	4,9%	0,2%	**6,1%**
Bacteroides fragilis - 295405	10,0%	**8,8%**	12,1%
Clostridium perfringens - 451752	6,5%	1,8%	**7,8%**
Enterococcus faecalis - 936153	5,9%	29,0%	**7,2%**
Escherichia coli strain 1 - 83334	10,7%	7,2%	13,9%
Escherichia coli strain 2 - 2778656	9,4%	3,4%	**7,8%**
Escherichia coli strain 3 - 941322	9,7%	0,9%	3,0%
Escherichia coli strain 5 - 83333	8,8%	1,5%	1,0%
Faecalibacterium prausnitzii - 657322	6,3%	8,7%	**7,4%**
Fusobacterium nucleatum - 469607	4,4%	6,8%	**5,2%**
Lactobacillus fermentum - 1381124	4,4%	**5,1%**	5,5%
Prevotella corporis - 28128	5,8%	8,8%	**7,1%**
Salmonella enterica - 59201	8,8%	11,0%	**10,6%**
Veillonella rogosae - 423477	4,4%	6,9%	**5,5%**

dom and chimeric reads, do not fully reflect the real situation due to the composition of the sample. However, even in such a less complex case, the results of the experiments show that the classification after the assembly of the metagenome significantly improves the estimation of the actual number of organisms in the sample, both at the species level and at the strain level. In addition, the experiments show that assembling the metagenome results in a smaller number of false positive classifications.

Further research of this approach is necessary, which would include simulation of more realistic datasets (larger number of species, more uneven true abundances, more strains of the same species, organisms that are not present in the database...) and testing other tools for classification and metagenome assembly. Also, it is necessary to investigate a method of evaluating the metagenome assembly to be used for further classification, including finding a correct way to detect contigs that do not contribute to the correct classification due to poor or incorrect assembly.

VII. ACKNOWLEDGMENTS

This research was funded by the Croatian Science Foundation under the grant IP-2018-01-5886 (SIGMA) and by the European Union through the European Regional Development Fund, under the grant KK.01.1.1.01.0009 (DATACROSS).

REFERENCES

[1] S. M. Bakhtiar, J. G. LeBlanc, E. Salvucci, A. Ali, R. Martin, P. Langella, J.-M. Chatel, A. Miyoshi, L. G. Bermúdez-Humarán, and V. Azevedo, "Implications of the human microbiome in inflammatory bowel diseases," *FEMS Microbiology Letters*, vol. 342, no. 1, pp. 10–17, 05 2013. [Online]. Available: https://doi.org/10.1111/1574-6968.12111

[2] T. Korem, D. Zeevi, J. Suez, A. Weinberger, T. Avnit-Sagi, M. Pompan-Lotan, E. Matot, G. Jona, A. Harmelin, N. Cohen *et al.*, "Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples," *Science*, vol. 349, no. 6252, pp. 1101–1106, 2015.

[3] A. Latorre-Pérez, P. Villalba-Bermell, J. Pascual, and C. Vilanova, "Assembly methods for nanopore-based metagenomic sequencing: a comparative study," *Scientific reports*, vol. 10, no. 1, pp. 1–14, 2020.

[4] A. L. Lapidus and A. I. Korobeynikov, "Metagenomic data assembly–the way of decoding unknown microorganisms," *Frontiers in Microbiology*, vol. 12, p. 613791, 2021.

[5] A. Rhoads and K. F. Au, "Pacbio sequencing and its applications," *Genomics, proteomics & bioinformatics*, vol. 13, no. 5, pp. 278–289, 2015.

[6] M. Jain, H. E. Olsen, B. Paten, and M. Akeson, "The oxford nanopore minion: delivery of nanopore sequencing to the genomics community," *Genome biology*, vol. 17, no. 1, pp. 1–11, 2016.

[7] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of molecular biology*, vol. 215, no. 3, pp. 403–410, 1990.

[8] J. Marić, K. Križanović, S. Riondet, N. Nagarajan, and M. Šikić, "Benchmarking metagenomic classification tools for long-read sequencing data," *BioRxiv*, pp. 2020–11, 2021.

[9] D. E. Wood, J. Lu, and B. Langmead, "Improved metagenomic analysis with kraken 2," *Genome biology*, vol. 20, no. 1, pp. 1–13, 2019.

[10] D. Kim, L. Song, F. P. Breitwieser, and S. L. Salzberg, "Centrifuge: rapid and sensitive classification of metagenomic sequences," *Genome research*, vol. 26, no. 12, pp. 1721–1729, 2016.

[11] R. Ounit, S. Wanamaker, T. J. Close, and S. Lonardi, "Clark: fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers," *BMC genomics*, vol. 16, no. 1, pp. 1–13, 2015.

[12] R. Ounit and S. Lonardi, "Higher classification sensitivity of short metagenomic reads with clark-s," *Bioinformatics*, vol. 32, no. 24, pp. 3823–3825, 2016.

[13] A. T. Dilthey, C. Jain, S. Koren, and A. M. Phillippy, "Strain-level metagenomic assignment and compositional estimation for long reads with metamaps," *Nature communications*, vol. 10, no. 1, pp. 1–12, 2019.

[14] D. H. Huson, B. Albrecht, C. Bağcı, I. Bessarab, A. Gorska, D. Jolic, and R. B. Williams, "Megan-lr: new algorithms allow accurate binning and easy interactive exploration of metagenomic long reads and contigs," *Biology direct*, vol. 13, no. 1, pp. 1–17, 2018.

[15] H. Li, "Minimap2: pairwise alignment for nucleotide sequences," *Bioinformatics*, vol. 34, no. 18, pp. 3094–3100, 2018.

[16] F. Beghini, L. J. McIver, A. Blanco-Míguez, L. Dubois, F. Asnicar, S. Maharjan, A. Mailyan, P. Manghi, M. Scholz, A. M. Thomas *et al.*, "Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3," *Elife*, vol. 10, p. e65088, 2021.

[17] K. D. Pruitt, T. Tatusova, G. R. Brown, and D. R. Maglott, "Ncbi reference sequences (refseq): current status, new features and genome annotation policy," *Nucleic acids research*, vol. 40, no. D1, pp. D130–D135, 2012.

[18] D. E. Wood and S. L. Salzberg, "Kraken: ultrafast metagenomic sequence classification using exact alignments," *Genome biology*, vol. 15, no. 3, pp. 1–12, 2014.

[19] J. Lu, F. P. Breitwieser, P. Thielen, and S. L. Salzberg, "Bracken: estimating species abundance in metagenomics data," *PeerJ Computer Science*, vol. 3, p. e104, 2017.

[20] H. Y. Simon, K. J. Siddle, D. J. Park, and P. C. Sabeti, "Benchmarking metagenomics tools for taxonomic classification," *Cell*, vol. 178, no. 4, pp. 779–794, 2019.

[21] S. McGinnis and T. L. Madden, "Blast: at the core of a powerful and diverse set of sequence analysis tools," *Nucleic acids research*, vol. 32, no. suppl_2, pp. W20–W25, 2004.

[22] C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner, "The silva ribosomal rna gene database project: improved data processing and web-based tools," *Nucleic acids research*, vol. 41, no. D1, pp. D590–D596, 2012.

[23] D. A. Benson, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler, "Genbank," *Nucleic acids research*, vol. 33, no. suppl_1, pp. D34–D38, 2005.

[24] D. M. Portik, C. T. Brown, and N. T. Pierce-Ward, "Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets," *BMC bioinformatics*, vol. 23, no. 1, pp. 1–39, 2022.

[25] B. Haider, T.-H. Ahn, B. Bushnell, J. Chai, A. Copeland, and C. Pan, "Omega: an overlap-graph de novo assembler for metagenomics," *Bioinformatics*, vol. 30, no. 19, pp. 2717–2722, 2014.

[26] T. Namiki, T. Hachiya, H. Tanaka, and Y. Sakakibara, "Metavelvet: an extension of velvet assembler to de novo metagenome assembly from short sequence reads," in *Proceedings of the 2nd ACM conference on bioinformatics, computational biology and biomedicine*, 2011, pp. 116–124.

[27] Y. Peng, H. C. Leung, S.-M. Yiu, and F. Y. Chin, "Idba-ud: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth," *Bioinformatics*, vol. 28, no. 11, pp. 1420–1428, 2012.

[28] D. Li, C.-M. Liu, R. Luo, K. Sadakane, and T.-W. Lam, "Megahit: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph," *Bioinformatics*, vol. 31, no. 10, pp. 1674–1676, 2015.

[29] S. Boisvert, F. Raymond, É. Godzaridis, F. Laviolette, and J. Corbeil, "Ray meta: scalable de novo metagenome assembly and profiling," *Genome biology*, vol. 13, no. 12, pp. 1–13, 2012.

[30] S. Koren, B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, and A. M. Phillippy, "Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation," *Genome research*, vol. 27, no. 5, pp. 722–736, 2017.

[31] M. Kolmogorov, D. M. Bickhart, B. Behsaz, A. Gurevich, M. Rayko, S. B. Shin, K. Kuhn, J. Yuan, E. Polevikov, T. P. Smith *et al.*, "metaflye: scalable long-read metagenome assembly using repeat graphs," *Nature Methods*, vol. 17, no. 11, pp. 1103–1110, 2020.

[32] X. Feng, H. Cheng, D. Portik, and H. Li, "Metagenome assembly of high-fidelity long reads with hifiasm-meta," *Nature Methods*, pp. 1–4, 2022.

[33] S. Nurk, D. Meleshko, A. Korobeynikov, and P. A. Pevzner, "metaspades: a new versatile metagenomic assembler," *Genome research*, vol. 27, no. 5, pp. 824–834, 2017.

[34] C. Ye, C. M. Hill, S. Wu, J. Ruan, and Z. S. Ma, "Dbg2olc: efficient assembly of large genomes using long erroneous reads of the third generation sequencing technologies," *Scientific reports*, vol. 6, no. 1, pp. 1–9, 2016.

[35] D. Bertrand, J. Shaw, M. Kalathiyappan, A. H. Q. Ng, M. S. Kumar, C. Li, M. Dvornicic, J. P. Soldo, J. Y. Koh, C. Tong *et al.*, "Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes," *Nature biotechnology*, vol. 37, no. 8, pp. 937–944, 2019.

[36] D. D. Kang, F. Li, E. Kirton, A. Thomas, R. Egan, H. An, and Z. Wang, "Metabat 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies," *PeerJ*, vol. 7, p. e7359, 2019.

[37] A. Mikheenko, V. Saveliev, and A. Gurevich, "Metaquast: evaluation of metagenome assemblies," *Bioinformatics*, vol. 32, no. 7, pp. 1088–1090, 2016.

[38] A. Mikheenko, A. Prjibelski, V. Saveliev, D. Antipov, and A. Gurevich, "Versatile genome assembly evaluation with quast-lg," *Bioinformatics*, vol. 34, no. 13, pp. i142–i150, 2018.

[39] D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson, "Checkm: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes," *Genome research*, vol. 25, no. 7, pp. 1043–1055, 2015.

[40] C. Kingsford, M. C. Schatz, and M. Pop, "Assembly complexity of prokaryotic genomes using short reads," *BMC bioinformatics*, vol. 11, no. 1, pp. 1–11, 2010.

[41] S. C. Clark, R. Egan, P. I. Frazier, and Z. Wang, "Ale: a generic assembly likelihood evaluation framework for assessing the accuracy of genome and metagenome assemblies," *Bioinformatics*, vol. 29, no. 4, pp. 435–443, 2013.

[42] M. Kuhring, P. W. Dabrowski, V. C. Piro, A. Nitsche, and B. Y. Renard, "Surankco: supervised ranking of contigs in de novo assemblies," *BMC bioinformatics*, vol. 16, no. 1, pp. 1–7, 2015.

[43] N. D. Olson, T. J. Treangen, C. M. Hill, V. Cepeda-Espinoza, J. Ghurye, S. Koren, and M. Pop, "Metagenomic assembly through the lens of validation: recent advances in assessing and improving the quality of genomes assembled from metagenomes," *Briefings in bioinformatics*, vol. 20, no. 4, pp. 1140–1150, 2019.

[44] O. Mineeva, M. Rojas-Carulla, R. E. Ley, B. Schölkopf, and N. D. Youngblut, "Deepmased: evaluating the quality of metagenomic assemblies," *Bioinformatics*, vol. 36, no. 10, pp. 3011–3017, 2020.

[45] S. Lai, S. Pan, C. Sun, L. P. Coelho, W.-H. Chen, and X.-M. Zhao, "metamic: reference-free misassembly identification and correction of de novo metagenomic assemblies," *Genome Biology*, vol. 23, no. 1, pp. 1–21, 2022.

[46] R. R. Wick, "Badread: simulation of error-prone long reads," *Journal of Open Source Software*, vol. 4, no. 36, p. 1316, 2019. [Online]. Available: https://doi.org/10.21105/joss.01316

Hyperparameter Optimization of Graph Neural Networks for mRNA Degradation Prediction

Viktorija Vodilovska* Sonja Gievska* Ilinka Ivanoska*
* Faculty of Computer Science and Engineering, Ss Cyril and Methodius University, Skopje, N. Macedonia
viktorija.vodilovska@students.finki.ukim.mk sonja.gievska@finki.ukim.mk ilinka.ivanoska@finki.ukim.mk

***Abstract*—Graph Neural Networks (GNN) emerged as increasingly attractive deep learning models for complex data, making them extremely useful in biochemical and pharmaceutical domains. However, building a good-performing GNN requires lots of parameter choices and Hyperparameter Optimization (HPO) can aid in exploring solutions. This study presents a comparative analysis of several strategies for Hyperparameter Optimization of GNNs. The explored optimization techniques include complex algorithms such as the bio-inspired Genetic Algorithm, Particle Swarm Optimization, and Artificial Bee Colony. In addition, Hill Climb and Simulated Annealing as well as the commonly used methods Random Search and Bayesian Search have also been covered.**

The proposed optimization algorithms have been evaluated on improving the performance of the GNN architectures developed for predicting mRNA degradation. The Stanford OpenVaccine dataset for mRNA degradation prediction has been used for training and testing the predictive models. Finding mRNA molecules with low degradation rates is important in development of mRNA vaccines for diseases such as COVID-19 and we hope to benefit research on ML in this domain. According to the analysis's findings, Simulated Annealing algorithm outperforms other algorithms on both architectures. Furthermore, population based algorithms like Particle Swarm Optimization show promising results, with certain limitations related to the complexity of the algorithms which encourages further exploration of the subject.

Keywords*—*Hyperparameter Optimization, Random Search, Bayesian search, Hill Climbing, Simulated Annealing, Genetic Algorithm, Artificial Bee Colony, Particle Swarm Optimization, GCN, GAT, mRNA degradation, mRNA vaccines

I. INTRODUCTION

Graph Neural Networks (GNN) are state-of-the-art methods for deep learning on data with a complex graph structure. Graphs can represent a wide variety of problems, which allows these models to be very expressive and applicable to arbitrary graph data of different shapes and sizes

Due to the growing quantities of biological and medical data, Graph Neural Networks have become an important tool in bioinformatics over the past years [1]. Graphs have been used to represent various types of biological information like genetic disease association networks, protein interaction networks, structure of molecules and macromolecules (proteins, DNA, RNA), brain networks, etc. With the growing complexity of these graph models, the need for an optimization strategy to build effective models also increased.

A model performance is directly linked to how well it was built to fit the specific data and problem. This requires choosing the right values for a wide range of hyperparameters that lie in both continuous and discrete feature spaces. Hyperparameter optimization (HPO) or tuning is the problem of selecting a set of values for the parameters of a learning algorithm, with the expectation of achieving better model performance. A hyperparameter is a parameter whose value is used to control the algorithm's learning process. As the hyperparameter space gets broader so does the difficulty of finding a good-performing architecture for the given problem. An extensive grid search over all those values is expensive and not always an option.

This study presents an experimental comparative analysis over a larger selection of optimization methods for the Hyperparameter Optimization of GNNs on the task of mRNA degradation prediction. In doing so, the aim is to facilitate future research of GNN application in bioinformatics. The experiments include the following HPO algorithms: Random Search (RS) [2] for benchmark, Bayesian Search (BS) [3], Hill Climb (HC) [4], Simulated Annealing (SA) [5], Genetic Algorithm (GA) [6], Particle Swarm Optimization (PSO) [7] and Artificial Bee Colony (ABC) Optimization [8]. The GNN architectures probed in the HPO experiments are GCN [9] and GAT [10]. The analysis used the Stanford OpenVaccine Dataset [11] for mRNA degradation prediction which is an important issue when developing mRNA vaccines for diseases such as COVID-19, and a relatively novel field. The GNN prediction task involved regressive multi-target node prediction.

The rest of the paper is structured as follows: an overview of the existing literature on the topic is given in section II. The dataset, explored GNN architectures as well as the hyperparameter optimization algorithms are explained in detail in section III. The experimental setup with results and discussion of the experiments are laid out in section IV. Finally, section V concludes the findings, and a reference to supplementary material is given in section VI.

II. RELATED WORK

Hyperparameter optimization (HPO) is a crucial step in the development of any high-performing neural networks.

Supported by Faculty of Computer Science and Engineering, Skopje, N. Macedonia.

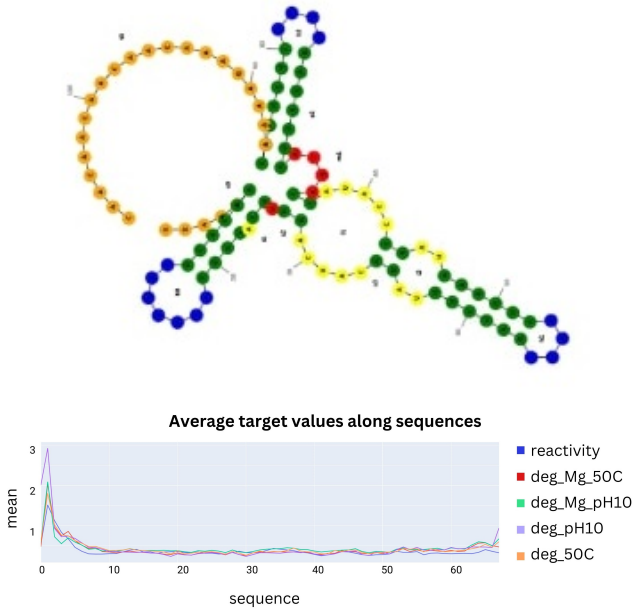


Fig. 1: (a) The 2D structure of an example mRNA molecule. Nodes represent the nucleotides of the mRNA. Image generated with the ViennaRNA package [12]. (b) Average of target values, with each line representing a different target, along all train molecules.

However, only a few studies have been undertaken to tackle the issue of finding a well-performing HPO strategy for GNNs.

A systematic comparison of three HPO algorithms, with RS for baseline, Tree-structured Parzen Estimator (TPE), and Covariance Matrix Adaptation Evolution Strategy (CMA-ES), was conducted in [13] on GNN models for molecular property prediction using three representative datasets from DeepChem's [14] MoleculeNet [15]. The search space includes values for batch size, learning rate, size of fully-connected layer, and size of a graph convolution layer. TPE outperforms RS and CMA-ES, achieving the best overall performance. Different studies such as DeepHyper [16], or [17], [18] suggest a preference for BS HPO techniques and also show promising results on GA for HPO. Moreover, other bio-inspired population algorithms like PSO have been effectively used in hyperparameter optimization tasks for Convolutional Neural Networks and Long Short-Term Memory (LSTM) networks (see [19], [20]).

Neural Architecture Search (NAS) framework can be considered as another automatic approach for choosing hyperparameters. For example the work on AGNN [21] proposes an AutoML framework NAS on GNNs, with the use of reinforcement learning for decisions on new actions for each class, while trying to optimize the objective function - the model performance. [22] proposes another AutoML framework for GNNs called DFG-NAS (Deep and Flexible Graph Neural Architecture Search) that employs an evolutionary algorithm and search space decoupling into propagation and transformation. Both of these works are evaluated on popular benchmarks like Cora, Citeseer, and PubMed, however their optimization frameworks are model specific to the chosen architecture compared to the model-agnostic optimization algorithms explored in this study.

Overall, while several studies have explored HPO methods for GNN-based models, our survey aims to provide a comprehensive comparison of seven different HPO algorithms (RS, BS, HC, SA, GA, PSO, ABC) on GNN models like GAT and GCN. This study focuses on the HPO of GNNs built specifically for prediction on mRNA molecules which we found to be a less explored topic. Our search space includes a wide range of micro and macro architecture hyperparameters such as learning rate, dropout rate, number of prediction layers, number of hidden units, number of graph layers. Since the optimization algorithms explored in this solution are model and problem agnostic we hope this study benefits research in other areas as well.

III. Materials and Methods

This section presents the dataset, GNN architecture, and hyperparameter optimization algorithms used.

A. Dataset

The Stanford Eterna Dataset [11] consists of mRNA molecules represented by it's sequence of nucleotides with structural features and targets. A visual representation of an mRNA molecule can be seen on Fig.1. During preprocessing each mRNA molecule was transformed into a graph $G = (V, E)$, with nodes V representing the nucleotides and edges E - the chemical bonds between them. Nucleotide types are used as node features while the primary structure bonds as well as the base pair probabilities are used as edge features, as in [23], [24].

The Eterna Dataset comes separated into a train set of 2400 molecules and a test set of 3634 molecules. The train set contains values for five different prediction targets (including reactivity, deg_Mg_pH10, deg_Mg_50C', 'deg_pH10', 'deg_50C') whereas the test set contains three of them (reactivity, deg_Mg_pH10, deg_Mg_50C'). For data visualization a line plot of the mean targets along all sequences in the training data can be seen on Fig.1. In this study we utilized the training set to perform model model HPO and training. The chosen models from the HPO process are trained on 80% of the train set and are validated on the other 20%. To evaluate each of the full models we utilized the test set from the Eterna Dataset. The data splits are available on the Kaggle link in Section VI along with our prepossessing code.

B. Graph neural networks models

The explored architectures are composed of two machine-learning stacks. A Graph Stack for creating node embeddings from the graph structured input and a Prediction Stack for getting output predictions. Furthermore, two variations of the Graph Stack are implemented using the models GCN and GAT, respectively. An overview of the

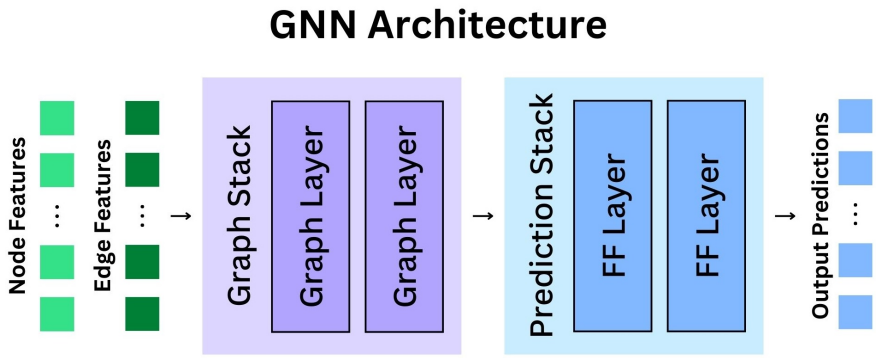


Fig. 2: High level overview of the GNN architecture tuned in the HPO experiments.

architecture can be found in Fig.2 and the hyperparameters of each part are listed in Table I. GNN models have been chosen to fit the problem inspired by other works that explored them [23], [24]. The suitability of the GNN architecture for the chosen data is beyond the scope of this study.

1) Graph Convolutional Network: Graph Convolutional Network (GCN) [9] emerged as an effective approach for semi-supervised learning on graph data. These models learn the features by performing convolution over neighboring nodes directly on the graph. Multiple layers of graph convolution can be added. Feed Forward layers have been added to get output node features. The various hyperparameters (graph layers, hidden size, dropout rate) chosen during HPO are detailed on Table I.

2) Graph Attention Network: The Graph Attention Network (GAT) [10] include attention [25] on top of the convolution. Similar to GCN multiple layers may be stacked on top of each other. Likewise, Feed Forward layers have been added to get output values. The hyperparameters explored in this study can be seen in Table I).

C. Optimization algorithms

The following optimization methods have been used in this study.

1) Random Search: Random Search is favored for hyperparameter optimization [2], and generally the most simple way to tackle this problem. It replaces the exhaustive enumeration of all combinations by selecting them randomly. Moreover, it is simple to apply to both discrete and continuous spaces. It performs better than Grid search [26] when only a few hyperparameters have an impact on the model performance. RS is a crucial benchmark for evaluating the effectiveness of new hyperparameter optimization techniques.

2) Bayesian Search: Another intuitive approach and a common choice for HPO is Bayesian Search [3]. It is a global optimization method for any noisy black-box functions. BS creates a probabilistic model that maps the values of the hyperparameters to the objective as determined by a validation set. By iteratively evaluating hyperparameter configurations, BS gathers observations about the optimum. This algorithm attempts to find the best solution by maintaining a balance between open exploration - hyperparameters for which the outcome is uncertain, and information exploitation - hyperparameters expected to be close to the optimum.

3) Hill Climbing: Hill Climbing [4], as a member of the family of local search algorithms, has an objective to minimize or maximize a target function by gradually updating the hyperparameter vector. The iterative process starts with a randomly chosen solution that is evaluated to find the direction of progress. Subsequently, in each iteration, HC adjusts a single hyperparameter and determines whether the change improves the performance. It differs from gradient descent methods, which adjust all of the values at each iteration.

4) Simulated Annealing: Simulated annealing [5] is a probabilistic technique for approximating the global optimum of a given function. Applied to hyperparameter tuning, SA is an iterative process which updates the solution while slowly adapting the exploration probability. Values for all hyperparameters are randomly chosen and an initial "temperature" is set. Alteration of the current state is done by updating one hyperparameter with a random value in the immediate neighborhood, with a certain probability based on the temperature. On every n^{th} step the temperature is decreased by a chosen rate. SA is preferred for discrete feature space problems when finding an approximate global optimum quickly is important.

5) Genetic Algorithm: The Genetic Algorithm [6] is an iterative, heuristic method inspired by the process of natural selection where the fittest individuals give rise to the following generation. In the context of HPO each individual is a proposed solution for the hyperparameter values. The population of solutions in each iteration is referred to as a generation. In each generation, the members of the population are evaluated in terms of their fitness, in this context the performance of the solution. The new generation is created by stochastically selecting the fittest solutions from the current population, recombining them, and introducing random mutations. The algorithm gradually generates better solutions by combining the best of each generation.

TABLE I: Summary table containing hyperparameters and explored ranges for optimization.

Configuration	Hyperparameter Name	Description	Hyperparameter Type	Explored Range
GCN Stack	**Graph Layers**	number of GCN layers	discrete	[1, 3]
	Hidden Channels	dimension of the hidden layers	discrete	[32, 128]
	Graph Dropout	dropout rate after GCN layers	continuous	(0.2, 0.5)
GAT Stack	**Graph Layers**	number of GCN layers	discrete	[1, 3]
	Hidden Channels	dimension of the hidden layers	discrete	[32, 128]
	Graph Dropout	dropout rate after GAT layers	continuous	(0.2, 0.5)
	Attention Dropout	attention dropout rate	continuous	(0.2, 0.5)
Prediction Stack	**Prediction Layers**	number of Feed Forward layers	discrete	[1, 3]
	Hidden Channels	dimension of the hidden layers	discrete	[32, 128]
	Prediction Dropout	dropout rate after Feed Forward layers	continuous	(0.2, 0.5)
Training	**Batch Size**	size of batches during training	discrete	[8, 32]
	Learning Rate	model learning rate	continuous	(0.00001, 0.01)
	Loss	loss function	categorical	[CrossEntropy, MSE, MAE]

6) Particle Swarm Optimization: Particle Swarm Optimization [7] is a well-known swarm intelligence technique, inspired by the navigation strategy of bird flocks in quest of food. The technique has been tested on a variety of high-dimensional real-world applications and has shown to be efficient and reliable [27], [28]. It starts with a random set of individual search 'particles', each representing a potential solution. The particles update states iteratively based on others in the swarm. PSO has an advantage in exploring a wide, multi-dimensional search space. While it cannot guarantee to find the global optimum solution it is likely to find a close-to-optimum solution in relatively few generations (iterations).

7) Artificial Bee Colony: Another swarm intelligence algorithm is Artificial Bee Colony [8], inspired by the foraging behavior of honey bee swarms. It contains three main components: employed bees, onlookers, and scouts. The position of the bees represents a set of hyperparameter values. The artificial bees cooperate to find better solutions using neighbouring search by choosing the best directions to move towards. By combining the gathered info, the population iteratively improves its performance.

IV. Experiments, results and discussion

A. Experimental details

The optimization algorithms are implemented to probe the hyperparameter space of two GNN architectures. A detailed list of hyperparameters tackled in the experiments can be found in Table I. The GNN architectures are implemented to perform multi-regression to predict five different numerical prediction targets. The predictions are performed on each node (nucleotide). Model performance is evaluated through the Mean Column-wise Root Mean Squared Error (MCRMSE) metric - following the Kaggle Competition [29] which uses this dataset. The optimization objective function is the training of the GNN configuration as given by the hyperparameters. The objective goal is minimizing the evaluation performance metric - MCRMSE.

In this experiment the GCN and GAT architectures have been optimized for 1h by each algorithm. In each optimization, model configurations were trained and validated for 5 epochs on a subset of 240 molecules with 107 nodes (nucleotides) and 5 targets per node, or a total of 25,680 labeled nodes. The optimal model configuration found by each algorithm was trained and validated on the full training set of 2400 molecules for 10 epochs. The developed models were additionally evaluated on the test set of 3634 molecules with 3 targets per node.

Due to the large differences in the internal process of each optimization method and following the recommendations of [13] the experiments are ran on a fixed time of 1 hour to assess the optimization performance fairly. The parameters of the HPO algorithms, also called meta-hyperparameters are left to default values. Exploration of the meta-hyperparameters is outside the scope of this study. The HPO algorithms are implemented with the NiaPy package [30], whereas the GNN models with PyTorch [31]. All experiments have been performed on an Intel(R) Core(TM) i7-10510U CPU with 1.80GHz and 16GB RAM. Additional details about the implementation can be found in Section VI.

B. Results and discussion

The results show that RS stands its ground as the most common HPO algorithm and manages to find the best performing model configuration for both GCN and GAT architectures on the HPO validation subset with MCRMSE of 0.55 and 0.60 respectively. Although, the fully trained models for RS do not hold the top performance on the train validation subset and are outperformed by ABC's model for GCN and HC's model for GAT. SA however, outperforms RS and all other algorithms for both GCN and GAT on the test set with Test MCMRSE of 2.86 and 3.05 respectively. Albeit the solutions found by the population based methods fall short compared to the simpler methods on the optimization set, PSO and GA offer the next best performances after SA on the evaluation set for GCN, while PSO comes close to SA on the evaluation set for GAT after RS.

By looking at the chosen hyperparameters from each algorithm, high heterogeneity in the values of the top architectures can be concluded (Fig. 3). Some hyperpa-

TABLE II: Summary table containing performance comparison between the HPO algorithms for both the GCN and GAT architectures. HPO MCRMSE is the top validation performance during the 1h HPO experiment. Train and Test MCRMSE are the validation and test performance respectively of the fully trained top architectures (best results with lowest MCRMSE).

HPO	GCN				GAT			
Algorithm	HPO MCRMSE	Train MCRMSE	Test MCRMSE	# Sol.	HPO MRCSE	Train MCRMSE	Test MCRMSE	# Sol.
RS	**0.552111207**	1.197424498	3.142875612	24	**0.606282744**	1.227286043	3.085664921	18
BS	0.663734097	1.397005282	3.73452739	24	0.701095777	1.237067464	3.607524834	17
HC	1.34278588	1.776945629	3.020120101	17	0.658518997	**0.754333032**	3.621118421	17
SA	0.594058272	1.568616626	**2.858786327**	28	1.19119597	1.19119597	**3.049026682**	18
GA	0.882215188	1.428922103	2.950592929	17	0.841650576	1.430237301	3.205829579	16
PSO	0.799015684	1.348786081	2.91169034	19	0.858330531	1.427626225	3.193924	19
ABC	0.751826537	**1.039043296**	3.686253486	12	0.716234768	2.027690987	3.159572088	18

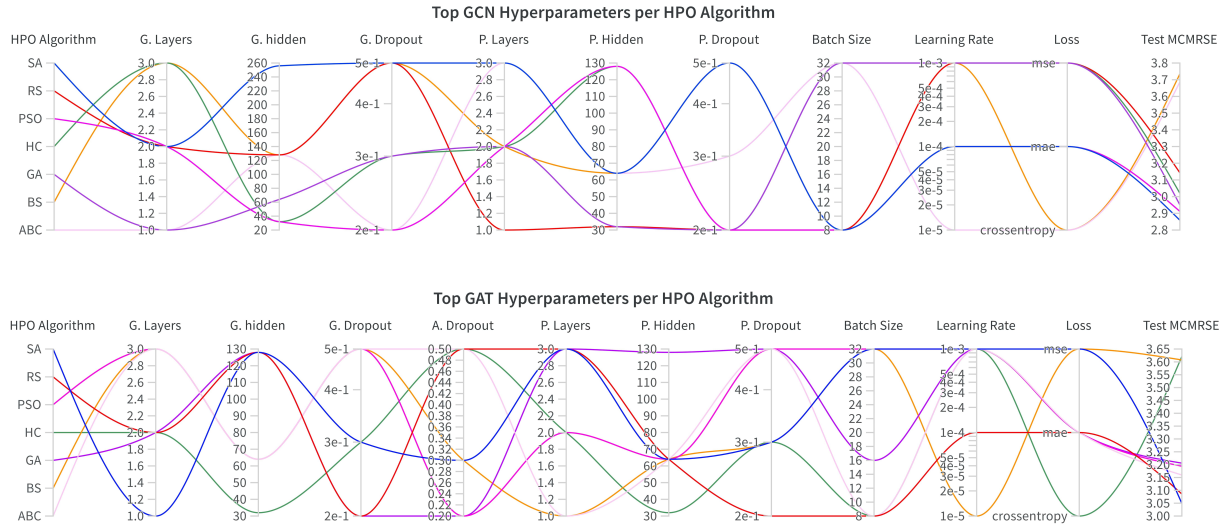


Fig. 3: Parallel coordinate plots showing the hyperparameter values chosen by each HPO algorithm during the 1h optimization experiment for a) the GCN and b) the GAT configurations, along with Test MCMRSE performance values.

rameters values as 2 for graph layers, a batch size of 8, learning rate of 0.001 or the MAE loss type are more dominant and agreed upon by most algorithms for both architectures. Other hyperparameter values for GCN with a majority vote are 0.5 for graph dropout, 2 for prediction layers, 32 for prediction hidden channels and 0.2 for prediction dropout, while there is no consensus on the choice of graph hidden channels. For the GAT architecture the algorithms generally agree more on the chosen values with a consensus on 128 for graph hidden channels, 0.2 for graph dropout, 0.5 for attention dropout, 3 for prediction layers, 64 for prediction hidden channels and 0.5 for prediction dropout. Additionally, the analysis of the optimization runs performance at each step during the 1h tuning shows there is high variation between performances of consecutive steps especially for the GAT. This points to a high level of complexity in the search space and large performance differences for small parameter changes.

Overall, the results suggest that SA might be a good choice for HPO on GNN in this application and also shows promising results for population based algorithms. The present findings confirm the state-of-the-art usage of these methods in line with the ideas of [16], [17], [18], [19], [20]. A summary of the results can be seen in the performance comparison table II. More details on the implementation and performance of each HPO algorithm separately can be seen in Section VI.

V. Conclusion

The objective of this study was to make an experimental comparison of HPO strategies for Graph Neural Networks implemented in an mRNA node-wise prediction problem. The performance analysis has shown promising results for HPO algorithms such as the SA algorithm as well as population based algorithms like PSO, which performed better than RS at choosing GNN hyperparameters. The SA algorithm found the best-performing GAT and GCN architectures, while being highly efficient and exploring the largest amount of configurations.

The population algorithms including GA, PSO and ABC show promising results and might perform even better given a bigger population size. It must be noted that this type of algorithms give significant memory complexity overhead and therefore had to be limited in population

size. The overhead also explains the smaller amount of configurations explored by them.

This survey offers a practical comparison of multiple algorithms (RS, BS, HC, SA, GA, PSO, ABC) with focus on the HPO of GNNs specifically for prediction on mRNA molecules. We attempted to cover a wide range of hyperparameters and problem agnostic algorithms in hopes of facilitating other areas as well. A more extensive exploration involving diverse meta-hyperparameter choices is needed for a more informed comparison. The overall effectiveness of the GNN architecture itself for the chosen task is a topic for another study. Additionally, this survey needs to be further expanded to other problems and datasets to get a more holistic review of the effectiveness of the HPO algorithms. Future work involves the aim to deepen the exploration of HPO algorithms on GNN architectures for mRNA modelling in order to support researchers in this and other related interdisciplinary fields.

VI. Supplementary Material

The code for all the experiments, GNN models and HPO algorithms, as well as all additional results, are available on the repository https://github.com/ViktorijaVodilovska/mRNA_pred. The data used in this study is available on https://www.kaggle.com/c/stanford-covid-vaccine/overview.

Acknowledgment

This work was partially financed by the Faculty of Computer Science and Engineering at the Ss. Cyril and Methodius University in Skopje.

References

[1] X.-M. Zhang, L. Liang, L. Liu, and M.-J. Tang, "Graph neural networks and their current applications in bioinformatics," *Frontiers in genetics*, vol. 12, p. 690049, 2021.

[2] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization." *Journal of machine learning research*, vol. 13, no. 2, 2012.

[3] M. Pelikan, D. E. Goldberg, E. Cantú-Paz *et al.*, "Boa: The bayesian optimization algorithm," in *Proceedings of the genetic and evolutionary computation conference GECCO-99*, vol. 1. Citeseer, 1999, pp. 525–532.

[4] B. Selman and C. P. Gomes, "Hill-climbing search," *Encyclopedia of cognitive science*, vol. 81, p. 82, 2006.

[5] M. Fischetti and M. Stringher, "Embedded hyper-parameter tuning by simulated annealing," *arXiv preprint arXiv:1906.01504*, 2019.

[6] K. Sastry, D. Goldberg, and G. Kendall, "Genetic algorithms," in *Search methodologies*. Springer, 2005, pp. 97–125.

[7] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, vol. 4. IEEE, 1995, pp. 1942–1948.

[8] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," 2005.

[9] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[10] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[11] J. Lee, W. Kladwang, M. Lee, D. Cantu, M. Azizyan, H. Kim, A. Limpaecher, S. Gaikwad, S. Yoon, A. Treuille, R. Das, and E. Participants, "Rna design rules from a massive open laboratory," *Proceedings of the National Academy of Sciences*, vol. 111, no. 6, pp. 2122–2127, 2014. [Online]. Available: https://www.pnas.org/content/111/6/2122

[12] R. Lorenz, S. H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, and I. L. Hofacker, "Viennarna package 2.0," *Algorithms for molecular biology*, vol. 6, pp. 1–14, 2011.

[13] Y. Yuan, W. Wang, and W. Pang, "A systematic comparison study on hyperparameter optimisation of graph neural networks for molecular property prediction," in *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, jun 2021. [Online]. Available: https://doi.org/10.1145%2F3449639.3459370

[14] B. Ramsundar, P. Eastman, P. Walters, V. Pande, K. Leswing, and Z. Wu, *Deep Learning for the Life Sciences*. O'Reilly Media, 2019, https://www.amazon.com/Deep-Learning-Life-Sciences-Microscopy/dp/1492039837.

[15] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "Moleculenet: a benchmark for molecular machine learning," *Chemical science*, vol. 9, no. 2, pp. 513–530, 2018.

[16] P. Balaprakash, M. Salim, T. D. Uram, V. Vishwanath, and S. M. Wild, "Deephyper: Asynchronous hyperparameter search for deep neural networks," in *2018 IEEE 25th International Conference on High Performance Computing (HiPC)*, 2018, pp. 42–51.

[17] Y. Liu, J. Liu, and Y. Li, "Automatic search of architecture and hyperparameters of graph convolutional networks for node classification," *Applied Intelligence*, pp. 1–16, 2022.

[18] C. White, W. Neiswanger, and Y. Savani, "Bananas: Bayesian optimization with neural architectures for neural architecture search," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, 2021, pp. 10 293–10 301.

[19] Y. Guo, J.-Y. Li, and Z.-H. Zhan, "Efficient hyperparameter optimization for convolution neural networks in deep learning: A distributed particle swarm optimization approach," *Cybernetics and Systems*, vol. 52, no. 1, pp. 36–57, 2020.

[20] P. Singh, S. Chaudhury, and B. K. Panigrahi, "Hybrid mpso-cnn: Multi-level particle swarm optimized hyperparameters of convolutional neural network," *Swarm and Evolutionary Computation*, vol. 63, p. 100863, 2021.

[21] K. Zhou, Q. Song, X. Huang, and X. Hu, "Auto-gnn: Neural architecture search of graph neural networks," 2019.

[22] W. Zhang, Z. Lin, Y. Shen, Y. Li, Z. Yang, and B. Cui, "Dfg-nas: Deep and flexible graph neural architecture search," 2022.

[23] A. Singhal, "Predicting hydroxyl mediated nucleophilic degradation and molecular stability of rna sequences through the application of deep learning methods," 2021.

[24] A. Muneer, S. Fati, N. Akbar, D. Agustriawan, and S. Tri Wahyudi, "ivaccine-deep: Prediction of covid-19 mrna vaccine degradation using deep learning," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, pp. pp. 7419–7432, 10 2021.

[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[26] P. Liashchynskyi and P. Liashchynskyi, "Grid search, random search, genetic algorithm: a big comparison for nas," *arXiv preprint arXiv:1912.06059*, 2019.

[27] M. Neethling and A. P. Engelbrecht, "Determining rna secondary structure using set-based particle swarm optimization," in *2006 IEEE International Conference on Evolutionary Computation*. IEEE, 2006, pp. 1670–1677.

[28] A. A. Esmin, G. Lambert-Torres, and A. Z. De Souza, "A hybrid particle swarm optimization applied to loss power minimization," *IEEE Transactions on power systems*, vol. 20, no. 2, pp. 859–866, 2005.

[29] "Openvaccine: Covid-19 mrna vaccine degradation prediction." [Online]. Available: https://www.kaggle.com/competitions/stanford-covid-vaccine/overview

[30] G. Vrbančič, L. Brezočnik, U. Mlakar, D. Fister, and I. Fister Jr., "NiaPy: Python microframework for building nature-inspired algorithms," 10 2018. [Online]. Available: https://github.com/NiaOrg/NiaPy

[31] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, Eds. Curran Associates, Inc., 2019, pp. 8024–8035.

Examination of Different Representations of Proteins Using Protein Ray-based Descriptor and Deep Learning Models

G. Mirceva, A. Naumoski and A. Kulakov
Faculty of computer science and engineering, Ss. Cyril and Methodius University in Skopje, Skopje, Macedonia
georgina.mirceva@finki.ukim.mk, andreja.naumoski@finki.ukim.mk, andrea.kulakov@finki.ukim.mk

***Abstract* - The study of proteins has been of high importance because it is needed to understand the processes in the living organisms in which these molecules are involved. Proteomics is the research area that studies the protein structures. One of the tasks on which proteomics is focused on is solving the protein classification task. Although there are many studies focused on this problem, it is still a popular task because there is still need for faster methods for protein classification. The aim of the study presented in this paper is to develop a fast and accurate protein classification model. For that purpose, for feature extraction we use our protein ray-based descriptor. We use a deep learning architecture for generating prediction model. Besides the standard form of the protein ray-based descriptor, we also consider several other representations of the proteins and make examination which is the most appropriate representation. Some experimental results are given and discussed.**

Keywords - protein structure; protein classification; feature extraction; deep learning

I. INTRODUCTION

Proteomics focuses on the study of protein molecules, which are among the most vital compounds present in the human body. They participate in a variety of cellular processes, making them indispensable for various biological functions. Proteins serve critical roles in biochemical reactions as enzymes, also play a role in the transport of oxygen to cells, have signaling role such as insulin, and serve as antibodies thus providing a defensive mechanism. Due to their importance in the functioning of organisms, there is a significant interest in the proteomics community to understand the structure and functions of proteins. By studying the structure and properties of proteins, researchers can gain valuable insights into their functional roles in different cellular processes.

The development of technology has provided various techniques to determine the structure of protein molecules. These techniques enable the determination of protein structures, which are subsequently deposited in the Protein Data Bank (PDB) [1], [2], serving as the primary repository for this purpose. Protein structures are stored in PDB files, which provide information about the primary, secondary, and tertiary structures of proteins. Despite the vast amount of data available on protein structures, there remains a significant need for research to discover the functional roles of proteins in living organisms.

There are various approaches for determining the functions of proteins. Some approaches rely on the assumption that proteins belonging to the same class share similar functions. Therefore, the protein classification task is of high importance in this regard. Numerous methods have been introduced for classifying protein structures. However, despite these efforts, there is still a significant gap between the number of known protein structures and the number of proteins that are functionally annotated. This highlights the high need for computational methods for fast and accurate classification of proteins.

The Structural Classification Of Proteins (SCOP) [3] is widely recognized as a very important method for protein classification. This method is highly accurate, as it involves visual inspections by human experts in order to make a decision. However, this process is not very fast, which highlights the need for automatic or semi-automatic methods. One such method is the Class, Architecture, Topology and Homologous superfamily (CATH) [4]. CATH uses a semi-automatic approach because it considers manual decision for a particular protein only if the automatic classification is not appropriate.

Also, there are methods that utilize protein sequence alignment to solve the protein classification task. For that purpose, methods like Needleman–Wunch [5], BLAST [6] and PSI-BLAST [7] could be applied. Nevertheless, these methods may not be suitable for proteins with similar structures whose sequences are not so similar. Therefore, it is preferable to align protein structures rather than sequences using methods such as CE [8], MAMMOTH [9] and DALI [10]. Additionally, some methods [11], [12] combine sequence alignment and structure alignment for improved accuracy.

Methods based on feature vector comparison offer a solution to the problem of long classification times using alignment. These methods extract the most relevant characteristics of protein sequences or structures and use them to create a prediction model. Feature vectors can be extracted from protein sequences [13] or structures [14], containing the most informative features for the analysis. The use of these feature vectors reduces the amount of data that's been processed, speeding up the creation of the classification model and reducing the time required for

testing for novel proteins. Different machine-learning algorithms can be used to create the classification model from the extracted feature vectors.

The goal of this research is to develop a fast and accurate method for protein classification. Our attention is on the last category of methods that utilize feature vectors for representing the key characteristics of protein structures. Our previous studies have concentrated on the extraction of feature vectors for protein tertiary structures. In [15], we conducted a comparative analysis of different methods for comparing protein structures. These methods are for solving protein retrieval task, where for a given target protein the most similar proteins from the database are identified. They can also be employed for solving the task of protein classification based on the extracted feature vectors. In this paper, we present a method for solving protein classification task. Our approach involves extracting feature vectors, and to accomplish this, we have selected the protein ray-based descriptor as it has demonstrated high accuracy, compactness, and ease of extraction in our previous research [15]. This descriptor captures information about the geometrical characteristics of the protein structure, specifically how the protein backbone is oriented in space relative to the center of mass. Once the feature vectors have been extracted, we then employ a deep learning architecture to generate classification models. Besides the standard form (variant) of the protein ray-based descriptor, additionally we employed three other variants of the protein ray-based descriptor that we previously used in [16], where several well-known classification methods were used for protein classification.

The rest of this paper is organized as follows. Section 2 provides an overview of the method that is used in this research. We outline the process for extracting the protein ray-based descriptor and present the deep learning architecture used for generating prediction models. We also present the four variants of the protein ray-based descriptor that are considered in this study. In Section 3, we present the experimental results obtained using various neural network settings and we discuss the effect of these settings. Section 4 gives conclusion of the paper and suggests future directions for enhancing this research.

II. The Method Used in This Study

In this paper, we present a method that solves the task for classifying proteins by considering only their tertiary structure. The method comprises two main steps. First, we extract feature vectors by employing our protein ray-based descriptor [15]. Besides its standard form, we also use three additional variants that are obtained from the standard form of the protein ray-based descriptor. Subsequently, we employ deep learning architecture to create a classification model for classification of new proteins. For that purpose, we employ fully connected neural network by using several hidden layers.

The training phase elucidates the process of generating the prediction model. In the testing phase, for a particular test (query) protein the respective class is determined. Both training and testing protein data are preserved in the corresponding PDB files, which are taken from the PDB database [2]. These files contain information regarding the primary, secondary, and tertiary structure of the proteins. In this study, the focus is on the tertiary structure of proteins, thus we inspect the proteins' geometry.

The training phase begins with the extraction of feature vectors for the training proteins. These feature vectors are utilized as samples in the process that is employed in the succeeding stage where the model is generated. Next, the model is created by using deep learning architecture where the weights in the neural network model are adjusted.

Upon creation of the model, it is possible to classify novel proteins. For the query protein, the feature vector is extracted in the same manner as the training proteins. The extracted descriptor is then presented to the model that generates corresponding output (class decision) for that protein. In this research, the classes relate to the SCOP domains used by the SCOP method. We consider only the domain level from the SCOP hierarchy as a level enabling the distinction between the proteins based on their functions. According to this, the query protein is classified into a corresponding SCOP domain.

A. Protein Ray-based Descriptor

Proteins are made of multiple chains, which are folded into specific 3D compositions. The SCOP database [3] provides information about the SCOP domains for the protein chains, so the protein chains are the samples within the dataset that is used. Each protein chain is comprised of amino acid residues, which are connected to form the protein backbone. The amino acid residues within each chain are folded in a particular way. The amino acid residues within each chain consist of various atoms. The Cα atoms within the amino acids connect two consecutive amino acid residues and form the protein backbone.

In a previous study, as documented in [15], we explored various methods for protein structure retrieval. While some of these methods take into account all atoms, the others are focused solely on the Cα atoms. As per the findings presented in [15], the accuracy of the protein structure retrieval process was observed to decline with the inclusion of the remaining atoms. It was concluded that considering only the Cα atoms is the more suitable approach. The protein ray-based descriptor was applied to extract information concerning the 3D coordinates of the Cα atoms, and as a result, the confirmation of the protein backbone is presented. These findings facilitated the creation of a 3D model for the protein, where the protein backbone was defined as a 3D object that occupies a position in the 3D space.

The 3D model created for the protein is scaled ensuring that the Euclidean distance between the most distant Cα atom and the center of mass is 1. With that, we provide scale invariance of the feature vectors. In 3D object retrieval, it is important to provide not only scale invariance but also invariance to translation and rotation. This implies that the same feature vector should be extracted for a given protein chain, regardless of any translations or rotations that are performed. The approach

by which this descriptor is extracted guarantees these properties.

Another challenge in protein structure retrieval is representing protein chains with feature vectors of the same length, considering their differing numbers of Cα atoms. To address this, the protein backbone is interpolated with a fixed number of interpolation points, ensuring that the same number of interpolation points is utilized for each protein chain, regardless of the number of Cα atoms. In [15], we employed two approaches for interpolating the protein backbone. The first approach entailed uniformly interpolating the backbone with interpolation points equidistantly spaced along it, while the second approach involved using more interpolation points in parts of the backbone where consecutive Cα atoms are spaced farther apart. The findings reported in [15] indicate that the feature vectors extracted using uniform interpolation are more accurate. Hence, in this study, we utilize the uniform interpolation of the protein backbone.

The uniform interpolation of the backbone of a given protein involves the initial step of determining the length of the backbone, which is achieved by applying Eq. (1).

$$L = \sum_{i=1}^{N_\alpha - 1} d_{Euclidean}(i, i+1) \tag{1}$$

Here, $d_{Euclidean}(i,i+1)$ represents the Euclidean distance between the i-th and (i+1)-th Cα atoms, and N_α denotes the total number of Cα atoms in the protein chain being analyzed.

In our previous work presented in [15], we conducted an analysis to identify the optimal number of interpolation points for uniform interpolation of the protein backbone. The findings indicated that the most effective number of interpolation points is 64. It was observed that increasing the number of interpolation points above this value did not significantly enhance retrieval accuracy, while using a lower number of interpolation points resulted in a notable decrease in performance. Consequently, in this paper, we have employed 64 interpolation points for uniform interpolation. The interpolation points are uniformly spaced along the curve of the protein backbone, with a distance between two consecutive points equal to $L/(N\text{-}1)$, where L denotes the length of the backbone calculated using Eq. (1), while N denotes the number of interpolation points (64 in this study).

Upon determining the interpolation points, the subsequent step involves extracting the feature vectors. The feature vector, as its name implies, draws inspiration from the ray descriptor [17], which was initially proposed for 3D objects retrieval. The name indicates that rays are "emitted" from the center of mass towards the points that represent the object, which in our case are the interpolation points. The features are obtained by calculating the Euclidean distances between the center of mass and the points that represents the object. This way, the feature vector provides invariance to translation and rotation. The protein ray-based descriptor obtained for a given protein chain elucidates how the backbone of the inspected protein chain is positioned in the space relative to the center of mass.

B. Four Approaches for Extraction of the Protein Ray-Based Descriptor

In the standard form of the protein ray-based descriptor described above, each element of the feature vector pertains to an individual interpolation point. In essence, the protein ray-based descriptor aims to elucidate the traversal of the protein backbone from one concentric sphere to another if we divide the 3D space with concentric spheres. In [18], it is illustrated visually how the protein backbone traverses between these concentric spheres as we go along the backbone.

In [16], we utilized three additional variants of the feature vector obtained based on the Euclidean distances for the interpolation points. In the standard form (standard variant) of the protein ray-based descriptor [15], the feature vector $f_{Eucl} = [f_1, f_2, \ldots, f_N]$ contains the Euclidean distances from the interpolation points towards the center of mass. For example, the feature for the i-th interpolation is $f_i = D_i$, for i=1, 2,…, N, where D_i is the Euclidean distance from the i-th interpolation point to the center of mass. In the other three variant, we consider pairs of consecutive interpolation points. With the second variant, we calculate the difference between the Euclidean distances for two consecutive interpolation points, so the i-th feature is calculated as $diff_i = f_i - f_{i+1} = D_i - D_{i+1}$, i=1, 2,…, N–1. The remaining two variant considers only the magnitude of this change (difference), or only the sign of this change. In the third variant, the i-th feature is calculated as $abs_i = |diff_i|$, i=1, 2,…, N–1. The fourth variant examines whether backbone goes towards the surface or towards the center of mass as we traverse along the backbone without reflecting the quantity of the increase or decrease of the Euclidean distance. With this variant, the i-th feature is calculated as $sign_i = sign(diff_i)$, i=1, 2,…, N–1, where the function $sign(x)$ returns 1 for $x > 0$, 0 for $x = 0$ and -1 for $x < 0$.

C. Deep Learning Models

This study is concerned with the development of classification models through the application of deep learning. Specifically, a fully connected neural network was utilized. The hidden layers are dense layers, wherein each neuron is interconnected with the neurons from the preceding and succeeding layers. This architecture enables a heightened degree of connectivity, resulting in a more intricate representation of the input data, and thereby leading to enhanced performance of the neural network.

The input layer of our neural network model consists of 64 neurons, corresponding to the length of the feature vectors. Each neuron in the hidden layers is activated using the rectified linear unit (ReLU) activation function. The output layer is composed of 150 neurons, with each neuron corresponding to one of the classes, namely SCOP domains in this context. The softmax activation function is employed in the output layer. The optimization algorithm used in this study is stochastic gradient descent (SGD), which is utilized for optimizing the objective

function. Moreover, SGD is also employed as a bias updater. The Adam optimizer is employed as an updater in this study and the learning rate equals 0.001.

In [19], we used the same deep learning architecture to build prediction models. However, in this study besides the standard form of the protein ray-based descriptor we also use the three remaining variants of this descriptor that were described before.

Based on the results obtained in [19], we used the best settings for the number of hidden layers and number of neurons per hidden layer. In the experiments we used 3 hidden layers with 100 neurons per layer. We trained the models for 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 epochs. In this way we wanted to find out if the models are underfitted or overfitted.

WekaDeeplearning4j package [20] was used in this study that utilizes the Deeplearning4j Java library. We used the default setting for the rest of the parameters.

III. Results and Discusion

To evaluate the efficacy of the approach used in this study, we employed data from the PDB database [2] regarding the confirmation of the protein backbone presented by the 3D coordinates data of the Cα atoms. As for the class labels, we utilized knowledge from the SCOP database [3]. It is worth noting that, in this study, we considered only the domain level, thereby establishing a correspondence between the domains of the analyzed protein chains and the classes to which they would be assigned.

The dataset used in this study comprises 6145 protein chains, evenly distributed across 150 SCOP domains, which serve as the potential classes for classification. To evaluate the proposed method, this dataset is split into training and testing sets where 90% of the samples are training data and 10% are test data, ensuring that the uniform distribution of the classes is maintained in both subsets. The resulting training set consists of 5531 protein chains, while the remaining 614 protein chains are used for testing the accuracy of the classification models.

In this section, we will present the results regarding the classification accuracy achieved with the classification models. Given that the test set is balanced, classification accuracy is suitable to measure the predictive performance of the models.

The results given in Table 1 present the classification accuracy of the models. The results revealed that by using the original form of the protein ray-based descriptor, as the number of epochs is increased, the accuracy also increases slightly, but when training for more than 80 epochs, the classification accuracy decreases that indicates that the model is overfitted. The best model is obtained when training the network for 80 epochs. However, with the other three variants, the same pattern is not shown. With the remaining three variants, the best model is obtained when training for 80, 100 and 20 epochs respectively. The best models obtained with the second and third variant achieved the same value for classification accuracy. The fourth variant has slightly lower accuracy (around 1% lower) due to the fact that only the sign of the change (difference) is considered. Although the original version of the protein ray-based descriptor outperforms the other models for almost 0.5% and 1.5%, it encapsulates more info, although the descriptor is longer just for one element. When comparing the second, third and fourth variant, we can note that using more info (magnitude and sign of the change with the second variant), the model is tuned for less epochs (80 in this case for the best model), while when the sign is not considered (with the third variant) a little bit more training is needed (the best model is obtained for 100 epochs). The fourth variant holds least info (only the sign of the change), but it still can provide accurate retrieval comparable to the other variants. Since the dataset is simpler in this case, there is a need for lower number of epochs. The best model in this case (with the fourth variant) is obtained for 20 epochs, and by training further, the model is overfitted, and more significant drop in the accuracy is observed.

In our earlier research [15], we conducted an analysis in which various approaches for comparing proteins were assessed in respect with DALI and CE methods. The results of this analysis demonstrated that, despite its simplicity and speed, the protein ray-based descriptor yields precise predictions of similar proteins and is competitive with time-intensive state-of-the-art methods.

In our study, we conducted an analysis to compare the performance of the models created in this research with those from our previous study [16], which utilized different classification methods and the protein ray-based descriptor. The results of this comparison are presented in Table 2. Our analysis revealed that the best model obtained in this study outperforms the majority of the models obtained in [16], with the exception of knn. However, it is worth noting that the testing time with knn is higher. If larger dataset is used, the difference in testing

TABLE I. The Results Obtained Using Different Number of Epochs

Number of epochs	Euclidean distance	Diff	Abs diff	Sing diff
10	97.07	95.77	95.60	94.63
20	97.39	96.91	94.79	**96.25**
30	97.23	97.07	95.77	95.44
40	97.23	96.91	96.42	95.77
50	97.23	96.42	95.77	95.60
60	97.23	96.74	96.91	94.79
70	97.56	95.60	96.42	94.79
80	**97.72**	**97.23**	96.58	96.09
90	97.56	96.74	96.42	95.77
100	96.91	97.07	**97.23**	94.95

time would be even more significant, as knn is an instance-based learning classifier where model is not created, rather it makes decision based on the similarity between the query and the training samples.

IV. CONCLUSION

In this research paper, we presented a two-steps approach that could be used to classify novel protein structures utilizing the data regarding their tertiary structure. Firstly, we extract the protein ray-based descriptors for the training protein chains. Besides the standard form of the protein ray-based descriptor, we also used three other variants where instead of the Euclidean distances between the interpolation points and center of mass, the change of the Euclidean distance was considered. Particularly, we analyzed the difference between the Euclidean distances for two consecutive interpolation points, as well as the absolute value and the sign of this difference. In this way, we examined what is the best variant to represent the confirmation of the protein backbone in the 3D space. After extracting the feature vectors, then the second step utilizes a deep learning architecture to create classification models. We utilized fully connected neural network models with 3 hidden layers, each containing 100 neurons, as these settings showed best performance in our former study. The evaluation is based on knowledge obtained from the SCOP database. As evaluation measure, the classification accuracy was used as it is appropriate when using balanced dataset. As standard of truth, knowledge about the belonging to SCOP domains from the SCOP hierarchy was used.

In this study we performed experiments by training the model using different setting for the number of epochs. The best model is obtained with the first variant (the original form of the protein ray-based descriptor) when training the network for 80 epochs. The accuracy increased by increasing the number of epochs up to 80, while when training further the model is overfitted. The other three variants showed other patterns, and they lead to the best model for 80, 100 and 20 epochs respectively. The accuracy with the fourth variant is slightly lower (around 1% lower) since it considers only whether the backbone goes towards the surface or center of mass (presented with the sign of the change). The first variant (the original version of the descriptor) slightly outperforms the others since it considers more data. It is interesting to note that this variant is longer just for one element compared to the others. When comparing the three other variants we can make conclusion that using both the magnitude and sign of the change (the second variant), the model is fine-tuned earlier (for less epochs) compared with the case when the sign is not considered (third variant). The fourth variant considers info only for the sign of the change but is comparable to the other variants. With this variant, the data are simpler, so less training is needed, leading to best performance when training the model for 20 epochs. If the model is trained further, the model becomes overfitted. Although this variant has lowest accuracy (less than 1.5% lower compared to the first variant), it is comparable to the others, and especially it is worth due to the faster training of the network. This would be more important if larger dataset is used for training where more time would be needed for training, so for less epochs an appropriate model would be obtained using this variant.

TABLE II. THE RESULTS OBTAINED BY UTILIZING DIFFERENT CLASSIFICATION METHODS

Classification model	**Classification accuracy**
This study	97.720
C4.5	92.997
Naïve Bayes	94.625
Bayes Net	96.417
knn	98.534
SVM	97.557

As regards potential improvements, there exist various possible directions that could be taken. Aside from the protein ray-based descriptor, other feature vectors could also be employed. In relation to the neural network model, other parameters such as activation function and optimization algorithm could be assessed. Also, the other settings could be examined in order to create more accurate model. Additionally, beyond the fully connected neural network, other deep learning architectures could be explored. It is also intended to apply deep learning architectures that exploits the fuzzy logic. Besides deep learning architectures, also other approaches based on other classification methods could be employed, taking into consideration algorithms that are established on classical sets and fuzzy sets.

ACKNOWLEDGMENT

This work was partially financed by the Faculty of computer science and engineering at the “Ss. Cyril and Methodius University in Skopje”, Skopje, Macedonia.

REFERENCES

[1] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The protein data bank,” Nucleic Acids Res., vol. 28, no. 1, pp. 235–242, January 2000.

[2] RCSB Protein Data Bank, http://www.rcsb.org, 2019.

[3] A. G. Murzin, S. E. Brenner, T. Hubbard, and C. Chothia, “Scop: a structural classification of proteins database for the investigation of sequences and structures,” J. Mol. Biol., vol. 247, no. 4, pp. 536–540, April 1995.

[4] C. A. Orengo, A. D. Michie, D. T. Jones, M. B. Swindells, and J. M. Thornton, “CATH – a hierarchic classification of protein domain structures,” Structure, vol. 5, no. 8, pp. 1093–1108, August 1997.

[5] S. B. Needleman and C. D. Wunsch, “A general method applicable to the search for similarities in the amino acid sequence of two proteins,” J. Mol. Biol., vol. 48, no. 3, pp. 443–453, March 1970.

[6] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool,” J. Mol. Biol., vol. 215, no. 3, pp. 403–410, October 1990.

[7] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs,” Nucleic Acids Res., vol. 25, no. 17, pp. 3389–3402, September 1997.

[8] I. N. Shindyalov and P. E. Bourne, “Protein structure alignment by incremental combinatorial extension (CE) of the optimal path,” Protein Eng., vol. 11, no. 9, pp. 739–747, September 1998.

[9] A. R. Ortiz, C. E. Strauss, and O. Olmea, “Mammoth: an automated method for model comparison,” Protein Sci., vol. 11, no. 11, pp. 2606–2621, November 2002.

[10] L. Holm and C. Sander, “Protein structure comparison by alignment of distance matrices,” J. Mol. Biol., vol. 233, no. 1, pp. 123–138, September 1993.

[11] S. Cheek, Y. Qi, S. S. Krishna, L. N. Kinch, and N. V. Grishin, “SCOPmap: automated assignment of protein structures to evolutionary superfamilies,” BMC Bioinformatics, vol. 5, pp. 197–221, December 2004.

[12] C. H. Tung and J. M. Yang, “fastSCOP: a fast web server for recognizing protein structural domains and SCOP superfamilies,” Nucleic Acids Res., vol. 35, W438–W443, July 2007.

[13] K. Marsolo, S. Parthasarathy, and C. Ding, “A multi-level approach to SCOP fold recognition,” IEEE Symposium on Bioinformatics and Bioeng., pp. 57–64, October 2005.

[14] P. H. Chi, Efficient protein tertiary structure retrievals and classifications using content based comparison algorithms, PhD thesis, University of Missouri-Columbia, 2007.

[15] G. Mirceva, I. Cingovska, Z. Dimov, and D. Davcev, “Efficient approaches for retrieving protein tertiary structures,” IEEE/ACM Trans. Comput. Biol. Bioinform., vol. 9, no. 4, pp. 1166–1179, July/August 2012.

[16] G. Mirceva and A. Kulakov, “Protein classification by using four approaches for extraction of the protein ray-based descriptor,” CIIT 2020, Macedonia, 2020.

[17] D. V. Vranic, 3D Model Retrieval, Ph.D. Thesis, University of Leipzig, 2004.

[18] G. Mirceva, M. Mirchev, and D. Davcev, “Hidden Markov Models for classifying protein secondary and tertiary structures,” Journal of Convergence, vol. 1, no. 1, pp. 57–64, 2010.

[19] G. Mirceva, A. Naumoski, and A. Kulakov, “Classification of Protein Structures Using Deep Learning Models,” 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO 2022), pp. 991-996, Opatija, Croatia, 2022.

[20] S. Lang, F. Bravo-Marquez, C. Beckham, M. Hall, and E. Frank, “WekaDeeplearning4j: a Deep Learning Package for Weka based on DeepLearning4j,” Knowledge-Based Systems, vol. 178, no. 15, pp. 48-50, August 2019.

Enabling Hyperparameter-Tuning of AI Models for Healthcare using the CoE RAISE Unique AI Framework for HPC

M. Riedel*†, C. Barakat*†, S. Fritsch†, M. Aach†, J. Busch†, A. Lintermann†, A. Schuppert‡, S. Brynjólfsson*, H. Neukirchen*, M. Book*

* School of Engineering and Natural Sciences, University of Iceland, Iceland
† Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany
‡ Joint Research Centre for Computational Biomedicine, University Hospital RWTH Aachen, Germany
c.barakat@fz-juelich.de, morris@hi.is, s.fritsch@fz-juelich.de, m.aach@fz-juelich.de, j.busch@fz-juelich.de, a.lintermann@fz-juelich.de, schuppert@aices.rwth-aachen.de, sb@hi.is, helmut@hi.is, book@hi.is

***Abstract*—The European Center of Excellence in Exascale Computing "Research on AI- and Simulation-Based Engineering at Exascale" (CoE RAISE) is a project funded by the European Commission. One of its central goals is to develop a Unique AI Framework (UAIF) that simplifies the development of AI models on cutting-edge supercomputers. However, those supercomputers' High-Performance Computing (HPC) environments require the knowledge of many low-level modules that all need to work together in different software versions (e.g., TensorFlow, Python, NCCL, PyTorch) and various concrete supercomputer hardware deployments (e.g., JUWELS, JURECA, DEEP, JUPITER and other EuroHPC Joint Undertaking HPC resources). This paper will describe our analyzed complex challenges for AI researchers using those environments and explain how to overcome them using the UAIF. In addition, it will show the benefits of using the UAIF hypertuning capability to make AI models better (i.e., better parameters) and faster by using HPC. Also, to demonstrate that the UAIF approach is indeed simple, we describe the adoption of selected UAIF building blocks by healthcare applications. The examples include AI models for the Acute Respiratory Distress Syndrome (ARDS). Finally, we highlight other AI models of use cases that co-designed the UAIF.**

***Keywords*—High-Performance Computing; Software Framework; Machine Learning; Deep Learning; Quantum Computing**

I. Introduction

Artificial Intelligence (AI) is a technology that is still relatively new in exploiting High-Performance Computing (HPC) systems for a wide variety of application domains compared to traditional simulation sciences using numerical methods based on known physical laws. Therefore, the European Centre of Excellence - Research on AI- and Simulation-Based Engineering at Exascale (CoE RAISE)[1] develops a Unique AI Framework (UAIF) to provide seamless solutions for a wide variety of complex AI applications. The timespan of CoE RAISE is from 2020 until the end of 2023 whereby several EuroHPC JU National Competence Centers (NCCs) have already agreed to maintain the UAIF beyond its project lifetime. Several NCCs (e.g., Germany, Iceland) use parts of UAIF already with different use cases while we highlight one particular use case in this paper. The UAIF aims at the convergence of HPC and innovative AI techniques making it easier for AI researchers to make use of cutting-edge hardware infrastructures. In addition to HPC hardware infrastructures, a seamlessly usable and versatile Software Infrastructure is critical for accelerating convergence through new AI toolsets that are ready to scale for enormous quantities of datasets. CoE RAISE considers AI requirements of Computing-driven Use Cases using numerical methods based on known physical laws. Use case examples include AI used in Computational Fluid Dynamics (CFD) problems (i.e., wind wheels, turbulence, hydrogen research, coating, etc.). CoE RAISE also addresses AI requirements of Data-driven Use Cases with large datasets of measurement devices. Use case examples include AI used in additive manufacturing (i.e., 3D printing quality assurance), dataset analysis from the Large Hadron Collider (LHC) at CERN, and remote sensing data analysis. The UAIF methodologies are co-designed by the above use cases to ensure its usage by a wide variety of scientific and engineering applications. The co-design activities leverage the Interaction Room Methodology (1, 2) using Mural Boards[2].

HPC environments getting more complex with a massive increase in technology (e.g., hierarchical storage and malleability) and heterogenity (e.g., different GPU vendors). Domain users and AI researchers are often overwhelmed when using HPC resources and cant focus on their specific domain problem. For example, the high amount of AI packages on various HPC machines in different versions and the vast availability of different module environments[3] for those is hard to understand. Healthcare experts such as those addressed in this paper are overwhelmed with these low-level environments and demand a lower barrier to use HPC/AI methods such as hyperparameter optimization (HPO) of AI models.

This work was performed in the Center of Excellence (CoE) Research on AI- and Simulation-Based Engineering at Exascale (RAISE) receiving funding from EU's Horizon 2020 Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733. Icelandic HPC National Competence Center is funded by the EuroCC-1 projects that has received funding from the EU HPC Joint Undertaking (JU) under grant agreement No 951732. Parts of the work have been also supported by the European Digital Innovation Hub (EDIH) of Iceland (EDIH-IS) funded in parts by the Digital Europe Programme. Finally, parts of this work were performed in the SMITH Project receiving funding from the German Federal Ministry of Education and Research.

[1] https://www.coe-raise.eu/

[2] Mural Boards https://mural.co

[3] https://modules.sourceforge.net/

This paper provides a comprehensive overview of the UAIF software layout design that addresses the challenge mentioned above, including major updates since its last publication by M. Riedel et al. (3). Its main contribution is to describe each component of UAIF and the comprehensive application co-design efforts of use cases over two years. One key component of the UAIF that is used in many use cases is the use of HPO tools. Therefore, we also show a short adoption example of the UAIF in healthcare leveraging this popular HPO component. While more healthcare-related AI models leverage HPC as shown in (4), this paper focuses particularly on the adoption of the HPO through the CoE RAISE UAIF.

The remainder of this paper is structured as follows. Section 1 starts with the introduction to CoE RAISE and the approach of co-designing the UAIF with distinct engineering and scientific use cases. A review of related work is given in Section 2, while Section 3 describes in detail the main contribution of this paper that is the result of the CoE RAISE UAIF co-design process. Section 4 then described one adoption of the UAIF components by use cases that were not involved in the original co-design process with a special emphasis on hyperparameter tuning. The paper ends with some concluding remarks.

II. Related Work

Despite the fact that the HPC environments are getting more complex for domain-specific researchers, there is quite a history of tools and platforms that make it easier for researchers to access HPC resources. One example is a HPC-driven Grid middleware Uniform Interface to Computing Resources (UNICORE)[4] (5). It offers a ready-to-run system including client and server software that makes HPC resources available in a seamless and secure way. In contrast to the UAIF, UNICORE is not focused on a streamlined set of components for AI on HPC resources. UNICORE is much more general and operates basically one level above the module environments and schedulers providing REST APIs or GUIs.

Other similar ideas of a framework or collection of tools are existing in the commercial space from key vendors. One example is the Intel oneAPI HPC Tools for Developers[5]. These tools enable developers to build, analyze, optimize, and scale HPC applications across multiple types of architectures more easily using the Intel oneAPI Base Toolkit and Intel oneAPI HPC Toolkit. The tools include approaches such as state-of-the-art techniques in vectorization, multithreading, multinode parallelization, and memory optimization so that one can more easily build software that is HPC-ready. In contrast to UAIF, these tools are not fully open source and the overall design of the oneAPI is not focused on HPC at Exascale but more general use by HPC applications. For the sake of completeness, several solutions also exist in Clouds like in Amazon Web Services (AWS) with SageMaker[6]. Those services do not work on EuroHPC JU nor EU HPC systems.

[4]https://www.unicore.eu/

[5]https://www.intel.com/content/www/us/en/high-performance-computing/hpc-software-and-programming.html

[6]https://aws.amazon.com/sagemaker/

III. CoE RAISE UAIF Co-Design Layout Results

This section provides a comprehensive overview of the framework software layout after the final co-design phase of CoE RAISE, including major updates since the last publication by Riedel et al. in (3). It briefly describes each component of the framework. Fig. 1 includes several components that are not directly under the control of CoE RAISE, but are identified as important dependencies for the UAIF running on HPC systems. For releases of a wide variety of UAIF components, we refer to the CoE RAISE Web page.

A. Applications

The application layer contains different use cases that contributed to the co-design of the Unique AI Framework (UAIF), but also initiatives that might adopt the UAIF. In Fig. 1, the large red arrow represents the co-design activities that influence the reference architecture components. The large green arrows represent the benefits and adoption potential for external project activities in the larger HPC ecosystem.

Component (A) in Fig. 1 represents the co-design efforts of the UAIF based on compute- and data-intensive use cases[7]. Hence, several different use case types contributed to benchmarking and proof of scalability of components of the UAIF on various HPC systems. During the project, and especially in the last few months, a clear picture is provided of what components are relevant for the UAIF.

Also, a wide variety of CoEs[8] have been funded in different domain-specific areas providing use cases that leverage simulation sciences or AI/HPC methods to utilize the emerging Exascale computing. As shown in Fig. 1 (B), UAIF adoption is recommended to CoEs as a way to prevent AI developers in domain-specific sciences from wasting a lot of effort in configuring the correct AI setup on HPC.

In addition, National Competence Centers (NCCs) has been created under the EuroCC-1 and EuroCC-2 project umbrella to enable industry and Small and Medium Enterprises (SMEs) to leverage HPC resources of EuroHPC[9]. Component (C) of Fig. 1 represent adoptions of the UAIF by NCCs and the significant potential to governmental, academic, industry, and SME partners to speed-up and scale-up their applications.

Finally, Digital Twins (DTs) and their workflows using HPC, e.g., in the Destination Earth[10] or Inter-Twin[11] projects, are becoming important for HPC users in Europe. Component (D) has been added to Fig. 1 to represent the processing-intensive applications of DTs that are also highly relevant for CoE RAISE, either the DTs adopting parts of UAIF components or including new use cases in CoE RAISE.

B. Reference Architecture Elements

This section describes the reference architecture components of the UAIF for Exascale HPC/AI methods, which are

[7]CoE RAISE Use Cases https://www.coe-raise.eu/use-cases

[8]EU HPC Centres of Excellence https://www.hpccoe.eu/eu-hpc-centres-of-excellence2/

[9]EuroHPC JU HPC Systems https://eurohpc-ju.europa.eu/about/our-supercomputers_en

[10]Destination Earth https://digital-strategy.ec.europa.eu/en/policies/destination-earth

[11]InterTwin https://www.intertwin.eu

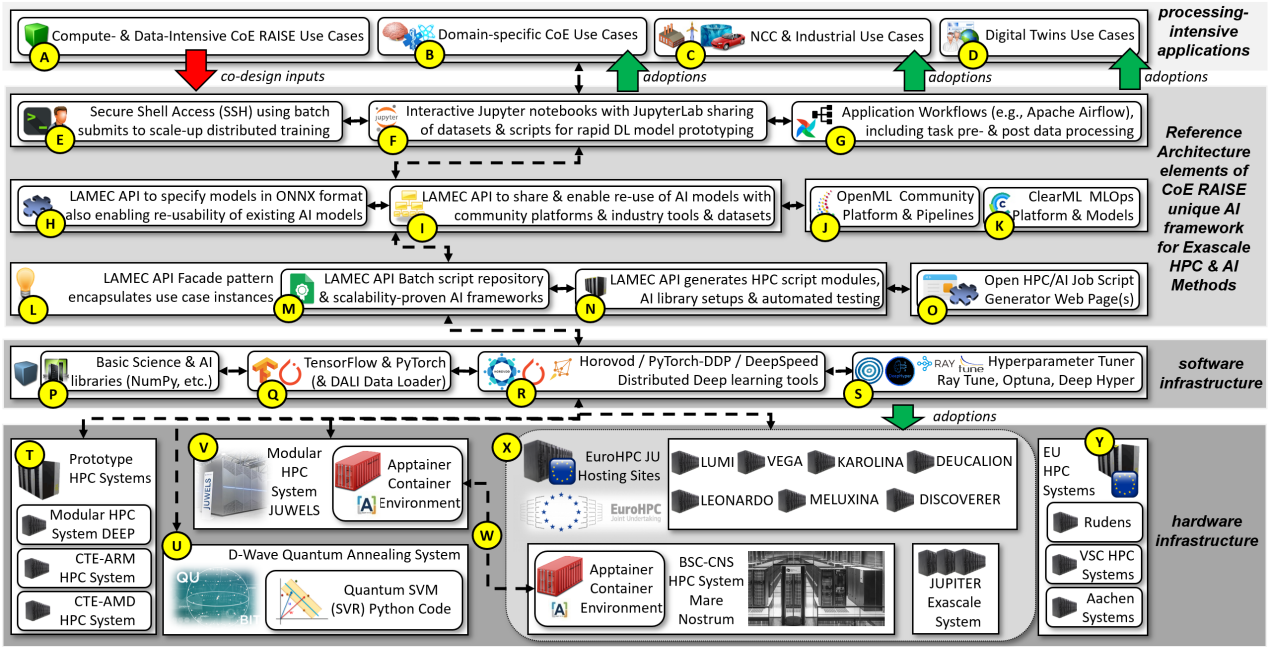


Figure 1. Application co-design result creating a unique AI software framework (UAIF) for HPC towards Exascale.

listed in Fig. 1 in the second layer. As shown in Fig. 1 (E), the first element includes the use of the Secure Shell (SSH) protocol. Principally, as a means to remotely log into HPC systems and submit batch scheduler scripts, e.g., via the Simple Linux Utility for Resource Management (SLURM) (6) tool, it remains one of the integral access methods for HPC applications. Also, AI researchers frequently require interactive access to HPC systems to facilitate quick and rapid prototyping of Machine Learning (ML) and DL models. Component (F) in Fig. 1 addresses this need in the UAIF by offering Jupyter notebooks and JupyterLabs[12]. The component (G) in Fig. 1 is a new addition to the UAIF and supports application workflows and workflow automation, including task pre- and post- data processing capabilities. The UAIF recommends the Apache Airflow[13] tool that is a platform to programmatically create, schedule, and monitor workflows.

As shown in Fig. 1 (H), a key component of the overall UAIF Load AI Modules, Environments, and Containers (LAMEC) API is the fast portability between different DL frameworks and reproducibility achieved by using the standard Open Neural Network Exchange (ONNX). The implementation of the overall UAIF LAMEC API using ONNX is still a work in progress. Another element of the overall UAIF LAMEC API shown in Fig. 1 is component (I), which represents a seamless integration with other tools. The goal is to use the LAMEC API to share and re-use existing AI models with community platforms (see below), industry tools, datasets, and to enable Transfer Learning (TL). While initial discussions with community platforms have taken place, the implementation of the overall UAIF LAMEC API integration and provisioning of AI models is still work in progress.

OpenML[14] is an open community platform for sharing datasets, algorithms, models, and experiments using a wide variety of traditional ML approaches. One ansatz to enlarge the user community of UAIF is to integrate its components into the OpenML platform such that experiments can be also run on cutting-edge HPC systems. Hence, component (J) in Fig. 1 represents how this community might leverage the LAMEC API integration components. On the other hand, ClearML[15] is an ML Operations (MLOps) platform that can be used to develop, orchestrate, and automate ML workflows at scale. CoE RAISE provides one installation of ClearML for its internal and external users. Hence, another approach to enlarge the user community of UAIF is to integrate components into MLOps platforms such as ClearML that are often used in the industry such that its tasks can be also run on cutting-edge HPC systems. Hence, the component (K) in Fig. 1 represents how this community might leverage the LAMEC API components through integration with MLOps platforms.

To map the abstract specifications of software and hardware needs by AI researchers to specific software and hardware HPC infrastructure elements, a facade pattern is used by the UAIF LAMEC API general design. Hence, as represented by component (L) in Fig. 1, the UAIF design employs an abstract wrapper functionality that maps the abstract specifications from users to specific software and hardware configurations. The LAMEC API core is split into the two following elements. The first core element is a batch script repository and the second is an API using this repository to generate new batch script elements, and both are described below. Thus the first core element of the UAIF LAMEC API is a batch script repository. It consists of batch scripts for specific HPC systems

[12]Project Jupyter https://jupyter.org/
[13]Apache Airflow https://airflow.apache.org/
[14]OpenML https://www.openml.org/
[15]ClearML https://clear.ml/

with a correct setup of modules needed for using specific UAIF AI tools As described above and represented by component (M) in Fig. 1, one idea is to use this repository with the UAIF LAMEC API as follows with the second core below. But it quickly becomes clear that the repository in itself is also a great resource for AI/HPC researchers that already know how to deal with changing HPC modules in batch scripts. The second core element of the UAIF LAMEC API, which is represented by component (N) in Fig. 1, is using the above-mentioned repository to generate new batch script segments. This lowers the barrier for entry to leveraging HPC systems for AI researchers that may not have much experience working with modules in HPC environments, as well as saving valuable time through automation for experienced users. Additional components beyond verified site AI modules and libraries, such as AI model scripts or datasets for training and inference, are planned for later addition (although these are usually elements of a job script that inexperienced AI researchers do not find challenging). Maintenance of the job script repository is needed to keep it up-to-date, but that is effort is lower than letting all users find the right modules themselves.

Finally, the open HPC/AI job script generator web page(s) shown as component (O) in Fig. 1 also uses the implementation of the UAIF LAMEC API. This concept is derived from existing job script generators available at the Swiss National Supercomputing Centre (CSCS)[16] or the National Energy Research Scientific Computing Center (NERSC)[17], where the difference to these existing tools lies in the use of UAIF AI toolsets.

C. Software Infrastructure

The software infrastructure layer components (P) – (S), which are depicted in Fig. 1. are presented in this section. Despite the increase of DL tools, and their uptake in the AI communities, there remains a core of basic science libraries heavily used by CoE RAISE communities. Examples are NumPy[18] and scikit-learn[19]. In addition, the building block (P) in Fig. 1 of the UAIF also includes simulation science codes, e.g., those using numerical methods based on known physical laws and that have the potential to benefit from coupling to AI models. Since CoE RAISE focuses on AI models, the various relevant simulation science codes have been kept out of the UAIF layout. As shown in Fig. 1 (Q), the UAIF recommends the use of PyTorch[20] and TensorFlow[21]. CoE RAISE has tested their scalability in depth using various applications. NVIDIA Data Loading Library (DALI)[22] further increases the performance of PyTorch and TensorFlow. This inclusion is represented by component (Q) in Fig. 1 in parenthesis due to the proprietary nature with NVIDIA GPUs. CoE RAISE investigates still other GPU vendors such as Advanced Micro Devices (AMD). Component (R) in Fig. 1 outlines three supported tools used for accelerating distributed AI model training by leveraging the large number of GPUs available at HPC sites today. PyTorch-Distributed Data Parallel (DDP)[23] and Horovod[24] are included in the UAIF software layout. More recently, the component also added DeepSpeed. One of the most successful aspects of the current adoptions of the UAIF are Hyperparameter Optimization (HPO) tools represented by component (S) in Fig. 1. The most adopted component is Ray Tune tool[25], and other UAIF components in that context are Optuna[26] and DeepHyper[27]. This paper describes more about the adoption of RayTune in the later adoption section.

D. Hardware Infrastructure

The hardware infrastructure layer components (T) – (Y) depicted in Fig. 1 are presented in this section. The benchmarking and porting activities of Coe RAISE have been performed on a number of interesting prototype HPC systems that feature new and emerging technologies. The Dynamical Exascale Entry Platform (DEEP)[28] system has been used to experiment with the Modular Supercomputing Architecture (MSA) (7, 8) type of HPC architecture. This component (T) in Fig. 1 also includes two new prototype systems, the Advanced Reduced Instruction Set Computer Machine (ARM)-based CTE-ARM and CTE-AMD, hosted at the Barcelona Supercomputing Centre (BSC) in Spain. The CTE-ARM[29] is a supercomputer based on 192 A64FX ARM processors, with a Linux Operating System (OS) and an Tofu interconnect network (6.8GB/s). CTE-AMD[30] is a cluster based on AMD EPYC processors, with a Linux OS and an Infiniband interconnection network. It includes two AMD MI50 GPUs per node. Quantum Computing (QC) is gaining momentum as the EuroHPC JU recently funded, together with national contributions, several QC systems [31]. Multiple use case applications (9, 10) have successfully engaged in QC by utilizing the D-Wave QA system available via the Juelich UNified Infrastructure for Quantum computing (JUNIQ)[32] at JSC in Germany. As represented by component (U) in Fig. 1, the quantum AI models implemented were Support Vector Machines (SVMs). They were used for regression tasks via Support Vector Regression (SVR).

The MSA-based HPC system JUWELS is massively used for co-designing the UAIF and performing necessary speed-up and scaling benchmarks of its components, see component

[16] CSCS job script generator https://user.cscs.ch/access/running/jobscript_generator/
[17] NERSC job script generator https://my.nersc.gov/script_generator.php
[18] NumPy https://numpy.org/
[19] scikit-learn https://scikit-learn.org/stable/
[20] PyTorch https://pytorch.org/
[21] TensorFlow https://www.tensorflow.org/
[22] DALI https://developer.nvidia.com/dali
[23] PyTorch Distributed Data Parallel https://pytorch.org/tutorials/beginner/dist_overview.html
[24] Horovod https://github.com/horovod/horovod
[25] Ray Tune https://www.ray.io/ray-tunel
[26] Optuna https://optuna.org/l
[27] DeepHyper https://deephyper.readthedocs.io/en/latest/l
[28] DEEP Prototype HPC System hosted by JSC https://www.fz-juelich.de/en/ias/jsc/systems/prototype-systems/deep_system
[29] CTE-ARM HPC System https://www.bsc.es/innovation-and-services/technical-information-cte-arm
[30] CTE-AMD HPC System https://www.bsc.es/innovation-and-services/technical-information-cte-amd
[31] EuroHPC JU Quantum Computers https://eurohpc-ju.europa.eu/selection-six-sites-host-first-european-quantum-computers-2022-10-04_en
[32] JUNIQ https://www.fz-juelich.de/en/ias/jsc/systems/quantum-computing/juniq-facility

(V) in Fig. 1. It is an ideal HPC system for AI workloads as described by Kesselheim et al. in (11). Container technologies are an important tool within larger AI communities to facilitate porting of applications and datasets between systems. One example is shown as component (W) in Fig. 1, where the porting operation of a containerized application from JUWELS at JSC to the MARE NOSTRUM 4 system at BSC is shown. This transparent deployment of containerized code is made possible by the support of Apptainer[33] (previously named Singularity) available at both sites (e.g., see container runtime on JUWELS[34]). Initial test have been performed with containers on HPC platforms at scale, but more application use cases are still work in progress. This component of the UAIF is crucial to support industry applications that have not used HPC before.

Component (X) in Fig. 1 covers the EuroHPC JU hosting sites[35] that may adopt the UAIF. Several European HPC systems contributed to co-design with applications to the UAIF design. It is the goal of the UAIF developer community to support many EuroHPC JU systems. Initial discussions with some of these sites have been started by CoE RAISE partners to encourage the adoption of the UAIF. They reveal however that many of the components of the UAIF are already partly adopted by EuroHPC JU hosting sites. The broader adoption strategy is in its initial stages, while components such as the LAMEC API are considered to be further developed adding more EuroHPC JU systems support over time. One highlight of the adoption will be the integration to the first European Exascale system JUPITER[36], which will be installed in 2024.

It is observed that new users of the UAIF are often starting using regional or university-level systems before scaling up to larger systems. Component (Y) in Fig. 1 contains examples such as the university-level systems Rudens[37] of the Riga Technical University (RTU), or the HPC systems of Rheinisch-Wesfälische Technische Hochschule Aachen - RWTH Aachen University (RWTH)[38]. These sites are in the process of adopting parts of the UAIF framework through users in CoE RAISE. Another example are Belgium regional HPC systems such as the Vlaams Supercomputer Centre (VSC)[39] that are in use by CoE RAISE. There is a wide variety of other HPC systems, such as industrial systems in Iceland (e.g., Responsible Compute[40]) that are not shown in Fig. 1 (Y), but are in progress to adopt elements of the UAIF.

IV. Framework Adoption Example in Healthcare

Critically ill patients who require treatment on an intensive care unit (ICU) in hospitals are at high risk of developing respiratory diseases. Therefore, the Simulation and Data Lab (SDL) Health and Medicine[41] of the Icelandic National Competence Center (NCC) for HPC and AI performs research in that area with a special emphasis on leveraging cutting-edge HPC systems. One of the main research areas of the SDL, in cooperation with the Smart Medical Information Technology for Healthcare (SMITH)[42] project, is Acute Respiratory Distress Syndrome (ARDS), a condition that was first described by Ashbaugh et al. (12) and which has a high mortality rate among affected ICU patients due to its heterogeneity (13).

This section describes short insights on the evolution of using HPC for that research based on earlier work by C. Barakat in (4) with an emphasis on adopting the UAIF for HPO (i.e., component S in Fig. 1). More recently, the SDL Health and Medicine developed and improved a deep learning-based surrogate model of one tool for modeling ARDS onset in a virtual patient called the Nottingham Physiology Simulator (NPS) (14). The model development process takes advantage of current ML and data analysis techniques, as well as efficient HPO methods using the UAIF deployment of the DEEP MSA system at JSC (i.e., component T in Fig. 1). Also, this healthcare use case adopts the basic science libraries like NumPy (i.e., component P in Fig. 1) and DL tools with distributed training (i.e., component Q and R in Fig. 1).

As shown in Fig. 1 (S), the concrete HPO tool used is Ray Tune (15), which in turn employs different scheduling algorithms in order to simplify the task of finding the optimal hyperparameters for training the final ARDS model. DL model hyperparameters are the variables that affect the way in which a model is built and its training process, and can be altered either through a process of trial and error, or automatically using optimization algorithms. The schedulers used by Ray Tune in the optimization process are HyperBand, Asynchronous HyperBand, Population-Based Training (PBT), and the default First-In, First-Out (FIFO). The whole process of finding hyperparameters and using the aforementioned schedulers and algorithms are computationally expensive and thus HPC is required to perform it. Medical experts in the SDL are not experts in HPC and thus leverage the seamless access of HPO tools on HPC via UAIF components such as Jupyter (i.e., Fig. 1 F) and the LAMEC API (i.e., Fig. 1 N).

In order to achieve results, these algorithms distribute the tuning task over the available HPC resources and may interfere with the process by introducing perturbations, or by shutting down under-performing tasks. The comparison of the different algorithms is enabled to highlight the most efficient in terms of resource use and accuracy of ARDS model results. Figure 2 showcases the performance of the different schedulers. It is clear where the most efficient schedulers, namely HyperBand (Figure 2b) and Asynchronous HyperBand (Figure 2c), begin stopping trials that seem to underperform, thereby reducing unnecessary resource use. On the other hand, PBT (Figure 2d)consumes the most resources, but does so while intro-

[33] Apptainer https://apptainer.org/
[34] JUWELS Container Runtime https://apps.fz-juelich.de/jsc/hps/juwels/container-runtime.html
[35] EuroHPC JU Hosting Sites https://eurohpc-ju.europa.eu/about/our-supercomputers_en
[36] Path to JUPITER https://www.fz-juelich.de/en/ias/jsc/news/news-items/news-flashes/2023/path-to-jupiter
[37] Rudens HPC System https://www.rtu.lv/en/research/science-and-innovation-centre/scientific-equipment-unit/hpc-center
[38] RWTH Aachen University HPC Systems https://help.itc.rwth-aachen.de/en/service/rhr4fjjutttf/
[39] VSC HPC Systems https://www.vscentrum.be/
[40] 52Responsible Compute https://responsiblecompute.com/
[41] https://ihpc.is/simulation-and-data-lab-health-and-medicine/
[42] https://www.smith.care/en/

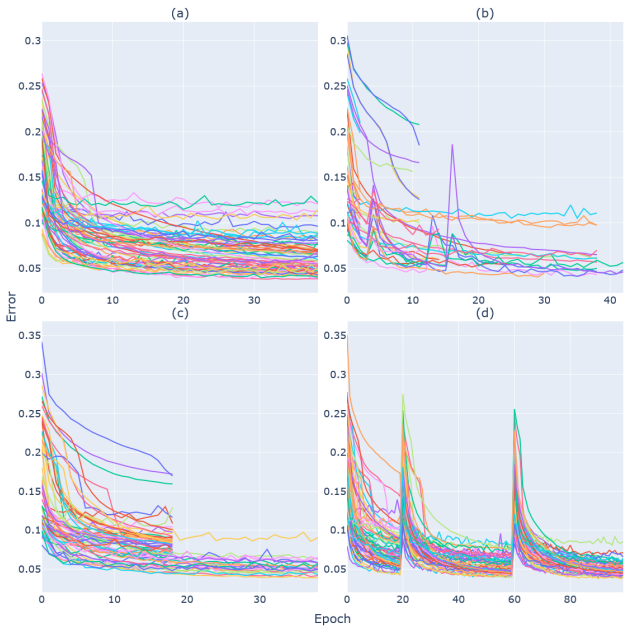


Figure 2. Mean Absolute Error of the ARDS model hyperparameter tuning process using different schedulers.

ducing perturbations to the tuning process, through which more insight into the best parameter combination could be achieved. The tuner is therefore adaptable to different system architectures, where researchers can choose the approach that best suits the available hardware. The UAIF and HPO tools are used to tune parameters like the learning rate, the batch size, the dropout rate, the loss function, and the presence of an intermediate fully-connected layer before the output layer in the network architecture. While concrete medical results are out of scope of this paper, we conclude that the results have been not only faster achieved with UAIF and HPC speed-ups, but also enabled better AI model accuracies through HPO.

V. Conclusion

The implementation process of the UAIF goes forward as planned with respect to scalability tests and benchmarking of UAIF components, but also with the implementation and design of its LAMEC API to make it easier for non-technical users to adopt it. We conclude that the co-design use cases of CoE RAISE for the UAIF helped to be able to adopt the UAIF in other scientific or engineering domains such as healthcare. For the mentioned healthcare use case application of UAIF in ARDS research, the UAIF components enabled the researchers to create better AI models faster than before. While there are many benefits, one drawback of the framework is its maintenance and keeping the repository of job scripts up-to-date which requires some development efforts. Initial discussions with partners in CoE RAISE and NCCs reveal that many want to contribute to an open-source solution with development efforts that keeps the UAIF alive even beyond project lifetime. There is work to be done on the broader deployment of the UAIF across EuroHPC JU hosting sites and other EU HPC resources. Also the adoption of the UAIF in EuroCC NCC industrial use cases, other CoEs, and selected Digitial Twins is work in progress.

References

[1] Book, M. et al, "Facilitating collaboration in high-performance computing projects with an interaction room," in Proceedings of the 4th ACM SIGPLAN International Workshop on Software Engineering for Parallel Systems, 2017, pp. 46–47.

[2] Book, M. et al, "Facilitating collaboration in machine learning and high-performance computing projects with an interaction room," in 2022 IEEE 18th International Conference on e-Science (e-Science). IEEE, 2022, pp. 529–538.

[3] Riedel, M. et al, "Practice and experience using high performance computing and quantum computing to speed-up data science methods in scientific applications," in 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2022, pp. 281–286.

[4] Barakat, C. et al, "Lessons learned on using high-performance computing and data science methods towards understanding the acute respiratory distress syndrome (ards)," in 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2022, pp. 368–373.

[5] Streit, A. et al, "Unicore 6—recent and future advancements," Annals of Telecommunications-annales des Télécommunications, vol. 65, pp. 757–762, 2010.

[6] Yoo, A.B., Jette, M.A. and Grondona, M., "Slurm: Simple linux utility for resource management," in Job Scheduling Strategies for Parallel Processing: 9th International Workshop, JSSPP 2003, Seattle, WA, USA, June 24, 2003. Revised Paper 9. Springer, 2003, pp. 44–60.

[7] Eicker, N. et al, "The deep project an alternative approach to heterogeneous cluster-computing in the many-core era," Concurrency and computation: Practice and Experience, vol. 28, no. 8, pp. 2394–2411, 2016.

[8] Suarez, E., Eicker, N. and Lippert, T., "Modular supercomputing architecture: from idea to production," in Contemporary high performance computing, pp. 223–255. CRC Press, 2019.

[9] Riedel, M., Cavallaro, G. and Benediktsson, J.A., "Practice and experience in using parallel and scalable machine learning in remote sensing from hpc over cloud to quantum computing," in 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, 2021, pp. 1571–1574.

[10] Pasetto, E. et al, "Quantum svr for chlorophyll concentration estimation in water with remote sensing," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022.

[11] Kesselheim, S., Herten, A. et al, "Juwels booster–a supercomputer for large-scale ai research," in High Performance Computing: ISC High Performance Digital 2021 International Workshops, Frankfurt am Main, Germany, June 24–July 2, 2021, Revised Selected Papers 36. Springer, 2021, pp. 453–468.

[12] Ashbaugh, D. et al, "Acute respiratory distress in adults," The Lancet, vol. 290, no. 7511, pp. 319–323, 1967.

[13] Barakat, C. et al, "An HPC-Driven Data Science Platform to Speed-up Time Series Data Analysis of Patients with the Acute Respiratory Distress Syndrome," in 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, pp. 311–316, IEEE.

[14] Hardman, J., Bedforth, N. et al, "A physiology simulator: validation of its respiratory components and its ability to predict the patient's response to changes in mechanical ventilation.," British journal of anaesthesia, vol. 81, no. 3, pp. 327–332, 1998.

[15] Moritz, P. et al, "Ray: A distributed framework for emerging {AI} applications," in 13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18), 2018, pp. 561–577.

CTI

International Conference on

TELECOMMUNICATIONS & INFORMATION

Steering Committee

Chairs: Željka Car, University of Zagreb, Croatia
Stjepan Golubić, Zagreb, Croatia
Dragan Jevtić, University of Zagreb, Croatia
Branko Mikac, University of Zagreb, Croatia

Members: Slaviša Aleksić, Leipzig University of Applied Sciences – HTWK, Germany
Krešo Antonović, Ministry of Sea Transport and Infrastructure, Zagreb, Croatia
Gordan Gledec, University of Zagreb, Croatia
Miran Gosta, Croatian Regulatory Authority for Network Industries, Zagreb, Croatia
Darko Huljenić, Ericsson Nikola Tesla, Zagreb, Croatia
Oliver Jukić, Consulten, Slatina, Croatia
Erich Leitgeb, Graz University of Technology, Austria
Velimir Švedek, Croatian Regulatory Authority for Network Industries, Zagreb, Croatia

Program Committee

Members: Željka Car, University of Zagreb, Croatia
Vlado Delić, University of Novi Sad, Serbia
Saša Dešić, Ericsson Nikola Tesla, Zagreb, Croatia
Renato Filjar, Krapina University of Applied Sciences, Croatia
Tihana Galinac Grbac, Juraj Dobrila University of Pula, Croatia
Stjepan Golubić, Zagreb, Croatia
Vlasta Hudek, Rezultanta, Zagreb, Croatia

Robert Inkret, Ericsson Nikola Tesla, Zagreb, Croatia
Dragan Jevtić, University of Zagreb, Croatia
Admela Jukan, Institute of Computer and Network Engineering, Braunschweig, Germany
Ozren Jureković, VERN' University, Zagreb, Croatia
Marko Lacković, Boston Consulting Group, Singapore, Singapore
Dražen Lučić, Croatian Chamber of Economy, Zagreb, Croatia
Branko Mikac, University of Zagreb, Croatia
Mladen Sokele, Zagreb University of Applied Sciences, Croatia
Dina Šimunić, University of Zagreb, Croatia
Darko Švelec, Zagreb, Croatia
Antonio Teixeira, University of Aveiro, Portugal
Miroslav Vrankić, E-Glas d.o.o., Rijeka, Croatia
Mario Weber, Lamaro Digital, Zagreb, Croatia

PAPERS

NEW INFORMATION AND COMMUNICATION TECHNOLOGIES IN PRACTICE

Smart Work Order

Zoran Bosić, Pavao Gusić, Miroslav Zubčić, Marija Marković, Tomislav Žitnik
Ericsson Nikola Tesla d.d., Zagreb, Nimium d.o.o., Zagreb, Croatia
zoran.bosic@ericsson.com, pavao.gusic@ericsson.com, miroslav.zubcic@nimium.hr, marija.markovic@ericsson.com, tomislav.zitnik@ericsson.com,

Abstract: **Smart work order (SWO) is a technology-driven approach that utilizes augmented reality (AR) smart glasses and knowledge management data base (KMDB) to improve field services (FS). A work order (WO) is a specific job, repair, or any other FS team's physical activity. SWO leverages identified opportunities for automating operations in the field. It makes use of machine learning and artificial intelligence (ML/AI) processes to access knowledge captured in KMDB and generates options for improved hardware repair, installation management, or predictive maintenance. Consequently, the efficiency and sustainability of field services and work orders are raised due to shorter time to deliver, lower travelling costs and less rework required. SWO developed in Ericsson Nikola Tesla Group (ENTG) is configured by augmented reality technology, and machine learning algorithms supported by ITILv4 best practice framework where each SWO part creates added value in the field services chain. Additional integration with operations wisdom logging (OWL) toolset empowers FS team members with advanced knowledge in daily work. All SWO parts can access real-time data which enables FS teams to utilize knowledge management systems to determine first-time right actions. SWO automatically generates all required steps which walk FS teams through required actions, from capturing an alarm, over identifying the problem, finding a solution, and visualizing required actions to confirming the problem has been cleared up.**

Key Words: smart work order, augmented reality, knowledge management, field services, real-time data, smart technology, green technology

I. INTRODUCTION

Smart work order (SWO) has been developed in Ericsson Nikola Testa Group to enable field service operations (FSO) teams to maximize the value and efficiency of their daily work. The mission of field service operations process [1] is to leverage predictive field service management capabilities which enable delivery of field services even before a fault arises.

Field service operations process defines the execution of on-site activities related to predictive, preventive, and corrective maintenance, or deployment of new functionalities in the customers' ICT infrastructure at any of their locations. The main objective of SWO is to utilize augmented reality (AR) capability throughout the operations to empower operators by enhancing their digital efficiency, which supports operations cost reduction, enhanced productivity, and high-class customer experience.

To support smart work order functionalities the Ericsson Nikola Tesla Group's solution integrates two main platforms: augmented reality modeling support (ARMS) and information technology service management (ITSM). ARMS is a mixed reality platform developed in ENTG which uses Microsoft's HoloLens for augmented reality together with virtual reality to deliver powerful tools for daily installation engineering, hardware support, space construction work and learning. ITSM as a complete IT operations and automation tool suite that supports and enables Ericsson's IT managed services business. Ericsson's ITSM tool suit is based on ITIL [2] where ITSM is implementation and management of quality IT services that meet the needs of the business. Ericsson's ITSM uses internally developed process engine for actively moving any incident, problem, change or other request rapidly through the key process steps stipulated by the service level agreement (SLA) according to the ITILv3 best practices framework. The integration of those two platforms ensures end-to-end data and process flow, starting with collecting alarms about issues on field hosted equipment, and identifying faults via monitoring and comparison with already captured known errors. The next crucial step is to find a suitable solution for a fault where ITMS internal engines search through the knowledge database and generate a work order with specified series of steps, such as tracking down equipment location, the execution of fix actions and verifying the fault resolution.

There are two main SWO guiding principles for design: making smart AI/ML driven decisions and sticking to the green technology to reduce the human impact on the environment. SWO solution design supports multiple types of field service operations (FSO) work orders for administrative, corrective, planned or preventive activities. ML/AI concept based on predefined events and user profiles can enable synergy between SWO and Operations Wisdom Logging (OWL) [3], enabling additional value to customer experience during node or equipment maintenance. OWL toolset is based on ITIL best practices and integration with ITSM key components, such as configuration management data base (CMDB) and KMDB. All relevant information can be extracted from already available configuration, the known error and knowledge management database or provided via ARMS to field support people. Such approach is very beneficial when the customer does not have local field support nor suitable technical knowledge how to handle faults on any equipment. The major benefit is the reduction of travelling costs as with SWO delivered on HoloLens the remote support can be easily and quickly executed without the need for FSO employees to travel to the site. Since every

FSO employee's business trip includes varied means of transportation, from cars, over buses, to airplanes, SWO can be considered environmental-friendly technology as the implementation and use of SWO in the field reduces the need to travel and thus reduces carbon emissions from transportation.

II. PROBLEM STATEMENTS

Work order is a specific job, repair or any other physical activity executed by a field support team. This includes all the activities required to manage and execute on-site services for end-customer. To achieve full field operation support delivery there are common communication requirements. A field support team needs to communicate frequently with subject matter experts (SME) to resolve faults or possible issues which occur on field equipment. As telco or IT equipment is specific and requires certain operations knowledge for problem resolution, any troubleshooting or resolution problem session may require one or more SMEs to travel to the site.

A frequent problem which FSO team members meet is how to organize ad-hoc travel when hardware or any other type of fault occurs on site equipment requiring manual intervention. The last-minute travel arrangements not only incur significant travelling cost, but also come with a possibility that suitable transport and accommodation are not available at the required moment. Additionally, some world countries require a visa and additional approvals which can take days.

The second issue is the time to resolve the problem which is defined by the service level agreement (SLA) and the key performance indicators (KPI) as mean time to repair (MTTR) and mean time to restore the service (MTRS) [4]. Once a fault occurs on the equipment, the time measuring starts, and every operations team needs to meet the requirements of the signed SLA. In case ad-hoc travelling arrangements cannot be organized promptly, any delays may cause missing KPI targets, which has a negative effect on operations team's efficiency.

The third issue which FSO team members may experience is lack of on-site equipment details, like the exact location, physical position in the data center and in the rack. Also, data about the type of the equipment or previous faults, errors and work orders executed by other operations teams in the past may be missing. In this case work order rejections are mainly due to insufficient information, incorrect information or wrongly created work orders, which generates significant but unnecessary operational costs.

The fourth problem, and that is the one that has the most impact on an operations team, is utilization of SME and FSO team members. While they travel, team members are not utilized as they do not have available nor appropriate working conditions to work remotely, for example during the flight, while changing transports or during any other travelling activity.

Every operations organization tries to optimize the combination of all issues with properly defined processes and measures their overall efficiency and effectiveness. The performance standards and specifications must be met by efficient delivery of services with respect to time, cost, quality, and other parameters to satisfy the customers and stakeholders. Smart work order combines tools and platforms which operations teams can use to undertake daily operations activities with a reduced risk of missing information and with operations cost reductions. SWO can be used for on-site equipment route optimization, as it helps find the fastest way to reach the target equipment, intelligently schedules preventive maintenance or real time status update, etc.

III. SWO ARCHITECTURE DESCRIPTION

SWO architecture integrates two main platforms, ARMS and ITSM which both consist of different modules exchanging data in real time. The key ARMS and ITSM components of SWO integrations are the following:

- ITSM monitoring module
- ITSM configuration management data base
- ITSM known error data base
- ITSM knowledge management data base
- ITSM Zabbix-ARMS media type connector
- ARMS message gateway
- ARMS work order.

A) ITSM monitoring module

ITSM monitoring module is based on Zabbix open-source solution [4]. The principal metrics from end equipment are collected via Zabbix proxy component by the means of its own Zabbix agent, or standard monitoring protocols such as simple network management protocol (SNMP) and intelligent platform management interface (IPMI).

Zabbix, an open-source enterprise grade network monitoring system solution, is integrated with ITSM tool suit as a core component. The use of Zabbix in large and distributed IT environments provides reliability, scalability, and flexibility, regardless of it being a standalone, or integrated solution.

B) ITSM Configuration Management Data Base

ITSM's configuration management data base (CMDB) has been build based on IT best practices according to ITILv3 Service Transition [5]. CMDB is a repository that acts as a data warehouse for information technology installations. It holds data relating to a collection of IT assets. CMDB is a series of tables used to keep track of the state of assets such as products, computers, systems, software contracts and licenses, devices on the network, facilities, and people as they exist at specific points in time, as well as the relationships between such assets. CMDB contains and records data called configuration

items (CI). It also provides details about the important attributes of CIs and the relationships between them. In the context of ITILv3 standard the use of CMDBs is a part of infrastructure operations and support, defined as a set of tools and databases used to manage IT service provider's configuration data, and information about incidents, problems, known errors, changes and the release. For every CI that is created the federation process into Zabbix monitoring system is mandatory. An item is fully configured when the relation with a CMDB CI is defined, for example, when the type of the host group, IP address and similar IT CI attributes are specified. Zabbix and CMDB federation will ensure that all configured alarms and events are properly mapped to the appropriate CMDB CI, in our case study, any physical equipment hosted in the field.

C) ITSM known error data base

Known error database (KEDB) contains all known error records identified by operations teams in previous fault events. This database is created by problem management and used by incident and problem management. The KEDB is a part of ITSM knowledge management system and ITSM service operation problem management process [6]. The term "problem" refers to the unknown cause of one or more incidents. Problem management consists of two major processes: reactive or proactive management. Reactive problem management is generally executed as part of service operations daily business and proactive problem management is usually initiated by service operations as part of continued service improvement [7]. Smart work order leverages all existing captured knowledge and provides summarized relevant information for any problem resolution. The main intention of SWO is to recognize potential events which may become an incident as error or fault that occurred on the maintained equipment.

The problem management process starts with the event and incident management. Once the problem is detected by the local operations team or proactively, the evaluation of incident patterns and alerts takes place. Due to the event management a full historic record of incidents exists. Then the problem is moved to the next phase where it is diagnosed and investigated in detail by the operations team. Once the solution of the problem has been defined and actions to restore service are successful, the problem is considered solved. The key step before the problem is closed is to log it into the error record. In this way all relevant details about the problem will be captured and used for the future operations and maintenance support. In case of SWO the KE details will be used to identify history incidents for specific CI, explore KMDB with the existing knowledge items and map a list of actions which will be used for problem resolution. The list of action will be used by ARMS message gateway to send detail instructions via HoloLens on how to fix the problem.

In next step, integrated CMDB, KMDB, problem management with SWO and OWL [3] provide the information that answer who, what, when, where and how questions, with reference to the performed or required actions.

D) ITSM knowledge management database

Knowledge management data base (KMDB) is responsible for maintaining the total body of knowledge within the service management organization. To deliver service successfully, it is necessary that relevant knowledge is captured, organized, and made available to all with a need-to-know technical details. ITSM knowledge management supports the process of authoring, updating, searching, collecting, organizing, structuring, and sharing the knowledgebase in an IT organization. For successful KEDB and KMDB integration it is mandatory that CMDB CI are federated within ITSM platform. To ensure accurate and up-to-date knowledge about fault resolution, there are knowledge analyst and knowledge reviewer roles where users with those roles needs to align their content with operations, L3 support teams or any other members who have worked on the problem resolution. Clear fault resolution instructions for each CI form a key prerequisite for successful SWO creation. Fault resolution instructions come in the form of a list of actions which are already verified and confirmed as a problem solution, and they will be sent to ARMS message gateway. Complete SWO process depends on accuracy of all CMDB and KMDB data as they will end up as instructions on HoloLens screen and be used by FSO engineer during the fault resolution action.

E) ITSM Zabbix-ARMS media type connector

Owing to Zabbix Media Type facility plug-in system, and configurable actions, alert and recovery operations messages can be distributed according to company defined rules among different responsible departments - network, systems infrastructure, applications, security.

The purpose of the media type connector is to route all events of hardware type which demand field engineer intervention on ARMS message gateway.

The plugin written in Python is defined in Zabbix as ARMS message gateway media type. This media type is bounded to HoloLens action type which uses groups of hosts and portable operating system interface.2 (POSIX.2) regular expressions to cover all hardware events. Zabbix is feeding Python based plugin with all relevant parameters about the location, equipment, IP address and other attributes during the new event processing phase and instructing it to deliver this message as JavaScript object notation (JSON) structure to ARMS message gateway for further data enrichment in KMDB and a work order creation.

With this approach, the monitoring component of the ITSM triggers recovery and fix procedure of the problem as a smart work order creation.

F) ARMS message gateway

The ARMS message gateway is an integral component of the SWO, and it serves as the API endpoint for data transfer between Zabbix and KMDB hosted in ITSM and ARMS work order. The ARMS message gateway is implemented using the C# programming language and

.NET 6. ARMS message gateway is using the standard hypertext transfer protocol (HTTP/HTTPS protocol) for its own message management. ITSM Zabbix-ARMS media type connector is therefore implemented as HTTP(S) client for the event and alarm message delivery. The ARMS message gateway is a centralized repository of information that contains all the information related to issues and their resolution, including the issue name, location, description, status, and steps required to resolve it.

The ARMS message gateway operates through a set of defined flows. When ITSM signals an event, the ARMS message gateway merges the event with the resolution data from the KMDB and stores the information in the ARMS database called ARMS open SWO queue (AOSQ). When ARMS work order requests information on a specific event, the ARMS message gateway retrieves the event and its resolution data from AOSQ database, enriches the original JSON structure received from ITSM Zabbix-ARMS media type connector with KMDB resolution and returns it to the client in the same session.

G) ARMS work order

The ARMS work order is a comprehensive solution that leverages both augmented reality (AR) technology and knowledge management (KM) to streamline the issue resolution process for Ericsson field services. The solution was designed specifically to assist field service engineers on site, providing them with an intuitive and interactive AR interface to resolve issues quickly and efficiently. Utilizing the latest AR technology, the solution seamlessly integrates virtual and real-world elements, providing a unique and engaging experience for the engineers.

The ARMS work order solution is developed using the C# programming language and the Unity3D engine, which is a leading game engine and platform for creating 3D and 2D interactive experiences. The solution has been developed using the Unity3D engine due to its versatility, scalability, and ability to create high-quality AR experiences, making it the perfect fit for the ARMS work order solution.

The solution runs on the Microsoft HoloLens 2 device, providing field service engineers with a hands-free AR interface to interact with. One of the key benefits of AR HoloLens device is that it leaves field engineers with free hands to access faulty equipment, for example to climb to different heights and then to execute any manual work on the equipment. The workflow starts with a field engineer scanning a QR code v2 placed in a physical room. QR code v2 is about 7,000 alphanumeric characters which is sufficient to store all relevant information about the room. QR code retrieves the related issues from the ARMS message gateway. After the successful QR canning, a virtual space is then calculated where ARMS work order solution contains a 1 to 1 copy of each physical room's layout recorded in digital format. This digital representation of each physical room enables the solution to accurately detect the location of the engineer and presents the required steps to resolve the issue in AR view.

As the engineer moves to the area near the issue, the solution detects their presence and begins to present the steps on HoloLens as AR windows required to resolve the issue. The engineer can mark each step as finished, and in the end, signal the ARMS message gateway to resolve the issue.

IV. SWO USER EXPERIENCE

The target audience for smart work order are teams working within field service operations. Based on the scope of FSO teams there are different types of work orders, such as:

- administrative and site engineering maintenance as site surveys, audits, inspections.
- corrective and preventive maintenance with changes on the equipment due to the identified faults, equipment installation, acceptance of services.
- planned maintenance activities which are required to be scheduled in advance as hardware replacement.

This chapter will cover HoloLens user experience for administrative maintenance during the site engineering and corrective maintenance actions. Site engineering is the process of installing new hardware equipment in a field which can be any remote location in a different town or country. The role of the subject matter expert is the most important during this process. He is responsible for every part of the equipment installation. SME is required to create the site installation documentation that contains all blueprints related to the installation. During the creation of the documentation, it is important to include the equipment installation in the cabinets, the connections between the equipment, the power supply, and the cooling of the room. All these inputs need to be documented and stored in CMDB as CI to fully understand the relationship among them.

To start working on the documentation, SME previously had to visit the site location to get an insight into all aspects of the installation issues. Now, owing to HoloLens we can do a site survey remotely. Instead of sending SME to the remote location HoloLens can be used by any FSO or non-FSO personnel located at a remote site. During the site survey, SME, who is located at the central service delivery spot, communicates with the FSO team member who is on the site via HoloLens using the ARMS application. ARMS allows SME to see the situation at the location in real time and to get the necessary information that will help him in creating the documentation.

Both, SME and FSO team member can make a complete analysis that will later be used for the creation of the documentation and CIs in CMDB. Using the ARMS application, which measures the dimensions of the room, SME is receiving information about where to place equipment cabinets, cooling systems, power supply and how much cable length is needed. With all the above information, they are ready to create the documentation

and prepare CI attributes for CMDB update. CMDB data accuracy is key for any future maintenance and operations.

After the installation of the hardware equipment has been completed, SME is responsible to perform the final part of the site survey: they will check whether everything has been done correctly as specified in the documentation. At this moment HoloLens will be used by FSO under remote SME surveillance. During this activity, SME is using the ARMS application implemented inside the HoloLens and can see in the real time whether all equipment installation actions have been done correctly according to the documentation. At the end, through the communication with the FSO team member using the ARMS application SME can get all the answers he needs related to the site survey.

There are many major benefits of ARMS usage identified by end users: significantly reduced travelling cost, reduced MTTR and MTRS times, increased customer delivery and operations satisfaction and very promising return of investment (RoI).

SME's trip to the remote location in the field includes increased costs caused by travel and accommodation. Usually, a person travelling to the location generally spends 80-90% time on the trip. By using the ARMS HoloLens solution, the mentioned disadvantages are eliminated. Travelling cost and time reduction may vary from case to case, depending on the field support service KPI and agreement. Many cases can be identified; however, we will describe three different characteristic scenarios to review travelling cost and resolution time reduction. All cost defined in below scenarios are illustrative to support cost calculation examples.

Scenario 1

The remote site with equipment is in the same county but geographically dislocated from FSO or SME's central location. Once a problem on the equipment occurs and work order is created SME needs to travel by public transport or car to remote location between two and five hours.

TABLE I. Scenario 1

SME Activities	Related cost	Time/average
From SME office to remote location by car	40 €	3,5 hours
Problem resolution – men/hour rate	100 €	1 hour
From remote location to SME office/home	40 €	3,5 hours
Total	180 €	8 hours

Local person activities	Related cost	Time/average
SME remote assistance/optional – m/h	100 €	1 hour
Problem resolution by local person – m/h	40 €	1 hour
Total	40-140 €	1 hour

Table 1. shows the list of SME activities where they need to travel around 3,5 hours to the remote location by car. After arrival, they will spend 1 hour to resolve the problem and return to the office or home. During the scenario 1 activities, which last one working day, SME will not be utilized 87,5% of his working time as during the traveling time they cannot be used for any other business activity. In this case, MTRS is around 4,5 hours after the info about the problem has been received and a work order created. The alternative way is to use SWO where the list of problem resolution activities will be sent to ARMS HoloLens and performed by local staff in the field with the hosted equipment. In this scenario, the local staff requires only HoloLens and does not need to be SME or equipment domain knowledge expert to resolve the problem as SWO will provide the list of required actions to fix the problem. Usually, local staff men/hour (m/h) rate cost is significantly lower than SME m/h rate. This way SME can be optionally used to support the local person and SME will not have any gaps in utilization. MTRS time is within 1 hour after the info about the problem has been received.

Scenario 2

The remote site with equipment is in different country from FSO or SME's central location. SME requires a suitable transport to reach the remote location as soon as possible to meet service KPI. The transport includes taxi and plane with one connection flight. SME's business trip includes one night in the hotel and travel allowance.

TABLE II. Scenario 2

SME Activities	Related cost	Time/average
From SME office to the airport by taxi	30 €	1 hour
Connection flight 1	210 €	2 hours
Connection flight 2	190 €	2 hours
From airport to remote location by taxi	40 €	1 hour
Problem resolution – men/hour rate	100 €	1 hour
Hotel accommodation – 1 night	100 €	12 hours
From hotel to airport by taxi	30 €	1 hour
Return connection flight 1	190 €	2 hours
Return connection flight 2	210 €	2 hours
From the airport by taxi to SME office	30 €	1 hour
Total	1130 €	25 hours

Local person activities	Related cost	Time/average
SME remote assistance/optional – m/h	100 €	1 hour
Problem resolution by local person – m/h	40 €	1 hour
Total	40/140 €	1 hour

Table 2. shows the list of SME activities where they need to travel to a different country by taxi and plane to reach the remote location as soon as possible. During this travel they will have connection flights as direct flight may not be available. After the arrival they will spend 1 hour to resolve the problem and return to the hotel for one night as there are no available return flights. During scenario 2 activities, which last 2 working days, SME will not be available most of his working time as travelling time cannot be used for any other business activity. In this case, MTRS time is minimum 6 hours, depending on public traffic and airport conditions after the info about the problem has been received and a work order created. The major factor in scenario 2 is a significant cost increase due to SME's travel activities. The alternative is to use SWO where a list of problem resolution activities will be sent to ARMS HoloLens and performed by the local staff in the field with the hosted equipment. This option is the same as in scenario 1 where MTRS time is within 1 hour after the info about the problem has been received without loss of SME working time utilization.

Scenario 3

The remote site with equipment is in a different country which has restricted entering country process and includes visa approval. Visa application and approval at the embassy may last minimally 1-2 days, depending on the country and emergency visa process.

TABLE III. Scenario 3

SME Activities	Related cost	Time/average
Embassy visa application process	100 €	24 hours
From SME office to the airport by taxi	30 €	1 hour
Connection flight 1	210 €	2 hours
Connection flight 2	190 €	2 hours
From airport to remote location by taxi	40 €	1 hour
Problem resolution – men/hour rate	100 €	1 hour
Hotel accommodation – 1 night	100 €	12 hours
From hotel to airport by taxi	30 €	1 hour
Return connection flight 1	190 €	2 hours
Return connection flight 2	210 €	2 hours
From the airport by taxi to SME office	30 €	1 hour
Total	1230 €	49 hours

Local person activities	Related cost	Time/average
SME remote assistance/optional – m/h	100 €	1 hour
Problem resolution by local person – m/h	40 €	1 hour
Total	40/140 €	1 hour

Table 3. shows the list of SME activities where they need to travel to a different country which has restrictive entering process. In this case, travelling preparation starts with a visit to the embassy to apply for a visa. The visa approval process may vary from country to country and is an uncertain process where the visa may not be approved. During the scenario 3 activities which take 3 working days minimally SME will not be utilized most of his working time as during travelling arrangement actions they cannot be used for any other business activity. In this case, MTRS time is minimum 30 hours, depending on the visa approval process after the info about the problem has been received and a work order created. The travelling cost is additionally increased. The alternative to SME activities is the same as in the scenarios 1 and 2 where local staff can give support via HoloLens and where MTRS time is within 1 hour after the info about the problem has been received without any loss of SME working time utilization.

All three scenarios show that MTRS time and SME or local staff support costs are quite flat. On the other hand, SME travelling cost can vary from case to case, which has an impact on MTRS. The average cost of one Microsoft's HoloLens 2 device is around 3,500 EUR. As HoloLens cost is comparable to a couple of travelling arrangement cost, ARMS SWO solution becomes the first choice for every company that wants to reduce operational costs. For full RoI calculation SME working hours utilization and MTRS KPIs need to be taken into consideration as they come with additional cost.

V. GREEN TECHNOLOGY

As environmental-friendly technology has strong effect on the environment, the carbon footprint can be measured for each performed activity. The carbon footprint is total amount of greenhouse gases that are generated by the human actions, in this case human traveling activities.

Table 4. shows carbon emissions calculation for travelling scenario 2 and 3 as they include different means of transport. The calculation is based on carbon emissions calculator [8] where connection flights are withing 2 hours of flight distance inside Europe. If SME travel to the remote site, their carbon footprint in one way trip would be 0,55 MT (metric tons of CO2). With the return trip the carbon footprint would increase, and their travelling activities would produce 1,1 MT of carbon emissions.

TABLE IV. Carbon emissions calculation

Travel routes	Carbon footprint
From SME office to the airport by taxi	0,01 MT
Connection flight 1	0,13 MT
Connection flight 2	0,40 MT
From airport to remote location by taxi	0,01 MT
Total	0,55 MT

After all those considerations we can say that SWO certainly is an environmental-friendly technology. By using SWO, SME can easily communicate remotely with FSO team member who is in the field without having to go to the site location. This eventually leads to the reduced emission of greenhouse gases.

The additional component which reduces the carbon footprint and travelling cost is "the first-time-right" (FTR) approach. FTR concept finds its origins in Six Sigma and it is defined as a principle where all actions will be performed correctly the first time, thereby eliminating the need of any rework. Thus, FTR minimizes the operational cost and optimizes FSO and other supporting organization's efficiency. SWO, as many technology layers driven approach, offers multiple tools, controls, and a mindset to maintain FTR concept.

VI. CONCLUSION

In conclusion, the augmented reality modeling support (ARMS) work order solution is a cutting-edge technology that streamlines the issue resolution process for Ericsson field services, providing several benefits over traditional manual processes. Its innovative approach, intuitive interface, and seamless integration of virtual and real-world elements make it an efficient and effective solution for Ericsson engineers and field service operations teams. When it comes to the user experience of the HoloLens smart work order, we need to look at both ends of the process: end customers have their equipment maintained according to the signed service-level agreements while FSO teams have the benefit of reduced delivery time and lower operations costs with "the first time right" concept in place. The key contributors for successful implementation of SWO FTR are the following: properly defined CI within CMDB, federation and relationship with KEDB and KMDB where product development teams, SMEs and operations must work together to ensure the most accurate data and instructions for SWO content. As physical field operations activities are very delicate and require full field staff's concentration on the execution of problem resolving actions, SWO offers smart instruction checks and real-time resolution confirmation to avoid any potential rework. An additional benefit of SWO is related to preventive actions which are prepared through advanced equipment monitoring logs analysis and recognition of potential events which may turn into incidents and problems. With ARMS HoloLens smart work orders can be prepared for each site so that local staff can perform preventive actions, which makes travelling for problem resolution obsolete.

REFERENCES

[1] Field Service Operations (FSO): FSO Process, Ericsson internal document, Stockholm:2019.

[2] Office of Government Commerce (OGC), ITIL - Service Design, TSO, 2007, ISBN 9780113310470.

[3] T. Žitnik and Z. Bosić, "Operations Wisdom Logging," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 506-510, doi: 10.23919/MIPRO55190.2022.9803373.

[4] Office of Government Commerce (OGC), ITIL - Service Transition, LAG, 2007, 9780113310487.

[5] Zabbix software, Available at: Zabbix documentation, [Accessed: 06-Feb-2023]

[6] Office of Government Commerce (OGC), ITIL - Service Operation, LAG, 2007, 9780113310463.

[7] Office of Government Commerce (OGC), ITIL – Continual Service Improvement, LAG, 2007, 9780113310494.

[8] Carbon Footprint Ltd., Carbon Calculator, Carbon Footprint Calculator for Individuals and Households - https://www.carbonfootprint.com/calculator.aspx, [Accessed: 06-Feb-2023]

Bringing Children Closer to Science and the Universe with New Technology

P. Lahovsky, A. Ljubek, K. Mišura, M. Žilak
University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{patrik.lahovsky; ana.ljubek; katarina.misura; matea.zilak}@fer.hr

***Abstract* - In the era of technology and social networks, it is difficult to attract ch[ildren to science. To change this, three applications have been developed whose primary goal is not pure entertainment, but to educate and encourage scientific exploration in children. These applications were developed using new technologies, i.e. augmented reality, virtual reality and Ultraleap. Each application brings the intrigue of the universe to children in an interesting way by providing them with age-appropriate content and information about the universe. In these applications, interactive space travels can be taken to learn about the Sun and planets, and fun games can be played to stimulate interest in further observations and knowledge about the universe. The paper describes the applications and the multidisciplinary context of their development, as well as the potential of the new technologies in education.**

Keywords - universe; serious game; augmented reality; virtual reality; Ultraleap

I. INTRODUCTION

The advancement of technology has opened numerous opportunities for innovation, and one such area is game development. Serious games are both entertaining and educational, making them a perfect tool to introduce complex concepts to children [1]. This paper presents the development of three serious games, named SolARsys, TouchSpace and UniVRse. These games were created by a team of college students attending the University of Zagreb, Faculty of Electrical Engineering and Computing as part of a course named Project. The project was part of Project SPARK, which involved the construction of an observatory and cosmological center in Križevci, Croatia. The developed games will be available to people in the lobby of the cosmological center and observatory [2][3]. Each game utilizes a different technology - augmented reality (AR), Ultraleap and virtual reality (VR) - and serves as an interactive learning experience for children. The team was divided into three groups, each focusing on one technology and application. This paper includes a section detailing the technologies used, followed by a section describing the development process of all three games. The final section is a product description that highlights the unique features of each game. The ultimate goal of the students project was to spark an interest in science among children through the use of new and engaging technologies. The development of SolARsys, TouchSpace and UniVRse demonstrates the potential of new technologies in education. By integrating augmented reality, Ultraleap technology and virtual reality the team was able to create engaging and interactive learning experiences that could captivate and inspire children. These technologies have the potential to transform the way we teach and learn, providing new ways to explore complex concepts in a more intuitive and immersive way.

II. USED TECHNOLOGIES

In this chapter, the technologies used in the project are described. Subchapter A describes how augmented reality (AR) equipment works, i.e. Nreal Light Developer tool kit, NRSDK and Scrcpy. Subchapter B describes Ultraleap, while Subchapter C explains Virtual Reality (VR) technology, i.e. Meta Quest 2. Subchapter D explains why Unity was used to build the applications and its description is given.

A. Augmented reality – Nreal Light Developer tool kit, NRSDK and Scrcpy

The main components of the Nreal Light Developer tool kit are Nreal AR glasses, Nreal computing unit and Bluetooth Nreal controller. The glasses connect to the computing unit with a wire and use the computing unit as a power supply. They can be connected to the user's mobile phone or computer, in which case the screen from the mobile phone or computer is projected in front of the user, i.e., they have the function of an additional screen. The Nreal controller connects to the computing unit via a Bluetooth connection. When the user uses the controller, a white beam can be seen that represents the pointer used for selection and interaction with the AR elements [4].

The glasses have well-built spatial computing that includes motion tracking, plane detection, image tracking, hand tracking and tracking of real-time position of the glasses relative to the world and environment around them. Plane detection enables Nreal glasses to detect both horizontal and vertical flat surfaces in the environment. Image tracking enables recognition of physical images from the environment. Hand tracking enables tracking the position of key points of the user's hands and recognition of hand poses in real-time. Nreal glasses have 6 degrees of freedom achieved by identifying feature points and tracking how those points move over time. Combined with the sensor readings, the glasses accurately track both position and orientation of the glasses through space [5].

NRSDK package is Nreal's platform package for developing mixed and augmented reality for Nreal glasses. The package enables Nreal glasses functionalities and provides Mixed Reality (MR) and Augmented Reality (AR) features. The core features are spatial computing,

optimized rendering, and multi-modal interactions. The package enables full use of spatial computing and gives the developer easy access to all its abilities. Optimized rendering is automatically applied to the applications to minimize latency, reduce judder, and enhance user experience. The package also contains an example project that shows all Nreal glasses capabilities and provides the developers with example code. The written documentation only provides with tutorials on some of the features of Nreal glasses and are not written in depth. Because of the scarce documentation, the developers had to go through the examples given and inspect the packages core scripts that contain the code with all the functionalities written. NRSDK supports the development environment of Unity 2018.4.X and above [5].

Scrcpy application enables control of Android devices connected via USB or TCP/IP. It displays the screen of the device and enables interaction on an intuitive level. The application requires at least Android version 5.0 or API level 21 to work and is nonintrusive, i.e., nothing remains installed on the Android device, in this case Nreal computing unit [6]. This application was used for installing the developed application on to the Nreal computing unit.

Nreal's various features provide a different approach to user interactions. The user is placed in the real-world but can simultaneously interact with virtual elements. By merging virtual and real-world elements, it becomes possible to fill in the gaps of physical reality with digital enhancements, resulting in a complete and immersive experience. Plane detection provides the ability to place objects on detected flat surfaces in the real-world. Combined with 6 degrees of freedom, the user can easily walk through the real-world while looking at virtual objects through the glasses. This feature could be used for educational purposes, interior design, or general visualization of inaccessible objects. Image recognition can be used for educational purposes, museum tours and even advertisement. Hand tracking feature significantly changes user experience because it provides a simple and intuitive interaction with virtual objects. This feature can be combined with any other feature to expand on the user experience by making it more engaging and immersive. Hand tracking is best suited for education and simulations of jobs that require a hands-on approach.

B. Ultraleap

Ultraleap was formed when Leap Motion and Ultrahaptics came together in May 2019. Built for research and development, the STRATOS Explore development kit, which is used for development of TouchSpace serious game, is ideal for exploring how mid-air haptics can enhance innovative products and experiences. Particularly suitable for automotive, appliances/smart home, kiosks, gaming and computing applications. It consists of two parts, the Leap Motion Controller and Ultrahaptics. The Leap Motion Controller is a sensor that tracks hand movements, while Ultrahaptics projects tactile effects using ultrasound. Some of the tactile effects in mid-air that Ultrahaptics can project are virtual buttons and sliders, haptic pulses and alerts, immersive sensations such as textures, and presence for virtual objects, surfaces and shapes, magical sensations such as lightning, fireballs, ghosts, clouds, bubbles and force fields [7].

To start Ultrahaptics and Leap Motion Controller, it is necessary to download suitable drivers for connecting the device to the computer (described in the chapter Development process, subchapter Implementation, third paragraph). After installing the driver, the device must be physically connected to the computer. Ultrahaptics needs to be connected to its charger, and it needs to be connected with a USB cable to the computer before turning on the switch on it. The Leap Motion Controller also needs to be connected with a USB cable to the computer before launching the application. To start using the application and Ultraleap, the user should place his hand above the Leap Motion Controller and Ultrahaptics at a height of approximately 20 cm. After the sensor recognizes the hand, the hand (in 3D form) and its movements will be displayed on the screen, and the user will feel ultrasonic waves on the hand in the form of a light breeze.

Ultraleap technology controls ultrasound waves to project tactile sensations onto your hand, without the need to wear haptic gloves or use hand-held game controllers. It is completely hygienic and it improves gesture control, making the user experience more engaging and immersive. Ultraleap can be used in education, it could help students to learn shapes, textures and movements, such as the movement of tectonic plates. It also helps visually impaired people to feel objects they cannot see.

C. Virtual reality – Meta Quest 2

Meta Quest 2 is a virtual reality set made by tech giant Meta, formerly known as Facebook. It is a successor of Oculus Quest, former virtual reality set made by Oculus. Meta Quest 2 set consists of virtual reality glasses and two controllers, one for each hand. Glasses have four built-in infrared cameras that track the room, controllers, and hands. Tracking the room is important for creating guardian. Guardian is an area created in the room in which is player free to move. Tracking controllers is used for showing the controllers in virtual reality in same position they are in real space. Hand-tracking technology is used to interact and play using the user's own hands. The Meta Quest 2 glasses have a high screen resolution and strong hardware that enables a virtual reality experience without the need for a computer or console. The glasses have a resolution of 1832 x 1920 pixels per eye and a refresh rate of 90Hz, which makes visuals in virtual reality clean and smooth. It also has 6 degrees of freedom (6DoF) which enables more precise and accurate tracking of the user's movements. The Meta Quest 2 also has a feature called Quest Link, which allows the wearer to wirelessly connect the glasses to a computer and play virtual reality computer games. This feature requires a compatible computer with a high-quality Internet connection [8].

On the right side of the glasses there is a power button, volume buttons and a USB-C port for charging. The device also has built-in stereo speakers and a microphone, as well as a 3.5 mm audio jack for users who prefer to use their own headphones [8].

The device comes with two controllers, Oculus Touch, which are designed to imitate the movement of the user's

hands in the virtual world. Controllers have numerous buttons and sensors that allow the user to interact with objects and perform a wide range of actions in virtual reality. Each controller has two buttons on the top, one on the side and a trigger on the back. The buttons on top are labeled A and B on the right controller and X and Y on the left and are used for various interactions. The button on the side is used to access the battery. The controllers also have several sensors that allow tracking the movements of the user's hands, including a gyroscope, an accelerometer, and a magnetometer. These sensors allow the controllers to detect the movements of the user's hands and translate them into actions in the virtual world. The controllers are wireless and have a very long battery life. They are designed to be intuitive and easy to use, allowing the user to interact with the virtual world as naturally as possible [8].

Virtual reality enables the creation of a completely new, virtual world in which the user has a 360-degree 3D view. Because of this it is best suited for visualizations and simulations for educational purposes. It is ideal for games where the user stands in one position or games where the player moves using teleportation. The ability to track the movement of the controllers in the real-world provides a more immersive usage than traditional console controllers.

D. Unity

Unity was the logical choice of platform because of its support for 3D video games, with a lot of built-in functions for VR. The Oculus XR plugin was used to connect Unity and the controller. It provides support for hand tracking and motion controls, as well as the ability to use the Meta Quest glasses' built-in cameras to create augmented AR experiences. It also allows developers to use Quest's advanced features such as 6DoF and hand gesture recognition [9]. Unity version 2021.3.11f1 was used for development of SolARsys. Unity version 2021.3.13f1 was used for development of UniVRse and version 2019.3.13f1 for development of TouchSpace.

Unity is a cross-platform technology developed by Unity Technologies, first announced and released in June 2005 at Apple's Worldwide Developers Conference. Since then, it has been gradually expanded to support a variety of desktop, mobile, console and virtual reality platforms. It can be used to create three-dimensional (3D) and two-dimensional (2D) games, as well as interactive simulations and other experiences. In Unity, it is possible to develop applications in the programming languages C# and C++. As part of this project, all three serious games were developed in C# [10].

III. DEVELOPMENT PROCESS

This chapter describes the application development process. Each subchapter represents one stage of development, and there are four in total. The first phase of the project included the formation of a team and familiarization with the project task. The students' task in the second phase was to get to know the given technologies and explore their possibilities. The third and longest phase was the application implementation phase. The last phase was presentation of the created applications to the cadets of the Croatian Military Academy and to the initiators of the SPARK project.

A. Team forming

First phase of development process includes getting to know the team, familiarization with the project task, selection of the team leader and division into sub-teams (one team for each technology, i.e., serious game development).

B. Exploring technology and its possibilities

The second phase lasted a total of four weeks. At the beginning, the obtained devices (AR – Nreal Light Developers tool kit, STRATOS Explore development kit, VR – Meta Quest 2) were tested and the team members got acquainted with the possibilities of the technologies. After that, there was the launch and presentation of demo applications in between using the Unity system for each device. After that, the content of each serious game was defined in accordance with possibilities of its development technology (more about the content of serious games is described in section IV.). After the work was distributed among the members of the sub-teams, the implementation phase followed.

C. Implementation

The third phase lasted for ten weeks. Weekly meetings were held on which the features and functionalities of each product, i.e., serious game, were defined and/or tested. The development of each product was agile which means that requirements and solutions were incrementally improved between meetings. Sub-teams tested each other's products, commented on user experience and suggested improvements. This way mistakes were discovered and fixed much earlier during the development and the product was more flexible for content and user experience corrections. All sub-teams implemented their core features and functionalities very quickly and at the beginning of the implementation phase. The rest of the phase was spent on polishing the user experience and fixing minor problems that were discovered during continuous testing.

Nreal sub-team parallelly developed the Meet the Universe mode and Recognize the Planets game. The first iteration demonstrated the idea behind the application and its parts. Next iterations focused on implementing the main functionalities and user interactions. There were slight problems with the hand tracking feature, but after its resolve, the rest of the iterations focused on user experience and instructions. In those later stages of the implementation, testing and reactions from other team members were very important. Their feedback and suggestions helped in building better user experience.

To start Ultrahaptics and the Leap Motion Controller, it is first necessary to download the appropriate drivers for connecting the device to the computer: Ultrahaptics SDK (for Ultrahaptics) [1] and V2 Desktop SDK (for Leap

[1]https://developer.ultrahaptics.com/

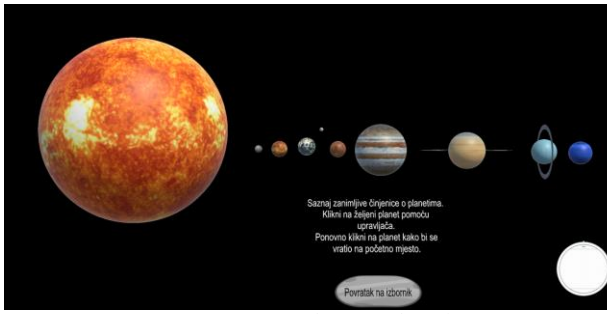


Figure 3 SolARsys: Mode „Meet the universe“

Motion)[2]. The version of Unity compatible with the mentioned drivers is Unity 2019.3.13f1.

The VR team implemented the application collaboratively using Plastic SCM. The team initially had problems building the application and running it on VR glasses. Once that problem was fixed, it was challenging to connect the functionality to the Oculus Touch controllers since it was the first time the developers were doing it. Apart from the above, the team did not experience any major problems in implementing the application.

D. Presentation

The applications that were developed were showcased to the cadets of the Croatian Military Academy, who received the opportunity to explore all the features and possibilities that they offer. The cadets were highly impressed with the innovative and user-friendly design of the applications and were eager to make the most of them.

Additionally, the creators of the applications also presented them to the initiators of the SPARK project. The initiators had the opportunity to try the developed applications. They discussed their potential improvements and applications, especially in working with children. During this discussion the idea of these applications and technologies being available in the observatory for the visitors to play with and learn something new while they wait for their turn in the observatory arose. The initiators were extremely satisfied with both the implemented applications and the possibilities offered by new technologies.

Overall, the presentation of the applications to the cadets and the SPARK project was a successful and productive endeavor that generated enthusiasm and positive feedback.



Figure 2 TouchSpace: Main menu

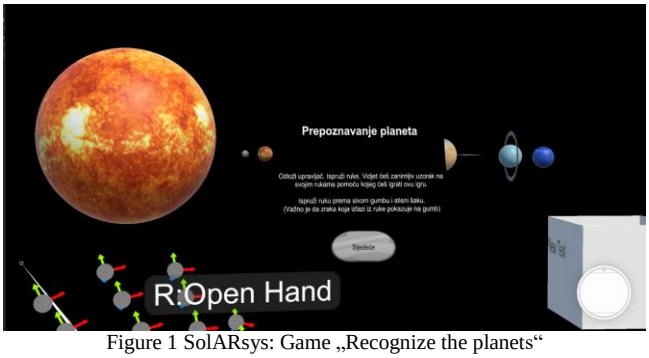


Figure 1 SolARsys: Game „Recognize the planets“

IV. PRODUCTS DESCRIPTION

This section describes three serious games - SolARsys, TouchSpace and UniVRse. SolARsys uses augmented reality technology, TouchSpace uses Ultraleap technology, and UniVRse uses virtual reality technology. The ways of using these three applications are also described. The images in this section contain texts in the Croatian language since all applications were created for use exclusively in Croatia.

A. SolARsys

When the application opens, the main menu is displayed, and the user has the option of selecting the Meet the Universe mode or Recognize the Planets game. The desired mode is selected using the Nreal controller.

By selecting Meet the Universe, the planets and instructions for interactions are displayed in front of the user (Figure 2). The goal is to learn interesting facts about the planets of the Solar System. The interactions with the planets are achieved with the Nreal controller. The user sets the pointer on the desired planet and presses the button on the controller. The selected planet will move closer to the user. On the left side of the planet an information panel will be displayed containing information about the planet. By clicking on the planet again, it returns to its original position and it's possible to select the next planet. The user can return to the main menu at any time by pressing the Back to Main Menu button.

Recognize the Planets game examines if the user has learned how to recognize each planet of the Solar System. After selecting the game, instructions are shown which explain to the user how to master the hand tracking functionality (Figure 4). The user must first put down the

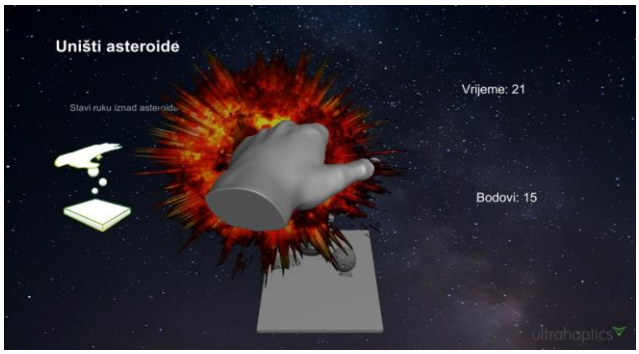


Figure 4 TouchSpace: Game „Destroy Asteroids“

[2]https://developer.leapmotion.com/

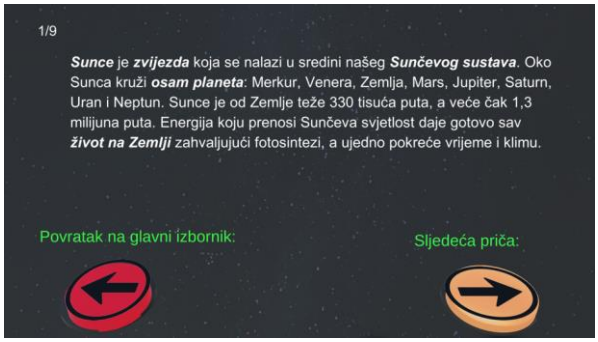


Figure 6 TouchSpace: Screen with informations about Sun



Figure 5 TouchSpace: Display of planets after the correct answer to the question asked as part of the quiz

Nreal controller, read the instructions and test the hand tracking functionality to get familiar with it. The hand tracking functionality recognizes when the users palm is opened or closed to form a fist, and with these gestures user selects the desired elements through the game. To select a button in the instructions, the user must extend their hand with the palm towards the button, and when the pointer coming out of the palm is on the button, the user must squeeze their hand into a fist. After the user has gone through the instructions the game begins. In front of the user is the Solar System, and on the right side is a box with the name of the given planet. The goal of the game is to recognize the planet of which the name is displayed on the box and drag the planet into the box. For the user to select a planet, they must walk up to it and form a fist as though they are grabbing an object in the real world. The user must keep his fist clenched to hold the planet and move their hand slowly to move the planet. When the user has placed the correct planet into the box, a new planet name is displayed. After the user has successfully identified three assigned planets, the game ends and a panel appears with the options to replay the game or to return to the main menu.

B. TouchSpace

At the very beginning of using the application, the main menu of the application will be displayed on the screen (Figure 2). The hand should be placed at a distance of approximately 20 cm above the Ultraleap. When the hand is correctly positioned, the hand model will be also visible on the screen. In the main menu it is possible to choose between three modes, each mode is represented by one button and corresponding color from the legend. Mode on the left (the green one, first button) is for learning about asteroids, mode in the middle (the orange one, second button) is for learning about planets, and the last mode on the right (the red one, third button) leads to a quiz about planets. To enter one of the mentioned modes, it is necessary to press the appropriate button with a hand gesture (with an outstretched fist).

Before describing the modes, it should be noted that the return to the main menu of the application is possible in every part of the application by selecting the space bar, while the application is closed by pressing the escape key.

Selecting the first mode (the green button) in the main menu opens a screen with information about asteroids. After learning about asteroids, it is possible to play the game "Destroy Asteroids" by selecting the orange button with a right arrow on the same screen with the previously described hand gesture. As soon as the screen on which the asteroids are moving upwards is shown, it is necessary to use a hand gesture to destroy as many asteroids as possible within 30 seconds in order to protect the Earth. While playing the game, points are displayed on the screen (the number of destroyed asteroids). After the playing time expires, the final number of points is displayed on the screen. On the newly displayed screen, it is possible to return to the main menu by pressing the red button with a left arrow.

Selecting the second mode (the orange button) in the main menu opens a screen with information about the Sun and the planets of the Solar System. Return to the main menu is enabled by pressing the red button with a left arrow, while getting to know the next planet is enabled by pressing the orange button with a right arrow.



Figure 7 UniVRse: Display of planet name



Figure 8 UniVRse: Instruction panel

Selecting the third mode (the red button) in the main menu opens a screen with a planet quiz, i.e., a question about the Sun or a planet of the Solar System. Below the question, there are four answers marked with the letters A, B, C and D. Depending on which answer you think is correct, you need to select the button with the corresponding letter using the same hand gesture as in the rest of the application. If the selected answer is incorrect, it will be possible to answer the same question again. When the chosen answer is correct, the planet to which the question was related will be displayed on the screen as a reward. It is possible to touch the displayed planet and "pass through it" with the appropriate hand gesture above Ultraleap. You can return to the main menu by pressing the back button (the red button with a left arrow).

C. UniVRse

UniVRse game, when started, puts the player above the Sun. While floating above the Sun, the player can enjoy a spectacular view of the universe. The player can also rotate 360 degrees and watch all the planets rotating around themselves and orbit around the Sun. For better orientation, the orbits of the planets are shown. The rotation and orbital velocity of the planets are scaled from actual velocities in universe. Size and distance from the Sun are also scaled, and the ratios are preserved. The only ratio that is not preserved is the size of the Sun, as it is disproportionately large compared to the planets to fit in the virtual reality environment. Using the left controller, the player can teleport to the planet by pointing the controller at the planet and pulling the trigger on the controller. While the controller is not pointed at any planets, the laser line coming out of the controller is red. The moment the controller starts pointing at a certain planet, the laser line changes color to white, and its name is displayed above the planet. When the player teleports to a planet, they will begin to move along with the planet through space. During this time, the player can observe the universe from that position above the planet. All the time, the player can press the Y button to see instructions for using the application. Also, pressing the X button shows interesting information about the planet the player is above, or about the Sun, if the player is above it.

V. CONCLUSION

The developed applications are a great addition to the observatory since they allow visitors to have fun and learn about the Solar System in an interactive way while waiting in line for stargazing through the telescope. Experiencing new technologies such as augmented and virtual reality, as well as haptic technologies, can not only make the wait more interesting, but also help popularize science, technology, engineering, and math (STEM) as children and other visitors are introduced to the possibilities of these technologies in engaging ways as they prepare to observe space.

Each of the developed applications takes advantage of the technology for which it was made, and they also complement each other. So SolARsys uses augmented reality, allowing the user to see virtual planets in the real world, and learn about the Solar System and assess knowledge by recognizing planets. TouchSpace uses Ultraleap technology, introduces the user to asteroids and planets of the Solar System, tests learning through a quiz, and as a reward allows touching the planets. UniVRse is an application that, with the help of virtual reality technology, tries to give the user an idea of the Solar System and the mutual relations between the planets and the Sun.

Based on the enthusiasm of the cadets and project initiators for the applications, there is no fear for the popularity of STEM in small towns like Križevci, but also in the whole of Croatia.

The success of this project highlights the exciting possibilities for the future of education as technology continues to evolve and improve.

REFERENCES

[1] "Serious games". [Online]. Available: https://grendelgames.com/what-are-serious-games/. [Accessed 16 February 2023].

[2] "SPARK project Križevci". [Online]. Available: https://www.krizevci.info/2022/08/07/o-kozmosu-prirodi-i-covjeku/. [Accessed 18 February 2023].

[3] "SPARK project - observatory and cosmological center in Križevci". [Online]. Available: https://www.poslovni.hr/vijesti/ovaj-mali-grad-imat-ce-prvi-kozmoloski-centar-u-hrvatskoj-te-zvjezdarnicu-4294660. [Accessed 18 February 2023].

[4] "Nreal glasses." [Online]. Available: https://nreal.gitbook.io/nrsdk/nrsdk-fundamentals/nreal-devices/readme. [Accessed 16 February 2023].

[5] "NRSDK". [Online]. Available: https://nreal.gitbook.io/nrsdk/nrsdk-fundamentals/core-features [Accessed 16 February 2023].

[6] "Scrcpy tool" [Online]. Available: https://github.com/Genymobile/scrcpy [Accessed 16 February 2023].

[7] "Ultraleap STRATOS Explore". [Online]. Available: https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf. [Accessed 16 February 2023].

[8] "Meta Quest 2" [Online]. Available: https://www.meta.com/quest/products/quest-2/ [Accessed 17.2.2023.]

[9] "Oculus XR Plugin" [Online]. Available: https://docs.unity3d.com/Manual/com.unity.xr.oculus.html [Accessed 17.2.2023.]

[10] "Unity (game_engine)". [Online]. Available: https://en.wikipedia.org/wiki/Unity_(game_engine). [Accessed 16 February 2023].

An Overview of 3D Holographic Visualization Technologies and Their Applications in Education

Ana Keselj, Krunoslav Zubrinic, Mario Milicevic and Martin Kuzman
Department of Electrical Engineering and Computing, University of Dubrovnik, Dubrovnik, Croatia
{akeselj; krunoslav.zubrinic; mario.milicevic}@unidu.hr, martin.kuzman@stud.unidu.hr

***Abstract* - Three-dimensional (3D) visualization is the process of creating graphical content using software technologies, where the object is displayed using a computer or mobile device with appropriate software, and the visualized 3D object can be viewed from different angles and sides. 3D objects can be visualized using a display on a computer screen with 3D graphics or using technologies such as augmented reality or holograms. This technology has significant potential and is used in education. Interaction with software solutions is recognized as one of the most important learning experiences. This paper provides an overview of holographic visualization technologies, their interaction capabilities, and technology integration. It also presents and describes an example of a holographic application for learning geometry that can be used in a 3D holographic pyramid projection system.**

Keywords – visualization technologies; holograms; interactions;

I. INTRODUCTION

The use of visualization tools in the learning process is not a novel concept; it has been practiced for a long time. The visualization provides an opportunity for learners to understand and accept complex concepts more easily and conveniently [1]. Nowadays, educators have access to a wide range of visualizations, including different technologies such as augmented reality, virtual reality, holograms, etc. Analysis of the use of visualization technology in education [2] revealed that it had a favorable impact on students' cognitive growth. Holographic projections draw students' attention because they enable them to examine a topic from various perspectives, which improves their comprehension of the subject. Usage of this technology also benefits the student with learning disabilities which usually presents unique challenges. Children with developmental disabilities often have limited attention spans, making it difficult for them to stay engaged with a task for long periods of time. New technologies for visualizing learning content, such as augmented reality, can capture their interest and direct their attention enough to ensure effective learning [3].

One of the most important aspects of educational applications is accessibility. Digital accessibility is the extent to which a computer program, website, or device is acceptable and suitable for use by people with disabilities and the elderly. This results in a greater degree of inclusion of people with disabilities in modern society. Digital inclusion stands for all efforts to increase the level of social inclusion of people with disabilities and older people by adapting existing or developing new services based on information and communication technology that enable more efficient communication, access to information, and assistance with education [4]. The need to develop accessible applications for educational purposes is reinforced by the increased inclusion of children with disabilities in the standard education system.

Given that visualization technologies are proven to help in the education of all children, especially children with disabilities, we were motivated to develop an application that will help them learn geometry. As a part of INNOSID consortium, we have developed application called HoloGeometry which prototype will be described in this paper. INNOSID, whose full name is "Innovative solutions based on new technologies to improve the social inclusion of people with disabilities"[1], was an Erasmus+ project led by the Faculty of Electrical Engineering and Computing at the University of Zagreb. The aim of the project was to improve the social inclusion of people with disabilities, particularly people with Down syndrome, by developing innovative IT solutions based on new technologies. The project involved six higher education institutions, two from Croatia and one each from Hungary, France, Spain and Portugal, as well as a non-governmental organisation, the Croatian Down Syndrome Association.

The rest of the paper is organized as follows: Section II describes visualization technologies and gives an overview of their application in education, Section III describes HoloGeometry application and, finally, Section IV concludes the paper.

II. VISUALIZATION TECHNOLOGIES

Three-dimensional (3D) visualization is the process of creating graphical content using software technologies, where the object is reproduced with the help of a computer or a mobile device with appropriate software and it is possible to view the visualized 3D object from different angles and from different sides [5]. 3D objects can be visualized using visualization technologies such as displaying them on a computer screen with 3D graphics or using technologies such as augmented reality or holograms [6]. All types of interactions that a user can perform in a 3D virtual world are called three-dimensional (3D) interactions [7].

[1] INNOSID, http://sociallab.tel.fer.hr/innosid/

3D visualization has significant potential and finds application in various fields such as science, art, health, etc. Among them, education stands out. Visualization and interactive manipulation of 3D objects can greatly enhance the learning experience [8]. 3D visualization and the possibility of interaction are associated with the concept of virtual learning environment. A virtual learning environment is a software platform available via the Internet that provides students digital solutions that enrich learning. Students can access required materials or data regardless of time and location and share knowledge through online communication applications [9]. There are at least three conditions that should be met for a virtual learning form to be successful: (1) the availability of the Internet, (2) the availability of support to maintain the system, and (3) the availability of a teacher who can help students in situations where they have learning difficulties [10]. In addition to the virtual learning environment, there is also a traditional learning environment with a teacher who teaches the material, i.e., explains the concept of knowledge while the students listen and take notes [9]. In this form of teaching, the teacher is the focus, and the information is delivered in only one way (through a traditional lecture), while in the virtual form of learning, the students are the focus, and the desired information can be delivered in different ways through lectures and various practical and visual representations [9]. In addition to traditional and virtual learning environments, the so-called combined form of learning can also be found in literature. The combined or hybrid form of learning is not just a simple mixture of both mentioned forms of learning, which complements the traditional form of learning with a virtual one and vice versa, but refers to their integration, whereby what is considered the best in both forms is combined, such as, for example, the integration of technology such as of augmented reality in classes that take place physically in classrooms [11].

A. Holographic technologies

Holographic projection is a technological novelty that will bring great changes in many fields: Education, science, art, and health are just a few. To understand how a holographic projector works, it is necessary to define a hologram. Holography is a method of recording patterns of light. These patterns are reproduced as a three-dimensional image called a hologram. There are three types of holograms: Reflection, Transmission, and Hybrid. Each type of hologram has its own characteristics, and its use depends on the type of object to be represented [12].

Authors in paper [13] describes the advantages of using visualization technologies with emphasis on holographic technology in education. The use of these technologies involves performing various activities in a risk-free environment, facilitates mutual collaboration and communication between students and teachers, and helps visualize abstract ideas and concepts. In addition, teachers are positive about the potential use of technologies to visualize classroom content. A study of 400 teachers in the United Kingdom examined the effectiveness of holographic technology as a teaching tool. The research results show that the majority of respondents believe that holographic technology will improve learning and will be an effective supplement to traditional teaching in the future [14]. While holographic technology offers positive learning experiences, more studies are needed in order to find ways to fully exploit their educational potential [15] .

Analysis of the application of holograms in education [2] showed that holograms have a positive effect on the cognitive development of students. Holographic projections attract the attention of students because they allow them to study a subject from different angles, which indirectly results in an increase in the understanding of the material. The research [16] mentions the lack of infrastructure, high costs of implementation and lack of experts as the main disadvantages of the application of this technology. For this reason, the authors in [17] present an example of a pyramidal hologram that would be used for educational purposes, and whose surface is made of glass with a base containing a mirror. The hologram is placed on the screen of the smartphone, which displays a three-dimensional object whose reflection will be reflected from the mirror of the pyramid and the object will be displayed as a hologram. In addition to holograms, they also propose the use of hologram control with the help of hand movements (Hand Gesture Control). Holograms often come in combination with augmented reality [18]. People who have learning or reading difficulties, using wearable devices such as HoloLens virtual reality glasses, can visualize and represent the described object or location in the passages of the books they are reading with the help of holograms.

B. Virtual and Augmented Reality

Virtual Reality (VR) is an environment that focuses on visual and immersive experiences. Through the use of interactive simulations, users are given the opportunity to immerse themselves in an environment where objects and events from the real world can appear, giving the user a subjective sense of presence [19]. Virtual reality has reached a broad user base that includes people with disabilities. In contrast to VR, augmented reality (AR) is based on augmenting the real objects around us using technology. Augmented reality can be interpreted as a subcategory of the virtual environment, where the image of the virtual world is blended with the image of the real world [20].

The use of VR and AR is most widely recognized in medicine, where it is most often used to learn anatomy [21]–[24] and to train skills such as resuscitation and, for example, as an aid in preparing for surgery [25]. The use of these technologies has also been recognized in music education , as an aid in learning material in the faculty of civil engineering [26] and graphics [27], and as a solution for learning computer security [28].

According to authors in [29], the goal of augmented reality is to use 3D virtual objects as a tool to enhance user perception and interaction with the real world by seamlessly inserting 3D virtual objects into the 3D environment of the real world. AR Technologies can be designed to interact across many sensory channels (e.g., auditory, visual, auditory, and haptic), so definitions that focus only on visual data are insufficient for future augmented reality development. Paper [30] offer an organizational chart that divides the applications of AR

into three main categories: Presentation and Visualization, Industry, and "Edutainment". The term edutainment was created by combining the words education and entertainment and includes anything that is both educational and entertaining [30].

Game-based learning with augmented reality has become a new trend in education because it has the potential to increase children's motivation to learn subjects such as mathematics. However, to be maximally effective for children's learning, the applications of AR must be appropriately designed, especially considering their new interaction and representation paradigms [31].

Teachers often use games to help students understand instructional concepts. Using technologies such as VR and AR, games set in the real world and enriched with digital data can provide powerful new ways for teachers to show relationships and connections between things. Games that use marker technology often consist of a flat game board or map that transforms into a 3D environment when viewed from a mobile device or webcam. A game like this could easily be used in a variety of subjects, including archaeology, history, anthropology, geography, and many others [30].

Related to the topic of accessibility and use of technologies for people with disabilities, in [32] the authors describe the development of a user interface for the system LIRKIS CAVE, which is used in the education of people with limited intellectual functions. CAVE stands for a room-sized space in which an image of the virtual world is projected on the walls and ceiling, where users can move freely. In such environments, users can often interact with the system through non-touch hand gestures or touch screens. This type of interaction can greatly facilitate use by people with limited intellectual abilities, as they are often unable to use a mouse and keyboard properly. The use of technologies for VR and AR has shown promising results in the education of children with autism spectrum disorders [33], where it has been demonstrated that these technologies help in independent learning and maintaining concentration.

III. HOLOGEOMETRY

HoloGeometry is application designed to help students learn geometric objects. In geometry, it is important to consider the phenomenon of ostensive, which is an obstacle for students. This phenomenon results from the representation of figures and geometric solids in textbooks from a single perspective, which means that students are not able to apply the knowledge they have learned when they move to a different spatial orientation, i.e., they are not able to extrapolate. This application provides them with a visualization of geometric solids in space with the possibility of interaction. It meets the needs of learners who process information primarily visually and enhances learning for all students. This product enables students to clearly see and understand geometric objects and improves tactile hand-eye-mind connections that enhance the ability to memorize facts and retain what they learn.

Target group are all people who like mathematics and want to learn geometry. The product is intended to help learning, which also leads to better social inclusion of people with Down syndrome. Mathematics education usually focuses on arithmetic because people with learning disabilities, including people with Down syndrome and autism. These mathematical concepts are undoubtedly important for understanding the world quantitatively, and they also enhance communication skills and the ability to think, reason, and learn. The abstract ideas that geometry conveys from direct experience make the world more understandable to them, so it is worth developing a method to teach them geometry. Mathematics, especially geometry, was developed by humans specifically to embody abstract ideas from experience. Therefore, choosing this subject as a way to better understand the world could particularly enhance children's cognitive strengths [34].

The application was developed for computers and mobile devices and uses holographic technology for easier visualization of geometric solids. Accessibility was considered during its development, e.g., every geometric solid has both a written and an audio description. It is also possible to change the font to one suitable for people with dyslexia. The Unity development environment[2] was used to develop the application. The Unity asset "Hologram pyramid"[3] was used to implement the hologram in the application. The application includes three levels: (1) “Learn geometrical shapes”, (2) “Recognize geometrical shapes” and (3) “Test your knowledge”. (Figure 1).

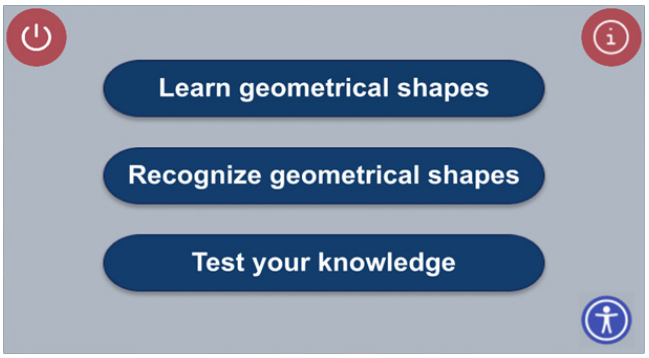


Figure 1 Initial screen of HoloGeometry application

The first level is about learning geometric solids. Users can familiarize themselves with 4 different geometric solids: a sphere, a pyramid, a cylinder, and a cube (see Figure 2)

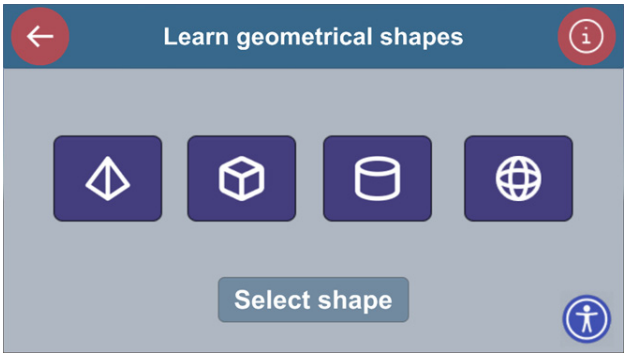


Figure 2 First level “Learn geometrical shapes”

[2] Unity, https://unity.com/download
[3] Unity Asset “Hologram Pyramid“, https://assetstore.unity.com/packages/tools/hologram-pyramid-61735

For each geometric body there is a written and audio description. The hologram can be zoomed in and out and rotated. It is also possible to change the color of the hologram. All these customization options serve to give the user a better idea of the 3D shape. Example of the design of description of one geometrical solid, in this case cube, is shown in Figure 3.

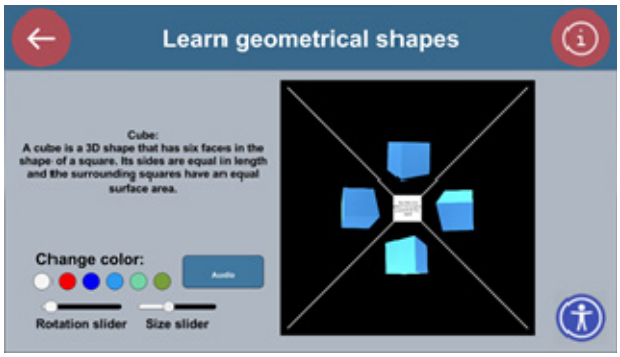


Figure 3 Example of description and interaction possibilities of Cube

The second level is similar to the first, but it is used to determine the knowledge acquired in the first level. The screen is designed to display the hologram in the center. On the left side of the hologram are buttons to interact with the hologram, and on the right side are the answers offered. The screen of this game is shown in Figure 4, and the appearance of the hologram can be seen in Figure 5.

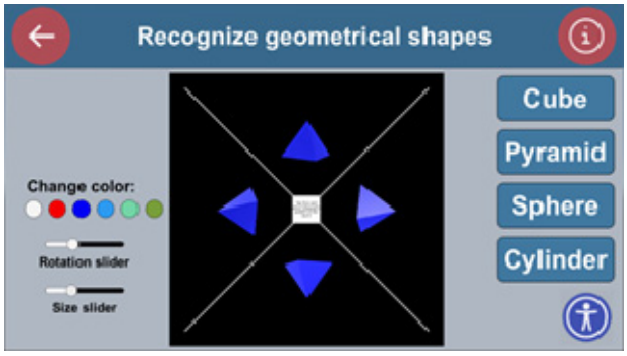


Figure 4 Second level “Recognize geometrical shapes”

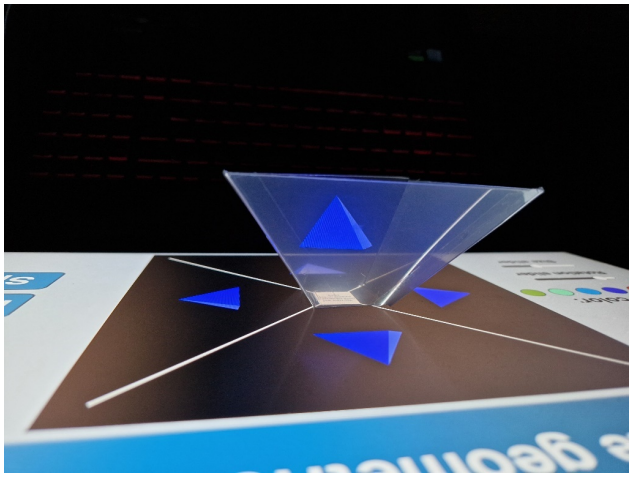


Figure 5 Holographic display of pyramid

Selecting the third level, "Test your knowledge", starts a quiz with ten random questions about geometry and geometric solids. The user is given one question and four suggested answers, only one of which is correct. If the user selects an incorrect answer, the selected answer turns red. If the user selects the correct answer, the selected answer turns green. At the end of the quiz, statistics are displayed with the number and percentage of correctly answered questions. The screen of this level of the game is shown in Figure 6. Figure 7 shows how the statistics look at the end of the third level.

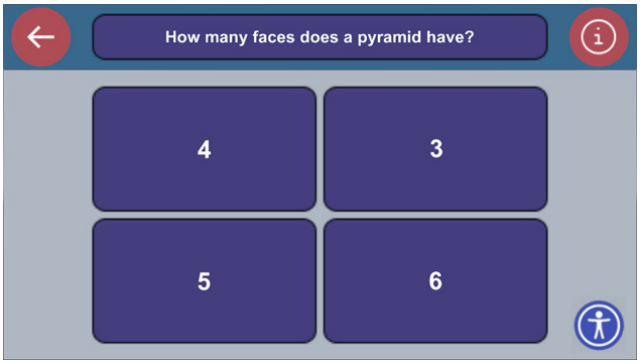


Figure 6 Third level “Test your knowledge”

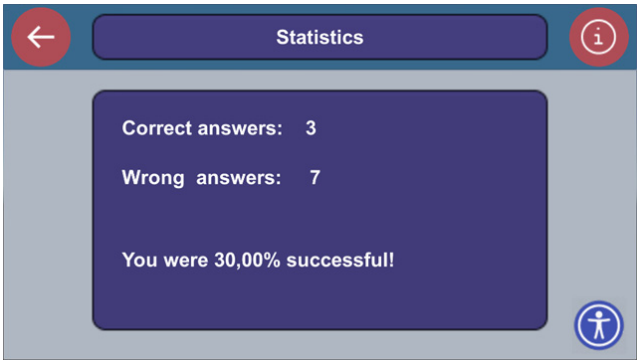


Figure 7 Example of statistics

Since the application was designed from the beginning to be accessible to as many people as possible, some accessibility options were introduced. It is possible to change the color of the hologram. The possible colors are white, red, dark blue, light blue, light green and dark green. It is also possible to rotate the hologram to get a better perception of the three-dimensional shape. The hologram can be zoomed in and out so that the application can be used smoothly on both small and large screens. It is also possible to listen to an audio recording with a description of a geometric body. In addition to these options, there is an accessibility options button in the lower right corner of all screens. When clicked, all text in the application is displayed in a font for people with dyslexia (see Figure 9).

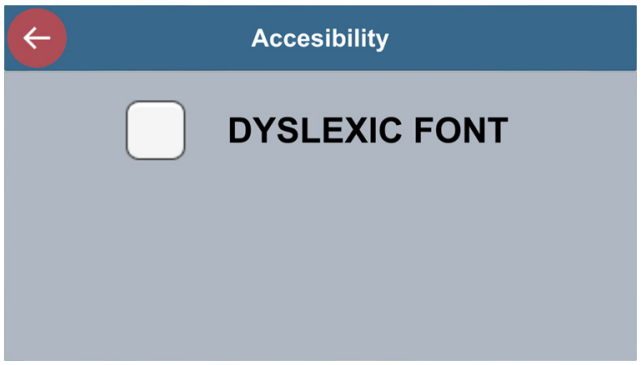


Figure 8 Accessibility option menu

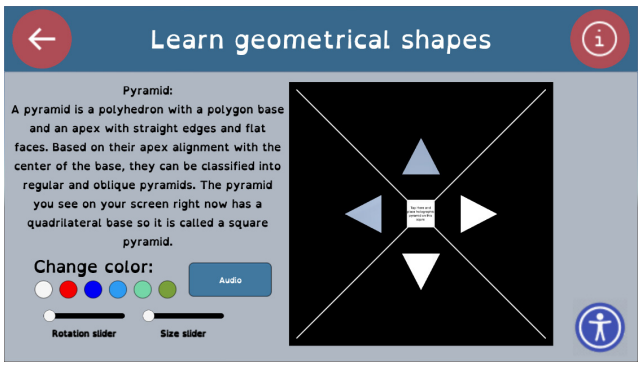


Figure 9 Screen design with dyslexic font enabled

IV. CONLCUSION

Visualization technologies have been around for many years, but only recently they have become prevalent in marketing, telecommunications, education, etc., with an emphasis on holographic technology. The application of this technology stands out in education, where it has been shown to enable students to understand concepts faster and better. Visualization technology in the classroom has a stimulating effect and attracts the attention of students. Similarly, these technologies have been found to positively impact the educational experiences of students with disabilities, increasing confidence, engagement, and interest, providing opportunities for independent learning, and increasing student satisfaction and motivation. Thanks to the development of these technologies, we now have the opportunity to learn and research in the virtual world without the possibility of our actions and decisions affecting the real world. Numerous industries use the virtual world to perfect their products or even experiment to the last detail before the process is carried out in reality. In this way, costs and various risks that could arise are significantly reduced. The development of holographic projection has also expanded its scope. Holograms are certainly an interesting group of visualization technologies that have proven to be an attractive and useful solution, especially in the educational field. An example of a holographic application that helps in learning geometric solids is HoloGeometry.

Given that HoloGeometry is currently in its prototype stage, in forthcoming iterations of the application's development, certain metrics that assess user performance will be incorporated, and the application will be evaluated with end users accordingly.

REFERENCES

[1] S. Dineva, "The importance of visualization in e-learning courses," *dalam 14 Int. Conf. Virtual Learn. ICVL*, vol. 1, no. 1, 2019.

[2] N. M. M. Barkhaya and N. D. A. Halim, "A review of application of 3D hologram in education: A metaanalysis," *2016 IEEE 8th Int. Conf. Eng. Educ. Enhancing Eng. Educ. Through Acad. Collab. ICEED 2016*, pp. 257–260, 2017.

[3] K. P. Vinumol, A. Chowdhury, R. Kambam, and V. Muralidharan, "Augmented reality based interactive text book: An assistive technology for students with learning disability," *Proc. - 2013 15th Symp. Virtual Augment. Reality, SVR 2013*, pp. 232–235, 2013.

[4] Ž. Car, "Ergonomija računalne opreme," *Nastavni Mater.*, 2020.

[5] A. R. Teyseyre and M. R. Campo, "An overview of 3D software visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 1, pp. 87–105, 2009.

[6] D. M. Elmahal, A. S. Ahmad, A. T. Alomaier, R. F. Abdlfatah, and D. M. Hussein, "Comparative study between hologram technology and augmented reality," *J. Inf. Technol. Manag.*, vol. 12, no. 2, pp. 90–106, 2020.

[7] N. Ali, S. Ullah, and Z. Musa, "The Effect of List-Liner-Based Interaction Technique in a 3D Interactive Virtual Biological Learning Environment," pp. 297–317, 2020.

[8] C. Neto, "Developing Interaction 3D Models for E-Learning Applications," pp. 0–15, 2009.

[9] S. H. Halili, N. H. A. Rahman, and R. A. Razak, "Traditional versus virtual learning: How engaged are the students in learning English literature?," *Indones. J. Appl. Linguist.*, vol. 8, no. 1, pp. 79–90, 2018.

[10] A. N. Cahyono and M. Asikin, "Hybrid learning in mathematics education: How can it work?," *J. Phys. Conf. Ser.*, vol. 1321, no. 3, 2019.

[11] D. Sonntag, G. Albuquerque, M. Magnor, and O. Bodensiek, "Hybrid learning environments by data-driven augmented reality," *Procedia Manuf.*, vol. 31, pp. 32–37, 2019.

[12] A. Elmorshidy, "Holographic Projection Technology: The World is Changing," vol. 2, no. 2, 2010.

[13] B. Eschenbrenner, F. F.-H. Nah, and K. Siau, "3-D Virtual Worlds in Education," *J. Database Manag.*, vol. 19, no. 4, pp. 91–110, 2011.

[14] H. Ghuloum, "3D Hologram Technology in Learning Environment," *Proc. 2010 InSITE Conf.*, pp. 693–704, 2010.

[15] E. Fokides and I. A. Bampoukli, "Are hologram-like pyramid projections of an educational value? Results of a project in primary school settings," *J. Comput. Educ.*, no. 0123456789, 2022.

[16] P. Kalansooriya, A. Marasinghe, and K. M. D. N. Bandara, "Assessing the Applicability of 3D Holographic Technology as an Enhanced Technology for Distance Learning," *IAFOR J. Educ.*, vol. 3, no. SE, 2015.

[17] A. H. Awad and F. F. Kharbat, "The first design of a smart hologram for teaching," *2018 Adv. Sci. Eng. Technol. Int. Conf. ASET 2018*, pp. 1–4, 2018.

[18] R. Klamma, *Perspectives on Wearable Enhanced Learning (WELL)*. 2019.

[19] M. Teófilo, J. Nascimento, J. Santos, Y. Albuquerque, A. L. Souza, and D. Nogueira, "Bringing basic accessibility features to virtual reality context," *Proc. - IEEE Virtual Real.*, vol. 2016-July, pp. 293–294, 2016.

[20] M. Pandžić, Igor Sunday; Pejša, Tomisla; Matković, Krešimir; Benko, Hrvoje; Čereković, Aleksandra; Matijašević, *Virtualna okruženja: Interaktivna 3D grafika i njene primjene*. 2011.

[21] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab, "Mirracle: An augmented reality magic mirror system for anatomy education," *Proc. - IEEE Virtual Real.*, pp. 115–116, 2012.

[22] W. Si, X. Liao, Q. Wang, and P. A. Heng, "Augmented Reality-Based Personalized Virtual Operative Anatomy for Neurosurgical Guidance and Training," *25th IEEE Conf.*

Virtual Real. 3D User Interfaces, VR 2018 - Proc., pp. 683–684, 2018.

[23] A. Heather, T. Chinnah, and V. Devaraj, "The Use of Virtual and Augmented Reality in Anatomy Teaching," *MedEdPublish*, vol. 8, no. 2, 2019.

[24] D. Eckhoff, C. Sandor, D. Kalkoten, U. Eck, C. Lins, and A. Hein, "TutAR: Semi-Automatic Generation of Augmented Reality Tutorials for Medical Education," *Adjun. Proc. - 2018 IEEE Int. Symp. Mix. Augment. Reality, ISMAR-Adjunct 2018*, pp. 430–431, 2018.

[25] J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, F. J. Toledo-Moreo, and H. Adeli, "Bioinspired Computation in Artificial Systems: International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2015, Elche, Spain, June 1-5, 2015, Proceedings, Part II," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9108, pp. 359–360, 2015.

[26] F. M. Dinis, A. S. Guimaraes, B. R. Carvalho, and J. P. P. Martins, "Virtual and augmented reality game-based applications to civil engineering education," *IEEE Glob. Eng. Educ. Conf. EDUCON*, no. April, pp. 1683–1688, 2017.

[27] H. Chen, K. Feng, C. Mo, S. Cheng, Z. Guo, and Y. Huang, "Application of augmented reality in engineering graphics education," *ITME 2011 - Proc. 2011 IEEE Int. Symp. IT Med. Educ.*, vol. 2, pp. 362–365, 2011.

[28] N. Kommera, F. Kaleem, and S. M. S. Harooni, "Smart augmented reality glasses in cybersecurity and forensic education," *IEEE Int. Conf. Intell. Secur. Informatics Cybersecurity Big Data, ISI 2016*, pp. 279–281, 2016.

[29] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Comput. Graph. Appl.*, vol. 21, no. 6, pp. 34–47, 2001.

[30] S. C.-Y. Yuen, G. Yaoyuneyong, and E. Johnson, "Augmented Reality: An Overview and Five Directions for AR in Education," *J. Educ. Technol. Dev. Exch.*, vol. 4, no. 1, 2011.

[31] M. Weerasinghe and A. Quigley, *Augmented Reality Games II*, no. September. 2019.

[32] B. Sobota, S. Korecko, L. Jacho, P. Pastornicky, M. Hudak, and M. Sivy, "Virtual-reality technologies and smart environments in the process of disabled people education," *ICETA 2017 - 15th IEEE Int. Conf. Emerg. eLearning Technol. Appl. Proc.*, 2017.

[33] C. S. Lányi, Z. Geiszt, P. Károlyi, Á. Tilinger, and V. Magyar, "Virtual Reality in Special Needs Early Education," *Int. J. Virtual Real.*, vol. 5, no. 4, pp. 55–68, 2006.

[34] A. Keselj, "Case Study 2: HoloGeometry," 2022. [Online]. Available: http://sociallab.tel.fer.hr/innosid/valencia-2022-hackathon/case-studies/hologeometry/. [Accessed: 16-Feb-2022].

COMMUNICATION PROTOCOLS AND SYSTEM PERFORMANCE

An analysis of the TR069 (CWMP) protocol

Ilija Basicevic
Faculty Of Technical Science
University Of Novi Sad
Novi Sad, Serbia
ilibas@uns.ac.rs

***Abstract*—This paper presents a short analysis of the TR-069 (CWMP) protocol. The protocol is widely used in consumer electronics and the Internet of Things (IoT), and this paper is an attempt to explain its success. The purpose of the protocol, its architecture, integration with other protocols, and the technologies that it relies on are discussed. A parallel is drawn with some other management protocols, primarily the Simple Network Management Protocol (SNMP), where possible.**

***Keywords*—device management, TR-069, remote monitoring, Internet management**

I. INTRODUCTION

TR-069 is an established protocol for device configuration and monitoring in the consumer electronics field. As this field has grown significantly in the last two decades, the spectrum of devices that can be managed by TR-069 has also widened. For example, TV operator probably uses this protocol to monitor set top box at the reader's home. The protocol is published by the Broadband Forum [1]. The first version appeared in 2004 and the current version is 1.4.

The two sides in the TR-069 communication are CPE (Customer Premises Equipment) and ACS (Auto-Configuration Server), see Fig. 1. CPE is a managed device that contains at least one CWMP endpoint and is located in the end user's network. The ACS performs auto-configuration and other management functions on the CPE over the broadband connection. The ACS is located in the operator network or datacentre.

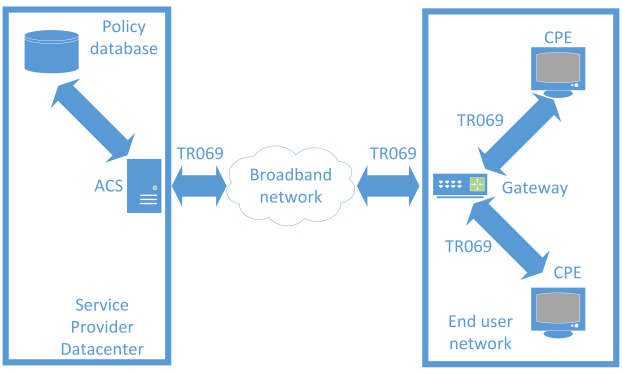


Fig. 1: TR069 network topology.

The protocol integrates the following functions: auto-configuration and dynamic provisioning, management of software and firmware, monitoring of device state and performance, and diagnostics. The protocol enables an ACS to provision a CPE (or a set of CPEs) based on different criteria. It also allows for the control of downloads of the CPE software/firmware image files. The download can be initiated by both sides. There are also mechanisms to manage modular software and execution environment on a CPE such as installation, update, uninstallation and inventory of software modules. The TR-069 protocol includes mechanisms which enable an ACS to monitor the status and performance statistics of a CPE. Also, it is important that this protocol enables a CPE to send information to an ACS which makes it possible for the ACS to diagnose and resolve different operational issues. The last feature is used when the aforementioned set top box malfunctions and the reader calls the operator's call center for technical support.

II. RELATED WORK

The Internet standard for device management and monitoring is SNMP [2]. Thus, the TR-069 features overlap with those of the SNMP, but TR-069 has been developed two decades later and consequently uses more modern technologies than SNMP, such as HTTP, XML, and SOAP. It is also less general than SNMP, as it is custom tailored for scenarios in consumer electronics systems. Its architecture is simpler compared to SNMP, which is probably one of the reasons for its success. Both TR-069 and SNMP are application level protocols. That fact, among other things, enables protocol operation (e.g. device management), when the managed device and the management agent are not in the same network (there can be several hops between them). (Early WAN management protocols have been situated on data link level, which has been possible in the environment with uniform network technology - levels 1 and 2 [3]).

Network Configuration Protocol (NETCONF) is a network management protocol developed by IETF. The current version is published in [4] and has the status of an Internet standard. The protocol has been developed as a result of a search in IETF for a device network configuration solution (see [5]), as it has been concluded that SNMP serves well for device monitoring, but for device configuration the network operators often use proprietary solutions.

YANG [6] is a data modelling language used for modelling configuration and state data. It is popular today and supports retrieving information using NETCONF and RESTCONF [7]. The language provides descriptions of network nodes and their interactions. In combination with NETCONF, YANG is a tool that administrators can use to automate configuration tasks in a heterogeneous network.

Open Mobile Alliance Device Management (OMA DM) [8] is a device management protocol designed for the management of mobile devices such as mobile phones, PDAs and tablet computers. The protocol is specified by OMA Device Management working group as well as the Data Synchronization working group. It supports the following operations: provisioning, data configuration, software upgrade and fault management.

OMA Lightweight Machine to Machine (LwM2M) [9] is a protocol from OMA alliance for machine to machine and IoT device management. Originally it used Constrained Application Protocol (CoAP) for transport but later versions included support for other protocols (UDP, TCP, TLS, Message Queueing Telemetry Transport - MQTT).

III. THE PROTOCOL STACK

The TR-069 protocol stack is presented in Fig. 2. The foundation is TCP/IP protocol, which provides transport functionality. (As mentioned later, for notification purpose and for sending the connection request, TR-069 can use UDP, too). Above TCP/IP is SSL/TLS which provides security. Residing on SSL/TLS is HTTP which provides client/server transactions that are the building elements of the TR-069 dialogue, corresponding to the session layer of the ISO OSI stack. Above HTTP is SOAP which provides transfer syntax – thus positioning it it in the presentation layer. Above SOAP are the remote procedure call (RPC) methods. Having in mind the ISO OSI application level architecture, the RPC methods would correspond to a Specific Application Service Element (SASE) in ISO OSI. On the top is the TR-069 management application (CPE or ACS) - the user application process. Thus the application is according to the ISO OSI in the local system environment (LSE), and the stack below is in the OSI environment (OSIE).

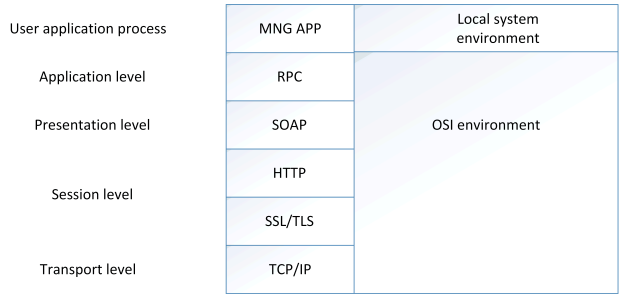


Fig. 2: TR-069 protocol stack.

IV. INTEGRATION WITH OTHER PROTOCOLS

Requests and responses are sent between a CPE and an ACS using HTTP or HTTPS, in the form of remote procedure calls (RPC), while SOAP 1.1 is used as the transfer syntax. Both CPE and ACS can take the role of client or server, but an interesting protocol design decision is that the communication is always started by a CPE. When needed, an ACS can send an asynchronous connection request which will result in the CPE reporting to the ACS, though. This is used when near-real-time reconfiguration of a CPE is required. For example, if the user subscribes to a service, that service becomes immediately accessible to them, without waiting for the next periodic contact. (But in that case, the TR-069 session is also started by the CPE.)

In certain cases, it is not possible for an ACS to send a connection request to a CPE directly. If a CPE is behind a firewall or a gateway, which prevents direct communication from an ACS to a CPE, TR-069 makes use of the XMPP protocol. Both the ACS and the CPE register on the XMPP server, and then the ACS can use it to relay the connection request (in the form of XMPP IQ Stanza) to the CPE. The CPE and the ACS do not have to be registered to the same XMPP server, it can be a cluster of XMPP servers instead.

Alternatively, if XMPP is not available, TR-069 can use Universal Plug and Play Internet Gateway Device framework (UPnP IGD). Using that framework a CPE can obtain a WAN IP address and add a port mapping on the gateway to enable traversal of the connection request from the ACS.

Still another possibility is the use of STUN, when a CPE is behind a NAT gateway. This is another case when an ACS cannot send the asynchronous connection request to an CPE (see also [10]). The standard prescribes the following procedure. When a CPE discovers (using STUN) that it is behind a NAT gateway with private addresses, it is required to keep open a NAT binding (by sending periodic STUN binding requests, based on the binding timeout it has discovered previously), and send to the ACS the public IP address and port associated with the binding. The ACS sends UDP Connection Request to the CPE using that address information when required.

In these cases, the gateway or firewall does not have to be managed by TR-069.

TR-069 can be integrated with DHCP, in a way that in the context of DHCP Offer message from the DHCP server, the CPE receives not only the new address it will use but also the address of the ACS server. Also, when sending the Inform message, the CPE can include the identification of the DHCP server, which can be checked later in communication between the ACS and the managed gateway. It is recommended that when a CPE device reports for the first time it is connected to a specific gateway that the ACS cross-checks that information either by explicitly soliciting the information from the gateway,

or by using its own table - in the case when it previously subscribed for event information from that gateway. The next time, and any subsequent time when the CPE reports that it has been connected to that gateway, the cross check is not necessary.

TR-069 has a proxy feature that further extends its applicability by allowing non-TR-069 devices to be managed by an ACS through a TR-069 CPE acting as a proxy. Two types of proxied devices are distinguished. In the case when simple ones, such as power switches and binary sensors are used, a CPE proxy contains an embedded object in its data model. More complex devices, such as routers, and set top boxes, are presented by having a TR-069 CPE device in the CPE proxy.

TR-069 communication is secured using the TLS protocol (as is the case with SNMPv3 and LwM2M as well). This protocol is also used for authentication. If TLS authentication is not used, an ACS must authenticate a CPE using HTTP authentication. If TLS is used for encryption, an ACS should use the basic authentication scheme whereas if TLS is not used, an ACS must use digest authentication. The ACS chooses the authentication scheme by providing a basic or digest authentication challenge.

Although two sides authenticate each other, the operators can additionally restrict the IP addresses from which a connection request can reach the CPE, by using a firewall or setting the routing table. Since linux operating system is common in this field, iptables can be used for this restriction.

V. COMMUNICATION LIFE-CYCLE

The life cycle of a TR-069 communication between the two endpoints is as follows. The first session is initiated by the CPE. We can say that the CPE reports for the first time to the ACS. The CPE uses the ACS URL it obtained through the configuration, or through DHCP address allocation (explained earlier in the text). In the first session, the CPE sends to the ACS the management URL (in the ManagementServer.ConnectionRequestURL parameter). The ACS can program future communication (using the ScheduleInform). Also, the ACS can afterwards at any time use the CPE management URL to send the connection request. If the CPE management URL changes, it is mandatory for the CPE to report that to the ACS.

The MSC of an example TR-069 session is given in Fig. 3. It can be seen that the CPE opens the TCP connection to the ACS. The SSL initiation follows. After that, the CPE sends Inform request in the HTTP Post method. The ACS sends an Inform response in the HTTP response. After the TR-069 session has thus started, the CPE sends an empty HTTP Post message, only to make it possible for the ACS to send a TR-069 request, in an HTTP response. In this example, the ACS sends a GetParameterAttributes request. In the next step, the CPE sends a TR-069 GetParameterAttributes response, in an HTTP Post message. It can be seen how HTTP and TR-069 transactions overlap - the CPE sends an HTTP Post request carrying TR-069 GetParameterAttributes response, and the ACS sends an HTTP response carrying the next TR-069 request, and so on. In the last step, the ACS sends an empty HTTP response, which is a signal to the CPE to close the TCP connection (and the TR-069 session).

To maintain the session state, the ACS uses a session cookie. It is recommended that for a sequence of transactions that comprise a single session, the CPE maintains a single TCP connection that persists throughout the duration of the session.

TR-069 can use port 7547 which has been assigned by IANA for the CPE WAN management protocol.

VI. THE PROTOCOL SYNTAX

When it appeared, SNMP inherited the fetch-store paradigm from an earlier Internet management protocol High-Level Entity Management System (HEMS) [11]. It was a novel paradigm in the field of device management protocols [3]. Instead of having a specific request type for each operation type, it has a set of generic requests and responses – primarily for getting and setting the values of data model parameters. The specific operation that will be initiated depends on the variable onto which the generic request is applied. For example, if a device reboot is required, a set request with the value 0 would be applied to the variable which contains the time till the reboot, and consequently, reboot will immediately be started. This paradigm results in a protocol design that is extendable and stable, with a small set of commands. Consequently, the command set of SNMP comprises the generic commands, such as GetRequest for retrieving a value, GetNextRequest for iterative retrieval, SetRequest for setting the value of a parameter, Trap, and InformRequest which are used for event notification, and Response which is a response to any of the aforementioned commands, except Trap for which no response is expected.

On the other hand, in [5] it has been noted that there is a semantic mismatch between the data centric model of SNMP and the task oriented model which is preferred by human operators. This resulted in the development of a non-trivial code which performs mapping between these two models and is integrated into management applications.

The TR-069 has a kind of hybrid model. There is a set of generic commands: Get/SetParameterValues, Get/SetParameterAttributes, Add/DeleteObject, but it also contains commands for some specific operations, such as Reboot, FactoryReset, Download, ScheduleDownload or Upload.

A CPE connects to the ACS on initial installation, after reset, after a defined timeout, by schedule defined by the ACS in ScheduleInform, upon a connection request from the ACS, upon a change of the ACS URL, when the values of parameters for which the ACS requested monitoring have been changed by a third party, upon the

termination of file up/download – if requested in the CPE configuration.

The CPE-ACS session starts with the CPE sending an Inform message. It contains connection reason, current time stamp on the CPE, the number of attempts, and the values of parameters for which ACS requested monitoring, and which were modified by a third party since the last connection.

The baseline data model template is given in the TR-106 standard [12], and it is specified in XML (YANG and OMA DM are also XML based). There are additional standards, specific to the field of the consumer system (e.g. TR-135 [13] for set top boxes). Also, as noted, TR-069 makes use of SOAP, and SOAP is XML based. On the other hand, in the case of SNMP, the standard for data is MIB (Management Information Base), which specifies that the information is encoded using ASN.1. An important difference between these two approaches is that the XML model in TR-069 and YANG is readable by humans, while ASN.1 binary coded data are not. A similar differentiation existed in the 2000s in the VoIP world [14], at the time of market competition between H.323 and SIP protocols, as H.323 uses ASN.1, while SIP messages contain text coded in UTF-8 and are based on HTTP.

TABLE I: TR-069 Methods

Mandatory Methods	
CPE Methods	
	GetRPCMethods GetParameterValues SetParameterValues GetParameterAttributes SetParameterAttributes AddObject DeleteObject Download Reboot
ACS Methods	
	Inform TransferComplete AutonomousTransferComplete
Optional Methods	
CPE Methods	
	ScheduleInform Upload FactoryReset GetAllQueuedTransfers ScheduleDownload CancelTransfer ChangeDUState
ACS Methods	
	RequestDownload DUStateChangeComplete AutonomousDUStateChangeComplete

VII. THE PROTOCOL METHODS

TR-069 methods are given in Table I.

A. Mandatory Methods

GetRPCMethods is a method that can be called either by a CPE or an ACS, to discover the set of RPC methods supported by the other side in the communication.

a) CPE Methods: The CPE methods can be used by an ACS to perform different operations, such as obtaining and modifying values of parameters on CPE (GetParameterValues and SetParameterValues), discovering which parameters are existing on a CPE (GetParameterNames), reading and obtaining attributes (which refer to the access control information including notification type) associated with specific parameters on a CPE (GetParameterAttributes and SetParameterAttributes).

There are also methods that are used to create or delete an instance of a multi-instance object (AddObject and DeleteObject). For example, if there is an object Top.Group.Object, AddObject can be used to create a new parameter Top.Group.Object.i.Parameter, where i is the return value of AddObject method. Afterwards, Top.Group.Object.i. can be passed as an argument to DeleteObject, to delete that instance of the object.

Download method is used by the ACS to cause the CPE to download a specified file from the designated location.

Reboot method causes the CPE to reboot.

b) ACS Methods: We have already mentioned the Inform method which a CPE must call to initiate a transaction sequence whenever a session with an ACS is established. Whenever a file transfer requested by the ACS (by an earlier Download, ScheduleDownload or Upload method call) is completed, the CPE calls the TransferComplete method. Whenever a file transfer which was not requested by the ACS is completed, the CPE calls the AutonomousTransferComplete method.

B. Optional Methods

a) CPE Methods: An ACS can request from a CPE to send an Inform method sometime in the future (independent from periodic Inform method calls) using ScheduleInform. Upload method is used by the ACS to request the CPE to upload a specified file to the specified location. FactoryReset method is used by the ACS to reset the CPE to its factory default state.

When necessary an ACS can request from a CPE to send the status of all queued downloads and uploads, including autonomous transfers, using GetAllQueuedTransfers method.

ScheduleDownload is an advanced version of Download method, as it allows an ACS to specify one or two time windows in which download should be performed. File transfers requested by the ACS (using Download, ScheduleDownload or Upload method call) can be cancelled with CancelTransfer method.

An ACS can request installing a new Deployment Unit (DU), updating an existing DU, or uninstalling an existing DU, using ChangeDUState method.

b) ACS Methods: Using the RequestDownload method, a CPE can request a file download from the ACS, which may result in the ACS calling the Download method to initiate the download.

There are also two methods related to the DU state changes, which are used by a CPE to inform the ACS about the completion (successful or unsuccessful) of a DU state change. DUStateChangeComplete method is used if the state was requested by an ACS, and AutonomousDUStateChangeComplete if it was not.

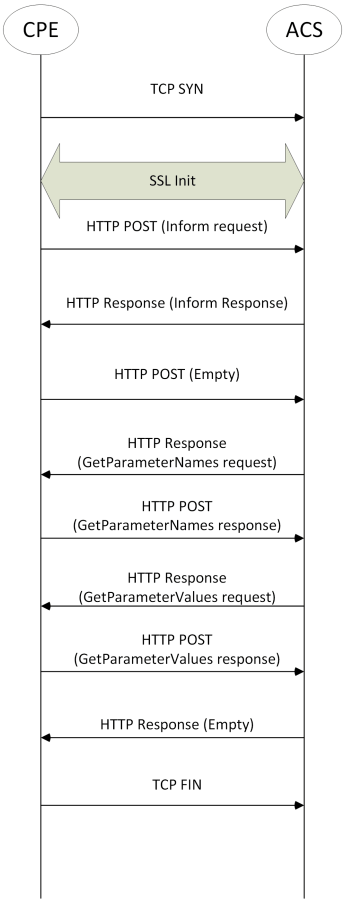


Fig. 3: Message sequence chart of a TR-069 session.

VIII. MECHANISMS FOR MONITORING OF CPE PARAMETERS

There are several levels of notification on a CPE, which is defined for each parameter. The notification can be disabled, or it can be passive, active, lightweight passive or lightweight active. Thus, lightweight notification can be combined with active/passive notification as well. Passive notification means that when a parameter value is changed, the new value will be sent to the ACS in the next Inform which is sent (and which was scheduled earlier or is a periodic request). Active notification means that when a parameter value changes that would cause the establishment of a new session and sending of the new parameter value in the Inform. Lightweight notification is a UDP based mechanism which is not reliable, and complements the existing notification mechanisms.

Two message types are defined in the SNMP protocol for asynchronous event notification. One is Trap, used for reporting without confirmation, and the other is InformRequest for confirmed notification delivery.

IX. THE PERSPECTIVES FOR TR-069 IN TODAY'S COMPUTING WORLD

As noted, TR-069 has already claimed an important position in the consumer electronics field in the previous period. The use of this protocol in the management of set top box devices was mentioned throughout the text, see also [15], [16] for more about this topic. Refs. [17] and [18] present an implementation of the protocol.

In a race with some other protocols (e.g. OMA-DM [8], LwM2M [9]) TR-069 is more and more often becoming the protocol of choice for device management in IoT systems ([19], [20], [21]) which is very important for the future of this protocol. An important segment of IoT is the smart home gateway, also named Machine-Type Communication Gateway (MTCG). MTCG acts as a bridge between smart objects and the Internet. Due to the significant expansion of the IoT domain, a separate configuration of each device is not viable anymore, as the number of devices is increasing. This is the place in IoT architecture where TR-069 can fit. There are also proposals for its use in LTE, [22], and Wi-Fi, [23]. In [23] a method is presented which ensures that all devices in a Wi-Fi mesh network have a consistent firmware version, which improves the stability of the system.

On the other hand, following the success of TR-069 and the sophistication of consumer network, the Broadband Forum came with its successor TR-369, [24] which allows for more complex communication scenarios (whereas in TR-069 there is one ACS that controls the CPEs, in TR-369 there can be several controllers controlling one device). TR-369 is sometimes referred to as TR-069 for IoT devices. The ability to have several controllers with different permissions controlling one device enables multiple providers, vendors and end-users to interact with the managed devices. Another difference from TR-069 is that TR-369 can be based on other protocols (Websockets, COAP, MQTT, Simple Text Oriented Messaging Protocol -STOMP), not just HTTP.

Also, an effort has been made to make less overhead in terms of communication cost, as for precise monitoring, a larger number of messages is required which incurs a network cost. With TR-369 more precise monitoring can be achieved and with less impact on network traffic than is the case with TR-069, as it is more lightweight in terms of communication cost. Among other elements, the number of communication handshakes has been reduced. The communication model used in TR-069 is based on the idea that a CPE would open a session to the ACS when required, exchange data and close the session as soon as possible. On the other hand, in TR-369, once opened a session remains open all the time.

X. CONCLUSION

This paper contains a short analysis of the TR-069 protocol, with respect to the protocol architecture, position in the protocol stack, and its reliance on the use of other Internet protocols. With respect to the protocol design, a

parallel is made to the SNMP protocol, which is the most widely known Internet management protocol.

Compared to SNMP, TR-069 uses more recent protocols and technologies, such as HTTP, SOAP and XML. Both protocols are application-level protocols. While TR-069 messages are coded in XML and readable by humans, SNMP messages are binary coded in ASN.1 and not readable by humans.

In the Internet ecosystem, there can be obstacles to the bidirectional communication between a CPE and an ACS. The standard acknowledges this situation and provides several mechanisms to overcome it.

The TR-069 protocol achieved significant success in the consumer electronics field due to a good custom tailored design and the use of relevant contemporary technologies, which among other allow for secure operation. The protocol entered the IoT field as well. However, its successor TR-369 carries some improvements that make it even more suitable for IoT applications. For many scenarios though, TR-069 is still the protocol of choice.

ACKNOWLEDGMENT

The author would like to thank Marko Vucicevic for providing advice on the integration of the TR-069 protocol.

This work was partially supported by the Secretary of Science and Technology Development of Vojvodina Province under Grant 142-451-2406/2022-01/01.

REFERENCES

[1] "TR-069 CPE WAN Management Protocol," 2020.
[2] J. Case, R. Mundy, D. Partain, and B. Stewart, "Introduction to Version 3 of the Internet-standard Network Management Framework," 1999, RFC: 2570.
[3] D. E. Comer, *Internetworking with TCP/IP Vol.1: Principles, Protocols and Architectures*. Upper Saddle River, NJ: Prentice Hall, 2000.
[4] R. Enns, M. Bjorklund, J. Schoenwaelder, and A. Bierman, "Network Configuration Protocol (NETCONF)," 2011, RFC: 6241.
[5] J. Schoenwaelder, "Overview of the 2002 IAB Network Management Workshop," 2003, RFC: 3535.
[6] M. Bjorklund, "The YANG 1.1 Data Modeling Language," 2016, RFC: 7950.
[7] A. Bierman, M. Bjorklund, and K. Watsen, "RESTCONF Protocol," 2017, RFC: 8040.
[8] "OMA Device Management V2.0," 2012.
[9] "OMA Lightweight M2M (LwM2M) V1.2," 2020.
[10] M. Savić, I. Papp, I. Rešetar, D. Majstorović, and D. Spasojević, "Implementation of server push mechanism based on tr-069 and tr-111 protocols for network device management," in *2013 21st Telecommunications Forum Telfor (TELFOR)*, 2013, pp. 361–364.
[11] C. Partridge and G. Trewitt, "THE HIGH-LEVEL ENTITY MANAGEMENT SYSTEM (HEMS)," 1987, RFC: 1021.
[12] "TR-106 – Data Model Template for CWMP Endpoints and USP Agents," 2022.
[13] "TR-135 Data Model for a TR-069 Enabled STB," 2012.
[14] I. Basicevic, M. Popovic, and D. Kukolj, "Comparison of sip and h.323 protocols," in *Proceedings of the 2008 The Third International Conference on Digital Telecommunications*, ser. ICDT '08. USA: IEEE Computer Society, 2008, p. 162–167.
[15] B. Trifunović, V. Mihailović, N. Ignjatov, R. Simikić, and I. Velikić, "Android development framework for tr-069-based services," in *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, 2014, pp. 492–494.
[16] D. Knezevic, I. Ostojic, I. Papp, and M. Savic, "Integration mechanism for live stream qos monitoring in android-based iptv set-top box," in *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, 2014, pp. 489–491.
[17] V. Mihailović, N. Ignjatov, M. Bojan, and T. Nikola, "Adaptive build system for tr-069 consumer device agent," in *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, 2014, pp. 314–317.
[18] M. Ćetkovic, N. Nemet, T. Samardžić, and N. Teslić, "Auto-configuration server architecture with device cloud cache," in *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, 2014, pp. 296–298.
[19] Y. Chahid, M. Benabdellah, and A. Azizi, "Internet of things protocols comparison, architecture, vulnerabilities and security: State of the art," in *Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems*, ser. ICCWCS'17. Association for Computing Machinery, 2017.
[20] L. Reinfurt, U. Breitenbücher, M. Falkenthal, F. Leymann, and A. Riegg, "Internet of things patterns," in *Proceedings of the 21st European Conference on Pattern Languages of Programs*, ser. EuroPlop '16. New York, NY, USA: Association for Computing Machinery, 2016.
[21] M. Štůsek, P. Masek, D. Kovac, A. Ometov, J. Hosek, F. Kröpfl, and S. Andreev, "Remote management of intelligent devices: Using tr-069 protocol in iot," 06 2016.
[22] M.-Y. Wu, Y.-H. Lin, T.-H. Tseng, C.-M. Hsu, K.-S. Hsu, and H.-C. Young, "A qos monitoring system for lte small cells," in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, 2018, pp. 1–4.
[23] K. S. Huang, "An enhanced tr-069 firmware upgrade method of wi-fi mesh system," in *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, 2019, pp. 655–659.
[24] "TR-369 – User Services Platform (USP)," 2022.

Implementation of Automated Testing Solution for Voice Services

M. Ivančić, N. Štokić, T. Žilić
Hrvatski Telekom, Zagreb, Croatia
marino.ivancic@t.ht.hr
nenad.stokic@t.ht.hr
tomislav.zilic@t.ht.hr

Abstract: Implementation of automated testing solution as a support in the projects and daily operational activities for voice services is of great importance in the optimization of time and human resources. Given the growing need in the field of telecommunications for maximum availability of the service to the end-user, it is important to have a tool that can perform system control activities as well as the ability to detect errors within network elements. Hrvatski Telekom has integrated one such system into the voice core network, which has several significant advantages for engineers who will improve the quality of service they provide to their customers using such a tool. The paper describes the technical solution of the Voice Automated Testing system as well as its implementation in the live network.

Keywords: automation, testing, voice, service, implementation, integration

I. INTRODUCTION

Technology changes in the telecom network environment are becoming more and more frequent over the last years. When talking about voice core networks and voice services, major technological evolution came with the introduction of the IMS (IP Multimedia Subsystem) based on the SIP (Session Initiation Protocol) protocol both for fixed and mobile voice services. Services are becoming more complex, hardware and software updates more frequent and system lifecycles shorter. All mentioned is impacting the increase of demand for resources needed for testing and verification processes [1]. Further modernization of the telecommunication network infrastructure and migration of platforms and services to the cloud environment even intensified such processes. A logical step forward is to speed up and automate the testing and validation process and optimize available resources to the greatest extent. This paper is describing one such approach of implementing an automated testing solution for telecom voice services.

Automation as a such is a much wider area tackling all aspects of the ICT universe. With technology progressing automation tools are becoming more available, and the need for automation for any sort of process is becoming increasingly important. As the Continuous Integration/ Continuous Delivery (CI/CD) concept already become standard in the software industry, the need for automation of the testing process is of great importance, and frameworks for automated testing are being developed [2]. Test automation frameworks are a set of rules and corresponding tools that are used for building test cases [3]. They are designed to help engineering functions work more efficiently. The general rules for automation frameworks include coding standards that you can avoid manually entering, test data handling techniques and benefits, accessible storage for the derived test data results, object repositories, and additional information that might be utilized to run the tests in a suitable manner. Examples of such automation frameworks are present in telecommunication networks as well, and they are focused on the design and development of a testing framework that automates the end-to-end configuration and control of a communication testbed [4].

Choosing the right automation framework and corresponding automation testing tool is not an easy task. Implementation of the selected approach is even more demanding, taking into concern not just technical efforts but also integration into the existing processes. However, the benefits of resource optimization should exceed those efforts. This paper will present an automated testing solution for voice services in a telecommunication provider network, based on the in-house developed automation tool.

II. VOICE AUTOMATED TESTING SOLUTION

A. Need for Voice Automated Testing

As part of the project that includes the modernization of the core network elements for voice services, Hrvatski Telekom decided to implement a new solution for automating the testing of scenarios in the field of voice services to speed up and optimize the process of testing equipment, platforms, and functionalities that support

voice services, including testing mobile and fixed call scenarios. Until now, test scenarios were mostly performed physically and manually by customizing the test environment for each test scenario, which resulted in a large expenditure of time and resources. With the increasing need for more complex scenarios when testing voice services, there was a clear need to implement a new tool for voice automated testing. As already mentioned, the selection of an appropriate test automation framework and test automation tool is not a trivial task [5]. Sometimes even a lack of adequate tools can be an obstacle to test automation. High initial investment and training for tool operation are also perceived as common challenges. Anyhow, in the telecommunication voice network and services area, there are a few established suppliers that offer test automation frameworks within their automation tools and solutions. Some of the most relevant ones are Segron Automation [6], Spirent Communications [7], Sigos (Mobileum) [8], Keysight (Ixia) [9], and many others.

Even though there are quite a few off-the-shelf test automation solutions, it was decided to use a new in-house solution developed within the Deutsche Telekom group, Voice Automated Testing (VAT) tool. It is a product owned by Deutsche Telekom, developed as an internal automated testing tool for the telecommunication service providers within the group. Hrvatski Telekom was the first one to choose it and became the first telecommunications operator within the DT group to implement and integrate such a system within a live network. This paper is focused on the VAT solution description and integration in the live voice network. It will be depicted how the tool works and what possibilities the new automated system for testing voice services has, and how it was implemented in the HT environment. Further solution details in the paper are based on internal technical documentation and firsthand experience.

B. Voice Automated Testing Solution Description

The implemented solution enables end-to-end testing of voice services, which include CS (circuit switch) mobile, VoLTE (Voice over LTE), and fixed services. This includes checking registrations and call establishment, as well as additional services such as call forwarding, call waiting, call retention, conference calls, etc. To achieve the above-listed services, it was necessary to implement equipment that can support this and integrate it with the live network.

In Figure 1 it can be seen what the VAT system hardware looks like integrated within the live network. For implementation, various hardware components were needed such as attenuators for signal management, shielded boxes for controlled conditions, mobile probes containing mobile devices, IP phones, VPN router, switch, patch panel, L1 switch, SIM array, and base station. In the presented environment, three shielded boxes were used with six mobile probes inside them (Samsung S9), in addition, there were six IP phones (VoIP technology) and six USB modems (analog technology) that are located outside the shielded boxes because they do not need signal management. With the attenuator connected to the base station and WiFi access point, it was enabled to manage the signal in shielded boxes (2G, 3G, 4G, WiFi).

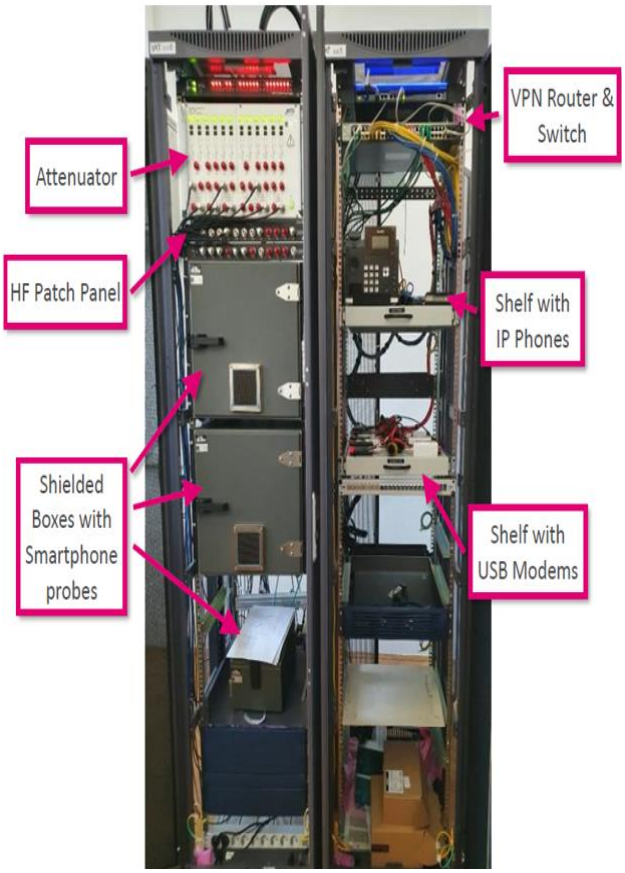


Figure 1. VAT system hardware

By integrating all these elements with live network elements, it was possible to achieve the conditions for testing different scenarios for the voice services. Users of such a system can perform different tests that can be performed multiple times without the need to set the same parameters before each test. Since the complex projects implemented in the telecommunications network are rapidly increasing, there is a need for tools that will help optimize resources in daily tasks. For the implementation of new services, regression testing, or monitoring network availability and stability (signaling, call flow, performance indicators, etc.), the VAT system should be very helpful in this area. Constant changes in the networks that are nowadays rather complex, taking into consideration cloud infrastructure lifecycle, core network applications lifecycles, new services, and features introduction, it is of vital importance to have an automated testing process to run the network appropriately.

VAT hardware equipment itself, also has its lifecycle and needs to be updated on regular basis, especially end devices like mobile phone probes which are changing very fast. In addition to that, the system should be scalable for expansion in case of new use cases implementation, new and additional probes integration, etc.



Figure 2. Shielded box with the mobile probe inside

Figure 2 shows how it looks when the shielded box is opened with a mobile probe inside (Samsung S9). VAT enables the sequential execution of several different tests individually and/or in a multi-defined block without the presence of the system user. The user of the system must be able to mark the tests to be carried out, as well as set the choice of test range that can be performed at any time of the day or night, and for each test a pass or fail result must be provided and all relevant information about the test scenario must be available. The test results together with the specified criteria are used to evaluate the test. This test may fail if the equipment test results are not good, but also if there is an error in the solution or issues in the network. After each test is performed, the solution must provide a test report - results, parameters used, test duration, etc. The report should be available in storage after the test along with screenshots of the phone during the test. It must be possible to export the network trace of the test scenario in *.pcap* format from the live network and *tcpdump* records from mobile devices. The solution must be able to test all mobile (2G, 3G, LTE), VoLTE, VoWiFi and VoIP technologies. For example, a VoLTE user calls a VoWiFi user registered on a WiFi network under a 2G signal. In addition, it can be enabled to use the feature for video recording of mobile devices during the test, which can be very helpful during the troubleshooting process. It can be configured to send email notifications to specific addresses if a series of tests fail, which can act as an alarm system. It is important to have various probes implemented to be able to perform interoperability tests within the whole network.

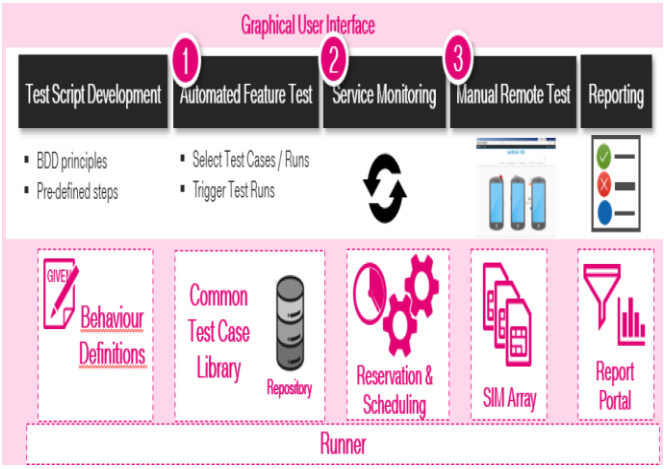


Figure 3. Automation logic

VAT could be used in several ways. Mainly it is used to automatically run a larger set of tests, but sometimes it is necessary to manually control the mobile device to achieve the desired goal. Manual control provides the same possibilities as if a mobile device is in a tester's hand. This possibility makes it easier to carry out specific test cases because it can be managed to use multiple devices at the same time and control their signal. To automatically run multiple tests within the VAT system, it is needed to have different tools to achieve this, and therefore multiple open-source platforms are used (Jenkins [10], Gitlab [11], Docker [12], ReportPortal [13]). Jenkins (automation server) is a tool used to create a "job" by selecting the probes, SIM cards and tests selected to be performed. Docker (containers) tool will use the parameters from the Jenkins "job" and will create an image with selected probes, selected SIM cards and run selected tests with them (before each test the device settings are reset so the test can be executed properly). Through this process, the system reserves selected mobile devices and connect them to selected SIM cards located in the SIM array, and then runs tests with selected parameters stored in the GitLab repository. After the test is finished, the results of the executed test scenarios and their contents are seen (*pcap* trace, *tcpdump*...) through the reporting portal as described in Figure 3. The picture visually depicts the automation logic of the VAT system that can easily be related to the test automation framework. In this way, the idea within the DT group is to have Test Automation as a Service (TAaaS) approach, and rapidly decrease manual interaction wherever applicable.

Figure 4 shows the execution of the test on mobile devices that can be monitored in real-time along with all the information about the progress of the test execution. The average time of performing one test is about 3-5 minutes depending on the test, so depending on the number of free probes and the execution time of each test, in one hour the system can perform 10-50 tests in a fully automated mode. If compared to the manual physical testing where it needed about 5 minutes to prepare the environment for the test scenario, execute the test and

collect the traces. The biggest advantage of a tool like VAT is that it can perform test scenarios execution without human interaction and testing can be scheduled anytime including during non-working hours.

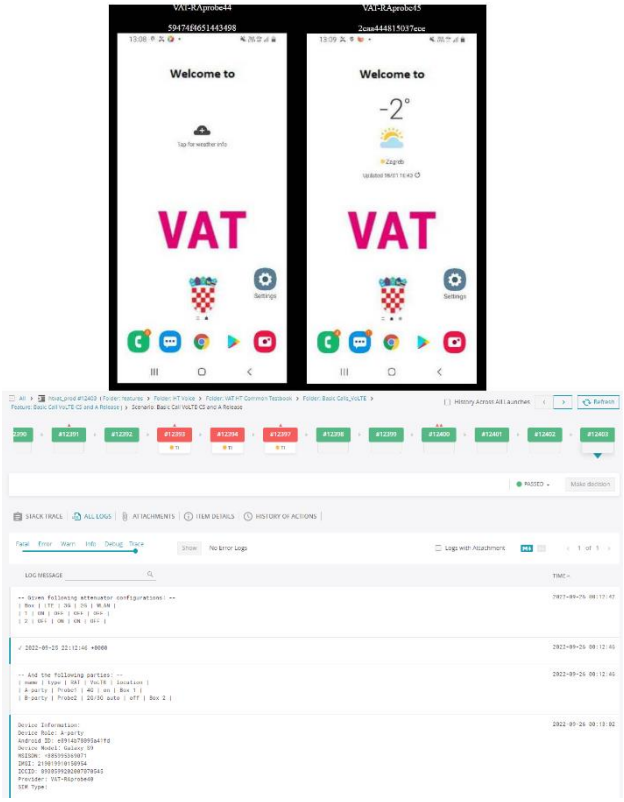


Figure 4. VAT mobile probes and ReportingPortal overview

III. Use Cases

VAT system could be used for various voice service testing scenarios like CS mobile voice services, VoLTE mobile services, fixed voice services, and should cover use cases such as:

- Availability and basic calls (registration, call establishment)
- Calling Line Identity Presentation/Restriction
- Additional services (call forwarding, call waiting, call hold, conference calls)
- HD Voice (VoLTE)
- Long-duration calls
- Interactive voice response (announcements)
- Ringtone sound detection
- SRVCC testing (loss of LTE signal during the call)
- SMS testing
- VoWiFi testing
- Inbound roaming testing
- Interoperability between CS mobile, VoLTE and fixed voice services

To provide a complete overview of calls from one end to the other, the system collects traces from the network (*pcap*) and network logs from mobile devices (*tcpdump*) if the user needs them for a more detailed analysis. The system should have full control over the devices so that it is possible to run test scenarios for voice services (registration, call establishment and call control) on them. For the devices to know exactly what they need to do, and what steps they must complete to execute a certain scenario, they use written commands that are found within the test scenarios. The test scenarios are written in the Gherkin programming language [14], which is convenient for automation because it uses plain English text that is understandable to almost anyone. For every step (command) written in Gherkin, there is a back-end code written in Ruby programming language [15].

```
@bcvolte001
Feature: Basic Call VoLTE-VoLTE and A Release

  Background:
    Given following attenuator configurations:
      | Box | LTE | 3G  | 2G  | WLAN |
      | 3   | ON  | OFF | OFF | OFF  |

    And the following parties:
      | name    | type   | RAT | VoLTE | location |
      | A-party | Probe5 | 4G  | on    | Box 3    |
      | B-party | Probe6 | 4G  | on    | Box 3    |

  Scenario: Basic Call VoLTE-VoLTE and A Release
    When A-party calls B-party
    And A-party verifies that the call state is 'dialing'
    And A-party verifies that the call state is 'alerting'
    And B-party verifies that the call state is 'incoming'
    And B-party verifies that the call is 'ringing'
    And A-party number is displayed at B-party
    Then B-party answers
    When B-party verifies that the call is 'established'
    Then A-party and B-party hear each other
    And all parties verify that they are camped on '4G'
    Then after 35 seconds A-party hangs up and becomes 'idle'
    And B-party verifies that the call state is 'idle'
```

Figure 5. Example of test in Gherkin

Figure 5 shows the test scenario under the unique tag @bcvolte001 with the configured signal settings inside the shielded box 3 containing probe 5 as the A-party and probe 6 as the B-party. A-party and B-party are VoLTE users, and A-party makes a 4G (VoLTE) call to B-party, which is also a VoLTE user. In the test, there are steps to verify call establishment where B-party answers the call. When the media check is verified on both sides during the call, it needs 35 seconds to pass so that the A-party ends the call.

Within this concept, a variety of voice service scenarios can be covered, including the most complex tests, and all the activities can be predefined and scheduled in advance, so no human interaction is needed while the

test is executed. As already stated, the VAT system uses a repository that contains most of the fixed and mobile voice services use cases present in the telecommunication network. It is up to the VAT system administrators and end devices availability how many use cases will be automated.

IV. Further Development

So far it was explained how the VAT system can help in daily operations, but there are still aspects of the solution that could be improved to make the system even better. The first point is to continue the development of the existing functionality and make the system even more robust, accurate, optimized and faster in task execution. Further on, support for the various types of end devices is needed as well, to be adaptable for any network environment. For example, at first only Android end devices were supported but now iOS support was introduced as well.

The next evolution step would be an introduction of some sort of AI (artificial intelligence) concept that would move automation to an even higher level. In that matter, to make the existing solution to be as high-quality as possible the results of the performed tests that are received in *.pcap* format should be verified automatically, so the system can decide whether the signalization in the call flow for a certain test is good or not (pass/failed). The execution of the test itself can be done successfully, it can look like the functional execution of the test scenario seems good, but it can happen that certain call signaling part is not done as expected. Because of such scenarios, in the current solution, human effort is needed to analyze the call flow for a specific use case in more detail to be able to confirm whether the performed test scenario is correct or not.

Another possible way forward is expanding the scope of the existing system which was intended for voice services only. With the introduction of the 5G network, there is a need to cover 5G-related voice services [16] but also the data network services as well. With additional functionalities for Data Automated Testing (DAT), automated processes to test data services will be implemented. Within the same concept, only different tests written in the same language would have the ability to test scenarios such as surfing the web, downloading, uploading, and measuring the performance indexes.

The solution has been recognized within the Deutsche Telekom group, and other national companies that are part of the group are starting to implement VAT solution in their networks. Even though this is an in-group (in-house) approach, it is still a commercial product that is offered to the group members in form of a shared services business model [17].

V. Conclusion

There is no doubt that automation in general is becoming an inevitable part of any ICT system introduction and maintenance, and telecommunication networks are no different in that matter. From all the above it can be concluded that the implementation of such a solution is fully justified in terms of time optimization and human resources. Furthermore, while continuously using the VAT system it can detect anomalies in the network if there is an active task where the same tests will be always performed. The system can also be set up to be used by non-technical people who can extract useful information for activities that are important to the company for some analysis and finding new business models that can be offered to the end user. The system is of great value for network and service regression testing when network changes, like system upgrades or new functionality introduction, are done and is very important to check and verify if everything is working as it should be. Automation concept evolution is not stopping there, with new use cases, scope enhancement, and new functionalities being introduced, this approach should have a bright future ahead.

Literature

[1] D. Gonçalves, A. Amaral, A. Costa, P. Sousa, “Towards automated test and validation of SIP solutions”, Telecommunication Systems, 61, 2016, pp. 579-590

[2] T. Fehlmann, E. Kranich, "A framework for automated testing", Systems, Software and Services Process Improvement: 27th European Conference, EuroSPI 2020, Düsseldorf, Germany, September 9–11, 2020, Proceedings 27. Springer International Publishing, 2020., pp. 275-288

[3] G. Hegde, A. Badkar, “Popular Test Automation Frameworks”, BowserStack web and mobile testing platform, 2020, retrieved from: https://www.browserstack.com/guide/best-test-automation-frameworks

[4] A. Díaz Zayas, B. García García, P. Merino, “An end-to-end automation framework for mobile network testbeds”, Mobile Information Systems, 2019.

[5] P. Raulamo-Jurvanen, M. Mäntylä, V. Garousi. "Choosing the right test automation tool: a grey literature review of practitioner sources", Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, 2017, pp. 21-30

[6] Segron Automation Website, retrieved from: https://segron.com/

[7] Spirent Communications Website, retrieved from: https://www.spirent.com/

[8] Mobileum (Sigos) Website, retrieved from: https://www.mobileum.com/

[9] Keysight (Ixia) Website, retrieved from: https://www.keysight.com/us/en/cmp/2020/network-visibility-network-test.html

[10] Jenkins Website, retrieved from: https://www.jenkins.io/doc/

[11] Gitlab Website, retrieved from: https://docs.gitlab.com

[12] Docker Website, retrieved from: https://docs.docker.com

[13] ReportPortal Website, retrieved from: https://reportportal.io/docs

[14] Cucumber (Gherkin programming language) Website, retrieved from: https://cucumber.io/docs/gherkin/

[15] Ruby (programming language) Website, retrieved from: https://www.ruby-lang.org/en/

[16] Z. Zhao, X. Liu, J. Miao, S. Zhou, X. Cheng, L. Xu, S. Liu, “Research on Automated Test and Analysis System for 5G VoNR Service”, Signal and Information Processing, Networking and Computers: Proceedings of the 8th International Conference on Signal and Information Processing, Networking and Computers (ICSINC), Singapore: Springer Nature Singapore, 2022, pp. 1396-1403

[17] T. Žilić, "Shared Services Business Model in ICT environment," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 1367-1372

RFID Inventory Management System Sampling Optimization Based on Zebra Android Framework

I. Benke, I. Heđi and E. Ciriković
Virovitica University of Applied Sciences, Virovitica, Croatia
ivan.benke@vuv.hr, ivan.hedi@vuv.hr, enes.cirikovic@vuv.hr

Abstract - ***In services with a high rate of input and output of items, the inventory data collection process often imposes high demands on staff in terms of reliable evidencing and overall repeatability. Certain business services, such as on-site laundry facilities in public hospitals, provide a good platform for the ad-hoc implementation of custom ICT (Information and Communication Technology) inventory management solutions, especially if they have already partially implemented proven technologies like RFID. Considering the relatively low cost and ease of design, development, and implementation of such systems in public institutions or other industry segments where accurate and reliable daily-based item inventory is crucial, it offers a viable alternative to commercially available inventory management systems. Processing operations for data elements are essential components of any information system. These operations are optimized with a focus on speed and reliability to meet the timing requirements of RFID sampling. This includes improvements to the database scheme, as well as the application and adaptation of existing algorithms for retrieving duplicate elements from data sets.***

Keywords – RFID; inventory; algorithms; optimization (key words)

I. INTRODUCTION

A. Data collection and processing in general

Data collection and processing using inventory means are common in work organizations, whether manual or automated. Inventory assets are marked with inventory codes, such as barcodes or QR codes, or more advanced RFID (Radio Frequency Identification) transponders. The interface to the database is typically realized through a screen form, where fields can be filled out using electronic readers or by entering inventory codes directly. However, the latter method is not suitable for recording a large number of inventory assets due to its inefficiency. In contrast, using electronic devices and supporting technologies, such as RFID, offers advantages in terms of speed, reliability, and mobility of the system. The basic architecture of an RFID-based system for recording inventory assets consists of a fixed or mobile RFID reader with an associated antenna and a remote transponder attached to the inventory asset. The codes of inventory assets are stored in the transponders' memory, such as the Electronic Product Code (EPC) memory or the Tag ID (TID) memory. The RFID reader sends an electromagnetic wave to the transponder, which powers the transponder to send back a response with the unique product code. The data is then saved on the RFID reader or sent to a remote terminal for further processing. The paper is structured as follows: it will describe the general system components and architecture, touch on basic implementation aspects of management system architecture and searching algorithms in a data structure after the short introduction.

B. Related work

RFID technology is well-known term described in many papers. Technology has been widely applied in inventory management, logistics tracking, identity recognition and so on. Most of papers focused on tag searching in large warehouses. In [1] authors highlight two major challenges in tag searching, one is how to exclude the interference caused by unexpected local tags, and the other is how to utilize the known IDs of wanted tags to improve the searching efficiency. They proposed a Bloom filter based tag searching protocol called BFSearch, which consists of two phases: non-wanted tag deactivation phase and target tag verification phase. Bloom filters also applied in [2]. Tags that carry similar information can be grouped into a category. To collect such information, most existing works have to query all tags in each category, which is time-consuming. In [3] authors proposed a new solution called arithmetic coding based sampling (ACS) protocol. They construct a sparse vector to sample only a subset of tags from each category, which can not only avoid repetitive information collection but also reduce interference from unsampled tags. Problem that occurs during RFID data capturing is false positives (i.e., tags that are accidentally detected by the reader but not of interest to the business process). In [4] machine learning algorithms described to filter false positives.

RFID scanner identifies all visible tags in each time interval. However, more than 90 percent of tag serial numbers are duplicates if you perform multiple scanning. These serial numbers need to be eliminated. Our problem is how to efficiently and quickly compare the field of existing tag serial numbers with the field of newly read tag serial numbers during each scan in order to obtain unique values. This problem specifically applies on our test device which is described in next chapter. Multiple scanning in one position is applied to ensure that all tags on a single stack are read.

II. ZEBRA RFID READER AND APPLICATION SUPPORT

The RFID reader is an essential component of RFID systems and comes in different versions, including fixed and portable types. The first type is mostly associated with industry and production plants where products or their components are moved through the reading zones of fixed-installed readers. A similar application can also be recognized in retail chains, where fixed detection systems at the entrances of stores are used to protect against item theft. Portable readers are used in applications where mobility is required, such as warehouses and other facilities. One such application of RFID reader used refers to hospital laundry management services, where incoming and outgoing articles of clothing, bed linen, and other laundry factory-marked with RFID transponders in the form of sewn-in labels are recorded and entered into the central management system. As part of the solution presented in this paper, the mobile RFID reader RFD40 manufactured by Zebra was used (Fig. 1).



Figure 1. Zebra RFD40 RFID UHF reader with Zebra TC21 management unit

In addition, the Zebra TC21 smart mobile device with the Android 10 system was also used as part of the control module, as shown in Fig. 1. Given that the said mobile device is part of the manufacturer's accessory equipment, it is added to the RFID reader using the appropriate integrated compatible docking station. Although the reader can be operated independently, without the additional element TC21, complete mobility of the entire system and extension of functions is possible only by using the entire package as shown in Fig. 1. The reader itself is designed to work in the UHF frequency range (865-868MHz) and is based on EPC Class 1 Gen 2 and EPC Gen2 V2 wireless access protocols. [5] Communication with the computer is provided via the USB-C interface, which provides the possibility for direct synchronization with management application solutions on the host computer side. In contrast, the built-in system memory buffer, according to the manufacturer, provides temporary storage of up to 40000 read RFID tags and thus enables excellent system autonomy and mobility. The RFID reader also has a built-in battery with a capacity of 7000 mAh, a reading speed of over 1300 RFID tags per second, and a range of about 6 m. It supports Windows, Android, and iOS development environments, and offers open support for upgrading and developing software solutions based on the offered source code and application programming interfaces (APIs). Ready-made applications such as 123RFID Desktop and Mobile provide functionality for searching and recognizing nearby or directly connected RFID readers, as well as antenna parameter adjustments, among others. [6] Both versions of the 123RFID application provide the functionality of searching and recognizing nearby or directly connected RFID readers as well as antenna parameters adjustment, and the latest firmware updates. The central applicative module of the system is responsible for managing the readings of RFID tags. This includes various tasks such as filtering, sorting, grouping, graphical presentation, and exporting raw data, as shown in Figure 2. The module also displays information such as the total number of readings, which is proportional to the default frequency or read rate, and the time of sampling or read time.

One important piece of information displayed in Figure 2 is the number of unique RFID tag values that have been read. This information is particularly valuable in inventory asset record-keeping systems. Keeping track of uniquely read tag values allows for accurate inventory management, as it helps identify individual items or assets that have been tagged and read by the system. This can be useful in tracking the movement of inventory, identifying missing or misplaced items, and reconciling inventory records with actual stock levels. By providing the number of unique RFID tag values, the system can offer a comprehensive overview of the inventory status, helping to streamline inventory management processes and improve overall operational efficiency.

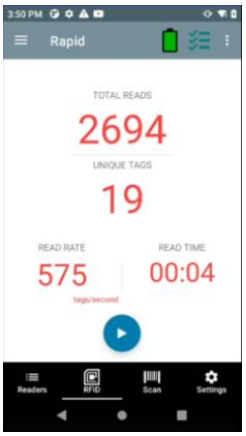
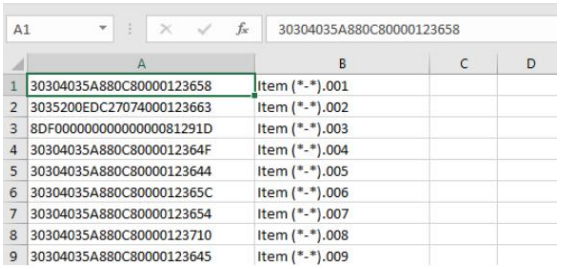


Figure 2. 123RFID application Rapid module window used for RFID tags display [7]

As part of the presented application, there are also options for writing unique identifiers on corresponding RFID transponders using the Tag Write program module, as well as the possibility of locating single or multiple RFID transponders according to the given identifier values. Regardless of the tag reading mode, all readings are written to a CSV file located in the previously created root directory of the Android operating system.



	A	B	C	D
1	30304035A880C80000123658	Item (*-*).001		
2	3035200EDC27074000123663	Item (*-*).002		
3	8DF00000000000000081291D	Item (*-*).003		
4	30304035A880C8000012364F	Item (*-*).004		
5	30304035A880C80000123644	Item (*-*).005		
6	30304035A880C8000012365C	Item (*-*).006		
7	30304035A880C80000123654	Item (*-*).007		
8	30304035A880C80000123710	Item (*-*).008		
9	30304035A880C80000123645	Item (*-*).009		

Figure 3. Scanned RFID tags displayed in corresponding CSV file

In the RFID Settings application module, there are specific options related to the operation of the reader's antenna and the type of communication link. In the Profiles configuration group, there are options for manual control of parameters such as the transmitted power of the reader (Power Level) expressed in units of dBm, and the link profile parameter (Link Profile) which determines the product of the symbol transmission rate and the Miller number of senders per symbol (2, 4, or 8), as indicated by the symbols shown in Figure 4.

The choice of the link profile affects the system's ability to successfully detect passive RFID tags, and it involves finding a compromise between data transfer speed, signal-to-noise ratio, and the number of incorrectly detected bits in relation to the number of transmitted ones. For example, the M2 320K profile was chosen for testing and production environment, as it proved to be the most reliable under the given conditions. This choice is based on the EPC Gen2 V2 standard and specific requirements of local regulators for electronic communications.

Furthermore, the system allows defining the Pulse Interval Encoding (PIE) and Type A Reference Interval (Tari) of the symbols of the discrete modulation procedure PR-ASK (Phase Reversal Amplitude Shift Keying), following the standards and requirements mentioned above [8]. These settings are important for ensuring compatibility with relevant standards and regulations, and for optimizing the performance of the RFID system in different operating environments.

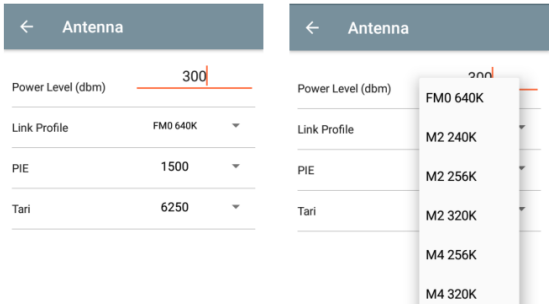


Figure 4. The display of available Zebra RFID reader communication parameters in 123RFID application [7]

The continuation of this paper will present a complete software solution that is based on the described hardware Zebra RFID platform and utilizes the freely available Zebra RFID Software Development Kit (SDK) package for Android systems.

III. MANAGEMENT SYSTEM ARCHITECTURE

Management system architecture like multilayered architecture is shown on Fig. 5. At the bottom of the system architecture is an RFID handheld scanner used for wirelessly scanning tags at a rate of more than 1000 tag reads per second. The RFID handheld scanner is controlled by a mobile device with an Android operating system and Zebra SDK. The Zebra Scanner SDK is integrated with a mobile application for Android, enabling cordless scanners to be connected and controlled by a tablet or smartphone without using a cradle, after pairing over Bluetooth. Scanning a pairing barcode that appears on a tablet or smartphone display will pair the device and connect it with the application. Each tag is identified by a 24-character string with numbers and capital letters, known as the serial number.

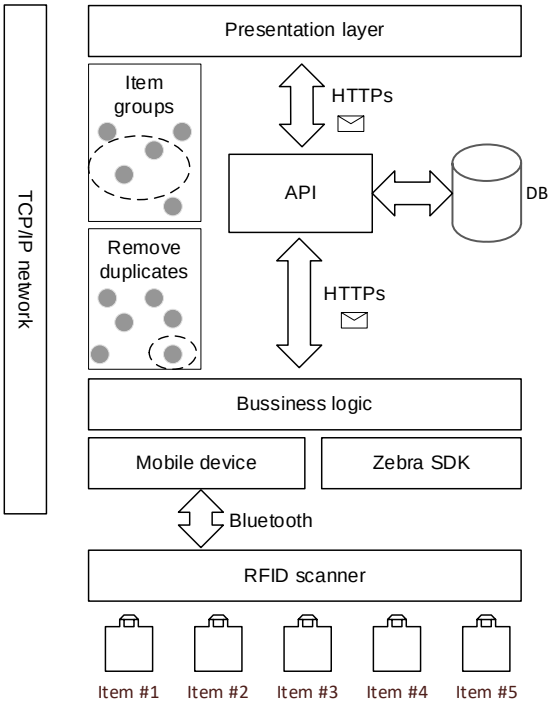


Figure 5. Management system architecture

An example of a serial number written on an RFID tag is:

"300ED89F335000B99A421245".

Given that the sampling frequency is 1kHz, the RFID scanner identifies all visible tags in each time interval. However, more than 90 percent of tag serial numbers are duplicates that need to be eliminated. This action is implemented in the Business logic layer, also known as the mediation layer, which is implemented in the mobile application. The main functions of the mediation layer include data collection from tags, aggregation of collected data on a low level, and preparation for storage in a real-time data warehouse. By removing duplicate values, every HTTP request of the API (Application Programming Interface) becomes lightweight and faster. The storage of the read data in the SQL Server Express database, processing of the results, and the complete web solution are located on a separate remote server location. The prepared data is then sent to the system backend for

further processing. Regardless of the tag reading mode, all readings are written to a CSV file located in the previously created root directory of the Android operating system. Our solution uses direct readings from the RFID reading device shared memory. To send data, it is only necessary to call a function in the API with data stored in JSON structure.

The published message looks like this:

```
{
   "datetime": 1581700066476,
   "operator": 1,
   "warehouse": 1,
   "direction": 1
   "serials": [300ED89F335000B99A421245,…],

}
```

The direction attribute is used to set if selected tags are coming to the warehouse or from the warehouse. The serials attribute is an array of unique values of serial numbers.

The RFID scanner and the mobile device are two interconnected objects with the ability to transfer data to the Internet. When devices are connected to the Internet, they can communicate with other devices or deliver information to specific endpoints [9]. These devices can connect to the Internet directly using standard technologies such as 3G, 4G, and 5G, or they can connect to a local area network that is connected to the Internet. On the other hand, devices can form M2M (Machine to Machine) networks, where devices are connected using radio communication standards and protocols such as Wi-Fi (based on the IEEE 802.11 standard), Bluetooth (based on the IEEE 802.15.1 standard), Zigbee (based on the IEEE 802.15.4 standard), or 6LowPAN over Zigbee (IPv6 over Low Power Personal Area Networks). In most cases, application layer protocols are used for handling communication. Some of the most representative application layer protocols are CoAP (Constrained Application Protocol), MQTT (Message Queue Telemetry Transport), XMPP (Extensible Messaging and Presence Protocol), RESTFUL Services (Representational State Transfer), AMQP (Advanced Message Queuing Protocol), and web socket [10].

If devices are connected to the Internet, there are several cloud providers that offer connectivity between devices and the cloud, such as AWS IoT (Amazon Internet of Things). The primary function of an IoT platform is to act as a middleware layer to connect devices or applications from one end to another end [11].

The top layer of a given architecture is the presentation layer. The presentation layer consists of various reports for controlling item flow. A high level of detection of inconsistency is an algorithm for generating item groups. Several types of serial numbers can form one group in one direction. If there is a group that is not complete, there is a high probability that something was not detected, or items were lost.

The most important problem in this architecture is the causal relationship between the sampling frequency and the speed of duplicates search. If the sampling frequency is increased due to a large number of items, optimization for duplicates search becomes necessary.

IV. SEARCHING ALGORITHMS IN DATA STRUCTURES

A search algorithm is an algorithm that provides a solution to a problem after evaluating a set of possible solutions. The set of all possible solutions to the problem is called the "search space" [12]. In data structures, two search algorithms are commonly used: sequential and binary. Sequential search is the simplest form of searching and is used on small, unsorted data sets. It goes through every element of the array, making it the slowest search algorithm with a worst case complexity of O(n). Sequential search can also have some subvariants, such as Probability Search, Sentinel Search, or Ordered List Search [13]. An example of sequential search is shown in Fig. 6

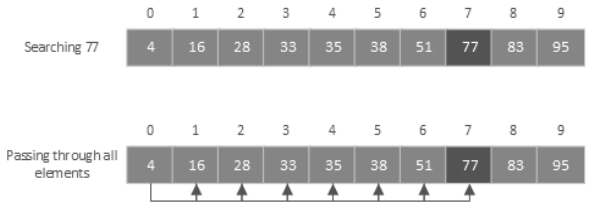


Figure 6. Sequential search

Binary search is a more advanced and faster search algorithm compared to sequential search. The search starts by testing the element in the middle of the array. The middle element is obtained as the average value of the first and last index of the array, mid = (fi+li)/2. If the middle element is the required element, the search ends. If the requested element is smaller than the middle, the search continues in the first part of the field, while the second part of the field is eliminated from further consideration, and vice versa. A simple example is shown in Fig. 7. The worst case complexity of the binary algorithm is O(log n). To use binary search, the main condition is that the array must be sorted. Since the elements of the array are not necessarily arranged according to some criteria, array indexing is used to make search and record sorting faster. Array indexing occurs when adding a new record to the array, and the main disadvantage of this search is worse performance in the case of a large number of additions. There are different variations of binary search (triple search, exponential search, Fibonacci search), and their effectiveness depends on the size of the array being searched.

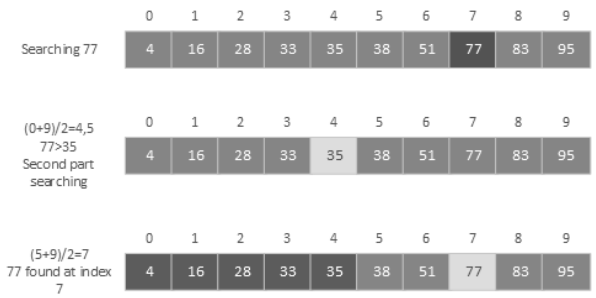


Figure 7. Binary search

A graphical overview of the measured performances of these two types of searches is shown in Fig. 8.

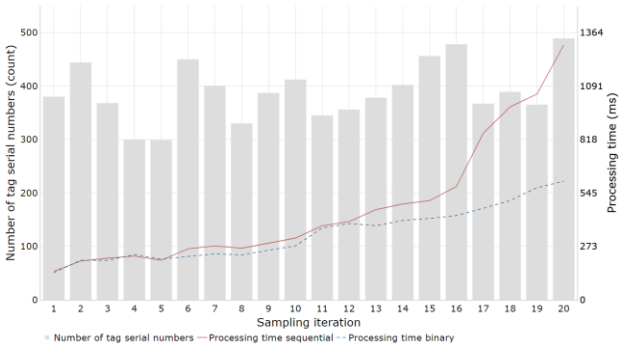


Figure 8. Statistical data comparison

The frequency of sampling serial numbers of tags in a time frame of 1000 ms limited the choice of algorithm in terms of speed. Therefore, binary search was found to be advantageous compared to sequential search. Even though sorting the existing serial numbers was required for binary search, it turned out to be 40% more efficient.

V. CONCLUSION

RFID technology has been widely adopted for recording inventory goods due to its benefits such as good industrial support, usage flexibility, and widespread standard adoption. The basic architecture of an RFID-based inventory recording system consists of a fixed or mobile RFID reader with an associated antenna, and remote transponders physically attached to the inventory assets. Codes of inventory assets, unambiguously associated with individual transponders, are generated based on a preset sampling frequency. The use of algorithms for handling large data sets is necessary to ensure efficient system performance. In conditions of high sampling frequency of RFID tag serial numbers, it has been demonstrated that the binary search algorithm can save enough time to meet strict time frame requirements before the arrival of a new set of data. Furthermore, given the technical limitations of the HTTP protocol, the application of more modern solutions, such as the MQTT protocol, could further optimize the system in terms of reliability and reading speed, which are critical parameters.

REFERENCES

[1] Na Yan, Honglong Chen, Kai Lin, Zhichen Ni, Zhe Li, Huansheng Xue, BFSearch: Bloom filter based tag searching for large-scale RFID systems, Ad Hoc Networks,Volume 139, 2023

[2] Z. An, Q. Lin, L. Yang, W. Lou and L. Xie, "Acquiring Bloom Filters Across Commercial RFIDs in Physical Layer," in IEEE/ACM Transactions on Networking, vol. 28, no. 4, pp. 1804-1817, Aug. 2020

[3] J. Liu, S. Chen, Q. Xiao, M. Chen, B. Xiao and L. Chen, "Efficient Information Sampling in Multi-Category RFID Systems," in IEEE/ACM Transactions on Networking, vol. 27, no. 1, pp. 159-172, Feb. 2019

[4] G. Alfian, M. Syafrudin, B. Yoon, and J. Rhee, "False Positive RFID Detection Using Classification Models," Applied Sciences, vol. 9, no. 6, p. 1154, Mar. 2019

[5] https://www.zebra.com/us/en/products/rfid/rfid-handhelds/rfd40.html (21.01.2023.)

[6] https://www.zebra.com/us/en/support-downloads/software/demo/123rfid-mobile.html

[7] Signal Coding in Physical Layer Separation for RFID, Tag Collision, Yi Li1, , Haifeng Wu1, and Yu Zeng1

[8] https://www.zebra.com/content/dam/zebra_new_ia/en-us/manuals/rfid/rfd40/rfd40-prg-en.pdf (15.01.2023.)

[9] F. Xia, L. T. Yang, L. Wang and A. Vinel, "Internet of Things", International Journal of Communication Systems, Vol. 25, pp. 1101-1102, 2012.

[10] V. Karagiannis, P. Chatzimisios, F. Vazquez-Gallego and J. AlonsoZarate „A Survey on Application Layer Protocols for the Internet of Things", Transaction on IoT and Cloud Computing 2015.

[11] O. Jukić, I. Špeh and I. Heđi, "Cloud-based services for the Internet of Things", Proceedings of the 41st International Convention MIPRO 2018, pp. 407-412, MIPRO, Opatija, 2018.

[12] N. Pavković, D. Marjanović, N. Bojčetić, Programiranje i algoritmi, skripta II, Zagreb, 2005.

[13] https://aits-tpt.edu.in/wp-content/uploads/2018/08/DS-UNIT-5.pdf (16.02.2023.)

DATA MANAGEMENT

Study of Data Transfer Nodes Infrastructure in Enabling Big Data Movement between Research and Cloud Storage Networks

Damir Regvart *, Miljenko Mikuc**, Vedran Dakić *
* Algebra University College, Zagreb, Croatia
** University of Zagreb, Croatia, Faculty of Electrical Engineering and Computing

***Abstract* -Data research projects generate huge amounts of data, which are usually stored in cloud networks for processing and storage. Transferring large amounts of data between multiple research sites, often within protected scientific DMZs (Demilitarized Zone) and cloud storage nodes, requires high-speed network connections.**

To ensure efficient data transfer and processing, parallel data streams are used between multiple research sites, with dedicated cloud infrastructure for processing and storage of a particular experiment. Since multiple nodes or sites are involved in data transfer, DTNs (Data Transfer Nodes) are used to solve the problems of fast data flow and performance between nodes that are not related to network throughput. DTNs include software transfer tools and high-end hardware components that are tasked with preparing data for transfer.

Several research groups are working to establish multiple DTN instances as part of a large-scale scientific data network. This paper provides a brief overview of the current state of the art in the use of DTN infrastructure and the challenges associated with preparing data for transfer.

***Keywords* - DTN (Data Transfer Node), Science DMZ (Demilitarized Zone), big data transfer, orchestration of data streams**

I. INTRODUCTION

A network transfer node is a point in a network where data, signals, or messages are transferred from one network to another. It can also be a device or system that serves as a connection point between two or more networks, enabling the transfer of data between them. Network transfer nodes can be used to connect different types of networks, such as local area networks (LANs), wide area networks (WANs), and the Internet. They can also be used to interconnect different types of devices such as computers, servers, storage, routers, and switches to enable communication and data transfer between them.

There are several factors that can limit the full network throughput on a network. Some common limitations are [1]:

- Bandwidth: The amount of data that can be transmitted over a network is limited by the available bandwidth. If the network is congested or bandwidth is limited, this can affect throughput.
- Latency: Latency or delay on a network can also affect throughput. A delay in the transmission of data can reduce the overall throughput of the network.
- Protocol overhead: Protocols such as TCP/IP are used to transmit data across a network. However, these protocols have overhead, which means that some of the available bandwidth is used for protocol-related tasks and is not available for data transmission. This can limit the throughput of the network. The delay between the conformance packet and the "ready to send next batch" or retransmission signal can also play an important role in controlling protocol overhead.
- Interference: Interference from other devices or sources, such as wireless interference or physical obstacles, can also affect the throughput of a network.
- Device capabilities: The capabilities of devices connected to the network can also limit throughput. For example, if a device has a slower processor or limited memory, it may not be able to process or transmit data as quickly as a more powerful device.

Data Transfer Nodes (DTN) [2] are defined as network nodes that are used to transmit data from one site to another and are used exclusively for data transmission. DTNs should also match the capabilities of the site and WAN infrastructure where the node is installed. They are also commonly used to transfer data over long distances or when data needs to be transferred between two networks that are not directly connected or immediately accessible. Some examples of where DTN can be used [3]:

- Satellite communications: DTN can be used to transmit data between ground stations and satellites or between satellites in different orbits.
- Space exploration: DTN can be used to transmit data from spacecraft to ground stations or between spacecraft in different parts of the solar system or beyond.
- Emergency or disaster relief: DTN can be used to transmit data in situations where conventional communications networks are unavailable or unreliable, such as after a natural disaster or during a military conflict.

- Remote or isolated locations: DTN can be used for data transmission in areas where there is no direct connection to the Internet or other communication networks.
- When a parallel data stream to one or more sites would be required.

DTNs can be used to transmit a wide range of data types, including text, images, video, audio, or raw data. They are often used in environments where data transmission presents problems, such as low bandwidth, high latency, or broken connections.

DTNs can also be used to transfer data between research centers and cloud storage. In this scenario, the DTN acts as a bridge link between the research center and the cloud, allowing researchers to transfer large amounts of data to and from the cloud without having to use the Internet or other traditional communication networks. This can be useful in cases where the research center does not have a high-speed Internet connection, the data to be transferred is sensitive and must be kept secure, or the destination of the data cloud is unknown (some examples: Put the data on a lower cost cloud storage or put the data on the first available HPC with the required processing parameters for processing).

Using DTN to transfer data between research centers and cloud services can allow researchers to take advantage of the large storage capacity and computational power of the cloud while still being able to work with the data locally at the research center. It can also help improve the speed and efficiency of data transfer, as DTN is designed to handle large amounts of data and can operate over long distances and in challenging environments.

II. DTN NODE DEFINITION AND USAGE

Data Transfer Nodes (DTN) are typically made up of a network of devices that are used to transfer data from one location to another. The specific components of a DTN infrastructure can vary depending on the specific requirements of the system, but some common components include:

- Nodes: DTN nodes are actual hardware devices that are used to transfer data within the DTN network. They can be located at the source and destination of the data transfer, as well as at intermediate locations along the way. DTN nodes can be standalone devices, or they can be integrated into other systems, such as routers or servers.
- Links: DTN links are the communication channels that connect DTN nodes and allow them to transfer data. Links can be physical, such as a cable or fiber optic connection, or they can be wireless, such as a radio or satellite link. In some cases, a DTN link may consist of multiple communication channels that are used in parallel to transfer data.
- Protocols: DTN relies on specialized protocols to transfer data between nodes and ensure that the data is delivered reliably. These protocols can include routing protocols, which determine the best path for the data to take through the DTN network, and error correction protocols, which are used to ensure the integrity of the data. Other protocols, such as flow control and congestion control, can be used to manage the flow of data through the DTN network and prevent the network from becoming overloaded.
- Management: DTN systems often include a management component, which is responsible for configuring and monitoring the DTN nodes and links, and for ensuring that the DTN network is operating efficiently. This can include tasks such as traffic management, load balancing and fault detection and recovery. The management component can be implemented as a standalone system, or it can be integrated into other components of the DTN infrastructure.
- Storage: DTN systems may include storage devices, such as hard drives or flash drives, to store data temporarily as it is being transferred. This can be useful in cases where the DTN network is experiencing intermittent connectivity or where the data transfer rate is slower than the rate at which the data is being generated.
- Security: DTN systems may include security measures, such as encryption and authentication, to protect the data being transferred from unauthorized access or tampering.
- User interface: Some DTN systems may include a user interface, such as a graphical user interface or command-line interface, to allow users to interact with the system and manage the data transfer process.

In addition to these components, DTN systems can also include other hardware and software elements, such as dedicated fast storage devices, security boxes, and user interfaces.

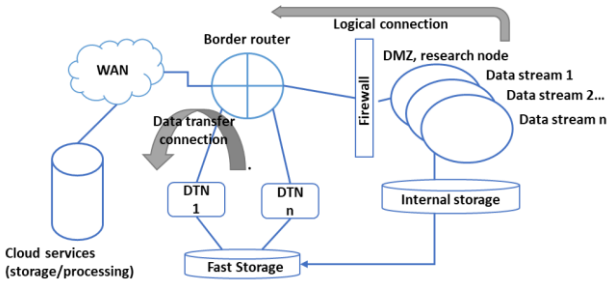


Figure 1. Simplified DTN architecture

The logical connection and real data transfer connection can take different paths within the network, especially in Science DMZ environment as shown in figure 1.

III. RELATED WORK

Currently there are multiple research projects within research academic network regarding implementation of DTN in parallel with study of science workflow regarding machine data generation and AI processing of the data in research environments. ESnet (Energy sciences network) has published "Deployment strategy" [4] and know-how documents for starting up DTN implementation, hardware selection and hardware tuning. Recently Paneuropean research network Geant [5] has published findings and testing of DTN deployment of different transfer tools.

Regarding security of inplementation the research paper in 2019 by Jorge Crichigno and others [6] has published a bottom-up approach, from the physical cyberinfrastructure to the application layer and security aspects of DTN in Science DMZ (Demilitarized Zone). The research concluded that DTN can mitigate the deliveries of malicious payloads and also disscussed the advantages and disadvantages of security appliances and techniques, such as ACL (Access Control Lists), IPS (Intrusion Protection System) , IDS (Intrusion Detection System), and related best practices to secure Science DMZs while avoiding performance degradation.

Recently more research was given for maintaining high-capacity, data-intensive network flows in a medical context [7]. Medical research centers have resulted in high-rise of biomedical big data that has to be shared between multiple sites (National Library of Medicine, the National Cancer Institute, European Bioinformatics Institute). The authors describe the ability to store and compute data by a combination of local, national, and industry resources that exchange large data sets using the possibility of DTN.

The paper from Wontaek Hong and others [8] discuss how to enhance the data transfer utilizing DTN between Cloud Service Providers. The paper proposed procedure for transferring VM images to a destination controller that includes local/remote copy processes and a DTN-to-DTN transfer process. The procedure improved throughput by 22%.

IV. CHALLENGES OF DTN IMPLEMENTATION

Implementing and maintaining a DTN network can be expensive, especially if it involves the deployment of specialized hardware and software, or the use of expensive communication channels, such as satellite links [9]. One of the potential solutions to this challenge is to use a hybrid approach that combines DTN with other technologies, such as the internet, to transfer data. This can help to reduce the cost of the DTN network by leveraging existing infrastructure and taking advantage of lower-cost and shared communication channels. Regarding data storage and processing, a possible cost effective solution is to use a cloud-based DTN service, which can provide access to DTN capabilities without the need to build and maintain a hardware and software infrastructure.

Challenge in using DTN to connect research centers with cloud storage or HPC systems is mostly the limited bandwidth of the DTN network. This can especially be demanding and challenging when transferring enormous amounts of data or when the DTN network is operating over long distances or in remote locations. A potential solution to this challenge is to use high-bandwidth dedicated communication channels, such as fiber optic cables or satellite links, to connect DTN nodes.

Data compression technique to reduce the amount of data can also be a reasonable solution (predictive algorithms, multicast, etc.). Most research centers nowadays are connected via high-speed fiber connection provided by National Research Networks (NREN) which are mutually inter-connected with high network speed.

An important challenge in establishing DTN infrastructure is hardware. These challenges include:

- Reliability: DTN nodes and links must be reliable to ensure that data can be transferred consistently and reliably. This can be especially challenging in environments where DTN nodes and links are subjected to extreme conditions, such as temperature fluctuations, humidity, or physical shock.
- Compatibility: DTN nodes and links must be compatible with each other and with the systems they are connecting to. This can be a challenge when different DTN nodes and links are manufactured by different vendors or when they use different communication protocols or technologies.
- Scalability: DTN networks may need to be scaled up or down as the volume or complexity of the data being transferred changes. This can require the deployment of additional DTN nodes and links, which can be a challenge in terms of cost and complexity.
- Maintenance: DTN nodes and links may require regular maintenance to function properly, including tasks such as firmware updates, hardware repairs, and cleaning. This can be a challenge in situations where DTN nodes and links are located in remote or inaccessible locations.
- Power: DTN nodes and links may require a reliable power source to operate. This can be a challenge in situations where there is no access to grid power or where the DTN network is being used in a mobile or portable setting.

V. SCENARIOS OF DTN ORCHESTRATION

Data Transfer Nodes (DTN) can use orchestration to manage the flows of data being transferred over the DTN network and to configure bandwidth on demand. Orchestration refers to the process of automating the coordination and management of DTN nodes and links to transfer data efficiently and reliably [10].

In the context of DTN, orchestration can involve tasks such as:

- Routing data: Orchestration can be used to determine the best path for data to travel through the DTN network based on factors such as the location of the source and destination, the volume of data being

transferred, and the capabilities of the DTN nodes and links.

- Scheduling transfers: Orchestration can be used to schedule data transfers to optimize the use of DTN resources, such as bandwidth and storage. This can involve tasks such as prioritizing certain transfers over others, or scheduling transfers to take place during periods of low network usage.
- Configuring bandwidth on demand: Orchestration can be used to configure the bandwidth of DTN nodes and links on demand, based on the needs of the data transfer. This can involve tasks such as allocating additional bandwidth to transfers that require it, or releasing bandwidth when it is no longer needed.

Orchestration can be implemented using a variety of technologies, such as software-defined networking (SDN), network function virtualization (NFV), or cloud-based orchestration platforms. These technologies can be used to automate the management and configuration of DTN nodes and links to improve the efficiency and reliability of data transfers over the DTN network.

VI. DTN NODE PERFORMANCE TESTING

Several DTN performance tests were performed on Geant4-3 project using GTS (GÉANT Testbed Service) [11]. GTS enables the research community to create automated platform on which to build virtual networks using dedicated fiber links between sites.

DTN research project was started as research activity to support current and future projects that require large-scale, high-performance transfers via NREN and Geant network. Examples of this include Square Kilometer Array (SKA), Worldwide Large Hadron Collider Computing Grid and AENEAS project.

The testbed scenario included simulated DTN nodes in Geant network located in data centers in London, Amsterdam, Prague, Paris and Hamburg. DTN simulated nodes were directly connected via 10Gbit/s dedicated dark-fiber links. DTN hardware node was evaluated in a configuration role as a virtual machine (VM) or a bare-metal server (BMS).

Maximum fiber length was 2093km on London-Prague link. VMs were generated using OpenStack GTS GUI interfaces and were connected via virtual circuits using Network Service Interface (NSI) across multiple sites on GTS network.

Different hardware was tested primarily Dell PowerEdge R430 (2x20C/40T Intel Xeon E5-2660v3 @ 2,6GHz, 128 Gb ECC DDR4 2133 MHz RAM) and Dell PowerEdge R520 (1x8C/16T Intel Xeon E5-2450 v2 @ 2,5GHz, 32 Gb ECC DDR3 1600 MHz RAM) in VM or BMS configuration.

Several file transfer tools were used in the performance testing:

- iperf - network performance tool used as a reference tool for maximum throughput.
- GridFTP - tool which is an extension to FTP developed for grid computing by Open Grid Forum (widely used tool for data transfers in science projects and supercomputer centers).
- FDT - Fast Data Transfer tool - open-source tool that can use and control multiple TCP data streams via different sockets.
- XtootD - open-source extended ROOT daemon that can utilize load-balancing for multiple data streams in client-server transfers.

In simulation all the hardware components were fine-tuned to give maximum throughput and performance require for high volume data traffic (mtu, packet pacing, I/O Scheduler, data-file system). For the VM the same Docker environment was used with additional tuning on kernel level. The different results were accomplished depending on hardware used or virtualization level.

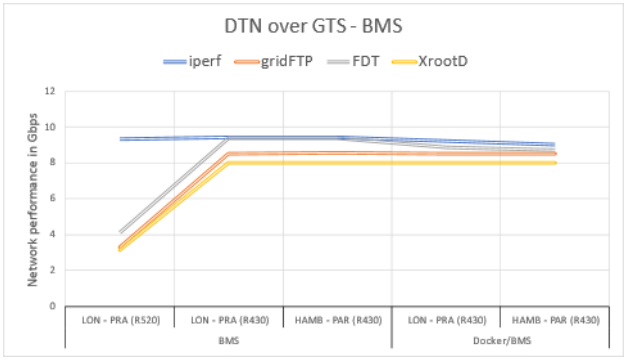


Figure 2 DTN test results of BMS on GTS

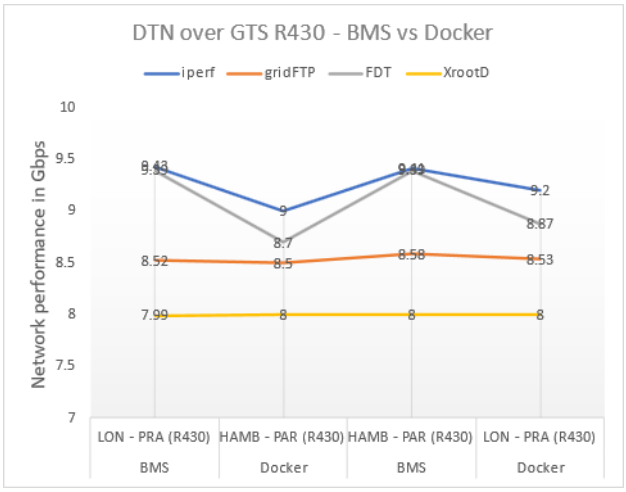


Figure 3 DTN test results over GTS network

Iperf tool was used as a reference point for maximum network throughput. Results of the performance testing of different BMS with same tunning properties, over different fiber channels are shown in figure 2. Results show better performance of FDT tool on all tested scenarios. Figure 3. shows difference in performance of BMS vs. VMs on Docker platform for the same transfer tools and same data file.

XrootD tool give deficient performance in virtual/contained environment since the tool is highly dependent on hardware resources and direct hardware access, while others show similar results and 5% less performance than what was results of iperf reference tool measurement.

VII. CONCLUSION

Large data transfer is becoming increasingly important as research projects generate vast collections of data that require fast and efficient transfer and processing.

DTN is an innovative solution for the transfer of large amounts of data between research centers and cloud services. The paper has presented several advantages and challenges over traditional data transfer methods in using DTN and constraints for development. Paper also presented performance results of different tools that can be incorporated inside DTN node. It is important that DTN is a separate piece of hardware dedicated only for preparing data transfer and actual data transfer. DTN must be tuned for maximum performance regarding the network interfaces and server hardware components.

DTN provides a versatile and efficient solution for data transfer that is rapidly gaining popularity among researchers and data scientists. With ongoing advancements in technology and research, DTN is set to play an even more crucial role in data transfer and processing in the future.

Future work will include enhancements of performance [12], [13] in regards of reliability of the system, security, increasing the efficiency and speed of data transfer, integration with emerging technologies such as Artificial Intelligence [14], Machine Learning and Big Data Analytics to provide even more powerful data transfer and processing capabilities for research communities.

REFERENCE

[1] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," in IEEE/ACM Transactions on Networking, vol. 2, no. 1, pp. 1-15, Feb. 1994, doi: 10.1109/90.282603.

[2] Eli Dart, Lauren Rotman, Brian Tierney, Mary Hester, Jason Zurawski, "The Science DMZ: A Network Design Pattern for Data-Intensive Science", Scientific Programming, vol. 22, Article ID 701405, 13 pages, 2014. doi:10.3233/SPR-140382

[3] Nathan Hanford, Brian Tierney, and Dipak Ghosal. 2015. Optimizing data transfer nodes using packet pacing. In Proceedings of the Second Workshop on Innovating the Network for Data-Intensive Science (INDIS '15). Association for Computing Machinery, New York, NY, USA, Article 4, 1–8. doi:10.1145/2830318.283032

[4] Science DMZ [on-line], Available: https://fasterdata.es.net/DTN/

[5] GÉANT DTN working group [on-line], Available: https://wiki.geant.org/display/NETDEV/DTN

[6] J. Crichigno, E. Bou-Harb and N. Ghani, "A Comprehensive Tutorial on Science DMZ", in IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 2041-2078, Secondquarter 2019, doi: 10.1109/COMST.2018.2876086.

[7] Peisert, Sean; Barnett, William; Dart, Eli; Cuff, James; Grossman, Robert L.; Balas, Edward; Berman, Ari; Shankar, Anurag; Tierney, Brian, "The Medical Science DMZ", Journal of the American Medical Informatics Association, svez. 23, br. 6, pp. 119-1201, 2016., doi:10.1093/jamia/ocw032

[8] W. Hong, J. Moon, W. Seok, and J. Chung, "Enhancing Data Transfer Performance Utilizing a DTN between Cloud Service Providers," Symmetry, vol. 10, no. 4, p. 110, Apr. 2018, doi: 10.3390/sym10040110.

[9] Lee, Chankyun, Minseok Jang, Minki Noh, and Woojin Seok. 2021. "Scalable Design and Algorithm for Science DMZ by Considering the Nature of Research Traffic." Journal of Supercomputing 77 (3). Springer: 2979–97. doi:10.1007/s11227-020-03383-y.

[10] Z. Liu, R. Kettimuthu, I. Foster, P. H. Beckman, "Toward a smart data transfer node", Future Generation Computer Systems, Volume 89, 2018, Pages 10-18, ISSN 0167-739X, https://doi.org/10.1016/j.future.2018.06.033.

[11] F. Farina, P. Szegedi i J. Sobieski, "GÉANT world testbed facility: Federated and distributed testbeds as a service facility of GÉANT", 2014 26th International Teletraffic Congress (ITC), p. https://doi.org/10.1109/ITC.2014.6932972, 2014.

[12] Lee, C., Jang, M., Noh, M., & Seok, W. (2021). Scalable design and algorithm for science DMZ by considering the nature of research traffic. Journal of Supercomputing, 77(3), 2979–2997. https://doi.org/10.1007/s11227-020-03383-y

[13] B. Tierney, E. Dart, E. Kissel and E. Adhikarla, "Exploring the BBRv2 Congestion Control Algorithm for use on Data Transfer Nodes," 2021 IEEE Workshop on Innovating the Network for Data-Intensive Science (INDIS), St. Louis, MO, USA, 2021, pp. 23-33, doi: 10.1109/INDIS54524.2021.00008.

[14] G. Kaur and R. S. Batth, "Edge Computing: Classification, Applications, and Challenges," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2021, pp. 254-259, doi: 10.1109/ICIEM51511.2021.9445331.

Using Low Power Wide Area Networks to Provide out of Band Management for Data Centers

J. Redžepagić, Z. Morić, D. Regvart
Algebra University College, Zagreb, Croatia
jasmin.redzepagic@algebra.hr, zlatan.moric@algebra.hr, regvart.damir@algebra.hr

Abstract - This paper presents a study on using low-power wide area networks (LPWANs) to provide out-of-band management for data centers. The research examines the potential of LPWANs to offer a cost-effective and secure alternative to traditional out-of-band management methods or supplement them to provide additional security and fault tolerance. The study also evaluates the technical capabilities of LPWANs in terms of range, throughput, and reliability and how these factors can affect the performance and security of data centers. Additionally, the paper examines the various LPWAN technologies available and their suitability for data center out-of-band management. The research concludes by recommending the best practices for using LPWANs for out-of-band management in data centers and providing guidelines for implementation.

Keywords – LPWAN, out-of-band management, data centers, data center, data center management

I. INTRODUCTION

In managing and overseeing data centers or any equipment in general, a significant challenge that must be addressed is effectively conveying the status of the equipment, even in the absence of a primary communication network. This paper primarily concentrates on the networking aspect that links management of data centers to the Internet. Commonly, communication connections are set up to ensure at least one alternative for each link. However, referring to these as backups is inaccurate since standby links are seldom employed as just backups, typically a data center maximizes the use of every available link to connect to the broader network.

All the available connectivity is normally used both for the regular service of the data center and to provide the management interfaces from the outside. Since the design of the data center demands that it will be able to continue functioning even if the primary communication with the Internet goes down, we configure a link failover connected to a different provider, sometimes even through a different medium. For example, if our primary link is on an optical cable, our backup link will also be on an optical cable but can be wireless or copper based.

Still, it should be running through different paths and be serviced by a provider other than the primary provider. This kind of redundancy is typical and is what most data centers do.

The main reason for using optical links is that they provide enough bandwidth for the primary service of the data center, while using them for management of the data center is secondary since the management functions require an order of magnitude less bandwidth, and management is done from inside the data center.

Historically it was common to have a backup link on copper lines. Since the price difference between copper and fiber links is now nonexistent, much higher bandwidth capabilities of data centers are the primary reason to use optical links.

There are, however, situations when every available link goes down, and they are more common than we think. Although this is the worst-case scenario for any data center, and in the design phase we do everything possible to mitigate this risk, sometimes it is unavoidable, especially in smaller data centers. The reason is commonly communal, or infrastructure works outside the data center perimeter that break or disconnect communication lines. In these cases, we rely on so-called out-of-band management strategies that enable connection to the data center itself through one of the wireless networks.

In most cases this is done through the commercial 3G, 4G or 5G network. Still, as soon as the center needs to switch to this link, we lose the ability to operate the data center and can just manage it to troubleshoot and solve the problem. In most cases for this we use commercial providers. Unfortunately, this sometimes leaves us with a single point of failure since the wireless connectivity provider can depend on the same link we used as our primary. Our provider will have the same communication problems as we have, and the redundancy of the backup link is then completely gone.

A. Commonly used wireless technologies

Total data center communication interruption is the worst nightmare of any operator. Since the main risk is physical interruption of the medium we use to connect, the commonly deployed strategy is to provide as many redundant wireless links to the outside world as possible. The technology we have seen used is mostly based on the 802.11 standards since the main idea is to provide the bandwidth needed for troubleshooting and keeping the data center online.

The problem here is that most of the standards for wireless communication were not designed for long-range

communications, which means that trying to connect the data center through a typical 802.11-based network requires a lot of additional infrastructure outside of the data center. We are talking about implementing access points or point-to-point links on locations outside the data center itself and then providing the link to the Internet from that outside location through the wireless network.

In enterprise computing, everything comes down to providing redundancy. As a last resort, creating a link that is going to enable us to have minimal control over a data center and infrastructure even if the primary and secondary providers go down is something that is extremely useful in all circumstances. Ideally, this communication provider will be completely independent both on the communication level and physically from the rest of the network so that the common failure points can be isolated as much as possible. For this, there are a couple of solutions that are, as of right now, underutilized in this context.

B. Low power wide area network breakdown

As the name implies, there is a complete set of modern technologies intended for communication in the modern Internet of Things world. We use the term low power wide area network to denote all of them since these are the main characteristics we look for in this context. Although being covered by one name, when we compare their different specifications, we will see that there are many characteristics that distinguish these networks apart.

So, what are LPWAN or Low Power Wide Area Networks?

The network part is self-explanatory, we need to concentrate on the rest of the words in the definition. The “wide area” denotes the capability of the network to cover not only points that are close together but instead to enable the communication to cover a wider area, measured in tens of kilometers or more. Inevitably this means that we are using some sort of hub and spoke configuration since we are trying to connect many devices in a single network.

In this paper, we are not interested in the network's low-power capabilities because we will presume that our equipment will run on “normal power” instead of batteries. When it comes to communicating with different sensors used in the Internet of Things applications, communicating using low power means that we can stretch a device’s battery life from hours to years.

Like all technologies, low power wide area networks face their own difficulties, and one of these challenges significantly impacts our intended goal. Networks typically facilitate direct bidirectional communication via a "channel" or a "circuit," contingent on the network type. Most of the solutions discussed in this paper will supply messages as distinct communication units. Consequently, our capacity to "connect" to the data center will be considerably constrained.

This, combined with bandwidth limitations of all the different standards, means that this proposed type of out-of-band communication will enable us only to monitor and have the minimal capability to issue commands to the data center.

II. LPWAN STANDARDS

LPWAN standards [1] on the market right now try to completely conform to the original idea of low power and wide area network. This means that different standards try to solve the problem of communicating over a large area using low power in diverse ways. This contrasts with the much more used standards like 802.11x, which have changed from their initial idea to something that tries to solve multiple different problems at once, and because of it fails its most common usage.

A. SigFox

Sigfox is one of the well-known standards for the Internet of Things created with the idea of connecting not only low-power devices but also devices in places unreachable by regular telecommunication networks, such as underground sensors (for example, water meters in urban areas). This comes at a cost since lowering the power levels required for communication also means reducing the bandwidth to a point where we are talking about 100 bps downlink speed on Sigfox. This standard was controlled by a French-based firm that filed for bankruptcy in January 2022 and is now acquired by another company that continues to run the French network and the rest of its operations.

From our point of view, SigFox is not a network we can use to communicate to data centers since the bandwidth is extremely low and the number of messages a single device can use is limited. We mention it because the network itself has a lot of available implementations, so many people first think about SigFox when LPWANs are mentioned.

B. LORA

LORA is an acronym coming from “LOng RAnge.” The name itself is proprietary name for a proprietary radio communication technique based on CSS technology and it is patented by another company from France.

LORA standard covers the physical protocol while LoRaWAN defines the communications and the architecture of the system[4]. LoRaWan is an official standard under ITU-T named Y.4480. The development is managed by the LORA alliance which has hundreds of members and is an open nonprofit organization.

The standard uses ISM license free part of the spectrum and is dependent on the geographic region. In Europe it works in the 868-megahertz part of the spectrum, in South America its 915 to 928 megahertz and in United States is 902-to-928-megahertz part of spectrum. For some applications 2.5 gigahertz spectrum is also used worldwide. Rates that can be achieved are slow - between 300 bps to 27 kbps. The exciting part of the protocol is that different gateways can receive data from one device simultaneously and then the data packets are forwarded to a centralized network server which means that reliability is high.

C. NB-IoT

Narrow-Band Internet of Things is LPWAN technology that is developed by 3GPP alliance and is mainly targeted towards cellular devices and services. This technology focuses on the low cost, long battery life and high density of the connections. Interestingly, this standard uses guard bands of the normal LTE standard to achieve speeds between 26 kilobits per second to 127 kilobits per second in downlink, or 16.9 to 159 kilobits per second uplink.

Together with LORA this standard is the one that is most heavily implemented in the field and is available across the world. The reason why it is interesting as a LPWAN solution is normally low cost and the low power requirements but the main disadvantage is that in most places it is directly connected to the mobile operators and if we require actual out-of-band communication this means that we are basically unable to use this standard since it is directly connected to being able to use the infrastructure of a given operator – if we decide that we are going to rely on something that is a de facto derivative of LTE, why not use LTE itself to establish a fully functional direct connection in the first place.

III. MESSAGING PROTOCOLS VS CIRCUITS

With all these different network technologies the main thing that is stopping us from directly connecting is their ability to provide us with a circuit or a channel of communication. All of them are based on the idea of the gateway or some other service providing a messaging queue the devices will use to deposit or send messages. Clients can connect and read messages and respond to them. Messages themselves are going to be either distributed or are going to wait in the queue to be read and the reason for all of this is that devices themselves need to be able to decide when to send data to save power .

Our proposed solution for this is to create a management network that is not only going to consist of backup communication loop using one of the standards but also that is going to be based entirely on the messaging running something like RabbitMQ or MQTT. Both of these are different standards of messaging protocols common in the Internet of Things space. When we compare these two, we will probably use MQTT since it is much closer to the idea of using the smallest number of messages to exchange data.

RabbitMQ, although also a similar messaging protocol, is primarily targeted to large systems that are online and exchange many messages. This is in stark contrast to what LPWAN networks are trying to accomplish.

This idea also has one unintended consequence that is extremely useful to us. Since a lot of the monitoring in the data center is based on monitoring different parameters of the data center itself - for example temperature, fire detection, monitoring current and power or monitoring access we can completely decentralize this part of the monitoring network. Instead of using a dedicated monitoring network inside the data center and a separate one to monitor devices from outside, we can use the low-power network to directly get the data into the local MQTT server and then provide synchronization to the servers outside of the data center. Add to that the ability to issue commands remotely and we have a working solution.

IV. INFRASTRUCTURE DESIGN

What is our proposed design? We will divide it into a few separate logical units, starting with the local infrastructure.

The idea is to install all the sensors that will monitor our local environment as we would normally do but to add to all the sensors the capability to send messages. Of course, sensors that are LORA capable from the start will be prevalent. Still, there are things we need to monitor that are incapable of messaging or connecting to anything other than a normal network.

Since we need messages to have a common platform for exchanging information, every sensor should be able to send them.

Ideally, this would be through a local LoRaWan Gateway. Still, it is also possible to use the local network since we are under assumption that we are designing a system to enable troubleshooting our data center from anywhere and that right now we are planning to mitigate the risk of all the links to the internet being disconnected. If we can create a local network of sensors that will be completely independent of the local LAN, this is just a nice bonus.

Next step is to create a local MQTT server that will collect all the information from the sensors. This server will be capable of forwarding all the messages to an external MQTT server outside our monitored network.

These steps cover the idea of monitoring our data center. To add functionality that will enable us to troubleshoot and possibly solve problems remotely we also need to create some sort of way to issue command and reconfigure equipment in the datacenter.

Our proposed way would be to create premade triggers triggered by messages containing commands we will issue from our control client. This part of the solution will have to be custom made, there are no solutions for this that are made for this particular usage case, but this capability exists in the protocol, messaging supports not only messages but also different ways of forwarding complex data.

Idea of premade commands would be to be able to issue specific commands with variable parameters to units in the data center, we could for example create a bridge to our virtualization environment to establish control over virtual machines or create a bridge to internal management network to reconfigure switches and routers.

The last part of the system would be the command-and-control client that will connect to the internal or external messaging server. Its main purpose would be to be able to display run-time parameters of our data center and to issue commands.

Unfortunately, for this paper we can only give a general idea how to design such a system. Further research and development is required to implement our ideas.

V. Limitations of MQTT and RabbitMQ

Main problem that we are going to have with this solution is security. MQTT as a server and protocol are inherently unsecured, and we would have to implement encryption and security for all the data in transit. Using RabbitMQ would solve some problems since it has built-in security for communication, but RabbitMQ has other difficulties when dealing with large numbers of IoT sensors. This dilemma will have to be resolved in the implementation itself.

VI. Conclusion

Implementing LPWAN solution to both monitor and control data centers is interesting from two standpoints: it provides a way of monitoring the data center utilizing a completely separate network, and thanks to a large number of inexpensive sensors it gives us the ability to monitor a lot of metrics that would otherwise require more complex solutions to implement, things like power monitoring, access monitoring and such.

We propose a straightforward design comprising of two separate servers that will enable both local and remote monitoring of data centers, with limited command capabilities. This solution can be created to be completely independent of existing networks, both in the data center and the Internet. The hardware required is already available and implementation can be done using existing software with slight modifications.

Result is a monitoring solution that enables both on-site and off-site monitoring, is resilient to network problems and is scalable to even the largest installations.

Reference

[1] Ayoub, W., Samhat, A. E., Nouvel, F., Mroue, M., & Prévotet, J.-C. (2019). Internet of Mobile Things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and Supported Mobility. *Communications Surveys and Tutorials*, *21*(2), 1561–1581. https://doi.org/10.1109/comst.2018.2877382ï

[2] Gill, P., Jain, N., Nagappan N., Understanding Network Failures in data centers: measurement, Analysis, and implications, Sigcomm 11 Conference Committee. (2011). *Proceedings of Sigcomm 2011 and Best Papers of the Co Located Workshops.* Association for Computing Machinery.

[3] Gall, J., & Gall, J. (2002). The systems bible is the beginner's guide to systems large and small , the third edition of Systemantics. General Systemantics Press. ISBN 0961825170

[4] https://lora-alliance.org/about-lorawan/ accesed on 20.2.2023.

Establishing the Guidelines for Using Refurbished Hardware in Creating New Data Centers to Lower the Amount of E-Waste

J. Redžepagić, V. Dakić, Z. Morić,
Algebra University College, Zagreb, Croatia
jasmin.redzepagic@algebra.hr, dakic.vedran@algebra.hr, zlatan.moric@algebra.hr,

***Abstract* - This paper presents a study on the guidelines for using refurbished hardware in creating new data centers to lower the amount of electronic waste (e-waste). The research examines the potential of refurbished hardware to provide a cost-effective and eco-friendly alternative to new hardware in data center construction. The study also evaluates the technical capabilities of refurbished hardware and how these factors can affect the performance and security of data centers. Additionally, the paper examines the various refurbishment standaorganizationsrds and certifications available and their suitability for data center use. The study also explores the implications of regulatory compliance and industry standards on the use of refurbished hardware. One of the main goals is to define guidelines on what hardware is viable as reusable in data center environments, especially given that most modern software infrastructure is designed with hardware failures in mind and supports hot-swapping parts of the hardware layer. The findings of this paper intend to create a guide for organisations looking to lower their environmental impact and reduce costs while building new data centers.**

Keywords – e-waste management, reuse, extending server lifecycle, reusing hardware, resilient system

I. INTRODUCTION

The typical design of data centers assumes that all the equipment, including infrastructure, servers, and network equipment, is brand new and unused. Green technologies and ecologically friendly designs for data centers aim to reduce carbon emissions and energy usage, but they don't typically consider the possibility of recycling equipment. This paper argues that used equipment should be included in data center design and explores cases in which reusing previously used hardware is feasible.

II. DATA CENTERS

The data center usually denotes the entire infrastructure that enables the Internet today. According to a few surveys, the number of data centers globally is estimated at around 8000, growing at 5% year-to-year [1],[5].

A typical perspective on data centers is that they consist of computers and other necessary infrastructure to perform a specific task or process. In some instances, specifically configured equipment may be used for particular case scenarios, but for the most part, data centers are initially designed as generic units. The primary focus during data center design is maximizing availability, which is achieved by ensuring the equipment is reliable and easily replaceable. This is accomplished in design; data centers are designed from the ground up to be resilient to failures in every conceivable way.

Design starts with choosing a site for the data center. The usual aim is to ensure that redundant power and connections to the rest of the network exist on the site before we begin building – redundant connections are expensive, especially power supply.

The process of designing a data center then tries to cover other essential things: power consumption needs to be in line with what can be provided by the existing network on site. If more power is needed, it may be required to implement additional power lines.

Cooling is also a significant consideration since it has to be able to control the entire heat output of our equipment, and by design, it also needs to be fault tolerant in the same way as the rest of the equipment. Cooling creates additional requirements in terms of space required for the cooling system and the data center structure since cooling equipment tends to be heavy.

Power and cooling are outside our scope of interest, mainly because there is almost no secondary market since the equipment exploatation period is much longer than for the servers.

The traditional design also involves installing and provisioning networking, storage, and computational infrastructure, which up until recently, was always purchased new. The main reason was failures which complicate data center operations. While this viewpoint is understandable, it has evolved over the past few years.

The main thing that has changed is that the software infrastructure running in the data centers is much more resilient to hardware failure than it was a couple of generations ago. In reality, this means that the hardware running in the data centers right now can experience even significant failures while the data center itself can and will simultaneously be completely operational.

While this may seem normal since resistance to hardware failure is ingrained in the fundamentals of the data center design, it has dramatically changed over time. At first, failures were dealt with as exceptions, and failures rarely happened. This led to the creation of infrastructure designed with a small number of points of failure. This meant a backup system was in place for each system, but the general idea was that double failures were reasonably rare.

A. *Fault Resiliency*

Standards for the data centers specify levels of fault resiliency. For the most part, smaller data centers consider redundancy only in critical systems. Usually, this redundancy is noted as “n+1”, meaning the system can resist a single failure before it ceases to function.

With the rising complexity of today's systems, this had to be changed since the sheer number of installed devices meant that failures became commonplace. This means that instead of functioning with zero failures, almost all of today's systems spend a lot of time in a degraded state. This is a concept that has been discussed for years[2].

This degradation can be seen from the user's perspective as either transparent or detrimental.

The term "transparent degradation" refers to a situation where some parts of a system are not functioning as they should. Still, the user is unaware of any issues and believes the system functions normally. The system can detect that specific infrastructure components are not working correctly, but can work around the problems, so the user is not impacted.

On the other hand, "detrimental degradation" refers to a situation where the user experiences problems using the system, either partially or entirely. This may involve the system being completely offline, or performance being so poor that users cannot complete their tasks. In this scenario, the user is aware of the issue, which is considered a problem.

When building a resilient system, the goal is to create one that can continue functioning despite infrastructure failures. By recognizing that failures are a normal part of the system's operation, rather than exceptions, we can design systems that can withstand more failures. We approach hardware and infrastructure failures as common occurrences rather than exceptions and design our data centers accordingly.

B. *Replacing equipment*

According to the surveys published by the Uptime Institute, servers in today’s data centers are getting older:

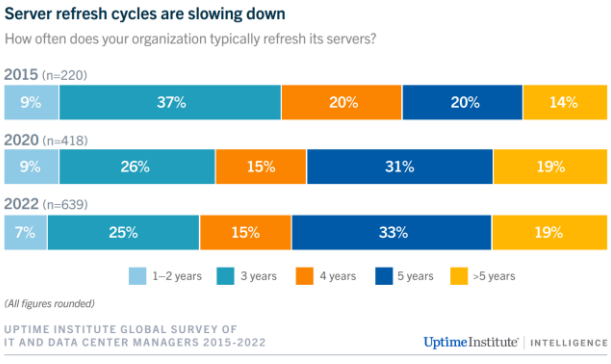


Figure 1 - Server refresh cycles in the last 7 years [1]

Data centers today are less inclined to buy new equipment but also to put older equipment out of commsion. Part of that has to do with “green” initiatives [3], and part is simple economic logic.

What are the reasons the equipment in the data center may be replaced in the first place? One reason for this is when it stops working, which is inevitable when dealing with a large number of units of any hardware. It doesn't matter if we are working with servers, disks, networking equipment, cables, or whatever; the large number of units we will use means that they will be failures and those failures will require replacement.

The second reason for replacing equipment is trying to optimize the process. This usually means that we are trying to either speed up the operations of the data center for a given power envelope or that we are trying to lower the power requirements or the power of the data center while keeping the operation speed at the same level[4]. By operation speed, we usually mean the number of operations or services the given data center can perform when operating normally. In this case, we usually call these replacements "upgrades," but in the end, we are swapping "old" for "new" equipment.

In addition to the other reasons for replacing equipment, a third one involves warranty considerations. When the standard warranty period for a particular unit has expired, repairing or replacing the equipment can become challenging if it fails suddenly.

This is why the equipment that runs out of warranty is usually removed from operation and either disposed of or sold. This is done primarily to ensure that the data center meets its operational agreements with its users. Using equipment without clear service support can create significant logistical issues that may call these agreements into question.

C. *Service levels*

Considering all of the aforementioned points, it is becoming increasingly common for IT infrastructures to require different levels of equipment availability. This is due to the complexity of many systems being used, which often necessitates multiple environments within a data center. One such environment is the production environment for clients, while one or more test environments with less strict availability requirements are also typically used. Even data center users are willing to pay a reduced cost for a test environment with lower availability if it is exclusively used for testing.

By recognizing the varying levels of availability required for different equipment in a data center, it may be beneficial to consider dividing the data center into two separate units. One unit would be dedicated to production, with a high level of availability, while the other would be used for testing and have a lower guaranteed level of availability. This proposal acknowledges the different needs of production and testing environments and aims to optimize the data center's overall efficiency and functionality.

In essence, this means that we will create something reliable but much more ecologically and power friendly with the idea of reusing the hardware that would otherwise be thrown out or not used in a data center environment.

D. Complexity requirements

Analyzing production environments becomes even more complicated when we examine them in depth. Some production environments can be easily replicated without any backups. In contrast, others require lengthy restoration procedures that can take hours, days, or weeks to bring the system back online after a complex failure.

We are going to deal with two examples:

The first one is the usual business-oriented production environment for the company that contains all the systems and services required to run financial systems, CRM (Customer Relationship Manager) systems, front and back-end web systems, and everything else the typical company has. All of this infrastructure needs to be online 24/7.

The other production environment is intended to serve students in their needs for testing and learning environments. Compared to the other environment this student setup is considered production but can be recreated from scratch with only a minimal set of data since all the data in the system is considered test data from the start.

The first environment requires regular backups and a lot of stored and secured data. This means we should design an environment to minimize downtime and avoid restore procedures since they have to move significant amounts of data around.

The second environment can be recreated from scratch in hours using only scripts.

E. System design

One approach that we can take in light of these considerations is to design two different systems. The first would be prepared to run on brand-new servers, while the second would utilize older servers to reduce the cost of acquiring and operating hardware. Typically, older servers are easier and cheaper to acquire and maintain, with more affordable spare parts readily available for upgrades. The price difference between new and old equipment can be significant, often more than enough to compensate for the potential risk of the equipment not performing as expected. If any server fails, it can be easily replaced with another used server at a fraction of the cost.

Until now we have only dealt with the hardware part of the equation since it is the most obvious. In our considerations, we must also take into account software and licensing. While designing a data center, this part of the design is crucial since it has a substantial cost associated with it while at the same time being almost entirely dependent on the hardware design.

Modern software licensing practices are evolving to become more closely tied to the hardware resources a given application utilizes. One notable example is virtualization environments. Licensing is typically determined by factors such as the number of CPU (Central Processing Unit) cores or the total amount of memory allocated for running the environment. This trend has emerged as modern servers continue to expand their capabilities, offering a significantly larger number of CPU cores and increased memory capacity to users.

As licensing fees are often levied on a per-active server basis, encompassing both production and standby servers, this can result in substantial costs for organizations. This financial burden is further worsened when the licensed software is run on refurbished hardware, which is often more affordable than new equipment. In such cases, the total cost of software licensing may even exceed the actual cost of the hardware on which the software is running.

Organizations can better anticipate and manage their expenses by understanding the increasing reliance on hardware resources for software licensing, ensuring a more cost-effective approach to deploying and maintaining software in both production and standby environments.

III. REFURBISHED HARDWARE

When dealing with your refurbished hardware, we must define some things. Other than its positive and negative sides we also need to define the sources we will get it from since we are dealing with the expanding market currently estimated at $5 billion annually. At the same time, the market is almost wholly avoiding equipment manufacturers and relying instead on proxies and second-hand dealers.

A. Sources

There are a couple of ways of acquiring refurbished hardware. The simplest is using one of the many dealers specializing in this business. In the European market, a couple of multimillion-euro companies deal exclusively with refurbished hardware, even offering warranties and spare parts. These companies usually have agreements with big data centers and other large users of enterprise-level equipment.

They usually take care of all the equipment that is either considered broken or old and provide a service where companies can sell their old equipment to the dealer instead of paying for the equipment to be destroyed.

Dealers then evaluate this equipment, check its condition, test it and then usually offer it on the market for a price that varies based on different criteria. Most commonly the most significant criterion is the overall state of the devices.

In many cases, if we are prepared to accept the devices that have scratches and minor dents but are in complete working order, we can acquire equipment for a price that can be measured in a single-digit percentage of the price of the new equipment.

The secondary market also consists of smaller dealers more focused on moving the boxes between companies and end users offering lower prices but providing no additional service and no testing.

There is also the third way of acquiring refurbished hardware by directly contacting companies and data centers, but this is out of the scope of this paper since it

requires lots of additional logistics and provides many additional problems.
Some data centers also rely on their inventory of used equipment since they can stretch the life cycle of the equipment they already have.

B. Reasons to buy

The primary reason for considering refurbished equipment is typically its lower cost, which can be orders of magnitude less than new units, depending on the age of the acquired equipment. Additionally, depending on the use case, refurbished equipment can offer hardware much more capable than new units at the same price point. This is especially applicable to network equipment in smaller environments. This provides the opportunity to purchase more hardware within a set budget or acquire equipment that would otherwise be unattainable at a given price point.

C. Reasons not to buy

Primarily, two things make refurbished equipment non attractive. The obvious one is that we are dealing with the equipment already used, which is pretty unpredictable when it comes to failures. This is a problem that can be dealt with by using systems able to withstand a lot of failures, as we described earlier.
The second issue is considerably more intricate because it involves outdated equipment which, depending on its type and the advancements in this specific market segment, may have a lower cost-to-performance ratio compared to newer equipment. Although we pay less for the older equipment, we receive considerably less performance than we would if we invested in new servers.
Measuring or determining what constitutes “better” equipment is not simple since performance involves different metrics. For example, we may be getting less CPU power but at the same time we may be buying a device that has much more memory and is much more suitable for the task at hand.

IV. Designing a refurbished data center

If we create a data center that relies on refurbished equipment, the primary focus should be the usage case. Typically, when designing a data center, the aim is to cover as many general cases as possible to reach the widest market. However, when working with refurbished hardware, we need to consider the likelihood of hardware failures from the outset and strive to target users or usage cases that can cope with such service disruptions. Therefore, the first step in designing the data center should involve understanding the user and their requirements for the desired service level. Additionally, licenses may pose a challenge since we may need to purchase more resources than necessary to account for the different failure modes of the equipment, and licenses themselves can be expensive since they are directly correlated to the amount of resources we use.

A. Software

Deploying to bare metal servers is almost entirely out of the question when designing with refurbished hardware, so we will probably go with one of the virtualization environments. This is going to be one of the biggest hurdles to overcome since we will pay for the licenses used to cover a lot of resources that will just be spares and hot standbys. Reason for this is that in a regular data center overprovisioning is considered waste of resources and is usually kept to a minimum, so the licensing usually has no provisions for services that are in standby and not in active use.
Our design must do the opposite and consider a lot of overprovisioned hardware to ensure that unexpected hardware failures can be replaced on a live system. That may mean buying licenses to cover the unused hardware simply to have it available if needed.
Aside from virtualization software, other licenses will be the same as in a regular data center.

B. Hardware

Once initial user requirement specifications are defined we should consider which hardware to order.
The easiest way is to find a reputable dealer with an explicit warranty and replacement policy since this will solve many problems with the hardware. Most of the dealers also have a strict return policy that covers hardware for at least a year so we can count on getting the spare units we need.
In our experience networking hardware is probably the easiest to replace with the refurbished one. Networking equipment is something that usually simply works[6].
If we are working with the dealer that tests the equipment before delivery we can almost be sure that the equipment will function correctly.
Depending on the network speed and the required adapters, we will also order the cables end connectors but when it comes to the access network, the patch cables should probably be bought new since they are most prone to damage.
When it comes to exact configuration in our experience when buying servers the best strategy is to contact the dealer and find out what servers they have in stock. Once you know that, we can define the services that will run on the server, for example, storage, databases, or virtualization, and then we can decide on which particular server to buy.
Almost all the dealers offer customization of the servers they have and the main thing we will usually do is maximise the amount of memory in the servers.
The only thing that we are not going to buy refurbished is the disks since they are the most vulnerable part of any system being at the same time the part that gets the most wear and tear. Remember to always order trays and different components needed to fix the disks and other adapters to the servers. Some dealers will offer things like this for free, but always check with them before agreeing on the server configuration.
Other than servers, we can also buy refurbished UPS (Uninterruptible Power Supply) devices and power

equipment, but we should pay attention to the batteries since they are basically spare parts and should be bought new.

In the limited time we are running such an environment we still haven't managed to collect all the empirical data needed for a serious conclusion, but we can say a few things as preliminary observations.

We tried running a virtual environment that students use for testing network and software configurations. First iteration used 40+ servers which was later reduced to 8 servers in total. Storage servers are separate entities.

In regards of budget, preliminary results say that we created a data server that is about 40% of the price of the equivalent center that would be created from new hardware. Performance wise we are more or less at the same perceived speed. Only thing is that power usage was much (almost order of magnitude) greater when using the oldest servers. We still have no numbers to relate last generation of servers we are using to new equipment.

V. CONCLUSION

It is completely feasible to create a working data center using mostly refurbished equipment. Depending on the usage scenario this may have two big advantages. First is reusing equipment and being completely eco friendly by not throwing away perfectly functioning devices and second is the price. While some of the devices cannot be reused, such as disks, most of the datacenter can be refurbished. Market for refurbished servers exists and is quite large making it even possible to buy devices with warranties, although almost exclusively we are talking about being able to exchange devices by returning the malfunctioning unit and getting another refurbished one in return. If we utilize this model it is possible to save money and help the environment while attaining almost the same level of service for the users.

On the other hand, some usage cases cannot support this model of refurbished data center and almost require new equipment.

Considerations for purchasing or othervise acquiring refrubished equeptmen should be:

- Are we getting equipment that was checked before coming back to market?
- Is there a "replacement" warranty in place?
- Is the vendor able to provide multiple units (usually the signal that replacements will be available later)
- Is the equimpent tested?
- Can the vendor provide his list of references? Big refurbished equpment vendors can provide you with the actual source of the equipment.
- Is equipment delivered cleanly reset and upgraded to the latest versions of the firmware?
- Is vendor providing guidance and advice on purchases?

Although not exaustive, this list of questions should help when deciding on a particular vendor in a market that is increasing each year.

REFERENCES

[1] Bizo, D. (2022). Uptime Institute Global Data Center Survey 2022 Resiliency remains critical in a volatile world

[2] Gall, J., & Gall, J. (2002). The systems bible : the beginner's guide to systems large and small : being the third edition of Systemantics. General Systemantics Press.

[3] Bashroush, R., Rteil, N., Kenny, R., & Wynne, A. (2022). Optimizing Server Refresh Cycles: The Case for Circular Economy with an Aging Moore's Law. *IEEE Transactions on Sustainable Computing*, *7*(1), 189–200. https://doi.org/10.1109/TSUSC.2020.3035234

[4] M. Dayarathna, Y. Wen and R. Fan, "Data Center Energy Consumption Modeling: A Survey," in IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 732-794, Firstquarter 2016, doi: 10.1109/COMST.2015.2481183.

[5] Daigle, B. (n.d.). Data Centers Around the World: A Quick Look The Exponential Increase in Data. https://www.usitc.gov/publications/332/executive_briefings/ebot_data_centers_around_the_world.pdf Accessed on 28.4.2023.

[6] Sigcomm 11 Conference Committee. (2011). *Proceedings of Sigcomm 2011 and Best Papers of the Co Located Workshops.* Association for Computing Machinery.

INFORMATION QUALITY AND SECURITY

The Challenges to Information Security in Croatian Chamber of Economy during COVID-19 Pandemic

Dražen Lučić
Croatian Chamber of Economy, Zagreb, Croatia
dlucic@hgk.hr

Abstract **- The paper deals with analysis of both challenges to information security Croatian Chamber of Economy prior and after COVID-19 pandemic and the results of the measures that were taken in order to counter fought with those challenges by means of state-of-the-art Information and Communication Technology solutions. COVID-19 pandemic has rapidly and significantly changed ecosystem for social and business activities. Croatian Chamber of Economy, as legal entity with public authorities, has been affected by numerous limitation which have directly impacted effectiveness of the business processes. The analysis of the measures and activities undertaken by Croatian Chamber of Economy in order to keep a high level of information security during COVID-19 pandemic and experience from that time period is described in the paper. Pre-emption of possible events with negative influence on business activities, as well as risk management and crisis management plan, are some of the most important prerequisites for a successful continuation of business activities during a crisis such COVID-19 pandemic has been. Beside the solutions based on ICT technology, permanent education of the employees is the activity that must not be omitted.**

Keywords – Information Security, Cyber Security, COVID-19 pandemic, legal entity with public autorities, Croatian Chamber of Economy

I. INTRODUCTION

As information and communication technology (ICT) have advanced in their capabilities, and especially with the greater availability of high-speed internet, remote working has grown in its use as a new mode of work in the past several decades [1]. However, prior to the COVID-19 pandemic, remote working was not a widely used practice [2]. Although the number of employees who worked from home at least half of their working time grew in last ten years, remote working at that time was below five percent of the total workforce in both US and Europe [3]. Remote working has, in fact, been a "luxury for the relatively affluent" [4], such as higher-income earners and "white-collar workers" e.g. executives, managers, or professionals.

Because of this situation, prior to COVID-19, most workers had little remote working experience; nor were they or their organizations prepared for supporting this practice. Now, the unprecedented outbreak of the COVID-19 pandemic in 2020 has required people across the world into being remote workers, inadvertently leading to a de facto global experiment of remote working [6]. Remote working has become the "new normal," almost overnight.

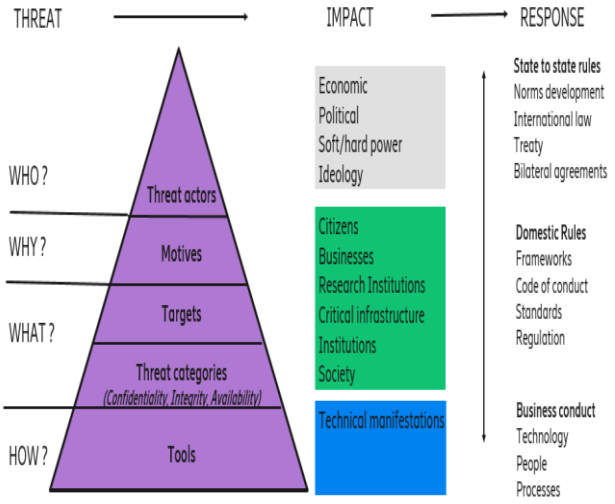


Figure 1. The cyber threats at electronic communications market

In order to cope with the new social and business situation, many companies, state and public institutions, legal entities, etc. have found own solutions for remote working and new business process. These solutions have been mostly based on ICT and electronic communications. Electronic communications, empowered with high speed data connections and high data rates, have shaped society and markets. Today we speak about digital/information society, digital transformation and new industrial revolution. Digitalization in now an intrinsic part of every country's sustainable economic and social development. However, digital transformation and the expansion of the digital ecosystem also comes with increased cybersecurity risks, especially for but not limited to the small and media enterprises and legal entities that may lack adequate cyber resilience against constantly evolving digital cyber threats. While the private sector and policy makers are continuously working to increase the resilience of digital infrastructure, software and devices top make harder for malicious actors to succeed in their objectives, this is not sufficient to break the growing trend and decrease cyberthreats (Figure 1) [6].

In parallel to this significant technology change in electronic communications there is an another interesting process ongoing. European Union (EU) commission published several documents that strongly influences data processing business as well as information security. In 2016 EU Parliament introduced General Data Protection Regulation (GDPR) [7]. GDPR is a data protection law which came in force in 2018 with influence on information security and EU single digital market as well. This has a huge impact not only huge impact on electronic communications market regulation in but also information security in the areas where new technology solutions are applied for business processes.

II. INFORMATION SECURITY AT LEGAL ENTITIES WITH PUBLIC AUTHORITIES

The transformation to digital society and digital economy is one of the results of ongoing process of digitalisation and globalisation. On one side this process creates enormous opportunities for business but on another side could become obstacle or even a huge hurdle in implementation of new technologies. The process is accompanied by a growing public awareness and concern for the importance of information security that also comprises cyber security [8].

Information security is one of the key issues to be addressed by legal entities It gives the importance of transparency and the risk of personal data breaches. To overcome possible information security problem, legal entities introduce numerous requirements regarding the confidentiality and security of personal data. This process implies also a number of changes that the legal entities need to implement, particularly in terms of security, privacy, personal data processing and protection as well as user services. Another threat is possible vulnerability user services because of their purpose, especially when legal entity has some public authorities. Some of user services are not applicable in existing telecommunications network due to satisfactory quality of service (QoS) and that might cause difficulties in providing such service by means of remote working. Wide use of mobile internet access via 4G and 5G public mobile telecommunications network might overcome the issue with QoS.

Due to the fact that almost nobody has been prepared for such unprecedented events and situation due to COVID-19, strengthened in Croatia by strong earthquakes in 2020, remote working at legal entities in Croatia in common haven't been prepared for neither organisational nor security challenges of remote working. In the same time, the users have expected the same QoS as it was the case prior to COVID-19 [9]. Lack of procedures and experience for such situation at legal entities has caused confusion and stoppage of usual business processes. Those legal entities that have had state-of-the-art ICT infrastructure and skilled employees at both IT and corporate security department, adapted quickly to "new normal" eco-system thus providing services with a good QoS and keeping a high level of information security at the same time. The complete new eco-system has been to cyber threat environment but that has been the challenge that the legal entities had to take on their shoulders (Figure 2).

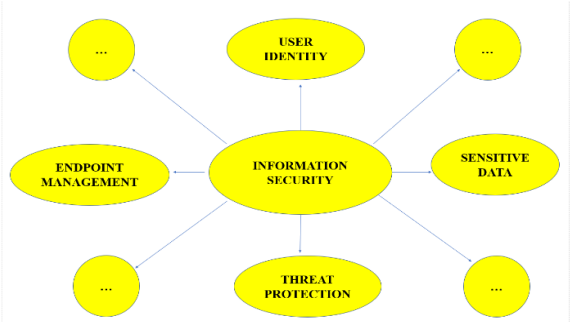


Figure 2. Information security at legal entities

III. INFORMATION SECURITY AT CROATIAN CHAMBER OF ECONOMY DURING COVID-19 PANDEMIC

Croatian Chamber of Economy (Hrvatska gospodarska komora – HGK). as legal entity with public authorities [10] with headquarter in Zagreb, has experienced severe challenges in the area of information security during COVID-19 pandemic, emphasized by strong earthquakes in Zagreb and its vicinity in March and December 2020. Given the constant threat and responses to cyber-attacks, it has been essential that HGK was technically prepared and with skilled employees who knew how to conceptualise cybersecurity threats, impacts, and responses [11].

Department for information security (OIS) at HGK already beginning of the year 2020 prepared enough lap-top computers, mobile phones, wi-fi routers and applications for a successful and efficient remote working. Applications have been mostly based on an industrial standard solution, Microsoft Office 365 with enterprise licence that includes, among others, MS Teams for video conference and collaboration as well as the tools like "SharePoint" and "OneDrive" for easy collaboration of remote workers. Accidentally just few months before COVID-19 pandemic officially was declared in Croatia, OIS conducted at HGK a survey in order to find out how employees, infrastructure, organisation and prepared are prepared for the challenges in the areas of both information and cybersecurity. The activity was a part of regular annual review of the Information Security Management System at HGK and the first step in preparation process for ISO 27001 certification.

Out of the results of that survey is shown in the figures 3 and 4, where 413 employees at HGK, who are working on 21 locations throughout Croatia and on 7 locations in Europe and China, answered to the question about the cyber threat(s) that they were exposed to. The result confirms general picture about the cyber threats in both business and private eco-system. The same survey, conducted in February 2023, shown that Spam and Phishing remain to be the most frequent cyber threats at HGK but without any negative impact on business due to educated employees. Meanwhile, the percentage of threats in internet shopping and investment fraud during COVID -19 pandemic decreased by roughly ten percent. "Courier fraud" increased during COVID-19 pandemic. It is interesting that "Romance fraud" has become a growing

problem, exacerbated by the COVID-19 pandemic, but hasn't been reported by any of the employees in the survey neither in 2019 nor in 2023.



Figure 3. Explanation for cyber threats

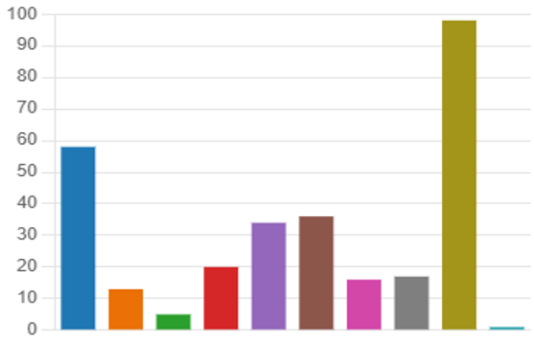


Figure 4. Cyber threats at HGK during COVID-19

IV. The Measures Taken in Order to Increase Level of Information Security at Croatian Chamber of Economy

Mitigating and preventing cyber-attacks are not a trivial task. There are practical approaches that can reduce the risk of cyber-attacks while working remotely. Security is an essential element of trust in new and emerging technologies and a factor that can impact any organization connected to the Internet, but it is not a one-size-fits-all solution and thus not suited to narrow top down prescription. As technologies continue to develop and new technologies continue to emerge, cybersecurity implications need to be considered on a case by case basis. EU has also recognized the importance of information and cyber security. They issued "Cybersecurity Act" in 2019. in order to counter fight with continuously increasing risk on both information security and cybersecurity [12]. The act is the base for many documents and guidelines prepared by EU agency for cybersecurity (ENISA), which include description of the threats to information and cybersecurity and measures how to cope with new challenges be means of state-of-the-art ICT solutions. Therefore, OIS at HGK performed many activities and undertook a lot of measures in order to increase level of information security at HGK during COVID-19 pandemic, especially due to many employees at HGK who over the night became remote workers on numerous locations throughout Croatia [13].

People are considered the weakest link in many security systems. Therefore, developing cybersecurity awareness among users by means of constant training is important to reduce the risks of cyber-attacks on an organization. HGK conducted online education in the area of information and cybersecurity for all employees with 20 to 40 terms yearly in the time period from 2020 to 2023. Number of employees in that period declined from around 550 to around 300.

HGK has implemented and provided to all employees as well as to the selected suppliers representatives Virtual Private Network (VPN), an encrypted communication channel between two points on the Internet to protect the data that is sent and received. HGK VPN provides two aspects of security: confidentiality and integrity and allows HGK to extend security policies to remote workers and suppliers. HGK has strengthened security by Enabling two-vector authentication (TVA), requiring a username and password plus a one-time code sent to user's mobile phone via an authentication app. TVA is an important factor to mitigate against password guessing and theft such as brute force cyber-attacks.

It has been ensured that all devices and equipment firmware/OS are up-to-date with the latest security patches implemented to inoculate them against known vulnerabilities. Regular and up-to-date patches may reduce the risk of a zero-day attack. It has been also ensured that up-to-date anti-malware software was activated in all network connected devices. regular and up-to-date anti-malware may reduce the risk of cyber-attacks caused by malware.

OIS has set a strong company online policy through Information Security Management System (ISMS). Organizational units at HGK had little or no time to prepare for remote work. Robust and comprehensive security policy is necessary to protect data and prevent cyber-attacks. Strong ISMS includes avoiding holding sensitive work conversations in public, use only company-approved video and audio conference lines, etc. The policies should also include a robust and proven recovery plan and backup strategy.

Physical security of home office has been emphasized in HGK ISMS as well. It is important to physically protect home office devices. Practical approaches include ensuring that work devices are not left unattended, use a

lock screen or lock the laptop, always log off devices after use, etc.

It is imperative that legal entities with public authorities protect their valuable data and assets from cyber-attacks by improving their defense. Two important components as regards detecting malicious behavior that can compromise the security and trust of a network are intrusion detection system (IDS) and security incident and event management (SIEM). Typically, an IDS employs anomaly detection, stateful protocol analysis (deep packet inspection), signature matching or a combination of all three techniques (hybrid) to analyze incoming cyber-attacks. Furthermore, it is important for legal entities with public authorities to take a comprehensive approach to cybersecurity and not to view security from a technological perspective only, but in the framework of their business processes.

V. Analysis of information security at Croatian Chamber of Economy

During COVID-19 pandemic, OIS at HGK performed yearly online education in the area of information and cyber security for all employees at HGK and selected representatives from the suppliers to HGK. In total more than 100 online sessions have been performed with more than 1 000 participants in total. At the end of each education cycles, OIS has conducted a survey in order to gain a better picture about behaviour, knowledge and experience of the employees at HGK in the area of information and cyber security. There were 364 participants of the survey conducted in the year 2019 and 213 participant of the survey from February 2023. Some of the results obtained from both survey and their comparison are presented in the table 1.

OIS also performed a risk assessment that should identify the typical and most significant threats and their relevance for the business process in HGK during COVID-19 pandemic. The assessment has included the risks due to circumstances of the strong earthquakes in the year 2020 when function of both HGK headquarter in Zagreb and regional office in Sisak were severely disrupted. The assessment presented in the Table 2 [14] includes the following high-level categories of information/cyber security threats and threat actors. The main threat actors are:

- Accidental (Acc.) threat actor, e.g. unintended impact or a side effect from an operation not targeting the operation of a mobile telecommunications network;
- An individual “hacker” (Ind.);
- A “hacktivist” group (Hack. group);
- An organized crime group (Crime group);
- An insider within a telecommunication operator or vendor;
- State or state-backed actor.

Table 1. An comparison of the activities and behaving of the employees at HGK prior and after COVID-19 pandemic

ACTIVITY	YES [%] November 2019	YES [%] February 2023
Have you used a remote access for business activities?	8	93
Have you already attended an (online) education in the area of information security?	2	90
Have you encountered with malicious computer network/software activities?	37	74
Do you back up your computer data regularly?	7	91
Do you take regular software update of your computer?	13	91
Do you take regular software update of your “smart phone”?	48	89
Do you use “OneDrive” for your daily work?	6	93
Do you use “SharePoint” or similar tool for document share?	15	69
Do you use “secure print” in your office?	0	27

The main threats are:

- Compromised confidentiality, including espionage;
- Compromised availability
- Compromised integrity of a service.

Relevance rating is from 1 to 5, i.e. very low, low, medium, high and very high respectively.

Table 2. Summary of findings on main threats to information security at Croatian Chamber of Economy (HGK)

Actors / Threats	Acc.	Ind.	Hack. group	Crime group	Insider	State
Confidentiality	2	4	3	3	4	3
Availability	3	3	3	2	4	2
Integrity	3	3	2	2	3	2

The main threats to information security in HGK don't differ significantly from the main threats in majority of legal entity with public authorities in Croatia. The threat of an intentional cyberattack is high or even very high in a case that main actor behind is a state or a state-backed group. In the case of HGK a threat of a cyberattack from individual hacker or a group of hackers is lower than in majority other legal entities in Croatia due to a preety high level of information and cyber syecurity in HGK, back by a strong ISMS. However, the threat from an insider at HGK or their suppliers is still significantly high and on a similar level to the other legal entities in Croatia.

VI. CONCLUSION

Information security, including cyber security, remains a foundational element of all business activities of the legal entities with public authorities as our societies continue on their path for rapid digitalization. The COVID-19 pandemic was a remarkable and unprecedented event which altered the lives of citizens globally as well as the business processes. Aside from the extraordinary impact on society and business as a whole, the pandemic generated a set of unique cyber-crime related circumstances which also affected society and business. In the same time the impact of cyberattacks and other threats to information security grows, especially because the actors increasingly target critical infrastructure and systems in order to disrupt their functions.

During COVID-19 pandemic, cyber criminals and APT groups have taken advantage of targeting vulnerable people and systems. Furthermore, it is a situation that is unlikely to change in the foreseeable future. Legal entities with public authorities are one the victims of cyber-attacks during the pandemic for various reasons. Hence, it is crucial that these organisations improve protection of their important data and assets from cyber-attacks by leveraging their defence by implementation of comprehensive mechanism and state-of-the-art ICT solutions in order to increase level of information security. Certain practical approaches to reduce the risks of cyber-attacks and possible mitigation techniques are presented in this paper with Croatian Chamber of Economy as the study case.

REFERENCES

[1] V. Di Martino, L. Wirth, „Telework: a new way of working and living", "International labour review", Vol. 129 (5), pp. 529-554, 1990

[2] E.E. Kossek, B.A. Lautsch, "Work–life flexibility for whom? Occupational status and work–life inequality in upper, middle, and lower level jobs", "The Academy of Management Annals", Vol. 12(1), pp. 5–36, 2018
https://doi.org/10.5465/annals.2016.0059

[3] "Living and working in Europe", Annual report 2017, 2018 https://www.eurofound.europa.eu/publications/annual-report/2018/living-and-working-in-europe-2017

[4] B. Wang, Y. Liu, J. Qian, S.K. Parker, "Achieving Effective Remote Working During the COVID-19 Pandemic: A Work Design Perspective", "Applied Psyhology", Vol. 70 (1), pp. 16–59, 2021

[5] L. Djurkin-Kőnig, A. Ostojić, A. Delić, "Korporativna sigurnost u okruženju pandemije COVID-19", Poslovno učilište integralna sigurnost i razvoj, Zagreb, Croatia, 2020

[6] D. Lucic, P. Misevic, "An Impact of Implementation of 5 G Technology on Information Security", 44th International Convention on Information, Communication and Electronic Technology – MIPRO, CTI, Opatija, Croatia, 2021

[7] „Protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)", REGULATION (EU) 2016/679, 27 April 2016
https://eur-lex.europa.eu/legal-content/HR/TXT/?uri=CELEX:32016R0679

[8] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, W. Zhao, "The Covid-19 Threat Landscape", "Computer Fraud & Security", vol. 2021, Issue 9, pp 10 – 15, 2021

[9] B. Pranggono, A. Arabo, "COVID-19 Pandemic Cybersecurity Issues", "Internet Technology Letters", Volume 4, Issue 2, 2021
https://doi.org/10.1002/itl2.247

[10] Hrvatska gospodarska komora (HGK)
https://www.hgk.hr/

[11] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act)
https://eur-lex.europa.eu/legal-content/HR/TXT/PDF/?uri=CELEX:32019R0881&from=EN

[12] H. S. Lallie, L. Shepherd, J. R. C. Nurse and A. Erola, "Cyber Security in the Age of COVID-19: A Timeline and Analysis of Cyber-crime and Cyber-attacks during the Pandemic", "Computer & Security", vol. 105, 2021

[13] J. Zhou, Z. Cao, X. Dong, A.V.Vasilakos, "Working from Home: Cybersecurity in the Age of COVID-19", "IEEE Communications Magazine", vol. 55, no. 1, pp. 26 – 33, 2017

[14] https://cdn.netzpolitik.org/wp-upload/2019/08/2019-07-02-FinalRiskassessmentguidelinesandtemplateforreporting.pdf

Do Croatian Citizens Want the Digitalization of Elections? Differences in Socio-Demographic Characteristics

V. Pavlović Vinogradac, B. Cerin, D. Vanđelić, and A. Prišćan
ZRIN Institute
zrin.udruga@gmail.com

***Abstract* - In modern society characterized by a fast-paced lifestyle, digitalization at all levels is certainly a tool that makes this fast-paced life easier. Elections, regardless of whether they are at the local or national level, are one of the key events in every society. The aim of this paper is to investigate whether the citizens of the Republic of Croatia want the digitalization of elections, and to see whether the attitudes of citizens differ according to sociodemographic characteristics. The research was conducted by ZRIN Institute on a representative sample of citizens of the Republic of Croatia (n=650) using the survey method. The results showed that most citizens want the digitalization of elections (69.5%). The results of multivariate statistical analyzes showed that there are no statistically significant differences in the desire of digitalization regarding sex and place of residence. On the other hand, there are statistically significant differences regarding age group and level of education. Apart from the scientific contribution, the results of this research have strong practical implications that will be discussed in the paper.**

Keywords - digitalization of elections, Republic of Croatia, e-voting, socio-demographic characteristics

I. INTRODUCTION

In modern society, digitalization has become a usual part of everyday life. We can do all activities online - from shopping to ordering food and paying bills. The digital system greatly facilitates the accelerated life of people. Digitalization of elections is one of the options that states can implement. Since the turnout for elections in Croatia is very low, about 46.9% (parliamentary elections) [1], it seems important to study this topic.

Prosser and Krimmer talk about four dimensions of the country's readiness to implement e-voting: (1) politics, (2) law, (3) technology, (4) society. The fourth dimension - society - refers to the degree of political participation and citizens' attitudes towards the implementation of e-voting, as well as their use of technology in general [2].

The aim of this paper is to investigate whether the citizens of the Republic of Croatia want the digitalization of elections, and to see whether the attitudes of citizens differ according to sociodemographic characteristics.

Croats are very active on social networks, about 70% of population actively use them every day, which is a high proportion compared to, for example, Germany, where this percentage is 55%, or France, where this percentage is 45% [3]. Likewise, the 2021 census, where 1,157,000 citizens were successfully registered [4], is also proof of the e-activity of Croatian citizens. Such data give the wind behind the idea of implementation of e-voting in the Republic of Croatia, considering the discrepancy between the turnout for elections using the pen-and-paper method at the polling station and the activities of Croats in the virtual world, especially when we talk about the participation of young people. The topic of e-voting in Croatia is also important because of the large number of Croats who live outside Croatia around the world and have the right to vote in elections [5].

Research has shown that the implementation of e-voting can significantly increase the participation of those who have not voted so far [6]. A survey conducted in the Netherlands in 2018 showed that 53.03% of people want to vote via computer, 1.19% via e-mail, and 18.66% want the implementation of voting via the Internet in general [7]. Research conducted on Polish citizens showed that Poles generally support the implementation of e-voting and have a desire to vote online, although the topic of e-voting is underrepresented in the media space [8].

II. HYPOTHESES AND METHODOLOGY

A. Hypotheses

The aim of this paper is to investigate whether the citizens of the Republic of Croatia want the digitalization of elections, and to see whether the attitudes of citizens differ according to sociodemographic characteristics. Considering the aim of the research, the following hypotheses were constructed:

H1: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding *sex*
H2: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding *place of residence*
H3: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding *age*
H4: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding *level of education*

B. Methods and procedure

The quantitative method of a telephone survey was used in the research. The data collection instrument was a questionnaire. The questionnaire consisted of the question "To what extent do you want electronic voting to be implemented in the Republic of Croatia?" (1-5 degree scale), and questions related to the socio-demographic characteristics of the participants - age, sex, place of residence, and level of education. The research was conducted in September and October 2022.

C. Sample

The sample in this research was random and representative for the general population of the Republic of Croatia. The company *Significo* conducted the survey according to the instructions of ZRIN Institute. The Sample size was n=650. The sample was representative according to **sex** (45.2% male; 54.8% female), **age** (18-24 years 6.5%, 25-34 years 14%, 35-44 years 21.4%, 45-54 years 17.8%, 55-64 years 12.8%, 65+ years 27.5%), **place of residence** (village 7.4%, town (up to 20k residents) 6.3%, a medium-sized city (20-50k residents), a bigger city (more than 50k residents) 48.9%, City of Zagreb 30.2%), **level of education** (elementary school 3.4%, high school 46.8%, bachelor degree 20.2%, master degree 27.7%, PhD 2%). Table 1 shows a comparison of the structure by age in the sample and the population.

Table 1 – *Age population and sample structure*

	% (sample)	% (population)
18-24	7,7	10,34
25-34	15,3	11,45
35-44	24,8	13,53
45-54	18,3	13,46
55-64	12,3	14,65
65+	21,6	22,34

III. RESULTS

In general, all groups want digitalization of elections (Mean>2.5). In percentage terms, 69.5% of the total number of respondents agree and fully agree with the implementation.

Descriptives

Tables 2-5 show the frequencies of responses to the item "To what extent do you want electronic voting in elections to be implemented in Croatia?" on a scale from 1 to 5, where 1 represents "I don't want at all" and 5 represents "I totally want to". According to the data presented, the highest percentages are at level 5, from 31.8% (age 65+) to 65.5% (age 35-44) (except in the group 'elementary school', Table 4). It can also be noted that for part of the groups, the number of responses at level "2" is 0%, that is, if there is no desire to implement electronic voting, then it is at level "1 - I don't want at all". The percentage of people who marked level 1 is from 0% (people with PhD) to 11.2% (age 65+).

Table 2 - Frequencies regard to sex - *To what extent do you want electronic voting in elections to be implemented in Croatia?* (%)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Male	7.8	2.0	17.3	20.7	52.0
Female	4.5	2.5	26.1	22.5	44.4

Legend: (1) - I don't want at all; (2) - Mostly, I don't want; (3) I neither want nor want; (4) - Mostly, I want; (5) - I totally want to

Table 3 - Frequencies regard to regard to place of residence - *To what extent do you want electronic voting in elections to be implemented in Croatia?* (%)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Village	10.4	0.0	22.9	22.9	43.8
Town (up to 20k residents)	4.9	0.0	22.0	17.1	56.1
A medium-sized city (20-50k residents)	6.4	0.0	23.4	23.4	46.8
A bigger city (more than 50k residents)	6.9	3.8	22.0	23.0	44.3
City of Zagreb	3.6	1.5	21.9	19.9	53.1

Legend: (1) - I don't want at all; (2) - Mostly, I don't want; (3) I neither want nor want; (4) - Mostly, I want; (5) - I totally want to

Table 4 - Frequencies regard to regard to age - *To what extent do you want electronic voting in elections to be implemented in Croatia?* (%)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
18 - 24 years	4.8	0.0	31.0	23.8	40.5
25 - 34 years	4.4	0.0	26.4	24.2	45.1
35 - 44 years	4.3	0.7	11.5	18.0	65.5
45 - 54 years	2.6	2.6	19.8	20.7	54.3
55 - 64 years	4.8	1.2	20.5	22.9	50.6
65+ years	11.2	5.6	28.5	22.9	31.8

Legend: (1) - I don't want at all; (2) - Mostly, I don't want; (3) I neither want nor want; (4) - Mostly, I want; (5) - I totally want to

Table 5 - Frequencies regard to regard to level of education - *To what extent do you want electronic voting in elections to be implemented in Croatia?* (%)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Elementary school	13.6	0.0	45.5	22.7	18.2
High school	4.9	3.3	26.0	23.0	42.8
Bachelor's degree	8.4	1.5	14.5	25.2	50.4
Master's degree	5.6	1.7	18.9	16.7	57.2
PhD	0.0	0.0	15.4	23.1	61.5

Legend: (1) - I don't want at all; (2) - Mostly, I don't want; (3) I neither want nor want; (4) - Mostly, I want; (5) - I totally want to

The statistical method of analysis of variance (ANOVA) was used to test the hypotheses.
In the variance analysis, the dependent variable was the particle "To what extent do you want electronic voting in

elections to be implemented in Croatia?" which was treated as a pseudo-interval scale in the test. The independent variables were sex, age group, place of residence, and level of education.

H1: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding sex

Table 6 shows average values (mean) for the desire for digitalization of elections in Republic of Croatia regarding to sex.

Table 6 – *Differences regarding to sex (descriptives)*

	Mean	Std. Deviation
Male	4.07	1.22
Female	4.00	1.10

The ANOVA results showed that there is no statistically significant difference between men and women in the desire for digitalization of elections, $F(1, 648) = 0.665$, $p = 0.415$, $1\text{-}\beta=0.144$. H1 hypothesis is rejected.

H2: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding place of residence

Table 7 shows average values (mean) for the desire for digitalization of elections in Republic of Croatia regarding to place of residence.

Table 7 – *Differences regarding to place of residence (descriptives)*

	Mean	Std. Deviation
Village	3.90	1.28
Town (up to 20k residents)	4.20	1.10
A medium-sized city (20-50k residents)	4.04	1.14
A bigger city (more than 50k residents)	3.94	1.2
City of Zagreb	4.17	1.05

The ANOVA results showed that there is no statistically significant difference between the desire to implement e-voting with regard to the size of the place of residence, $F(4, 645) = 1.616$, $p = 0.169$, $1\text{-}\beta=0.62$. H2 hypothesis is rejected.

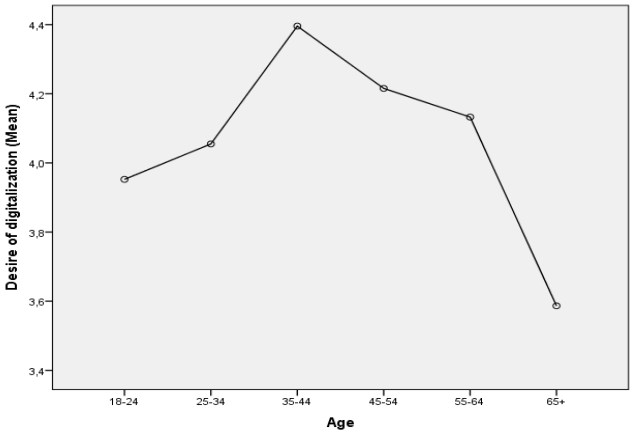
H3: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding age

Table 8 shows average values (mean) for the desire for digitalization of elections in Republic of Croatia regarding to age.

Table 8 – *Differences regarding to age (descriptives)*

	Mean	Std. Deviation
18 - 24 years	3.95	1.08
25 - 34 years	4.05	1.06
35 - 44 years	4.40	1.02
45 - 54 years	4.22	1.02
55 - 64 years	4.13	1.09
65+ years	3.59	1.23

ANOVA results showed that there is a statistically significant difference between the desire to implementation of e-voting regarding to age, $F (5, 644) = 9.416$, $p < 0.05$, $1\text{-}\beta=0.99$. H3 hypothesis is accepted. In order to see between which groups there is a statistically significant difference, the Games-Howell post hoc test was used. The test showed that there is a statistically significant difference in the desire of implementing e-voting between people in the 65+ age group and other age groups except for the 18-24 age group. H3 hypothesis is accepted. Differences by age groups are visible in Graph 1.



Graph 1 – *Desire for digitalization regarding to age*

H4: There is a statistically significant difference in the desire for digitalization of elections in Republic of Croatia regarding *level of education*

Table 9 shows average values (mean) for the desire for digitalization of elections in Republic of Croatia regarding to level of education.

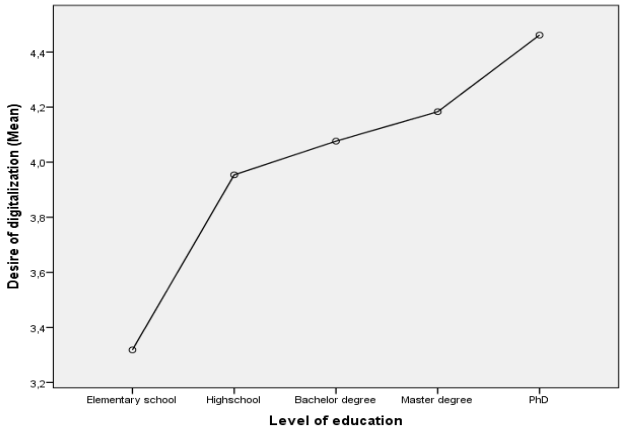
Table 9 – *Differences regarding to level of education (descriptives)*

	Mean	Std. Deviation
Elementary school	3.32	1.21
High school	3.95	1.13
Bachelor's degree	4.08	1.21
Master's degree	4.18	1.14
PhD	4.46	0,78

ANOVA results showed that there is a statistically significant difference between the desire of implementation of e-voting with regard to age, $F(4, 645) = 3.783$, $p = 0.05$, $1\text{-}\beta=0.96$. H4 hypothesis is accepted. The Games-Howell post-hoc test revealed a statistically

significant difference between people with primary school education and people with a master's degree or PhD.

Differences according to level of education can be seen in Graph 2.



Graph 2 – *Desire for digitalization regarding to level of education*

IV. DISCUSSION AND CONCLUSION

Two out of four hypotheses were confirmed in this paper. Differences were not found regarding the desire for digitalization of elections by sex, which means that both men and women equally want the implementation of e-voting. This finding is consistent with previous findings [9]. Also, no differences were found regarding the size of the place of residence. This finding can be interpreted as the fact that the Internet and the virtual world are mostly available in all parts of Croatia.

Statistically significant differences were found regarding the level of education and the age of the respondents. The finding regarding age is expected since younger people are more inclined to use modern technology than older people are. The findings are also consistent with previous findings [9]. One research showed that e-voting will be chosen to the greatest extent by people aged 18-24 [7]. A recent survey (2022) in Lithuania conducted on young people aged 19 to 29 showed that 73.1% of them choose the method of electronic voting via the Internet over other methods [10]. Research conducted in Greece on the student population showed that students are informed about what e-voting is and have positive attitudes about e-voting [11].

When we talk about the level of education, we can say that the findings are also expected since people with a higher level of education are generally ready to participate more [12] and are more open to changes than people with a lower level of education [13].

Limitations

This research comes with certain limitations. First, although the sample is representative of the general population of the Republic of Croatia in terms of socio-demographic characteristics, this data should be taken with a grain of salt since people who are more willing to participate politically are more willing to participate in survey research than others are. In addition, it should be emphasized that these results are an indicator of the public opinion of citizens at the time when the research was conducted (September-October 2022), and citizens' attitudes can change depending on the context and circumstances. For example, in the event of a major global hacker attack, citizens' trust in digitalization could significantly decrease, and thus the desire to implement e-voting. In addition to the mentioned limitations, further research of citizens' attitudes about the digitalization of elections is needed, with the included factors of socio-demographic variables, as well as other factors.

REFERENCES

[1] Gong, 2023. https://gong.hr/2020/07/20/hdz-je-osvojio-66-mandata-sa-samo-167-ukupnog-broj/ , accessed February 6th 2023

[2] Prosser, Alexander, and Robert Krimmer. "The dimensions of electronic voting–Technology, law, politics and society." In *Electronic voting in Europe-Technology, law, politics and society, workshop of the ESF TED programme together with GI and OCG*. Gesellschaft für Informatik eV, 2004.

[3] Eurostat. 2023. https://ec.europa.eu/eurostat , accessed February 6th 2023

[4] Croatian Bureau of Statistics 2023. https://dzs.gov.hr/en , accessed February 6th 2023

[5] Gjidara, Marc. "Općenito o pravu glasanja io biračkom pravu iseljeništva." *Zbornik radova Pravnog fakulteta u Splitu* 52, no. 2 2015. pp. 279-309.

[6] Houston, Andrea L., Yurong Yao, Chitu Okoli, and Edward Watson. "Will remote electronic voting systems increase participation?." *Electronic Government, an International Journal* 2, no. 3 2005.pp 353-368.

[7] Loeber, Leontine. "The E-voting Readiness Index and the Netherlands." Electronic Voting: *Third International Joint Conference*, E-Vote-ID 2018, Bregenz, Austria, October 2-5, 2018, Proceedings 3. Springer International Publishing, 2018.

[8] Musiał-Karg, Magdalena, and Izabela Kapsa. "Polish Mass Media Coverage and Public Opinion on E-democracy. The Case of Electronic Voting." *Media Studies* 12.23 2021.pp 3-18.

[9] Yao, Yurong, Chitu Okoli, Andrea Houston, and Edward Watson. "Demographic differences in attitudes toward remote electronic voting systems." *Journal of Computer Information Systems* 47, no. 2 2006.pp 34-45.

[10] Malužinas, Martinas, and Maciej Górny. "Electronic Voting in Opinions and Declarations of Young Lithuanians Aged 19–29 Years. Report on Preliminary Research." 2022.

[11] Salavasidis, Petros. "Exploring the attitudes of students in Tertiary and Upper Secondary Education towards Electronic Voting and employment of an electronic voting." 2012.

[12] Persson, Mikael. "Education and political participation." *British Journal of Political Science* 45, no. 3 2015. 689-703.

[13] Aslan, Mahire, Kadir Beycioglu, and Necdet Konan. "Principals' Openness to Change in Malatya, Turkey." *International Electronic Journal for Leadership in Learning* 12, no. 8 2008. n8.

PAPA for IoT - Role of Data Accuracy in IoT Deployment and Data-centric Decision-Making

Mikko Vermanen*, Juhani Naskali*, Ville Harkke†, Jani Koskinen*
* University of Turku/Information Systems Science, Turku, Finland
† Åbo Akademi/Information Systems, Turku, Finlad

***Abstract*—Data has become a central asset supporting decision-making in modern business environments. Internet of Things (IoT) solutions in particular are designed to gather information from a large variety of monitorable entities, providing companies with seemingly unlimited abilities to measure and analyze factors such as performance and productivity.**

Considering the reliability and trustworthiness in data-driven decision-making, the accuracy of the leveraged information is critical. While the development of IoT solutions makes them continuously more accurate, caution must be applied not only in terms of potential inaccuracy caused by technical failures but also by human error. Thus, some of the key questions to consider when investigating the accuracy and trustworthiness of data are: is the gathered information itself correct, is the leveraged data contextually correct, and what are the potential issues caused by human error.

From an ethical perspective, the justifiability of leveraging potentially incorrect data needs to be addressed. When decisions related to individual employees are being made based on data, it is necessary to not only consider the possible inaccuracy of the data but also to prevent manipulating or distorting the results by applying too limited or falsely selected data, creating a biased basis for analysis and decisions.

Keywords—Internet of Things, PAPA-model, Accuracy, Ethics

I. INTRODUCTION

The data economy, Digital society, Smart cities etc. are terms that are used to describe the change that society is experiencing by collecting and using information in large scale. Businesses and the whole society are more and more dependent on data and this phenomenon shows no sign of slowing down in the future. IoT has provided ways to collect information from different sources, offering possibilities to connect almost everything together, and thus turning things to data sources. This dependence of data, however, sets demands for the data and especially for its quality. As data is a central asset nowadays, the demand for data accuracy is crucial, if expectations for its use are to be met. Without data accuracy, we lose the fundamental benefit that technology offers us—automation of tasks—and are forced back to manual checks and routine data work conducted by humans; the very thing we were aiming to surpass with digital technologies.

While the correctness of the actual data is a central factor determining its quality [1], the impact of human factors appears to be a significant but easily overlooked issue in organisations [2]. Even if the concrete data gathered from monitorable objects are correct, the responsiblity for avoiding risks related to selecting the right type of data, conducting correct analysis, making justified conclusions etc. belongs to the human decision-makers and should be systematically governed [3]. This calls for not only understanding on how to correctly apply data in decision-making, but also how to retain ethicality in the related practices in business [4].

In addition to potentially causing harm to the company itself by drawing false conclusions based on inaccurate data, the individuals from whom the data is collected are at risk, and there is a demand for protecting users by ethical design that aims to empower users who are interacting with IoT [5]. For example, when IoT solutions are utilised for performance monitoring purposes and decisions are being made based on the gathered data, the accuracy of the information is one—but not the only—of the key factors to be taken into account. For this purpose, the PAPA (Privacy, Accuracy, Property, Accessibility) framework, developed for addressing the potential issues of information age [6], is leveraged. While the principle idea of PAPA framework still applies in modern technical environments, modifications and additions are required to thoroughly address all aspects of IoT data. This paper focuses on the effects of IoT on data accuracy. A more detailed examination of the other PAPA categorizations (Privacy, Property and Accessability) is outside the scope of this paper, and are examined in depth in other publications [7], [8].

In chapter II, the concept of IoT and its characteristics as a tool for producing data are explained. In chapter III, we introduce the ethical PAPA framework and its notions regarding data accuracy. In chapter IV we address four central factors defining the reliability of the conclusions driven by data. Firstly, the correctness of data and the factors defining it are introduced. Secondly, the role of selecting the right data is explained. Thirdly, we share our insight on the role of data analysis. Finally, the accuracy of governance is addressed. The chapter V provides views on how the research on IoT data accuracy should be widened to provide more thorough understanding on not only fundamental level, but also in form of concrete requirements needed to ensure accuracy. In chapter VI, we introduce our conclusions.

II. INTERNET OF THINGS

The Internet of Things (IoT) is a constantly evolving technology enabling its users to collect data from a large variety of monitorable objects. [9] The fundamental idea of IoT is built around three involved actors/aspects: the monitored entity as a data source, internet as a channel for transferring the data, and a human actor to whom the collected information is delivered in a readable form [10]. IoT solutions are widely used and found to provide concrete benefits in different professional environments, such as manufacturing, medical care, agriculture, retail, logistics, and transportation. [11], [12]

One of the beneficial measurement factors in business environments is performance, which can be monitored based on a variety of data types collected with IoT devices in differnt contexts — allthought adopting the possibilities is still challenging [13], [14]. The collected information is often connected to a certain individual operating or interacting with the object of data collection, such as machinery or vehicles, and thus, it is simultaneously possible to investigate personal productivity. Personal monitoring, including the collection of data and its use as a tool supporting decision-making is a central topic in this paper. Thus, we aim to address the potential risk factors related to this phenomenon and how companies can avoid negative outcomes.

The relatively low prices of IoT solutions compared to the amount and quality of data they are capable of producing has a large role when investigating the reasons behind its expansion. Almost unlimited amounts of information can be gathered from different data sources, enabling companies to observe and monitor the points of interest on a very detailed level. However, while the capabilities are vast and the technology itself is under constant development, accuracy of the gathered information cannot be taken for granted. Two main causes for potential issues are technical shortcomings, and perhaps even more importantly, human factors, which are described in more detail in chapter IV.

III. PAPA MODEL

The basis for investigating the role and impacts of data accuracy in decision-making, especially from the ethical standpoint, is built around the PAPA framework introduced by Richard O. Mason [6]. The original purpose of PAPA is to describe the ethical issues of the information age—it spans privacy, accuracy, property, and accessibility. In this chapter, each issue category and its relation to accuracy are described.

A. Privacy

Privacy is described as a right to seclude and protect information connected to an individual. It is suggested that individuals should be able to decide if, when, how, and to whom information about themselves can be revealed [6], [15]. Furthermore, central questions related to privacy are: what information should be revealed, under what conditions and with what safeguards? As the value of data grows alongside the developing potential for leveraging it for productivity and/or performance purposes, it is crucial that the companies aim to retain the privacy of employees when considering what information gets collected. Related to data accuracy and the possible lack thereof, the importance of limiting the visibility of individual data is emphasized further to avoid both privacy infractions and the distribution of false information.

Regulatory and legal measures have been taken to protect individual privacy, with the General Data Protection Regulation Act (GDPR) as a prominent example [16]. While regulations manage to provide some protection for individuals, a grey area does exist, where questionable approaches towards privacy can be found ethically questionable, yet legally allowed. [4] The connection between privacy and accuracy includes topics such as the justifiability of collecting and using potentially inaccurate and individually identifiable performance data, possibly leading to false conclusions and negative impacts towards the involved individual.

B. Accuracy

Mason associated the concept of accuracy with the responsibility of authenticity, fidelity, and accountability of information, as well as the accountability for errors and harmful events. Here, information accuracy is bound together with system accuracy, emphasizing that developers have a significant role in ensuring avoidance of errors. While this statement can be still be applied in modern business environments, it must be taken into account that the IoT ecosystems can be highly complex, which makes the distribution of responsibility more difficult, as we need to consider both the technological standpoint and the entire sosio-technical system. In this paper, we focus largely on the management's responsibility, but simultaneously underline that the whole network of actors must be taken into account.

Furthermore, a central topic of this paper is related to the potential harm related to the position of an employee caused by data inaccuracy. In this context, Mason used the terms 'misuse' and 'misinterpretation', which are bound to each of the accuracy-affecting areas introduced in the following chapter: accuracy of data, selection of data, accuracy of analysis, and the accuract of governance. Additionally, it covers situations where data is used falsely either intentionally of unintentionally, resulting in personal discrimination. According to Mason, employees should be aware and able to control data collection and distribution and to be protected from discrimination resulting from leveraging erroneous or unethically collected data—emphasizing that discrimination can appear in several different forms. Again, while regulations provide employees with a certain degree of protection, we claim that creating ethical guidelines is a necessary addition to thoroughly protect the individuals. The issue of combining data that is professionally irrelevant also needs to be brought up, including situations where the border between professional

and personal is being crossed [17], as this enables companies to conduct ethically questionable profiling [18].

In addition to traditional data accuracy, some thought needs to be given to system accuracy as well. Besides the data collection itself, organisational structures and information processes affect data and its accuracy. The surrounding business ecosystem, deployment planning and implementation, parameter selection, rule definitions and other implicit decisions made during the whole life-cycle of data collection and analysis all have major implications on what results data-driven decision-making will have.

C. Property

The concept of property involves matters related to the ownership of the collected data. The investigation of data property revolves around questions such as who owns the information, what are the just and fair prices for is exchange, who owns the channels through which the information is transmitted and how the access should be allocated.

It is under debate whether the rights to and ownership of the collected information attached to a specific individual should remain fully or partially as the observed individual's property. Again, if inaccurate data is collected from an employee and owned by the company - or an external party - issues affecting an individual's personal position may emerge. This binds together multiple issue categories introduced in PAPA: if the data is owned by the company (see property), is the data owner allowed to reveal (see privacy and accessibility) it without permission?

D. Accessiblity

Accessibility is a category partly intertwined with privacy, including further considerations related to who can access the data. Briefly explained, privacy focuses on data collection, while accessibility concentrates on data distribution.

Obvious issues can be found in situations where inaccurate information related to an individual is collected and distributed among any group of individuals. Without careful consideration regarding with whom the data should be shared with, the potential of negative information about the individual being leveraged to weaken their position raises. As the IoT ecosystems often involve not only the company's internal personnel, but also external actors such as service providers, partners and/or customers, it is apparent that significant issues can arise if corrupted information gets shared.

IV. FACTORS DEFINING IOT DATA ACCURACY

The original PAPA model has been further developed, as need arises, to envelop new technologies and environments of the digital age. For example, topics such as behavioral surveillance, interpretation, and governance have been identified as new ethical issues specifically surrounding big data [19]. Other emergent technologies, such as AI, also require their own additions [20]. In this chapter we examine the different dimensions of IoT data as it relates to data accuracy.

A. Accuracy of Data Collection

The accuracy of the actual data is a crucial factor, considering the reliability and justification of data-driven decision-making. For example, in cases where employees' performance is being measured and actions taken in terms of career development, recruitment or layoffs, leveraging inaccurate data will have a negative impact on the position of both the employees and the company.

From the company's perspective, inaccurate data can be used both intentionally and unintentionally. Intentional use of inaccurate data results in ethical issues, which can do significant harm towards the company's image and internal relations. On the other hand, unintentional use of inaccurate data will likely reduce or remove the aimed benefits of data-driven decision-making.

Considering the factors affecting the accuracy of the collected data, the quality of the utilised IoT solutions is important. The pricing of the solutions varies, and so does their quality. While multipurpose IoT devices do offer companies a cost-efficient opportunity to leverage a variety of beneficial features, compromises in quality tend to be inevitable. [21] In addition to a need for cost-efficiency, different types of optimization requirements, such as a need for smaller devices or low energy consumption may negatively affect the quality of the data. Additionally, reusing the same components in multiple different models of sensors will complicate cross-validating results, creating the possibility of systematic errors even when validating/calibrating different models of sensors against each other. The data quality issues can manifest themselves as problems in actual measurement accuracy, timeliness, completeness, relevance (utility), volume or concordance between data sources. [22]

B. Selection of Data

When selecting the type, source and collection time of data, companies have a responsibility to follow ethically just practices. Correct data collection decisions simultaneously enhance the potential to gain useful and reliable information about the performance of employees and reduce the possibility of harming individuals' through an incorrect collection of information and/or distortions to analysis.

Some of the key questions to consider in this context are: are the variables selected in a manner conducive to reliable and rational analysis? Do the variables actually measure the real-world phenomena tracked? Is the granularity of measurements correct? Can we rely on the accuracy of each selected and measured attribute? If measurements are utilized in tandem, are they comparable and compatible? The importance of accuracy is emphasized when multiple types of data are utilized in tandem, as one distorted attribute can eliminate the reliability of the entire analysis.

Wrong data selection practices may cause distortions in later analysis, even if the analysis itself is conducted perfectly. Simple mistakes, such as tracking daily averages instead of granular per-second values, may cause a complete loss of meaningful information, as the important bits are lost to miscalibration. Wrongfully selecting which attributes to track may create a false sense of accuracy, as everything from data gathering to data management and analysis is functional, but the measurements taken do not accurately represent the facts decision-makers think they represent.

C. Accuracy of Analysis

Even when the data has been gathered and selected based on correct practices, falsely conducted analysis may invalidate the results. Again, the issues can be a result of either intentional or unintentional practices. While good practices for data analysis have been covered in literature, it cannot be taken for granted that companies have the knowledge and expertise to conduct a practical and ethically correct analysis.

A thoroughly considered selection of data can reduce the potential for false analysis results caused by complexity, however, it may also dismiss relevant parameters. For instance, investigating solely the usage level of industrial machinery in a specific timeframe to observe individual performance levels while disregarding the lack of material available for processing is likely to result in false conclusions regarding the employee's performance. Furthermore, it is possible that maximizing machinery usage time is not the right goal to begin with, and deciding to optimize other metrics, such as the amount of produced end-product, product quality or amount of coordination between different employees, would result in bigger profits. This emphasizes that conducting reliable analysis often requires not only IoT data but also supportive information and effective control processes. It is also worth mentioning that leveraging artificial intelligence (AI) for data analysis is a growingly common approach. While this may reduce the number of potential issues caused by human error, AI cannot be trusted blindly, and thus cannot be considered a tool for eliminating issues caused by human behaviour.

D. Accuracy of Governance

As noted before, data does not reside in a vacuum, and many human factors affect the accuracy of data-driven decision-making. From lower-level data governance to higher-level organisational (or inter-organisational), governance processes can affect data accuracy in a multitude of ways [23], [24].

Data governance includes many information processes that ensure that data is managed in a way that ensures its high quality. Whether loss of data due to physical corruption or communication disconnections, alteration of data due to accidental edits or migration mishaps, or intentional deletion or manipulation by bad actors, bad data governance enables multiple avenues for data inaccuracies. The existing data quality management frameworks such as Wangs Total Data Quality Management [25] are suitable for governing IOT data but specific IOT data quality management frameworks are still in their infancy. [26]

Furthermore, organisation-level governance highly affects the outputs of data-centric decision-making [27], [28]. Whether the underlying data is an accurate representation of reality or not, the decision-making process can lead to unreliable or undesirable outcomes, if there are problems in the governance process. Therefore, governance should note the ethical issues of data use as data use may be affecting the individuals from where data is collected. Accuracy at all levels is an issue that should be governed to achieve the possibility to make the right decisions based on the right data instead of a false one.

V. DISCUSSION

This paper attempts to provide the reader with a fundamental understanding of the potential pitfalls related to data accuracy in IoT use, in terms of correctness and ethicality. To thoroughly understand the mechanics behind these issues and to find concrete actions needed to ensure accuracy, technical-level research is needed. That said, our principal intention is to investigate the risk factors involving human actors and to address the issues on which they have an impact on, namely decision-making. As the sensor data itself is produced by technical devices without the capability of logical reasoning, the accountability of information cannot be fully guaranteed by the companies implementing these solutions. The importance of this matter is especially high as the IoT is a rapidly developing and growing area of technology [4]. As the trustworthiness of the gathered data increases thanks to the development of the devices and solutions, we get closer to a situation where the mechanical pitfalls diminish, yet the impact of human factors likely retains its relevancy.

Another topic for further inspection is how to implement the suggested approaches in the companies applying IoT solutions. Especially considering the smaller companies, technical expertise cannot be taken for granted, and neither their ability to put the suggested approaches into practice. This leads to the question of how the responsibility of retaining accuracy and ethicality should be realistically distributed. While the IoT solution providers can direct this development to some degree, the human actor will always have a critical impact and the ability to harm individuals. That said, we see that providing all-encompassing guidelines is not currently a realistic goal, and thus we should aim to gradually decrease the realization of the introduced risks.

VI. CONCLUSIONS

While IoT solutions provide companies with significant benefits due to their seemingly unlimited data collection capacity with often moderable expenses, the reliability of the data cannot be taken for granted, nor the drawn conclusions. The lifecycle of information from unprocessed sen-

sor data to a trustworthy tool supporting decision-making involves a large variety of pitfalls. This paper contributes to the research gap by addressing the issues relevant specifically to IoT environments. A holistic overview of potential pitfalls regarding IoT data accuracy is introduced, as well as ideas on how to mitigate these issues.

While the accuracy of the actual data is the most obvious point of interest when considering the reliability of the gathered information, human factors have a large role as well. The two key areas in which human error can potentially damage the desired outcomes are the selection of data and the analysis phase. Firstly, the types of gathered data must be carefully selected to produce reliable conclusions, as either ignoring relevant parameters or including irrelevant ones can distort the results of the analysis phase. Thus, careful planning is critical before the solution implementation: what are the attributes necessary to cover to draw trustworthy conclusions? Can the accuracy of each selected data type be relied on considering the capability of the used solution?

It turns out that regulatory protection alone cannot safeguard employees to a sufficient degree, as opportunities for intentional and unintentional misuse are hidden throughout the decision-making lifecycle. This supports the idea that an ethical framework addressing the issues related to IoT use is needed, offering a fundamental understanding of the requirements of ethically sustainable practices covering the phases of design, implementation and actual use. This paper addressed the topic of accuracy, being a part of a larger study in which the applicability of the original PAPA categories is investigated in the IoT context. As a result, a separate publication will be released, in which the modernized and IoT-specific revision of the framework is introduced, including a modified representation of the four PAPA categories, accompanied by complementary issue categories enabling full coverage of the IoT ecosystem.

REFERENCES

[1] N. Côrte-Real, P. Ruivo, and T. Oliveira, "Leveraging internet of things and big data analytics initiatives in european and american firms: Is data quality a way to extract business value?" *Information & Management*, vol. 57, no. 1, p. 103141, 2020.

[2] P. Brous, M. Janssen, and P. Herder, "The dual effects of the internet of things (iot): A systematic review of the benefits and risks of iot adoption by organizations," *International Journal of Information Management*, vol. 51, p. 101952, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0268401218309022

[3] S.-I. Chang, L.-M. Chang, and J.-C. Liao, "Risk factors of enterprise internal control under the internet of things governance: A qualitative research approach," *Information & Management*, vol. 57, no. 6, p. 103335, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037872062030272X

[4] M. Vermanen, M. M. Rantanen, and V. Harkke, "Ethical framework for iot deployment in smes: individual perspective," *Internet Research*, vol. 32, no. 7, pp. 185–201, 2022.

[5] G. Baldini, M. Botterman, R. Neisse, and M. Tallacchini, "Ethical design in the internet of things," *Science and engineering ethics*, vol. 24, pp. 905–925, 2018.

[6] R. O. Mason, "Four ethical issues of the information age," *MIS quarterly*, pp. 5–12, 1986.

[7] M. Vermanen, M. M. Rantanen, and J. Koskinen, "Privacy in internet of things ecosystems–prerequisite for the ethical data collection and use by companies," *Human Choice and Digital by Default: Autonomy vs Digital Determination: 15th IFIP International Conference on Human Choice and Computers, HCC 2022, Tokyo, Japan, September 8–9, 2022, Proceedings*, pp. 18–26, 2022.

[8] M. Vermanen, J. Koskinen, and V. Harkke, "Internet of things (iot) data accessibility: ethical considerations," *Well-Being in the Information Society. Fruits of Respect: 8th International Conference, WIS 2020, Turku, Finland, August 26–27, 2020, Proceedings 8*, pp. 197–208, 2020.

[9] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," *Ad hoc networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[10] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through internet of things. ieee internet things j. 1 (2), 112–121 (2014)," *IEEE Internet of Things Journal April 2014*, 2013.

[11] J. Kaur and K. Kaur, "Internet of things: A review on technologies, architecture, challenges, applications, future trends." *International Journal of Computer Network & Information Security*, vol. 9, no. 4, 2017.

[12] P. Tadejko, "Application of internet of things in logistics–current challenges," *Ekonomia i Zarządzanie*, vol. 7, no. 4, pp. 54–64, 2015.

[13] I. Nappi and G. de Campos Ribeiro, "Internet of things technology applications in the workplace environment: A critical review," *Journal of Corporate Real Estate*, vol. 22, no. 1, pp. 71–90, 2020.

[14] H. Tran-Dang, N. Krommenacker, P. Charpentier, and D.-S. Kim, "The internet of things for logistics: Perspectives, application review, and challenges," *IETE Technical Review*, vol. 39, no. 1, pp. 93–121, 2022.

[15] H. J. Smith, T. Dinev, and H. Xu, "Information privacy research: an interdisciplinary review," *MIS quarterly*, pp. 989–1015, 2011.

[16] European Parliament and Council of the European Union, "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," *Official Journal of the European Union*, 2016.

[17] E. Oriwoh, D. Jazani, G. Epiphaniou, and P. Sant, "Internet of things forensics: Challenges and approaches," in *9th IEEE International Conference on Collaborative computing: networking, Applications and Worksharing*. IEEE, 2013, pp. 608–615.

[18] S. Wachter, "Normative challenges of identification in the internet of things: Privacy, profiling, discrimination, and the gdpr," *Computer law & security review*, vol. 34, no. 3, pp. 436–449, 2018.

[19] J. Young, T. J. Smith, and S. H. Zheng, "Call me big papa: An extension of mason's information ethics framework to big data," *Journal of the Midwest Association for Information Systems (JMWAIS)*, vol. 2020, no. 2, p. 3, 2020.

[20] B. C. Stahl, "From papa to papas and beyond: Dealing with ethics in big data, ai and other emerging technologies," *Communications of the Association for Information Systems*, vol. 49, no. 1, p. 20, 2021.

[21] M. Vermanen and V. Harkke, "Findings from multipurpose iot solution experimentations in finnish smes: common expectations and challenges," *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[22] C. Liu, P. Nitschke, S. P. Williams, and D. Zowghi, "Data quality and the internet of things," *Computing*, vol. 102, no. 2, pp. 573–599, 2020.

[23] R. Abraham, J. Schneider, and J. vom Brocke, "Data governance: A conceptual framework, structured review, and research agenda," *International Journal of Information Management*, vol. 49, pp. 424–438, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0268401219300787

[24] M. Janssen, P. Brous, E. Estevez, L. S. Barbosa, and T. Janowski, "Data governance: Organizing data for trustworthy artificial intelligence," *Government Information Quarterly*, vol. 37, no. 3, p. 101493, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0740624X20302719

[25] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," *Journal of management information systems*, vol. 12, no. 4, pp. 5–33, 1996.

[26] L. Zhang, D. Jeong, and S. Lee, "Data quality management in the internet of things," *Sensors*, vol. 21, no. 17, p. 5834, 2021.

[27] P. Brous, M. Janssen, and R. Vilminko-Heikkinen, "Coordinating decision-making in data management activities: a systematic review of data governance principles," in *Electronic Government: 15th IFIP WG 8.5 International Conference, EGOV 2016, Guimarães,*

Portugal, September 5-8, 2016, Proceedings 15. Springer, 2016, pp. 115–125.

[28] D. G. Broo and J. Schooling, "Towards data-centric decision making for smart infrastructure: Data and its challenges," *IFAC-PapersOnLine*, vol. 53, no. 3, pp. 90–94, 2020.

DATA PROCESSING AND SIGNAL TRANSMISSION

User Input Search - Custom Motion Estimation Algorithm Optimized for UAVs

J. Benjak, D. Hofman
Faculty of Electrical Engineering and Computing / Department of Control and Computer Engineering, Zagreb, Croatia
jakov.benjak@fer.hr, daniel.hofman@fer.hr

***Abstract*—Unmanned aerial vehicle technology is growing rapidly as it finds its application in the various industries, including military & defense, agriculture, logistics, transportation, healthcare, entertainment and many others. One of the fastest growing industries including drones is the entertainment industry. More specifically, First Person View (FPV) piloting has become a popular sport which attracts huge masses. In FPV systems, the pilot wears FPV goggles and controls the drone using a controller, and the drone transmits video data to the goggles in real time. Since drones usually fly very quickly, the video quality in terms of resolution, compression artefacts and end-to-end delay must be maintained. This paper explores the idea of optimizing the Motion Estimation (ME) algorithm of existing High Efficiency Video Coding (HEVC) algorithms by utilizing user input from the controller. For example, if the drone is directed to hover to the left, then it would make sense to search for the most similar blocks in the previous video frame only on the left side of the referent block. For the purpose of this research, a HEVC video coder was customized to receive additional input besides the video data – the user input from the controller, or in other words, drone movement directions. We call this ME algorithm User Input Search (UIS). Our UIS algorithm is compared with the standard ME algorithms and its efficiency is tested.**

Keywords*—*High Efficiency Video Coding, Custom Motion Estimation, Video Coding, Data Compression

I. INTRODUCTION

Unmanned aerial vehicles (UAVs), commonly known as drones, have witnessed a massive growth in their application in various industries [1]. The entertainment industry, in particular, has embraced the use of drones with First Person View (FPV) piloting systems. In FPV systems, the pilot controls the drone using a controller and receives real-time video data through FPV goggles. Such systems are commonly used in drone races and film making. Drone racing is a popular sport in which pilots control UAVs to compete in races through obstacle courses, and the goal is to complete the course as fast as possible, without crashing. However, such systems have their limitations which are yet to be solved. Due to the rapid movement of drones, maintaining high video quality, including resolution, low end-to-end delay and hiding compression artifacts is challenging.

To address this issue, this paper presents a novel approach to optimizing the Motion Estimation (ME) algorithms in existing High Efficiency Video Coding (HEVC) video coders. Our User Input Search (UIS) algorithm involves incorporating user input from the controller to improve the efficiency of the ME process. The main idea is to search for the most similar blocks only in the direction of the drone movement. For example, if the drone hovers to the left, UIS would only search for the most similar blocks on the current block's left side in the previous frame. UIS results in less searches, which means less operations in the encoding process. This optimization should maintain the video quality, but increase the encoding speed, or in other words, reduce latency.

This paper tests this idea only on the drone footage, but the idea can be expanded to other machines as well. For example, land vehicles like remotely controlled trucks, machines for remote medical procedures, remote security cameras, etc. Basically, any kind of remotely controlled device which transmits video. Another yet to be researched topic is using this idea on mobile phone cameras, by using the information from mobile phone's accelerometer in order to estimate device motion.

The UIS algorithm was implemented on an open-source High Efficiency Video Coding (HEVC) encoder and tested on several drone sequences. All sequences were filmed by the author's associates. At the end of this research, a performance comparison of our UIS algorithm and the standard ME algorithms is given.

II. RELATED WORK

Because ME can be the most power and time consuming part of the entire coding process, a lot of researchers tried optimizing it. Authors in [2] developed a low-complexity, adaptive fractional-pixel ME skipped algorithm for the HEVC encoder. Their algorithm reduced ME encoding time by an average of 63.22%, with encoding efficiency maintained. A method for optimizing the number of pixels in the current search window has been proposed in [3]. Their proposed method can save 56% of computations in comparison with the Test Zone Search (TZS) scheme - the default HEVC search scheme, with negligible decrease of coding quality. TZSearch algorithm improvements and other optimized ME algorithms have also been successfully implemented in [4], [5], [6]. Some more recent works implement artificial intelligence into ME algorithms in order to improve them. One example of optimizing an existing HEVC ME algorithm has been proposed in [7].

None of the above mentioned researches considered utilizing user input from any sort of input device. To the best of the authors knowledge, there are no published researches which discuss and implement custom ME algorithms which leverage user input from the controller.

III. USER INPUT SEARCH IMPLEMENTATION

For the purpose of this research, an open source HEVC video encoder Kvazaar [8] was upgraded to include our custom ME algorithm, UIS. Kvazaar source code includes a few different ME algorithms, such as TZS, Diamond Search (DS) and Hexagon Search (HS). UIS was implemented by tweaking the DS algorithm. DS aims to find the best prediction block from the previous frame for each block in the current frame by searching for the closest match. The search pattern forms a diamond shape where the center of the diamond is initially set to the location dependant on the search step, and the four points on the diamond correspond to candidate blocks. The four points are points above and below the center, and to the right and to the left of the center. The algorithm checks the cost of each candidate block in the diamond and moves the center of the diamond to the candidate with the lowest cost. The cost function is usually the Sum of Absolute Differences (SAD) or the Mean Square Error (MSE). SAD simply sums up all the absolute values of pixels differences between the current block and the candidate block. MSE squares each pixel difference and performs the mean operation afterwards. This process continues until no better candidate can be found or a stopping criterion is met (e.g., a maximum number of search steps). The best motion vector found during the search process is stored in the variable and returned as the result of the algorithm.

UIS, on the other hand, works similarly to the DS, but checks only for one candidate in each step, in the direction of the drone movement. For example, if the drone moves upwards, the DS would calculate the cost for all four candidates, in every direction, while the UIS would only calculate the cost for the upward candidate. This would be repeated until the best match is in the center, or the maximum number of search steps is reached. However, the maximum number of search steps was removed while performing tests.

Test were performed on four sequences, filmed using a DJI ZENMUSE X5S camera on a DJI Inspire 2 drone. Camera output was a set of .DNG frames, which were later losslessly compiled into a .yuv video using FFMpeg [9]. In each sequence, the drone moved or rotated straight in one direction, either upwards, left or right. Each sequence was different in terms of spatial complexity, meaning that obstacles - such as houses, bushes, trees, etc., were at different distances and generally the number of such obstacles was different in each sequence. Figure 1 and Figure 2 display sequences which were used for testing. Each sequence was exactly 240 frames long. During encoding, only I and P frames were used, because in real-time encoding, B frames do not make sense. However, some test were made using B frames as well, to show the UIS true advantages which can be used when coding any kind of drone footage which had a controlled movement, and which doesn't require ultra-low latency real-time encoding.

By inspecting the encoded video in the Elecard StreamEye software [10], it was very clearly seen that most of the motion vectors were indeed following the direction of the drone movement, when using the DS algorithm. This was the case with all the tested sequences. Of course, the results were the same when using the UIS algorithm. An example is shown in the Figure 3, and it can be seen that the generated motion vectors are very similar, which is exactly the result we were hoping for.

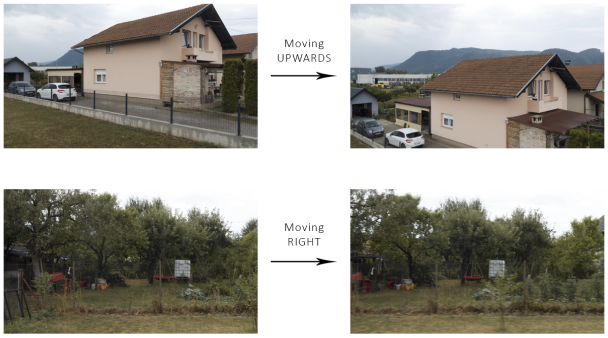


Fig. 1: Drone movement in the 1st and the 2nd sequence



Fig. 2: Drone movement in the 3rd and the 4th sequence

IV. EXPERIMENTAL RESULTS

All the test were performed on a 64-bit Windows personal computer with 16GB of RAM, 11th Gen Intel(R) Core(TM) i5-11400F @ 2.60GHz processor and an NVIDIA GeForce GTX 1650 graphics card. While performing the tests, no other processes were purposely started on the computer. Since some processes can start unnoticed during the encoding process, in order to do the UIS performance evaluation more precisely, we used the Intel VTune Profiler software [11]. To check exactly how

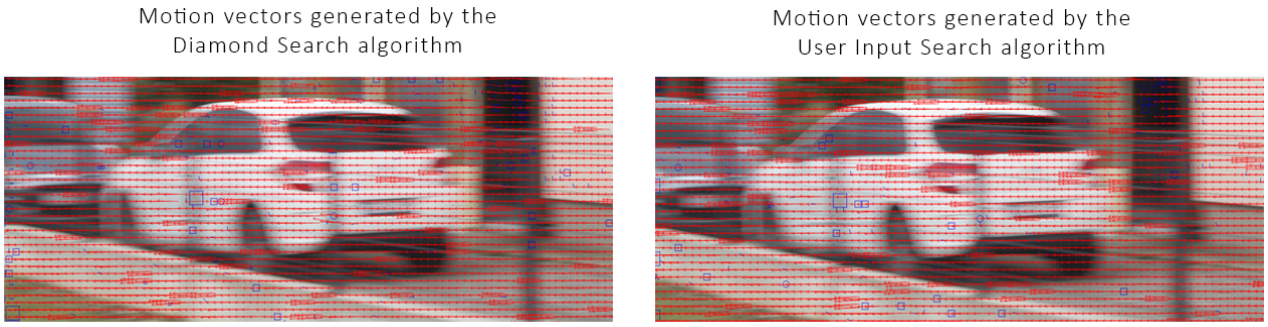


Fig. 3: Motion vectors generated by the DS and the UIS algorithms

much processor time each function had used, we used the VTune's Hotspots analysis and the Flame Graph (Figure 4). Using the Flame Graph tool, we were able to extract only the *diamond_search* and the *user_input_search* functions which perform the DS and the UIS algorithms. Table I represents crucial data in form of a table. Search time refers only to the CPU time that was taken by the *diamond_search* and the *user_input_search* functions. The obtained data shows that the goal was accomplished and that the UIS algorithm indeed improves the overall encoding process. The encoded file sizes were in a very close range, with the biggest difference being in the first sequence (10.2MB for DS, and 10.3MB for UIS). PSNR values were the same in all cases, but the search time was reduced significantly.

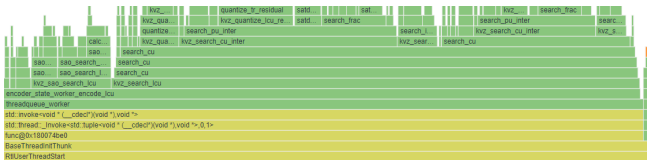


Fig. 4: VTune's Flame Graph from the Hotspots analysis

An noteworthy point is that the search time does not occupy a significant portion of the total encoding process time. The biggest percentage of the total encoding time was in case of the second sequence, where the search function took 1.4% of the total encoding time when using DS, and 0.4% when using UIS. This observation is visualised in Figure 5 and Figure 6, where the search function is highlighted with the pink color. The main reason behind the small percentage is because all sequences were encoded using the Kvazaar's default Group Of Pictures (GOP) structure, which is *lp-g4d3t1*, a low-delay p-frame only GOP. In such structure, each frame references only one frame, the previous frame. Figure 7 and Figure 8 display a Flame Graph of the second sequence, encoded using GOP *8: B-frame pyramid of length 8*. In such structure, each frame references up to four frames, which results in much more total searches, but in better compression as well. When using this structure, UIS shows it's real power, because the search function took 11.9% of the total encoding time when using DS, and only 2.7% when using UIS. In this case, UIS shortened the total encoding time by roughly 6.5%, while preserving video quality (in terms of PSNR), and increasing the encoded video file size by only 0.9%. In some cases, other than saving time, UIS resulted with video files of less size than the files generated using the DS algorithm.

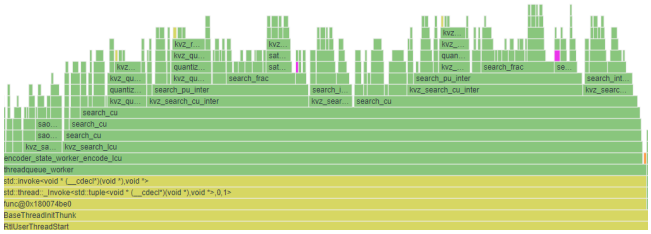


Fig. 5: DS algorithm portion of the total encoding process time (pink color)

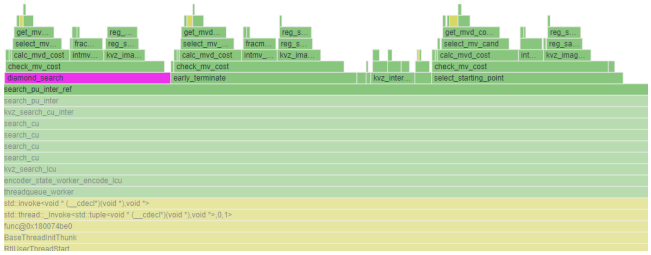


Fig. 6: Figure 5. zoomed-in

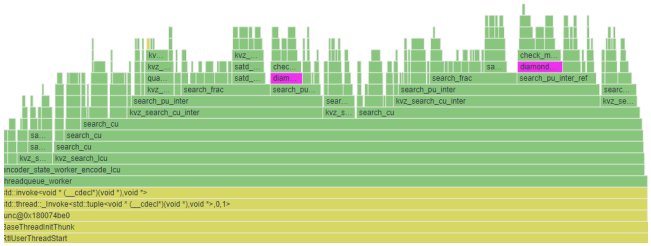


Fig. 7: DS algorithm portion of the total encoding process time (pink color), with using B frames

Algorithm	Sequence number	Encoded video file size [MB]	Search time [CPU s] (less is better)	PSNR [dB] (more is better)
DS	1	10.2	1.53	45.36
	2	38.1	38.70	43.08
	3	19.7	1.11	43.68
	4	9.87	12.37	44.92
UIS	1	10.3	0.37	45.36
	2	38.1	11.32	43.08
	3	19.7	0.39	43.68
	4	9.89	3.58	44.92

TABLE I: Comparison of encoded file size, encoding time and encoded file PSNR for DS and UIS algorithms

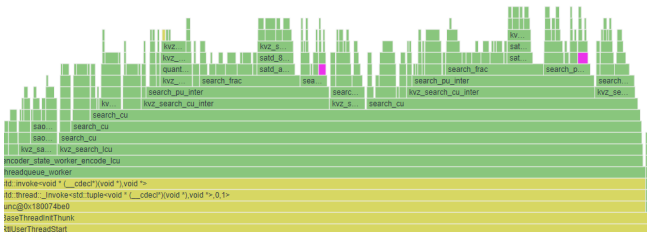


Fig. 8: UIS algorithm portion of the total encoding process time (pink color), with using B frames

V. CONCLUSIONS AND FUTURE WORK

This paper presents an optimization of the ME algorithm in the HEVC encoder Kvazaar by utilizing user input from the controller. An additional ME algorithm was implemented as part of the Kvazaar encoder, which we called UIS. Our results show that the UIS algorithm is more efficient compared to the standard DS algorithm. Video compression rate and quality is preserved, but the encoding speed is increased. When using the default Kvazaar's GOP structure, ME process was up to 4.14 times faster, and when using the GOP structure *8*, which includes B-type frames, the ME process was 4.72 times faster.

Future work includes modifying other ME algorithms besides DS, like HS or TZS. UIS algorithm could be upgraded in terms of number of search directions. Currently it offers only four directions, but could easily be expanded to also include diagonal directions. Besides hovering and rotation, UIS could be expanded to consider drone's forward and backward movement. Directions are not the only input the encoder receives from the user, it also receives the information regarding drone movement speed. Movement speed could be included in the UIS as well, to affect the search distance parameter. When the drone moves faster, it is expected that similar blocks will be a little more further in comparison when the drone moves slower. Of course, this is just a general rule which would have to be tested, because obstacles can always be either far or close to the drone. UIS could be implemented on the x264 encoder as well, since for some cases, it showed decent results in comparison with HEVC [12]. It could also be integrated in a system which offloads video encoding energy to the decoder side, performing ME only on the video G (green) component, instead of Y [13].

In summary, this research presents a promising solution to enhance video encoding speed controlled-camera video systems, offering ample potential for future research and growth.

VI. ACKNOWLEDGEMENT

This work has been supported in part by the project KK.01.2.1.02.0054 Razvoj uređaja za prijenos video signala ultra niske latencije (Development of ultra low latency video signal transmission device), financed by the EU from the European Regional Development Fund.

REFERENCES

[1] N. S. Labib, M. R. Brust, G. Danoy, and P. Bouvry, "The rise of drones in internet of things: A survey on the evolution, prospects and challenges of unmanned aerial vehicles," *IEEE Access*, vol. 9, pp. 115 466–115 487, 2021.

[2] Z. Pan, J. Lei, Y. Zhang, and F. L. Wang, "Adaptive fractional-pixel motion estimation skipped algorithm for efficient hevc motion estimation," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 14, 1 2018.

[3] L. Gao, S. Dong, W. Wang, R. Wang, and W. Gao, "A novel integer-pixel motion estimation algorithm based on quadratic prediction," *Proceedings - International Conference on Image Processing, ICIP*, vol. 2015-December, pp. 2810–2814, 12 2015.

[4] Z. Pan, Y. Zhang, S. Kwong, X. Wang, and L. Xu, "Early termination for tzsearch in hevc motion estimation," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 1389–1393, 10 2013.

[5] H. Kibeya, F. Belghith, H. Loukil, M. A. B. Ayed, and N. Masmoudi, "Tzsearch pattern search improvement for hevc motion estimation modules," *2014 1st International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2014*, pp. 95–99, 2014.

[6] P. Nalluri, L. N. Alves, and A. Navarro, "Complexity reduction methods for fast motion estimation in hevc," *Signal Processing: Image Communication*, vol. 39, pp. 280–292, 11 2015.

[7] E. M. Ibrahim, E. Badry, A. M. Abdelsalam, I. L. Abdalla, M. Sayed, and H. Shalaby, "Neural networks based fractional pixel motion estimation for hevc," *Proceedings - 2018 IEEE International Symposium on Multimedia, ISM 2018*, pp. 110–113, 1 2019.

[8] M. Viitanen, A. Koivula, A. Lemmetti, A. Ylä-Outinen, J. Vanne, and T. D. Hämäläinen, "Kvazaar: Open-source hevc/h.265 encoder," *MM 2016 - Proceedings of the 2016 ACM Multimedia Conference*, pp. 1179–1182, 10 2016. [Online]. Available: https://dl.acm.org/doi/10.1145/2964284.2973796

[9] "FFmpeg," https://www.ffmpeg.org/, accessed: 2023-02-04.

[10] "Video analysis software, mpeg analyzer for video quality tests automation, batch execution | elecard: Video compression guru," https://www.elecard.com/products/video-analysis/streameye, accessed: 2023-02-04.

[11] "Fix performance bottlenecks with intel® vtune™ profiler," https://www.intel.com/content/www/us/en/developer/tools/oneapi/vtune-profiler.html#gs.ohbu13, accessed: 2023-02-04.

[12] J. Benjak, D. Hofman, J. Knezović, and M. Žagar, "Performance comparison of h.264 and h.265 encoders in a 4k fpv drone piloting system," *Applied Sciences 2022, Vol. 12, Page 6386*, vol. 12, p. 6386, 6 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/13/6386/htmhttps://www.mdpi.com/2076-3417/12/13/6386

[13] D. Hofman and J. Benjak, "Offloading video encoding energy consumption to the decoder," *2022 7th International Conference on Smart and Sustainable Technologies, SpliTech 2022*, 2022.

FOGO - an Optimized Fog and Edge Computing Method for VANETs

Renato Ćorluka*, Ivona Matoš†, Davor Damjanović‡ and Josip Balen§
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b, 31000 Osijek, Croatia
*rcorluka@etfos.hr, †imatos@etfos.hr, ‡davor.damjanovic@ferit.hr, §josip.balen@ferit.hr

***Abstract*—Vehicular Ad-hoc Networks (VANETs) have many specific issues and challenges due to a heterogeneous and highly dynamic environment, and therefore traditional solutions need to be improved and adopted in order to satisfy the networking and processing requirements. Fog and edge computing principles enable VANETs to achieve a more realistic and dependable architecture by using several layers for information processing. In this paper, the developed FOGO method is improved and customized for VANETs. The proposed solution is using an enhanced infrastructure in which additional mobile fog nodes are added to the network together with the existing stationary nodes. Small and medium sized messages are processed by mobile fog nodes, which then disseminate the results across the network. Roadside units (RSUs) are fixed, highly-capable devices positioned throughout the network to assist the processing of increasing volumes of data. Additionally, the cloud is taken into account, but only when processing large amounts of data. Along with the presented method flowcharts, operating algorithms, and message structure, the proposed system architecture is described. Furthermore, potential use cases are suggested together with the metrics for the system performance evaluation. The LuST scenario was used to evaluate the proposed architecture and the results showed improved message processing efficiency.**

***Index Terms*—Fog and Edge Computing, VANET optimization, LuST scenario, On-Board Units (OBU), Traffic Management**

I. Introduction

Large-scale improvements in computation and communication technologies are leading the automotive industry towards connected vehicles in order to improve traffic safety, efficiency, and user experience. Based on short-range radio, vehicles are forming a special class of mobile networks (MANET) called VANET networks. Due to the highly dynamic topology of VANET, there are lots of challenges to overcome. One of the solutions for that problem and integration of new technology such as 5G in VANET is software-defined vehicular networks. These networks aim to improve security, QoS, routing reliability, and delay in VANET networks [1], [2].

To unlock the full potential of these networks researchers are making models in which VANETs cooperate with today's cutting-edge technology like cloud computing which is used to compute huge amounts of data. Due to VANETs high mobility nodes, scalability, and frequent topology changes, it is difficult to meet real-time traffic requirements using only cloud computing. Fog computing brings required computing power near VANETs nodes, it is strong enough for computing time-sensitive applications and it has enough storage capacity to gather local traffic information to send it to the cloud. There are different ways to improve intelligent transportation systems (ITS) with fog computing, as well as a variety of cloud computing capabilities and applications are available when VANET and fog computing are combined [3]. As it is described in [2], one of the most promising ways is to use fog computing for a dynamic traffic light system to optimize traffic light junctions. As vehicles are becoming more and more interconnected and generating huge amounts of data, there is a need for fog computing in order to meet time-critical requirements. However, it is not enough to just introduce processing at the edge of the network, it is important to optimize it.

Fog and edge computing reduces response time and brings computing closer to edge nodes compared to cloud computing, but also uses cloud computing when it is needed for big data computation which fog nodes can not compute in a predefined time window. In most methods, fog nodes are stationary nodes that bring the additional cost of maintaining infrastructure. Fog nodes can be deployed to collect trust evaluations from vehicles, allowing them to rely on local vehicles to perform certain tasks. Furthermore, fog nodes can be used to keep track of its local vehicles, reducing the need for cloud usage [4]. The proposed method in this paper includes mobile fog nodes, which have reduced maintenance costs and can cover a wider area because of their mobility but with added reliability and connectivity issues because there will be a time when some city districts will not be covered with mobile fog nodes. The performance of the proposed method is evaluated in the simulation of the LuST scenario. The rest of the paper is organized as follows: Section II describes fog and edge computing architecture. Section III presents our innovative method for applying fog computing in VANETs, whereas in Section IV the simulation results are discussed. Finally, conclusions and future work are given in Section V.

II. Fog and Edge Computing Architecture

Nowadays vehicles are becoming smarter as they are equipped with various sensors that can provide driving assistance. We are reaching peak private car years and entering the era of new market mobility-as-a-service (MaaS) because

the total cost of ownership (TCO) is high. As conventional transport develops into ITS, cars have to be equipped with numerous sensors that generate an immense amount of data that need to be processed and transmitted to other vehicles and infrastructure. At this stage, computational power is too expensive for regular cars and high computational capabilities are reserved for high-end cars, so at this stage, a transitional solution before the era of fully autonomous vehicles is needed. Cloud computing is a well-proven and reliable technology that can provide processing and storage power to vehicles. However, vehicles are highly dynamic nodes and require low latency, and clouds are centralized centers, far away from vehicles. Considering that, to achieve ITS goals, a mediator in our architecture between vehicles and the cloud is needed. Fog computing is an alternative to overcome these problems and challenges. It brings resources on the network edge closer to the vehicles so not all the data needs to be sent to the cloud. Most of the data can be processed on the edge and only the necessary data is forwarded to the cloud. As described in [5], fog computing can perform processing and storage tasks on the most logical resources and reduce the overall impact on the network and processing resources. There are three possible fog connections in VANETs: vehicle to fog, fog to fog, and fog to cloud. In VANETs, fog nodes could be RSU, public transportation, or taxis. As described in [2], [6], and [7], three-level architecture is proposed where the bottom level is composed of end-user vehicles, the middle of edge computing, and cloud computing on the top. Grover et-al. in [2] proposed this architecture for four different applications in VANET: smart traffic lights, parking systems, content distribution, and decision support system. Dolui et-al. in [7] compared three different implementations of edge computing: mobile edge computing (MEC), fog computing (FC), and cloudlet. Every application has its requirements and there are always trade-offs, therefore every implementation has advantages and disadvantages in the considered use case. Edge computing has an N-tier architecture where end devices, smartphones, cameras, and vehicles, are in the lower tier; intermediate tiers consist of FC, MEC, or cloudlet nodes; and the cloud is located at the highest tier. FC can use Bluetooth, Wi-Fi, and mobile networks as an access mechanism. On the other hand, MEC uses only mobile networks, and cloudlet uses only Wi-Fi. They proposed a decision tree for finding the best implementations based on six parameters such as physical proximity, power consumption, computation time, context awareness, logical proximity, and non-IP support. The decision tree is given in Table I.

TABLE I: Edge computing decision tree

	FC	**MEC**	**CC**
Physical Proximity	HIGH	LOW	HIGH
Power Consumption	LOW	HIGH	LOW
Computation Time	HIGH	LOW	LOW
Context Awareness	LOW	HIGH	LOW
Logical Proximity	MAYBE	MAYBE	ENSURED
Non-IP Support	YES	NO	NO

Depending on the proposed intra-fog architecture all the vehicles or some of them could be connected to the cloud. RSUs are the most expensive fog nodes and it is too expensive to use only them. Furthermore, finding the ideal number, locations, and computational capabilities of the RSUs is one of the main issues, especially in urban areas where there are many obstacles inside the RSUs' coverage area [8]. Public transportation like buses and trams can have large computational and storage capabilities because they are spacious, but their problem is that they have fixed routes. To cover other regions taxis can be utilized as fog nodes. They do not have huge computational and storage power, but they go on many streets that do not have considerable traffic flow and their fleet is a few times bigger than public transportation. As described in [9], the cloud paradigm is shifting from locally managed software on physical hardware toward virtual services. The cloud brings all different virtualized services which we can distinguish by cloud orchestration into two approaches: orchestration of hardware-as-a-service (HaaS) and orchestration of software-as-a-service (SaaS). The orchestration enables enterprises to use multiple clouds for numerous services and syncs them into a single workflow. Orchestration in ITS will play a huge role because services will probably be deployed in multiple cloud vendors and they will have to be represented as a single workflow to the end-user.

III. Description of Proposed FOGO Method

In this paper, a decentralized system distributed throughout the network with stationary and dynamic fog nodes is presented. In subsection A, a proposed system architecture along with the messaging structure is described. Subsection B describes the workflow of the proposed system. In subsection C, possible use cases are presented. Finally, in subsection D, evaluation metrics are proposed.

A. System Architecture

In this particular case, stationary nodes, i.e. fRSUs, would be placed in an area close to bus stops as they cover most of the city road traffic. In addition, usage of dynamically distributed nodes is proposed, which would be set up in buses and taxi vehicles as those are high-performance vehicles with more computing power and storage. Furthermore, it is worth mentioning that buses have been quite thoroughly researched and simulated, as in [6], [10], [11], in order to provide the best solution for fog message distribution. Therefore, these vehicles would be suitable for running real-time, complex applications that edge-level vehicles could not perform due to the lack of resources. The complete architecture of the proposed FOGO system is shown in Figure 1.

There are two types of nodes in FOGO architecture: the first type consists of edge nodes that represent vehicles with low computational power OBUs; the other type of nodes are fog nodes that represent buses and taxi vehicles along with RSUs which are placed at the public transportation stations. This category is further subdivided by the mobility of these

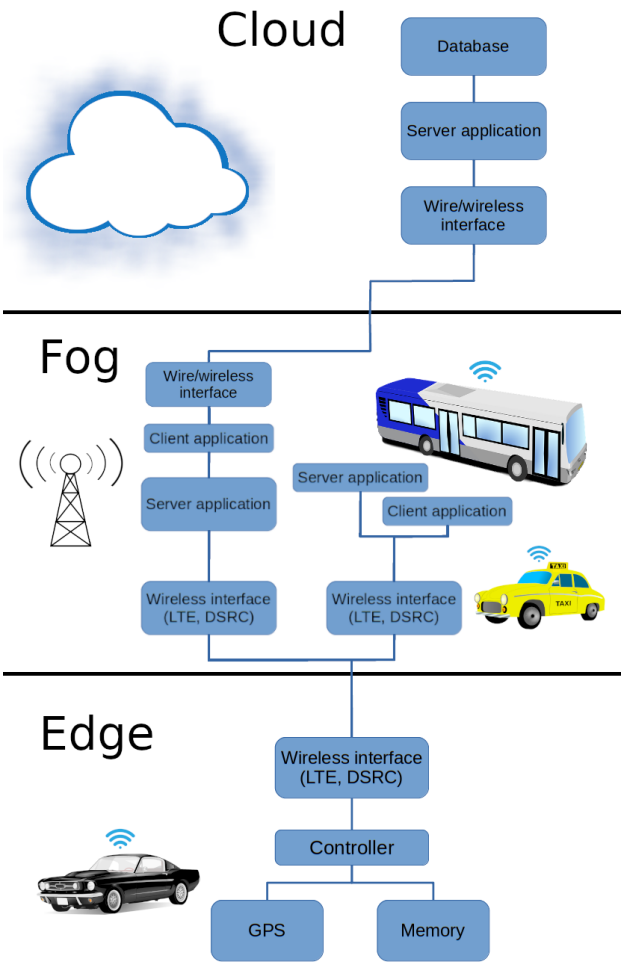


Fig. 1: Proposed FOGO architecture

nodes, taxis, and buses are moving nodes, and RSUs are stationary nodes. Arrows are representing possible types of communication in FOGO architecture. vehicle-to-vehicle (V2V) communication is performed in the DSRC band and it is represented with the white arrow. The blue arrow represents infrastructure-to-cloud (I2C) communication RSUs and cloud, and in FOGO architecture that can be performed wirelessly with a 5G network or by optical fiber infrastructure. The last type of communication is vehicle-to-infrastructure (V2I) in which mobile fog nodes send data to the RSU and RSU sends back results to the mobile fog nodes, which is represented by the green arrow.

In addition, four types of messages are proposed: urgent, essential, application, and media message, each identified by its special header character. Urgent messages refer to messages connected to sudden traffic situations such as congestion, collision, accidents, roadwork, etc. Those messages should be distributed in the network with high priority by using appropriate flags and by sending them to all vehicles on the map of the fog node. Furthermore, these messages contain the following information: message ID, sender ID, GPS coordinates of the location where the traffic situation has occurred and the type of the traffic problem. Moreover, these are the only messages that are being sent without requests from the edge nodes.

Furthermore, essential messages refer to the request messages in which a certain vehicle requests information from the node. The essential message consists of message ID, node ID, GPS location, and type of the information needed (for instance weather or traffic report) or response to the request.

Moreover, there are application messages which refer to edge requests for additional resources for time-critical applications. They consist of message ID, node ID, application ID, resources needed to perform a task, and a short description of a task.

Lastly, some media messages are referred to as audio and video content for passengers and they consist of message ID, node ID, type of media content, requested resources, and media content for processing.

Considering the size of the described messages, we could easily divide them into three classes: small-sized, medium-sized, and large-sized messages. Urgent and essential messages would be considered small-sized messages as they contain basic ID information and short notice about the surrounding traffic situation. Further, application messages and some of the media messages (that are smaller than 20 MB) could be classified as medium-sized messages as they usually consist of some detailed information about the application that needs to be executed such as ID, needed resources, and description. Moreover, media messages greater than 20 MB are considered large-sized messages. That includes transferring large files with video and/or audio content.

B. Proposed System Workflow

In the following section workflow of the proposed architecture will be explained. The proposed FOGO architecture consists of the following components:

a) Composing the vehicle map
b) Scan the vehicle map for the available fog resources (edge vehicles)
c) Processing requests
d) Handling the idle areas
e) Distribution of urgent messages (fRSU)

The algorithm used to simulate the fog message manipulation is presented in Algorithm 1.

Algorithm 1: FOG RSU/TAXI/BUS aplication

```
OnMsgRcvd();
if Msg.GetType() == SendBeaconMsg then
    SendToAll(Msg);
    SendToCloud(Msg);
else
    ProcessMsg();
    SendResultBack();
end
```

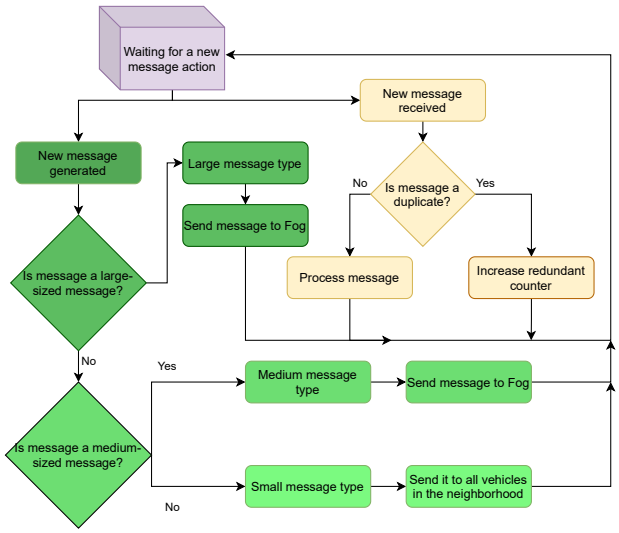


Fig. 2: Flowchart for handling messages on the edge vehicle

It is necessary to distinguish messages by their size to process them on the appropriate system. The flowchart for handling messages on the edge vehicle is presented in Figure 2. The algorithm for generating messages on edge vehicles is shown in Algorithm 2. The algorithm for processing messages on edge vehicles is presented in Algorithm 3.

Algorithm 2: EDGE CLIENT APPLICATION

```
OnGenMsg();
DiscNeighb();
if Msg.GetSize() > 20MB then
    Msg.SetType(TypeLarge);
    SendToFog(Msg);
else if Msg.GetSize() > 2MB then
    Msg.SetType(TypeMedium);
    if RSUinRange() and ProcessingTimeOnEdge
      <TransferProcessingTimeFog then
        SendToFog(Msg);
    else
        ProcessMsg();
    end
else
    Msg.SetType(TypeSmall);
    if Msg.GetPurpose == request then
        SendToFog(Msg)
    else
        SendToAllNeighbours(Msg);
    end
end
WaitForResult();
```

Algorithm 3: EDGE CLIENT APPLICATION

```
OnMsgRcvd();
MesageStatus = OK;
RedundantMsgCounter = 0;
for each item i in sizeof(KnownIDs) do
    if Msg.GetID() == KnownIDs[i] then
        RedundantMsgCounter++;
        break;
    end
end
if Msg.GetType() == TypeSmall then
    ProcessMsg();
    if distanceVehicle >=
     (maxDistanceNeighbourVehicle / 2) then
        Forward(Msg);
    if Msg.GetPurpose == response then
        MeasureMsgTransferTime();
    end
end
```

C. Proposed Metrics

In order to precisely determine the advantages of using FOGO, the following metrics will be taken into consideration: an average number of successfully sent messages (aV), edge and processing ratio, fog processing efficiency, and cloud processing efficiency. To acquire metrics, it is necessary to measure the following parameters during the simulation: data network traffic overhead ($dNtO$), which stands for the total number of messages sent during the simulation; the number of received messages (nRm); and the average time for processing messages ($avTime$). The number of received messages will be calculated considering message ID. Redundant messages will not be taken into the calculation.

In order to provide generally useful and realistic data, both urgent and essential messages will be consolidated to determine the coverage of small-sized messages. Coverage by small-sized messages will be described using the following equation:

$$aV_{small} = \frac{nRm_{small}}{dNto_{small}} \tag{1}$$

Furthermore, when talking about medium-sized messages, application requests will be counted to determine the edge and fog processing ratio, PRmedium. It is necessary to count medium-sized messages processed in the fog, as well as all the medium-sized messages generated during simulation.

$$PR_{medium} = \frac{nRm_{medium}}{dNto_{medium}} \tag{2}$$

Moreover, when talking about large-sized messages, media requests will be counted to determine the coverage by large-sized messages. Cloud delivery and processing efficiency by large-sized messages will be described using the following expression:

$$aV_{large} = \frac{nRm_{large}}{dNto_{large}} \tag{3}$$

As was stated above, the influence of clustering on the number of redundant messages will be considered along with

message ID which will be used to leave out redundant messages. In that manner, the overall network coverage will be measured. This parameter is described with the following equation:

$$aVrM = \frac{rM}{dNto_{total}} \tag{4}$$

Furthermore, fog processing efficiency Ef (fog) stands for the quotient of average time needed for processing medium-sized messages ($avTime_{med}$) on fog and edge, where time for processing on fog also includes transfer time to fog.

$$E_f(FOG) = \frac{avTime_{med}(EDGE)}{avTime_{med}(FOG)} \tag{5}$$

Simulation results are expected to show that the proposed algorithm efficiently reduces the network congestion by processing messages with a unique ID, consequently decreasing the load on the core network as well as improving the efficiency of data distribution in the fog layer.

IV. Performance Anaysis and Discussion

In simulation conducted via Veins, SUMO, and OMNET++ we compared stated metrics in case of using the aforementioned FOGO architecture. In this section, the results of a simulation performed on the Luxembourg map are presented. Simulations are performed in Veins, an open-source simulation framework, and are based on the LuST scenario which represents realistic 24h traffic on the Luxembourg map [12], [13]. Furthermore, by analyzing the LuST vehicle movement, we deployed 35 RSU fog nodes across different parts of the city, and 70 dynamic fog nodes which are representing taxi and bus vehicles.

There were six conducted simulations. Every simulation lasted for one hour. Every 10 minutes of the simulation, proper statistics were generated to determine an average of all measurements.

a) Network coverage In Figure 3, an increase in the total number of messages generated during simulation is shown. Every 10 minutes several thousand messages were generated. A large proportion of those messages were successfully processed in every step of the simulation which is shown in Figure 4. Furthermore, it is important to say that an average of 35,057 messages were generated in each simulation and an average of 29,132 were successfully processed which resulted in network coverage of 83.1 %. Moreover, for small-sized messages, it was measured that an average of 27,598 small-sized messages were generated, and 20,153 were successfully processed, which leads to 73 % of successfully processed small-sized messages.

b) Edge and fog processing ratio During the simulation, it was measured that an average of 4242 generated medium-sized messages were sent to fog, whereas an average of 10122 medium-sized messages were generated in total. Therefore, it is calculated that the edge and fog processing ratio is around 41.91 %, which leads us to the conclusion that fog is the practical solution for message management, and also, it is very beneficial in a financial manner since it can process more than 40 % of all messages in the network which will certainly reduce the load of the cloud.

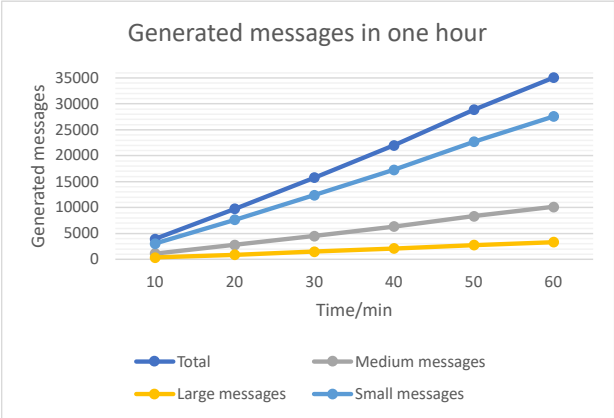


Fig. 3: Trend of generated messages during simulation

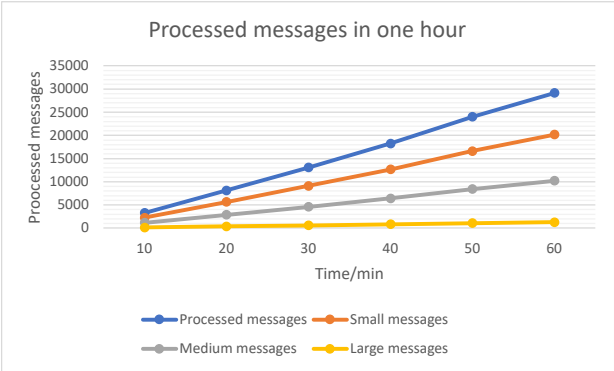


Fig. 4: Trend of processed messages during simulation

c) Fog processing efficiency Fog processing efficiency is a very important parameter since it tells how much faster a certain fog system can process messages that demand high computing power and could easily overload the edge vehicle. In our research, it is measured that the average time for edge vehicles to process a message is 46.18 s whereas the average time for the fog to process a message is 14.482 s. This measurement did not include any of the mechanisms for deleting messages that were not processed in a certain time, all medium-sized messages that did not get lost in the transfer were processed. It was established that fog is close to 320 % more efficient compared to edge vehicles. For example, if a fog can process a video sequence of 1 minute in 20 seconds, it would take more than a minute for an edge vehicle to do the same.

d) Cloud delivery and processing efficiency Cloud delivery and processing efficiency are counted based on successfully processed large-sized messages on the cloud compared to all generated large-sized messages in simulation. In our simulations, there is an average of 3348 large-sized messages created and 1273 messages that were successfully processed. That leads to the final result of

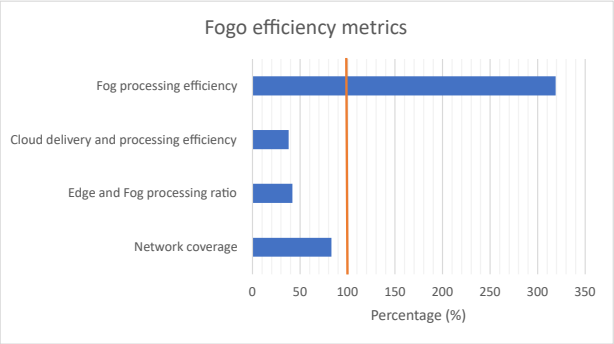


Fig. 5: Overall look on measured parameters

38.02 % of successfully processed large-sized messages. The rest of the messages were lost in the transfer or deleted due to incompetence in finding an RSU. Simulation results can be seen in Figure 5.

e) Discussion When we consider all measured and discussed parameters that are presented in this paper, we managed to prove that edge and fog computing systems are a great way to improve VANETs. Furthermore, we proved that these systems are very efficient, and considering our results in fog processing efficiency and network coverage, this area of research is truly an interesting field that will be one of the key elements toward completely autonomous vehicles. Nevertheless, considering the finances needed to develop VANETs, it is a relatively cheap solution compared to the price of the cloud. Our proposed solution is useful in various road traffic situations such as collisions, congestions, as well as everyday big data distribution in VANETs which appears to be one of the biggest problems that needs to be tackled with in process of developing autonomous driving.

V. Conclusion

In this paper, we concluded that fog and edge computing can be efficient in spreading messages across the city. As infrastructure is expensive, and most of the cities have public transportation that covers predetermined routes and taxi vehicles which can cover non-predetermined routes, with a good placing strategy of RSUs and using public transportation and taxi vehicles, most parts of the city could be covered cost-efficiently by using our proposed method.

Since there are a lot of messages being generated in the network with each minute and some of them are waiting in queue to be processed for some time, the time needed for medium-sized messages to be processed both on fog and edge is increasing drastically. As a solution to this problem, there is a possibility of implementing a mechanism that will delete messages that are waiting to be processed more than a certain, realistically set time. In such a manner, loss of important information will not happen due to the number of neighboring vehicles that will send similar requests to the fog and forward the response to other edge vehicles in the essential message.

Even though most of the medium-sized messages are still being processed on edge vehicles, by further improvement of our algorithm, as well as increasing the number of RSUs in simulation, it is possible to increase the efficiency of fog. Furthermore, adding more vehicles and infrastructure as fog nodes, such as trams and tram stations, and changing the parameters of message sizes could be fine-tuned to improve results. The drawback of this method is that most taxi vehicles and public vehicles could be on the same route, over-covering some parts of the city and not covering other parts. But looking at the economic benefits of this solution this could be a necessary trade-off between fog quality of service and infrastructure price.

References

[1] W. B. Jaballah, M. Conti, and C. Lal, "A Survey on Software-Defined VANETs: Benefits, Challenges, and Future Directions," *arXiv:1904.04577 [cs]*, May 2019.

[2] J. Grover, A. Jain, S. Singhal, and A. Yadav, "Real-Time VANET Applications Using Fog Computing," in *Proceedings of First International Conference on Smart System, Innovations and Computing*, ser. Smart Innovation, Systems and Technologies, A. K. Somani, S. Srivastava, A. Mundra, and S. Rawat, Eds. Singapore: Springer, 2018, pp. 683–691.

[3] H. Singh, V. Laxmi, A. Malik, and Isha, "Fog Computing as Future Perspective in Vehicular Ad hoc Networks," in *Fog, Edge, and Pervasive Computing in Intelligent IoT Driven Applications*. IEEE, 2021, pp. 177–192.

[4] R. J. Atwa, P. Flocchini, and A. Nayak, "A Fog-based Reputation Evaluation Model for VANETs," in *2021 International Symposium on Networks, Computers and Communications (ISNCC)*, Oct. 2021, pp. 1–7.

[5] J. C. Nobre, A. M. de Souza, D. Rosário, C. Both, L. A. Villas, E. Cerqueira, T. Braun, and M. Gerla, "Vehicular software-defined networking and fog computing: Integration and design principles," *Ad Hoc Networks*, vol. 82, pp. 172–181, Jan. 2019.

[6] G. Sun, Y. Zhang, D. Liao, H. Yu, X. Du, and M. Guizani, "Bus-Trajectory-Based Street-Centric Routing for Message Delivery in Urban Vehicular Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7550–7563, Aug. 2018.

[7] K. Dolui and S. K. Datta, "Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing," in *2017 Global Internet of Things Summit (GIoTS)*, Jun. 2017, pp. 1–6.

[8] R. Shahin, A. A. El-Moursy, S. M. Saif, H. M. Abbas, and S. M. Nassar, "Fog Node Optimum Placement and Configuration Technique for VANETs," in *2020 International Conference on Communications, Signal Processing, and Their Applications (ICCSPA)*, Mar. 2021, pp. 1–6.

[9] K. Bousselmi, Z. Brahmi, and M. M. Gammoudi, "Cloud Services Orchestration: A Comparative Study of Existing Approaches," in *2014 28th International Conference on Advanced Information Networking and Applications Workshops*, May 2014, pp. 410–416.

[10] F. Zhang, H. Liu, Y.-W. Leung, X. Chu, and B. Jin, "CBS: Community-Based Bus System as Routing Backbone for Vehicular Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 8, pp. 2132–2146, Aug. 2017.

[11] P. H. Cruz Caminha, F. Ferreira da Silva, R. Gonçalves Pacheco, R. de Souza Couto, P. Braconnot Velloso, M. E. Mitre Campista, and L. H. M. K. Maciel Kosmalski Costa, "SensingBus: Using Bus Lines and Fog Computing for Smart Sensing the City," *IEEE Cloud Computing*, vol. 5, no. 5, pp. 58–69, Sep. 2018.

[12] C. A. R. L. Brennand, A. Boukerche, R. Meneguette, and L. A. Villas, "A novel urban traffic management mechanism based on FOG," in *2017 IEEE Symposium on Computers and Communications (ISCC)*, Jul. 2017, pp. 377–382.

[13] L. Codeca, R. Frank, and T. Engel, "Luxembourg SUMO Traffic (LuST) Scenario: 24 hours of mobility for vehicular networking research," in *2015 IEEE Vehicular Networking Conference (VNC)*, Dec. 2015, pp. 1–8.

On Signal Attenuation in 5G Higher Frequencies Transmissions with Different Area Scenario

T. Iliev*, I. Stoyanov**, I. Penkov*, G. Mihaylov***, A. Fazylova****,
G. Balbayev*****, I. Beloev****** and S. Kosunalp*******
* University of Ruse, Department of Telecommunication, Ruse, Bulgaria
** University of Ruse, Department of Electrical and Power Engineering, Ruse, Bulgaria
*** University of Telecommunications and Post, Department of Telecommunications, Sofia, Bulgaria
**** Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan
***** Academy of logistics and transport, Almaty, Kazakhstan
****** University of Ruse, Department of Transport, Ruse, Bulgaria
******* Bandırma Onyedi Eylül University, Department of Computer Technologies, Bandırma, Türkiye
tiliev@uni-ruse.bg

***Abstract* - Energy and cost efficiency are becoming criteria of increasing importance in the design of 5G radio relay lines, especially for areas close to water surfaces, where there is pronounced fading, which is characteristic of radio relay paths that pass on the surface of the water. In this paper, the research was carried out on two radio relay links with different frequencies – the first one at 8 GHz, and the second at 18 GHz, working in frequency diversity mode. To troubleshoot, the measurements were made during the shortest possible time, with fixed emissions to account any fluctuations in signal levels.**

Keywords - network planning; 5G; radio – relay line; fading

I. INTRODUCTION

Communication and data transmission systems need a constant increase in bandwidth and data transfer rates. This need is most pronounced in mobile networks. Since 2017, the volume of data traffic in mobile networks has increased several times and is expected to increase in the future as well [1].

Contemporary mobile networks, such as LTE (4G) and 5G, provide users with ever-higher data transfer rates, which inevitably requires ever-increasing bandwidth from transport networks [2]. The fundamental limitations of data rate are due to the bandwidth of the communication system. International organizations, such as CEPT / ECC and ITU-R, dealing with the allocation of the frequency range, allocate a wide frequency range from 6 GHz to 86 GHz for operation of radio relay lines. As the frequency increases, the cost of equipment increases significantly, and the attenuation of signals on surface paths also increases [3]. Therefore, most radio relay communication lines operate in the frequency range up to 30 GHz.

The standard frequency bands for trunk lines are 3.5, 7, 14, 28 and 56 MHz [4].

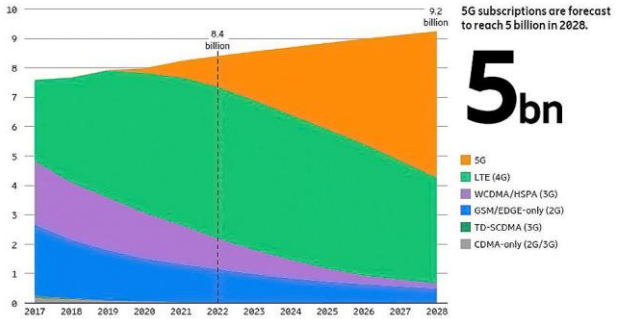


Figure 1. Mobile subscriptions (in billion) [1]

Due to the limited frequency resource, a more efficient use of the allocated frequency band is required; furthermore, the telecommunication operator must be able to increase the bandwidth of the communication line without changing the frequency planning [5].

This paper deals with various ways to improve the efficiency of the use of the frequency resource, as well as the speed of data transmission.

II. RADIO WAVE PROPAGATION

Radio communication is distinguished by the complexity and non-stationarity of the conditions for the propagation of radio waves in inhomogeneous dispersion anisotropic media and interference situations.

Signal fading is characteristic of the HF radio channel. In the ionospheric channel, the signal components propagate along several paths. First, rays are detected at the point of reception, which are propagated by multiple or single reflection. Second, the signal is emitted from the transmitting antenna within the width of the radiation pattern (RP), therefore we can assume that not a single beam falls on the ionosphere, but a beam of sub-beams that reach the receiving antenna with different amplitudes and phases. Third, a propagation medium with large-scale and small-scale inhomogeneous structures leads to

multipath interference. Multipath at high speeds and common types of tampering limit the ability to improve communication reliability by increasing transmitter power [6]. If the difference in the path of the rays is comparable to the duration of one element, then the information is received along the ray that currently has the highest level, and when moving from one ray to another, desynchronization of cycles occurs, which leads to loss of individual symbols of the message. Multipath with a large difference in the path of the beams leads to interference of the message elements arriving at the receiving point with different beams, and the signals in beams with a lower level in this case play the role of additive interference.

Free space is usually considered as a homogeneous, boundless and non-absorbing medium in which the relative dielectric permittivity is equal to one. Thus, many factors affecting the propagation of radio waves are not taken into account. In the Earth's atmosphere, there are always atmospheric gases, water vapor, hydrometeors (cloud, rain, snow, hail, fog), as well as dust particles raised from the Earth's surface, which cause a weakening of the power of the received signal. It is accepted to consider two types of attenuation - absorption and scattering [7]. When absorbed, part of the energy of radio waves is converted into heat energy, and when scattered, it is redistributed in space in directions other than the given one. Attenuation in the troposphere of centimeter range signals is of appreciable value.

$$\overline{P} = \overline{P}_0 10^{-2\int_0^R \alpha(R)dR}, \tag{1}$$

where R is the distance between transmitter and receiver; $\overline{P}$ is the average power at the receiver input without accounting for attenuation; $\overline{P}_0$ is the average power at the receiver input, taking attenuation into account; $\alpha(R)$ characterizes the change of the attenuating properties of the medium in the direction of propagation of the radio beam (wave) [8].

In the case of a homogeneous propagation medium, when the attenuation coefficient does not change along the propagation path of the radio waves, the power of the received signal is expressed as:

$$\overline{P} = \overline{P}_0 10^{-2\alpha R}, \tag{2}$$

where α is attenuation coefficient (dB/km).

When using a logarithmic scale, (2) will convert to:

$$10 \lg \frac{\overline{P}}{\overline{P}_0} = -2\alpha R. \tag{3}$$

To characterize the total loss along the propagation path, the coefficient Γ is introduced. The value of α is measured in decibels (dB) and is the sum of the attenuation coefficients in gasses and water vapor α_g, in clouds α_c and precipitation α_r:

$$\Gamma = 2\left(\alpha_c \Delta r_g + \alpha_c \Delta r_c + \alpha_r \Delta r_r\right), \tag{4}$$

A. Free Space Communications

The transmission in free space can be represented by the Friis formula (5):

$$\frac{P_r}{P_t} \approx \left(\frac{\lambda}{4\pi d}\right)^2 \tag{5}$$

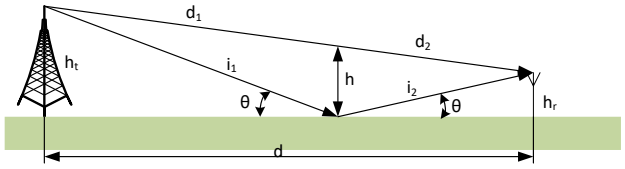


Figure 2. Direct and reflected signal on flat terrain

The attenuation of the signal when reflected from the earth's surface can be represented by:

$$\frac{P_r}{P_t} \approx \left(\frac{h_t h_r}{d^2}\right)^2, \tag{6}$$

where P_r is received power; P_t is radiated power; d is the distance between receiver and transmitter; h_t, h_r are the heights of transmitting and receiving antenna respectively.

When propagating in free space, attenuation of the signal may be present at sharp edges (knife - edge diffraction loss) - a wedge-shaped obstacle. These obstacles can be due to hills, buildings, trees, etc. obscuring the radio path and line-of-sight (LoS) beam. In this way, a "shadowing" effect appears on the receiver [9].

There are several different variations of signal attenuation from multiple wedge-shaped obstacles:

- Bullington model - attenuation by 2 obstacles. N obstacles are replaced by one of equivalent height h, reflecting the effect of all wedge-shaped obstacles (some obstacles are ignored) [6].

The diffraction parameter v is defined as:

$$v = h\sqrt{\frac{2}{\lambda}\left(\frac{1}{d_1} + \frac{1}{d_2}\right)}, \tag{7}$$

where h is the height of the obstacle, d_1 is distance transmitter – obstacle; d_2 is distance receiver – obstacle

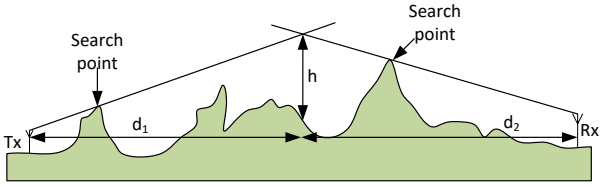


Figure 3. Bullington model

- Epstein-Peterson model – it considers the attenuation from each obstacle sequentially, with the total attenuation being the sum of the individual attenuations [6].

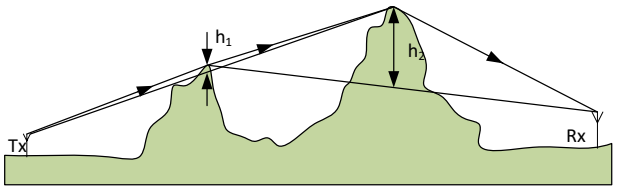


Figure 4. Epstein-Peterson model

- Deygout model – the model initially considers the damping effect of the main obstacle, and the effect of other obstacles relative to it [6].

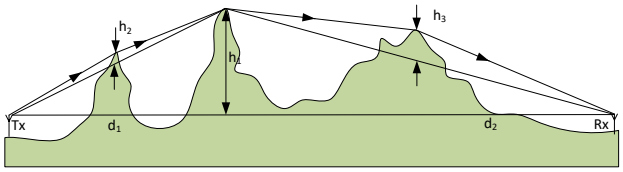


Figure 5. Deygout model

B. Attenuation of centimeter radio waves in atmosphere and vapor

Attenuation is due to the scattering and absorption of signal energy due to the conversion of some of the energy into thermal energy by gas particles (oxygen and nitrogen) and water vapor.

In the propagation of radio waves with a wavelength smaller than λ = 3.2 cm, the attenuation in the troposphere is mainly due to absorption, the scattering effect is insignificant and can be neglected [9]. Then the total attenuation coefficient α_Γ is the sum of the attenuation coefficients of the gas and water vapor:

$$\alpha_\Gamma = \alpha_g + \alpha_w, \tag{8}$$

where α_g and α_w are the attenuation coefficients (dB/km) of gas and water vapor, respectively.

C. Attenuation of centimeter radio waves in presence of rain

Rain consists of water particles of various sizes, shapes and densities. Due to the fact that the size distribution law is different for different particle sizes, the most appropriate method for calculating the attenuation coefficient is the empirical method, where it is assumed that the dependence of the attenuation coefficient on the intensity of precipitation is an exponential law:

$$\alpha_c = aI^b, [dB/km], \tag{9}$$

where the parameters a and b are empirical values and which depend on the wavelength (e.g., at λ=3.2 cm a=0.0074 and b=1.31); I- precipitation intensity [mm/h].

D. Attenuation of centimeter radio waves in presence of snowfall

The snowflakes that make up snowfall vary in shape, orientation, density, size, and moisture content. Most snowflakes are hexagonal and range from 0.5 to 3 mm in diameter. The refractive index is the combined index for a mixture of air, water and ice [10]. If we assume that the snowflake has a spherical shape and its mass is equal to the mass of the same water sphere, it is possible to calculate the damping coefficient of liquid snow by the snowfall intensity I by means of:

$$\alpha_c = 0.094 \frac{I}{\lambda}. \tag{10}$$

For dry snowfall at 0°C, the attenuation coefficient is described by:

$$\alpha_n = 0.035 \frac{I^2}{\lambda^4} + 0.0022 \frac{I}{\lambda}. \tag{11}$$

In conclusion, it can be said that the magnitude of the attenuation coefficient in the millimeter range is much larger than in the centimeter range; when operating at short distances, attenuation in gases and clouds without precipitation can be neglected; it is necessary to take into account only the attenuation of the signal in heavy rain, hail and snowfall [11].

III. STUDY OBJECTS

Three routes differing in type of terrain and altitude were investigated.

A. Trace route 1: Dospat - Sarnitsa.

The length of the track is 15.6 km. The landscape is mountainous, highly dissected and rounded. The average altitude is 1,300 m. For the lower parts of the territory, the annual amount of precipitation is from 620 to 700 l/m2, and for the higher ones from 780 to 1100 l/m2. The average snow cover is about 97 cm and lasts between 80 and 150 days. The average atmospheric pressure for the Smolyan region is 737.0 hPa. The average annual temperature ranges from 9 to 13°C. The forms of the relief in this part have a significant impact on the radio relay connection and the relatively high altitude (1000 - 1600 m) in the territory creates conditions for almost daily temperature inversions and the air temperatures throughout the year are much lower than places with a similar height in the rest of the mountains.



Figure 6. Dospat – Sarnitsa [12]

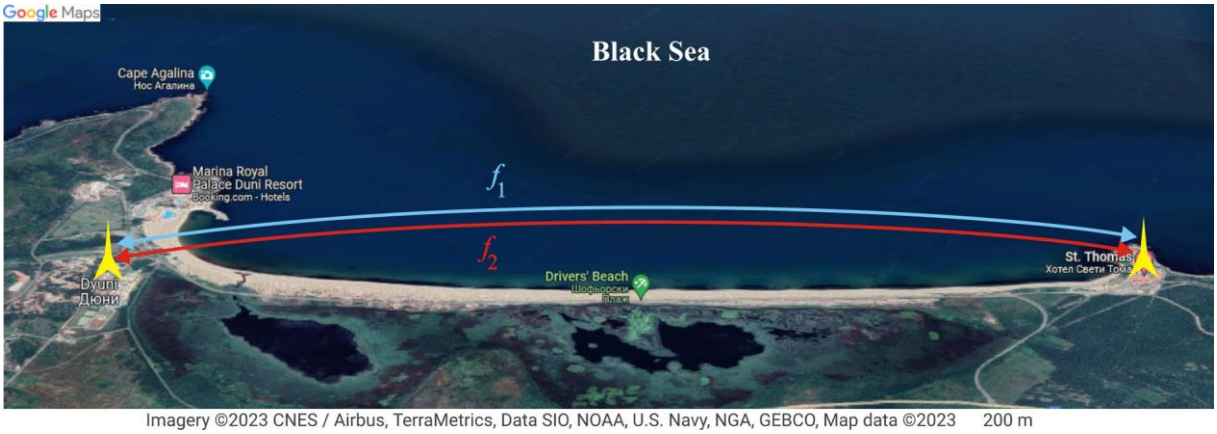


Figure 7. Dyuni – St. Toma [13]

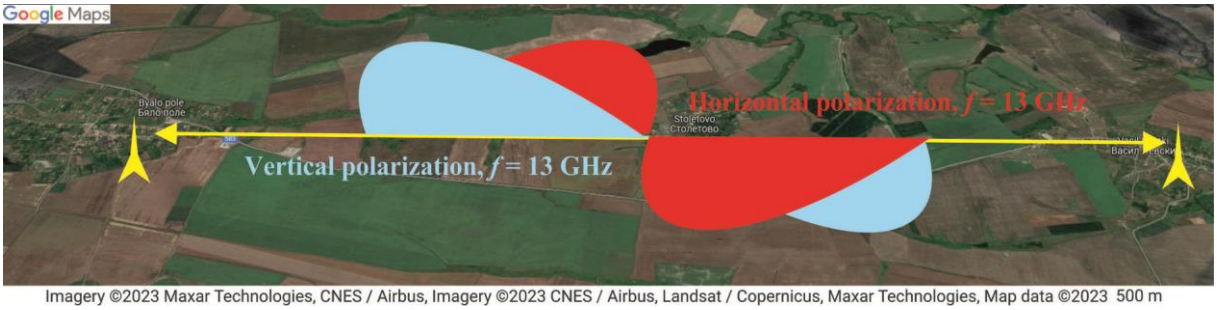


Figure 8. Opan – Vasil Levski [14]

B. Trace route 2: Dyuni – St. Toma

The second route is located on the coast of the Black Sea with geographical coordinates 42.3704473, 27.705076. The track is deployed entirely above sea level. The coast is highly fragmented, with a large part occupied by sand dunes. The region's climate is midway between humid subtropical and Mediterranean, with relatively warm but windy winters (average January temperature +6 °C) and long, warm summers (average July temperature +27 °C). The Black Sea is the main formative factor. The adjacent sea basin causes year-round high air humidity.

C. Trace route 3:Opan – Vasil Levski

The geographical position of the track is 42◦10' north latitude and 25◦23' east longitude, which shows almost equal proximity to the Equator and to the North Pole, which in terms of climate implies moderate values. The average altitude is 157.8 m, with the highest point being 192 m in the land of the village of Vasil Levski. The landscape is characterized as flat and slightly hilly. The route area falls within the transitional-continental climate sub-region of the European-continental climate region. The climate is warm, moderately humid. The area is characterized by good radiation characteristics. The average annual air temperature in the area is +11.4 °C, and all winter average monthly temperatures are positive. The average January temperature is + 1 °C.

In the atmosphere, especially during the winter months, there are significant sources of emissions that pollute the air - an increased concentration of soot in the air.

IV. Main Results and Analysis

In order to troubleshoot the three tracks in question, measurements were taken at short time intervals. The levels of the broadcast signals are fixed on both sides of the track, the aim being to register more clearly the fluctuations in the signal levels. The automatic transmit power control (ATPC) in our investigations is stopped.

Dospat - Sarnitsa is a radio relay route that is of utmost importance for the construction of a reliable and secure mobile network that provides mobile services to small but important regions.

The following solutions were initially used for the construction of the route: dipole antennas, at a frequency of 7GHz with a size of 0.6 meters. Operating in vertical polarization mode. Adaptive modulation, with the minimum modulation being 45Mbit/s- 16QAM and the maximum being 80 Mbit/s- 256QAM. The channel spacing is 14MHz. The broadcasting frequency from Dospat in the direction of Sarnitsa is 7.779750 GHz and the broadcasting frequency from Sarnitsa in the direction of Dospat is 7.928250 GHz. Transmission (Tx) is 20dBm and receive levels of -47dBm are expected under normal conditions. The maximum transmit power is 26dBm and -55dBm far end input power is set for the route target (Fig. 9).

After upgrading the generations of the mobile network from 4G to 5G, a parallel link was deployed at 18 GHz, which works in Frequency Diversity mode. Ultra high performance dual polarized 0.6 meter antennas are used. The size of the antennas is adapted to the length of the route. Adaptive modulation is used, with the minimum modulation being 4QAM - 21Mbit/s and the maximum being 64QAM - 64 Mbit/s. The channel spacing is 14MHz; The broadcast frequency in one direction from Dospat to Sarnitsa is 17.906250 GHz, and the broadcast frequency in the opposite direction from Sarnitsa to Dospat is 18.916250 GHz.

The measured reception levels of RRL Dospat - Sarnitsa are presented in Fig. 9. The radio relay route passes over the Dospat Dam (Fig. 6) and has pronounced fadings, which are characteristic of radio relay routes that pass on the water surface (Fig. 9). Transmit (Tx) is 19dBm and under normal conditions expected receive levels are around -49dBm.

There are two links deployed above the dam. One link operates at 8 GHz, and the other at 18 GHz, in frequency diversity mode, and in practice a separate link is established to carry the same traffic in the presence of fading.

Trace routes Sarnitsa – Dospat and Dyuni – St. Toma are deployed using frequency diversity, and radiolink (RL) Vasil Levski (Tx) – Opan is deployed using XPIC technology.

On trace route 3 Opan - Vasil Levski, data traffic is transmitted at 13 GHz using XPIC technology (Fig. 11). The observed fadings are at the same time for both horizontal and vertical polarization of the signal.

In the diagrams (Fig. 9, Fig. 10 and Fig. 11) distinct fadings are observed. Fadings are observed in the morning when there is an inhomogeneous environment, moisture, vapors, etc. at which the radio waves decay. Most often, this phenomenon is observed in the time range from 3 AM to 7 AM. This signal dropout depends on the carrier frequency.

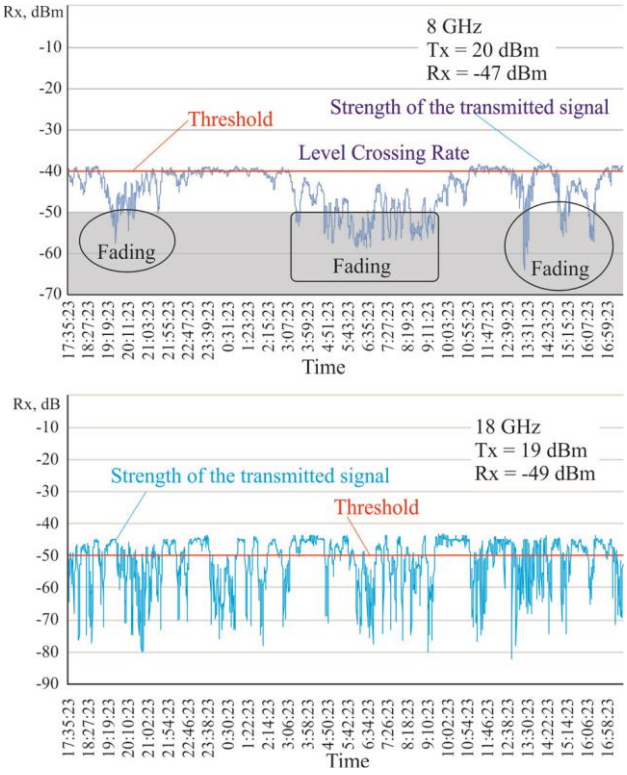


Figure 9. Fading casses of radiolink (RL) Sarnitsa (Tx) – Dospat (Rx)

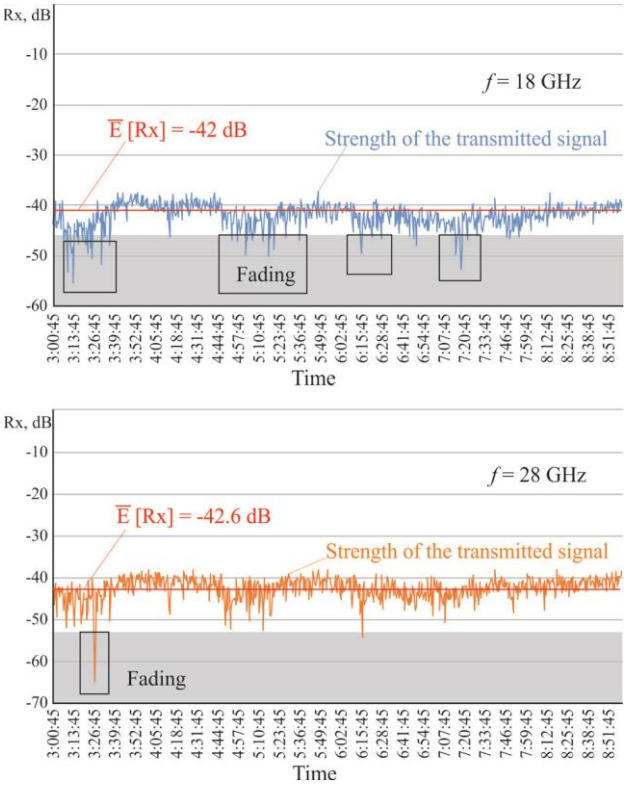


Figure 10. Fading casses of frequency diversity RL Dyuni (Tx) – St. Toma (Rx)

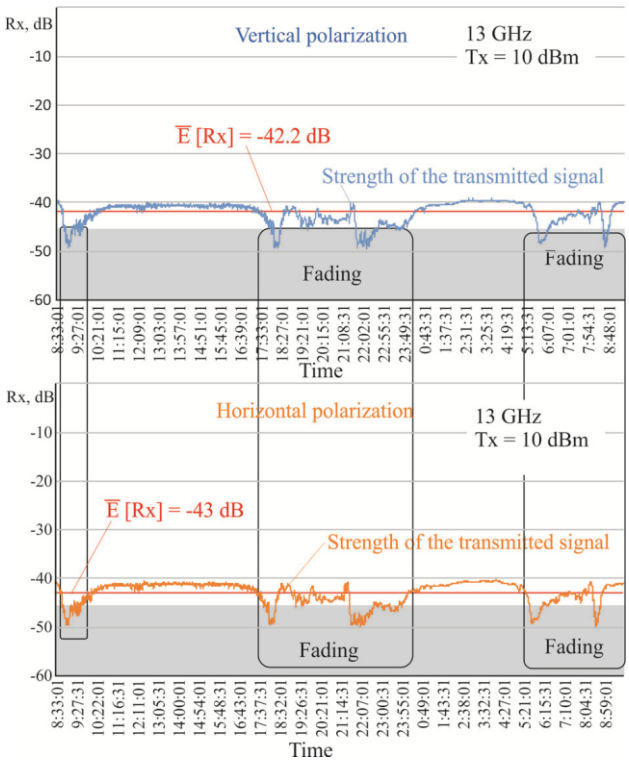


Figure 11. Fading casses of XPIC RL Vasil Levski (Tx) – Opan (Rx)

The difference in the signal drops for the trace routes above water surface (trace route 1 and trace route 2) using frequency diversity is greater compared to the fading for trace route 3 (built on flat terrain). For the 8 GHz traces, the frequency of observed fading is higher than the traces operating at a higher frequency (18 and 28 GHz).

V. Conclusion

Three different radio relay routes for a 5G network, differing in type of terrain and altitude, are considered in the work. By troubleshooting the three RRL routes in question and measuring the signal level in the shortest possible time. Low signal reception levels lead to loss or reduction of path capacity. 5G generation networks aim for 99.99% track availability. For that reason, frequency diversity technology has been applied, which uses the more stable frequency for transmission, which can be different from the one for reception. In such a system, attenuations are a rare event. If each link has a drop/loss probability of 1% or 0.01, the probability of both links dropping at the same time would be 0.01 x 0.01 = 0.0001 or 0.01%.

It would be good to have one of the tracks at a higher frequency, as higher frequencies such as 24, 26, 28, 30, 32 GHz are more resistant to fading. On the other hand, radio frequency spectrum is regulated by national and international regulations to ensure that social and economic benefits are maximized, therefore it is limited.

Acknowledgment

This work is also partially supported by the EC Project 2020-1-PL01-KA226-HE-096196, project title: “Holistic approach towards problem-based ICT education based on international cooperation in pandemic conditions”.

References

[1] Ericsson Mobility Report, November 2022, available at: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2022 [Access on 15 February 2023]

[2] S. Kostadinova, V. Markova, N. Kostov, I. Balabanova and G. Georgiev, "Latency Analysis for 5G Optical Transport Network," 2021 International Conference on Biomedical Innovations and Applications (BIA), Varna, Bulgaria, 2022, pp. 9-12, doi: 10.1109/BIA52594.2022.9831249.

[3] S. Sadinov, P. Kogias, K. Angelov, M. Malamatoudis and A. Aleksandrov, "The Impact of Channel Correlation on the System Performance and Quality of Service in 5G Networks," 2020 7th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE), Ruse, Bulgaria, 2020, pp. 1-4, doi: 10.1109/EEAE49144.2020.9278982.

[4] ETSI EN 302 217-1 V2.0.0, Fixed Radio Systems; Characteristics and requirements for point-to-point equipment and antennas; Part 1: Overview and system-independent common characteristics, 2012-09, available at: https://www.etsi.org/deliver/etsi_en/302200_302299/30221701/02.00.00_20/en_30221701v020000c.pdf [Accessed 10 February 2023]

[5] H.-Yi Chen and D.-P. Lin, "Volume integral equation solution of microwave absorption and scattering by raindrops," IEEE Antennas and Propagation Society International Symposium, Orlando, FL, USA, 1999, pp. 2676-2679 vol.4

[6] M. J. Conway, C. J. Payne, S. G. Bilén and E. N. Koski, "Ground-wave propagation characterization and prediction for HF cognitive radio," 2015 IEEE Military Communications Conference, Tampa, FL, USA, 2015, pp. 1643-1649

[7] T. Iliev, T. Bikov, G. Mihaylov, E. Ivanova, I. Stoyanov, V. Keseev, Algorithms for Congestion Control in LTE Mobile Networks. In: Silhavy, R. (eds) Software Engineering and Algorithms in Intelligent Systems. CSOC2018 2018. Advances in Intelligent Systems and Computing, vol 763. Springer, Cham. https://doi.org/10.1007/978-3-319-91186-1_49.

[8] A. Haka, R. Vasilev, V. Aleksieva and H. Valchanov, "Simulation Framework for Building of 6LoWPAN Network," 16th Conference on Electrical Machines, Drives and Power Systems (ELMA), Varna, Bulgaria, 2019, pp. 1-5, doi: 10.1109/ELMA.2019.8771633.

[9] G. Mihaylov and E. Ivanova, “Analysis and Estimation of the Field Strength of Digital Terrestrial Television Broadcasting”, The Journal of CIEES, vol. 1, no. 1, pp. 17–22, Jun. 2021. DOI:https://doi.org/10.48149/jciees.2021.1.1.3.

[10] H Valchanov, J Edikyan, V Aleksieva , A Study of Wi-Fi Security in City Environment, IOP Conf. Ser.: Mater. Sci. Eng. 618 012031, DOI: 10.1088/1757-899X/618/1/012031

[11] K. Angelov and S. Sadinov, “Modelling and Simulation Analysis of Routing Algorithms in Multichannel Optical Communication Networks”, The Journal of CIEES, vol. 1, no. 1, pp. 29–33, Jun. 2021. DOI:https://doi.org/10.48149/jciees.2021.1.1.5.

[12] Google Maps. "Dopat" Satellite image. 2023, https://www.google.bg/maps/place/4831+%D0%94%D0%BE%D1%81%D0%BF%D0%B0%D1%82/@41.6916978,24.0538651,10703m/data=!3m1!1e3!4m5!3m4!1s0x14ac785d454f4b53:0x85a9948f2f675ac4!8m2!3d41.645973!4d24.1595237?hl=bg&authuser=0. [Accessed 17.02.2023]

[13] Google Maps. "Dyuni" Satellite image. 2023, https://www.google.bg/maps/place/8130+%D0%94%D1%8E%D0%BD%D0%B8/@42.3587088,27.6966507,6672m/data=!3m1!1e3!4m5!3m4!1s0x40a6c892563a9ed3:0x40f61ec479a734d2!8m2!3d42.3704473!4d27.705076?hl=bg&authuser=0. [Accessed 17.02.2023]

[14] Google Maps. "Opan" Satellite image. 2023, https://www.google.bg/maps/place/6078+%D0%9E%D0%BF%D0%B0%D0%BD/@42.1699582,25.6643101,15024m/data=!3m1!1e3!4m5!3m4!1s0x40a808978eb7a279:0x400a01269bf5af0!8m2!3d42.2108979!4d25.6882893?hl=bg&authuser=0. [Accessed 17.02.2023]

Outage Probability Determining for Wireless Systems in the Presence of Beaulieu-Xie Fading and Co-channel Interference Rayleigh Modeled

D. Krstic*, S. Suljovic **, D. S. Gurjar***, and S.Yadav ****
* University of Nis, Faculty of Electronic Engineering, Nis, Serbia
** The Academy of Applied Technical Studies Belgrade, Belgrade, Serbia
*** National Institute of Technology Silchar, Cachar, Assam, India
**** Indian Institute of Information Technology, Allahabad, India
dragana.krstic@elfak.ni.ac.rs

***Abstract* - For realistic assessment of any wireless communication system performance an accurate modeling of this system is necessary. Till now, no many distributions were found to be able to characterize a channel with both, the line of sight (LOS) and the non-LOS (NLOS) components. One of such distribution unifying the non-central chi, generalized Rician and κ – μ distributions into a single fading distribution is Beaulieu-Xie distribution. Beaulieu-Xie fading, as disturbance in wireless system, will be considered in this work. Beside fading, co-channel interference (CCI) modeled by Rayleigh distribution, is present. An analysis of outage probability for selection combining (SC) receiver with multiple branches, used to mitigate these disturbances in the system, will be carried out. The expressions for probability density function (PDF), cumulative distribution function (CDF) and outage probability will be derived. A few graphs will be plotted and based on them an impact of fading and CCI parameters will be highlighted.**

Keywords - Beaulieu-Xie fading; outage probability; Rayleigh co-channel interference; selection combining

I. INTRODUCTION

The fading is the phenomenon of fluctuation in the amplitude of the sent signal [1]. Multipath is the appearance of non-line-of-sight (NLOS) signals at the receiver with a sufficiently short delay that it leads to an undesired distortion of useful signal. Furthermore, it is important to find as accurate a way as possible to describe this phenomenon in wireless systems.

Known distribution as are the Ricean and Nakagami-*m* have benefits as fading models, but also some lacks [2], [3]. For example, lacks are because of reflections and shadowing in the case of Ricean model and both, line-of-sight (LOS) and NLOS components in the event of Nakagami-*m* model. Because of that is necessary to develop new, powerful models in order to constrain the errors induced by fading.

To exceed the gap between theoretical considerations and measurements performed in a fading environment, intensive investigations of new general fading distributions are being carried out. That's why new fading model was introduced by Beaulieu and Xie [4] that overcomes these requirements and simultaneously includes both these models. BX fading model contains advantage of Nakagami-*m* model thanks to its flexible fading parameter and advantage of the Ricean model since its use of a non-central chi-distribution allowing description of both, LOS and NLOS fading channels.

Finally, this is general distribution because it includes some others known distribution as κ − μ distribution, generalized Rician, non-central chi, and a few others included in the said [4]. The BX distribution includes some special cases for different values of *m* and λ. Figure 1.3 in [5] shows the relationship between the BX distribution and other unimodal distributions [6].

It is still important to stand out that BX model can successfully describe the indoor propagation scenario in femtocells as well as high-speed train channels applied in 5GB networks. The statistics of the BX fading model, and later shadowed BX fading model, can be of significant application in land-mobile- satellite (LMS) and unmanned aerial vehicle (UAV) channels, which communicate via LOS signals. From that point many papers deal with this distribution for introduce fading influence.

So, in [7], the effect of randomness provoked by the Beaulieu-Xie fading on the performance of physical layer security is considered and that to the secrecy outage probability (SOP) and average secrecy capacity (ASC). Further, in [8], the second order characteristic of the BX fading channel were obtained, while in [9], [10] the asymptotic tight bound for correlated BX channel and different diversity combining techniques were determined.

The space diversity is very efficient method for reduction fading [1]. In [11], the first order system performance (outage probability (P_{out}), amount of fading (AF), average symbol error probability (ASEP) and channel capacity) for coherent and non-coherent modulation schemes) were derived for the Beaulieu-Xie fading channel with maximum ratio combining (MRC) diversity.

In wireless systems, the co-channel interference (CCI) can occur. CCI appears when a few devices are working on the same frequency channel. Consequently, it also should be introduced in modelling of wireless systems and after that in combating their influence. In paper [12], by determining the outage probability of several diversity models, they are compared analytically for an environment in the presence of Rayleigh fading and limitations from interference.

Even though selection diversity combining is inferior to optimum - maximal ratio combining, and equal gain combining, it is surely simpler for realization and implementation, and also cheaper. With SC combiner with more branches is used, quite good results are achieved in reducing the influence of fading and CCI.

In this work, we will consider Beaulieu-Xie fading, disturbing wireless system in the presence of CCI modeled by Rayleigh distribution. The outage probability for selection combining (SC) receiver with multiple branches will be determined, and the effect of diversity and the fading parameters on the performance measure will be demonstrated.

First section is introduction into new fading model with a review of the first works introducing BX fading. In the second section, system configuration is considered and main performance derived. In the third section obtained graphics are presented and parameters' analysis done. At the end is conclusion with promised future work.

II. Model of Considered System Configuration

We consider L-branch SC receiver in a channel disturbed by BX fading and Rayleigh CCI. The block-model of observed system is shown in Fig. 1. There are L input branches in SC receiver with input signals: x_1, x_2, …, x_n, $n \in \{2,\ldots, L\}$, and output signal is x. The CCI signals appearing at the receiver inputs are denoted by: $y_1, y_2, \ldots, y_n$, while the corresponding output signal is y.

Here, a SIR-based mathematical modeling of SC receiver will be made. The desired signal to CCI ratio (SIR) at each of L input branches in SC receiver is $z_i = x_i / y_i$. By working algorithm of multi-branch SC combiner, the output SIR z will be maximal value of all z_i: $z = \max(z_1, z_2, \ldots, z_i, \ldots, z_L)$.

Each desired signal at the input of SC receiver has the probability density function (PDF) modelled by the BX distribution [4]:

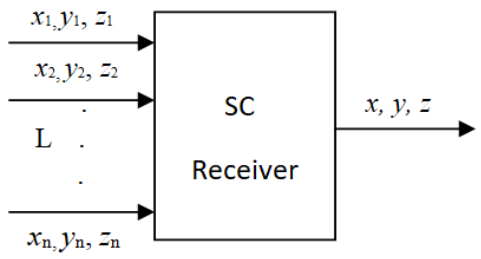


Figure 1. The block-diagram of the L-branch SC receiver

$$p_{X_i}(x_i) = \frac{2m x_i^{\,m}}{\Omega_i \lambda^{m-1}} \mathrm{e}^{-\frac{m}{\Omega}\left(x_i^2+\lambda^2\right)} I_{m-1}\left(\frac{2m\lambda}{\Omega_i} x\right). \tag{1}$$

To represent the PDF of the signal at each of the receivers' inputs as a sum, for the modified Bessel function of the first kind and order ν, $I_\nu(\cdot)$, can be used the shape [13; 17.7.1.1]:

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{\nu+2k}}{k!\Gamma(\nu+k+1)}, \tag{2}$$

where Γ(γ) stands for Gamma function. This form leads to the signal's PDF:

$$p_{X_i}(x_i) = 2\mathrm{e}^{-\frac{m}{\Omega}\left(x_i^2+\lambda^2\right)} \sum_{i=0}^{\infty} \frac{\lambda^{2i} x_i^{\,2i+2m-1}}{i!\Gamma(i+m)} \left(\frac{m}{\Omega_i}\right)^{2i+m}, \tag{3}$$

where m present the fading severity parameter; λ is non-centrality parameter, and Ω_i is the power.

The CCI is modeled by Rayleigh distribution [14]:

$$p_{Y_i}(y_i) = \frac{2y_i}{s_i} e^{-\frac{y_i^2}{s_i}}, y_i \geq 0. \tag{4}$$

In (4), the CCI power is marked by s_i.

Let's derive the PDF of the SIR z_i from [15]:

$$p_{z_i}(z_i) = \int_0^\infty y_i p_{x_i}(z_i y_i) p_{y_i}(y_i) dy_i . \tag{5}$$

By replacement (3) and (4) in (5) and some mathematical manipulations, it is obtained:

$$p_{z_i}(z_i) = 2 \cdot \sum_{i=0}^{\infty} \frac{\lambda^{2i} z_i^{\,2i+2m-1} m^{2i+m} \mathrm{s}_i^{\,i+m} \Gamma(i+m+1)}{i!\Gamma(i+m)\Omega_i^{\,i-1}\left(\Omega + \mathrm{s}_i m z_i^{\,2}\right)^{i+m+1}} \mathrm{e}^{-\frac{m\lambda^2}{\Omega}}. \tag{6}$$

Now, it is necessary to calculate the cumulative distribution function (CDF) of the SIR z_i. Based on formula for CDF [14]:

$$F_{z_i}(z_i) = \int_0^{z_i} p_{z_i}(t) dt, \tag{7}$$

the CDF of z_i will be:

$$F_{z_i}(z_i) = 2\sum_{i=0}^{\infty} \frac{\lambda^{2i} m^{2i+m} \mathrm{s}_i^{\,i+m}}{i!\Omega_i^{\,i-1}} \cdot \frac{\Gamma(i+m+1)}{\Gamma(i+m)\mathrm{e}^{\frac{m\lambda^2}{\Omega}}} \int_0^{z_i} \frac{t^{2i+2m-1}}{\left(\Omega + \mathrm{s}_i m t^2\right)^{i+m+1}} dt . \tag{8}$$

The integral in previous expression can be solved by dint of the incomplete Beta function $B_w(a, b)$ [13] based on [16]:

$$F_{z_i}(z_i)=\sum_{i=0}^{\infty}\frac{\lambda^{2i}\Gamma(i+m+1)}{i!\Gamma(i+m)}\cdot\left(\frac{m}{\Omega_i}\right)^i e^{-\frac{m\lambda^2}{\Omega}} B_{\frac{ms_iz_i^2}{\Omega_i+ms_iz_i^2}}(i+m,1). \tag{9}$$

Now, we should derive outage performance. Outage probability is defined as the moment at which the receiver power value goes under the threshold (where the power value relates to the minimal SIR). At that moment, the receiver is out of the range of base station in cellular communications.

The P_{out} can be simply calculated through PDF if we know PDF of SIR. The outage arises if the signal falls below the CCI power level. The derivation is based on integration over the PDF of SIR, and mathematically is equal to CDF from (7).

On the SC receiver output, P_{out} is [17]:

$$P_{out}(z)=F_{z_i}(z)=\left(F_{z_i}(z_i)\right)^L. \tag{10}$$

For our case, after replacement of (9) into (10), we obtain P_{out} in the form:

$$P_{out}(z)=\left(e^{-\frac{m\lambda^2}{\Omega_i}}\sum_{i=0}^{\infty}\frac{\lambda^{2i}}{i!}\cdot\frac{\Gamma(i+m+1)}{\Gamma(i+m)}\left(\frac{m}{\Omega_i}\right)^i B_{\frac{ms_iz_i^2}{\Omega_i+ms_iz_i^2}}(i+m,1)\right)^L. \tag{11}$$

III. ANALYSIS OF THE SYSTEM PERFORMANCE

To analyze the impact of fading and CCI parameters on the concerned measure, here P_{out}, we have plot some curves in the next Fig. 2. We assumed the minimal correlation between the input branches in the SC receiver. The powers are equal for all curves. The figure is made using Wolfram Mathematica and Origin.

We can conclude from the graph that due to the increase of the parameters m and λ, the outage probability decreases, and the system has better performance.

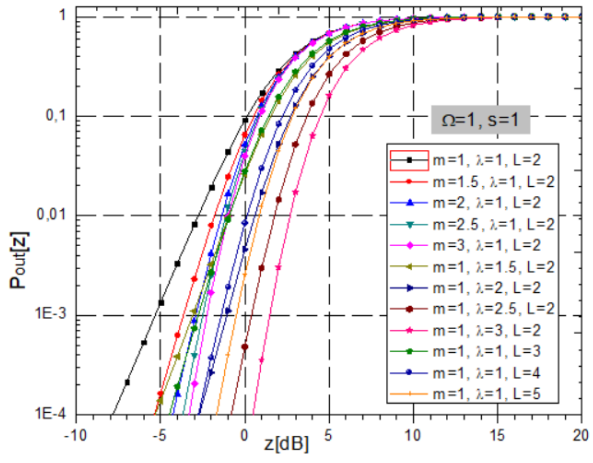


Figure 2. Normalized P_{out} of L-branch SC receiver versus output SIR, with variable values of fading parameters m and λ, and number of branches L

TABLE I. NUMBER OF TERMS THAT SHOULD BE ADDED IN EXPRESSION (11) FOR POUT TO REACH ACCURACY AT 5TH SIGNIFICANT DIGIT

	z=-10 dB	z=0 dB	z=10 dB
m=1, λ=1, L=2	5	5	8
m=1.5, λ=1, L =2	5	6	9
m=2, λ=1, L =2	5	7	11
m=2.5, λ=1, L =2	5	8	12
m=3, λ=1, L =2	5	10	14
m=1, λ=1.5, L =2	5	6	11
m =1, λ=2, L =2	5	8	16
m =1, λ=2.5, L =2	5	9	20
m =1, λ=3, L =2	5	6	24
m =1, λ=1, L =3	5	5	8
m =1, λ=1, L =4	5	5	8
m =1, λ=3, L =5	5	5	9

Also, with an increase of number of input branches in SC receiver, L, the P_{out} decreases what improves system performance. In massive MIMO systems, with a huge number of antennas at the base station, the usage of spatial diversity leads to phenomenon named "channel hardening" [18]. This means that a fading channel behaves as a non-fading channel. But, the fact is that biggest improvement is achieved with enlarging the number of branches from 2 to 3, and the improvement diminishes as L increases, and soon cannot be considered justified.

The number of terms that should be added in expression (11) for P_{out} in order to reach accuracy at 5th significant digit is presented in Table I. The parameters m and λ, as well as the number of branches L are changing, while powers take values: Ω=1, s=1.

From this table one can notice that for small signals, i.e. for z=-10db, the series converges quickly, and for any value of m, λ, or L, it is enough to add only 5 additions in sum in (11). On the other side, for z=0 and z=10 dB, the number of elements in the sum increases and depends on the value of the fading parameters. When the parameter m and parameter λ increase, it is necessary to add a larger number of terms, because the series converges more slowly. The most additions need to be added is 24 for large signals, z=10db, when λ increases to 3.

IV. CONCLUSION

One of very important measure showing system reliability is outage probability. We determined P_{out} for severely degraded transmission conditions influenced by multipath BX fading and Rician CCI in mobile systems. Mitigation of fading influence was made by L-branch SC diversity receiver.

Since the mobile system performs better for the lower P_{out}, based on the analytical results and graphics of P_{out} versus SIR, we have noticed that wireless system works better for bigger number of branches. This means possibility of receiver to chooses the antenna with the biggest SIR to forward to the user, what matches the theory.

On this spot, one can notice that increasing of fading parameter m to infinity means a channel without fading,

and leads to better system performance. Increasing of λ means increasing of useful signal power, what also improve system performance.

ACKNOWLEDGMENT

This work is made within projects of Serbian Ministry of Science, Technological Development and Innovation as well the project "Development of Secure and Spectral Efficient Simultaneous Wireless Information and Power Transfer Systems for Large-Scale Wireless Networks" under Science and Technological Cooperation Projects between the Republic of Serbia and the Republic of India.

REFERENCES

[1] M. K. Simon and M. S. Alouini, Digital Communication over Fading Channels, 2nd ed. Wiley, 2005.

[2] S. O. Rice, "Statistical properties of a sine wave plus random noise," The Bell System Technical Journal, vol. 27, no. 1, pp. 109–157, 1948.

[3] M. Nakagami, "The m-distribution, a general formula of intensity of rapid fading", In William C. Hoffman, editor, Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium, held June 18–20, 1958, pp. 3–36. Pergamon Press., 1960, doi:10.1016/B978-0-08-009306-2.50005-4

[4] N. C. Beaulieu and X. Jiandong, "A novel fading model for channels with multiple dominant specular components", IEEE Wireless Communications Letters, vol. 4, no. 1, pp. 54–57, Feb. 2015. doi:10.1109/lwc.2014.2367501

[5] A. Olutayo A Novel Fading Model for Emerging Wireless Communication Systems. Doctoral dissertation. Okanagan: The University of British Columbia, 2021.

[6] S. Glen, "Unimodal distribution in statistics", From StatisticsHowTo.com: Elementary Statistics for the rest of us! https://www.statisticshowto.com/unimodal-distribution-2/

[7] P. S. Chauhan, S. Kumar, and S. K. Soni, "On the physical layer security over Beaulieu-Xie fading channel", AEU - International Journal of Electronics and Communications, 2019, doi: https://doi.org/10.1016/ j.aeue.2019.152940

[8] A. Olutayo, H. Ma, J. Cheng, J. F. and J. Holzman, "Level crossing rate and average fade duration for the Beaulieu-Xie fading model", IEEE Wireless Communications Letters, 6 (3), pp. 326-329, 2017.

[9] A. Olutayo, J. Cheng, and J. Holzman, "Asymptotically tight performance bounds for equal gain combining over a new correlated fading channel", 15th Canadian Workshop on Information Theory (CWIT), 2017, Quebec City: 1-5.

[10] A. Olutayo, J. F. Cheng, and J. Holzman, "Asymptotically tight performance bounds for selection diversity over Beaulieu-Xie fading channels with arbitrary correlation", IEEE International Conference on Communications (ICC), Paris, pp. 1-6, 2017.

[11] M. K. Rajesh and K. Yadav, "Performance analysis of Beaulieu-Xie fading channel with MRC diversity reception", Trans Emerging Tel Tech. 2020; e3949. P. 13, https://doi.org/10.1002/ett.3949

[12] Y. Song, S. D. Bl, and J. Cheng, "Outage probability comparisons for diversity systems with cochannel interference in Rayleigh fading", IEEE Transactions on Wireless Communications, 4(4), 1279–1284. doi:10.1109/twc.2005.852141

[13] I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series and Products Academic. New York: 1980.

[14] S. Panić, M. Stefanović, J. Anastasov, and P. Spalević, Fading and Interference Mitigation in Wireless Communications. Taylor & Francis Publishing group, CRC Press, New York, USA, 2013.

[15] S. Suljović, D. Milić, S. R. Panić, "LCR of SC receiver output signal over α-κ-μ multipath fading channels", Facta Universitatis Series: Electronics and Energetics Vol. 29, No. 2, pp. 261– 268, June 2016. http://www.doiserbia.nb.rs/img/doi/0353-3670/2016/0353-36701602261S.pdf

[16] S. Suljović, D. Krstić, D. Bandjur, S. Veljković, and M. Stefanović, "Level crossing rate of macro-diversity system in the presence of fading and co-channel interference", Revue Roumaine des Sciences Techniques, Publisher: Romanian Academy, vol. 64, pp. 63–68, 2019.

[17] D. Milić, S. Vasic, N. Petrović, S. Koničanin, S. Suljović, "Outage probability simulation of a MIMO L-branch SC diversity system with k-μ fading and Rayleigh co-channel interference", 20th International Symposium INFOTEH-Jahorina, 17-19 March 2021, doi: 10.1109/INFOTEH51037.2021.9400682

[18] M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," IEEE Transactions on Information Theory, vol. 50, no. 9, pp. 1893–1909, 2004.

CE

International Conference on

COMPUTERS IN EDUCATION

Steering Committee

Chairs: Marina Čičin-Šain, Cybernetics Society, Croatia
Jaak Henno, Tallinn University of Tehnology, Estonia
Hannu Jaakkola, Tampere University, Pori, Finland
Snježana Babić, Polytechnic of Rijeka, Croatia

Members: Michael E. Auer, Carinthia Tech Institute, Villach, Austria
Marjan Krašna, University of Maribor, Slovenia
Nadica Kunštek, Education and Teacher Training Agency, Zagreb, Croatia
Duško Lukač, Siegburg, Germany
Robert Repnik, University of Maribor, Slovenia

Program Committee

Members: Lejla Abazi-Bexheti, SEE University, Tetovo, North Macedonia
Nikola Kadoić, University of Zagreb, Varaždin, Croatia
Ivan Kaštelan, University of Novi Sad, Serbia
Mario Konecki, University of Zagreb, Varaždin, Croatia
Božidar Kovačić, University of Rijeka, Croatia
Gorana Mudronja, University of Rijeka, Croatia
Dana Paľová, Technical University of Košice, Slovakia
Libuša Révészová, Technical University of Košice, Slovakia
Frano Škopljanac-Mačina, University of Zagreb, Croatia
Andreja Špernjak, University of Maribor, Slovenia

PAPERS

Applying TAM (Technology Acceptance Model) to Predict Effective ICT Use of Preschool Teachers during the COVID-19 Pandemic

T. Velki* M. Miočić**
* J.J. Strossmayer University/ Faculty of Education, Osijek, Croatia
tena.velki@gmail.com
** University of Zadar, Croatia

***Abstract* - During the COVID-19 pandemic, educational institutions were challenged to organize online activities, and it was especially challenging to organize them in preschool institutions. Based on the modified TAM (Technology Acceptance Model), the study aimed to examine the potential predictors of effective preschool teachers' use of information-communication technology (ICT) during the COVID-19 pandemic. Additionally, the main effect of previous work experience was tested. The participants were 104 employed preschool teachers. They completed two online questionnaires (Use of ICT and Digital competencies) created for this study's purpose and gave some demographic data. Hierarchical linear regression was performed in order to test TAM, i.e. to predict the effective use of ICT during the COVID-19 pandemic. The results showed that only the positive attitudes and recent experience were significant predictors of effective ICT use during COVID-19 pandemic, explaining overall 38% variances, while users' characteristics did not play a role. Furthermore, preschool teachers with less work experience self-assessed possessing higher levels of specific ICT competencies necessary for work during the COVID-19 pandemic. Previous experiences with ICT and a specific set of ICT competencies were crucial for preschool teachers to successfully transform activities into a virtual environment during the COVID-19 pandemic.**

Keywords – ICT; TAM; preschool teachers; digital competencies; COVID-19 pandemic

I. INTRODUCTION

During COVID-19 emergency remote education has added a different dimension to the change in the use of ICT (information and communication technologies), especially in early childhood education [1-3]. Online learning, which requires the use of ICT, provided an important learning opportunity during the pandemic by offering educational content and materials and allowing young children to meet and interact with their teachers and friends [4-7].

Although increased ICT use was a necessary to provide children with online learning content and materials, negative consequences of overuse of technology in early childhood emerged [4, 7]. Numerous recent studies have shown the same problems for preschool emergency distance education [8-15]. The main problems for preschool educators were the absence of appropriate technological equipment and an internet connection, the lack of technological knowledge, the lack of knowledge to use educational technology and produce digital educational materials, as well as to make good use of the platforms available [8, 16], and limited time of due to their personal and professional obligations or older children who also attended emergency online education [17]. Another major issue derives from the developmental stage of young children. Preschool children in the online learning environment cannot be autonomous learners, leading them to be distracted and uninterested in ICT distance education [18]. In addition, some studies reported that more than half of preschool teachers said they proceeded with distance education without receiving any training at all in using digital technologies [19]. Overall, it was concluded that the education process during the COVID-19 pandemic was inefficient and that teachers have negative views on distance education [20, 21].

A. *Technology Acceptance Model (TAM)*

The Technology Acceptance Model (TAM) and extended TAM have been widely implemented in the field of technology acceptance and used to determine the potential predictors of successful implantation of ICT [22, 23]. TAM is used to explain the relationship between humans and technology through Perceived Usefulness (PU) and Perceived Ease of Use (PEU). TAM was well applied in the field of technology in education with various modifications [24]. TAM includes Perceived Ease of Use and Perceived Usefulness, which are the important determinants of technology acceptance and user behavior. Extended TAM (Figure 1) includes another essential factor that indirectly influences user behavior, i.e., external variables that fundamentally lead to attitudes toward using a particular technology [25].

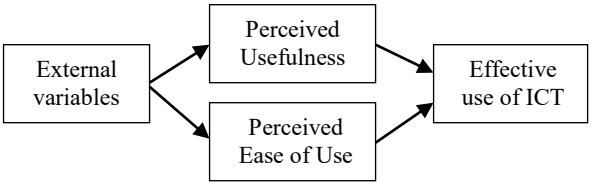


Figure 1. Modified Technology Acceptance Model

B. *Application of TAM on effective ICT use of preschool teachers*

This study evaluates the TAM's main variables for the effective use of ICT during the COVID-19 pandemic, such as Perceived Usefulness (the intention to use, positive attitudes towards ICT use), Perceived Ease of Use (usability and behavioral intention, ICT competencies, experience with ICT) and external variables (age, gender, and specific knowledge).

C. Previous resreach from TAM perspective in preschool education

To a large extent, Perceived Usefulness and Perceived Ease of Use directly influence user behavior (i.e., use of ICT during the COVID-19 pandemic). PU is described as users' attitudes toward ICT that determine the use of it and has proven to be a very important factor for technology adoption in numerous studies [24, 26]. For example, the positive attitudes among users, such as teachers, that use of ICT will make their work or that of their learners easier [27]. Preschool teachers must have a positive attitude toward the use of ICT for online learning and possess adequate skills in a digital environment in order to facilitate and encourage their own behavior, i.e., effective use of ICT in distance education [28]. Teachers' positive views and attitudes toward ICT integration in distance education are vital because they affect their actual use of ICT and therefore affect children's motivation and learning [29-31]. Only a little bit above half of the preschool teachers had positive attitudes toward the educational use of technology [32], and similarly, another study found only fairly positive attitudes toward integrating technology in kindergarten [33]. Another major issue influencing teachers' attitudes toward ICT is the lack of online platforms and digital educational technology optimized for children under six years old [34].

The perceived ease of use of new or existing ICT refers to users, i.e., teachers, view on technology as one that does not require much effort to learn how to use [35], i.e., that using ICT would be effortless and hassle-free [36]. Teachers would be more inclined to use ICT if they consider it easy to learn and use with minimal need for an expert consultation, i.e., if they already possess some specific competencies and experience. Preschool teachers have the most important role in online learning, and their competency in the use of ICT increases the effectiveness of online learning [28]. Some studies have found that preschool teachers did not have a clear understanding of the benefits of young children's technology use [37], and therefore, it is necessary for early childhood teachers to acquire ICT competencies in order to use ICT in distance education [38] effectively.

External factors are described as variables that can explain the procedure people go through in acquiring ICT skills [39]. The most important is the user`s personal characteristics, i.e., age, gender, computer literacy, et cetera. [40]. Previous studies have shown that despite the growth of ICT usage, there is a significant decline among the elderly, lower-income earners than the younger, well-educated, and high-income earners [25]. Age, education, and socio-economic status were associated with beliefs of the importance of ICT, which can influence attitudes towards ICT use and directly influence behavior, i.e., effective use of ICT. Furthermore, gender, age, and technology experience (i.e., specific knowledge) had an influence on the intention to use technology [35]. Men, younger, with more previous computer knowledge, will be more inclined to use ICT professionally and privately. More recent studies have shown that younger preschool teachers have a high perception of ICT usefulness, whereas senior preschool teachers lack knowledge and need courses to help them more confidently manage children's learning activities using ICT during distance education. [41]. For effective use of ICT in online learning in kindergarten, preschool teachers must have the knowledge and skills to use technology and deliver the lessons [42]. Furthermore, one of the most important factors affecting preschool teachers' attitudes was their ICT knowledge [32]. However, preschool teachers pointed out the urgent need to train in the use of technology, especially the use of ICT in educational settings [38, 43].

Although some previous studies have expanded and specified determinants of behavioral intention to use technology, they did not go a step forward to measure the actual effectiveness of ICT use. Moreover, there is a lack of knowledge in the literature about the antecedent variables of education technology's effective use for preschool teachers, especially during the COVID-19 pandemic. The presented paper emphasizes the importance of positive attitudes, familiarity and ease-of-use in the successful implementation of ICT in remote learning.

II. METHOD

A. Study aim

Based on the modified TAM, our study's main aim is to explore preschool teachers' effective ICT during the COVID-19 pandemic (Figure 2). According to the proposed modified TAM, in the prediction of primary school teachers' effective use of ICT during the pandemic COVID-19, three groups of factors are essential: external variables such as users' characteristics that shape experience in educational settings (years of work experience, level of ICT education which is indirectly measured by preschool teachers' need for additional ICT education i.e. preschool teachers with lower level of ICT education would self-assess necessary additional ICT education), attitudes which are a measure of perceived usefulness of ICT (in our study attitudes toward the usefulness of ICT and importance of ICT competencies) and previous and recent experience with ICT, i.e., the measure of ease of use (in our study previous use of ICT during the regular educational process and possession of specific ICT competencies necessary for work during the COVID-19 pandemic).

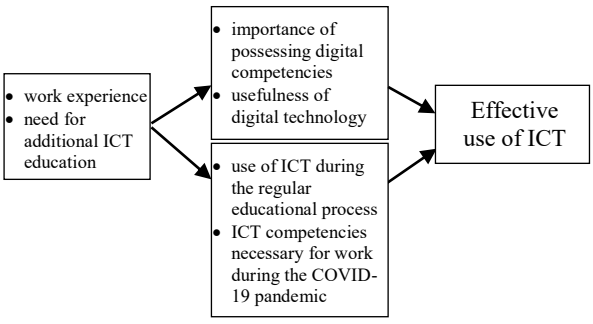


Figure 2. Tested TAM on preschool teachers effective ICT use

According to modified TAM, to predict the effective preschool teachers' use of ICT during the COVID-19 pandemic, work experience and the need for additional ICT education were introduced in the first step of the analysis. It was assumed that preschool teachers with a few years of work experience (due to inexperience) and with many years of work experience (due to their age) and preschool teachers with a higher self-evaluated need for additional ICT education (i.e., with a lack of ICT knowledge) would be less effective in the use of ICT during the COVID-19 pandemic. These variables present the antecedent that shapes users' attitudes and,

consequently, their ICT behavior; therefore, we introduce them in the first step of analysis as control variables. In the next step, the usefulness of digital technology and the importance of possessing digital competencies were introduced as measures of users' attitudes, i.e., indirect influence on ICT behavior. It was assumed that preschool teachers with more positive attitudes toward the importance and usefulness of digital technology would be more successful in using ICT during the COVID-19 pandemic. Finally, in the last step, the use of ICT during the regular educational process (use of ICT in previous work) and ICT competencies necessary for work during the COVID-19 pandemic were introduced as direct measures of effective preschool teachers' use of ICT during the COVID-19 pandemic. As proposed in the modified TAM, individuals who already used ICT during the regular educational process and who possess specific ICT competencies necessary for work during the COVID-19 pandemic would more easily and effectively use ICT during the COVID-19 pandemic.

B. Participants

The study involved 104 employed preschool teachers (95% of women) with the most frequent age span between 36 and 50 years old. All participants were from Croatia. Most of them had between 11 and 20 years of work experience (39.4%), less between 1 and 10 years (34.6%), and the least of them had more than 20 years of work experience (26%).

C. Procedure

The research was conducted during the school year. Principals and professional associates were asked to include preschool teachers in research and to forward an online survey to them. Participants were asked to fill out two online questionnaires (Use of ICT Questionnaire and Digital competencies Questionnaire) and to give some socio-demographic data (gender, age group, work experience).

D. Instruments

Two questionnaires (Use of ICT and Digital competencies) were created for this study's purpose.

Use of ICT Questionnaire (ICT) consisted of 20 items measuring the use of information-communication technology during the COVID-19 pandemic in educational work. On a 5-point Likert scale, participants must self-evaluate their use of ICT. Confirmatory factor analysis (KMO=0.788, Bartlett's Test=1229.45, p=0.000) shows the existence of three factors, i.e., subscales, namely: use of ICT during the pandemic COVID-19 (k=5, α=0.67; e.g., *During the COVID-19 pandemic (using modern technology) I sent recorded activities to parents as ideas, which they could carry out with their children at home*); use of ICT during the regular educational process (k=8, α=0.86; e.g., *During the regular educational process, I use digital technology in my work with children and parents*) and the ICT competencies necessary for work during the COVID-19 pandemic (k=7; α=0.93; e.g., *We received instructions from the authorities on how to conduct the online educational process during the pandemic COVID-19*). All three subscales had satisfactory internal reliability (Cronbach α greater than 0.66) and Eigen values greater than 2, which explained an overall of 58.51% (each of them more than 10% of variance).

The Digital competencies Questionnaire (DC) consisted of 10 items measuring preschool teachers' attitudes toward digital competencies necessary for educational work. On a 5-point Likert scale, participants must self-evaluate their attitudes about digital competencies. Confirmatory factor analysis (KMO=0.705, Bartlett's Test=627.72, p=0.000) shows the existence of three factors, i.e., subscales, namely: the importance of possessing digital competencies (k=5, α=0.83; e.g., *My digital competences are sufficient for daily educational work with children*); the usefulness of digital technology (k=2, α=0.89; e.g., *Digital technology is very useful for the implementation of educational practice in kindergarten*) and the need for additional ICT education (k=3; α=0.72; e.g., *You need additional education in the development of digital competencies necessary for everyday educational work*). All three subscales had satisfactory internal reliability (Cronbach α greater than 0.66), and Eigen values greater than 1, which explained an overall of 73.60% (each of them more than 10% of variance).

TABLE I. DESCRIPTIVE STATISTICS FOR SUBSCALES OF ICT AND DC

Subscales	Min	Max	M	SD	Skewness	Kurtosis
effective use of ICT during COVID-19 pandemic	1.00	5.00	2.99	0.77	-0.11	0.28
use of ICT during the regular educational process	1.00	5.00	3.64	0.74	-0.49	0.31
ICT competencies necessary for work during the COVID-19 pandemic	1.00	5.00	3.27	0.99	-0.20	-0.29
importance of possessing digital competences	2.80	5.00	3.91	0.69	0.01	-1.23
usefulness of digital technology	1.00	5.00	3.83	0.88	-0.88	1.44
need for additional ICT education	1.00	5.00	3.31	0.86	-0.19	0.28

Legend: Min – minimal value; Max - maximal value; M – mean; SD – standard deviation; Skewness & Kurtosis - asymmetry indices

III. RESULTS AND DISCUSSION

Before selecting the appropriate statistical procedures, pre-analysis was done (Table 1). It was concluded that inferential statistics could be applied based on descriptive indicators (asymmetry indices do not exceed value +/- 2).

Mean values of use of ICT during the COVID-19 pandemic indicate that participants self-evaluated only the average level of effective use of information-communication technology during the COVID-19 pandemic in their educational work and also the average level of their ICT competencies necessary for work during the COVID-19 pandemic. These results confirmed the problem, same as worldwide, with emergency distance learning in preschool education, lack of ICT competencies,

and specific knowledge in order to effectively use ICT in distance education [8-16, 37, 38]. However, their use of ICT during the regular educational process was slightly above average. Furthermore, their attitudes toward digital competencies were mostly positive, realizing the importance and usefulness of possessing digital competencies. After two years of distance education, preschool teachers have realized how important it is to have ICT competencies. However, that does not mean their overall view of distance education is positive, rather negative, but necessary in specific conditions such as the COVID-19 pandemic, as shown in previous studies [20, 21]. Nevertheless, preschool teachers self-evaluated only the average need for additional ICT education, which is in contrast with some previous study reports about preschool teachers' emergency distance education without any training at all in the use of digital technologies [19].

TABLE II. PREDICTING EFFECTIVE PRESCHOOL TEACHERS' USE OF ICT DURING THE COVID-19 PANDEMIC.

predictors	B	SE	β	p
1. step: external variables (preschool teachers characteristics)				
work experience	-0.042	0.133	-0.033	0.751
need for additional ICT education	0.161	0.119	0.139	0.179
First regression model	$R = 0.137$; $R^2 = 0.019$; $F_{(2,96)} = 0.92$; $p > 0.01$ (n.s.)			
work experience	-0.021	0.122	-0.016	0.863
need for additional ICT education	0.133	0.012	0.115	0.236
2. step: Perceived Usefulness (positive attitudes)				
importance of possessing digital competencies	0.682	0.172	0.449	0.000
usefulness of digital technology	-0.048	0.135	-0.041	0.722
Second regression model	$R = 0.447$; $R^2 = 0.200$; $\Delta R^2 = 0.181$; $F_{(2,94)} = 10.62$; $p < 0.001$			
work experience	0.119	0.139	0.091	0.393
need for additional ICT education	0.099	0.101	0.085	0.330
importance of possessing digital competencies	0.101	0.195	0.066	0.608
usefulness of digital technology	-0.053	0.120	-0.046	0.660
3. step: Perceived Ease of Use (experience in effortless use of ICT)				
use of ICT during the regular educational process	0.322	0.150	0.241	0.035
ICT competencies necessary for work during the COVID-19 pandemic	0.632	0.163	0.462	0.000
Final regression model	$R = 0.613$; $R^2 = 0.376$; $\Delta R^2 = 0.177$; $F_{(2,92)} = 13.03$; $p < 0.001$			

Legend: B – unstandardized regression coefficients; SE – standardized error for unstandardized coefficients; β – standardized regression coefficients; R – explanation of variance; F – test for ANOVA (significance of regression model)

In order to test modified TAM the hierarchical regression analysis with 3 steps was performed (Table 2). The results show that the selected predictors explain only a small proportion of the preschool teachers' successful use of ICT during the COVID-19 pandemic (about 38%). The characteristics of preschool teachers (work experience and need for additional ICT education) did not at all predict their effective use of ICT during the COVID-19 pandemic. These external variables have been shown to have an important influence on individuals' attitudes and behavior in previous studies [35, 39, 40]; however, they did not influence preschool teachers' effective use of ICT during the COVID-19 pandemic. One possible explanation is that preschool teachers self-evaluated only the average need for additional ICT education and not a lack of specific knowledge and ICT training like the ones in previous studies [38, 42, 43]. Results of ANOVA (Table 3) confirmed that years of work experience did not influence the need for additional ICT education, i.e., regardless of years of work experience, all preschool teachers estimated that they do not need urgent training or additional education in the use of ICT.

However, the perceived usefulness, i.e., preschool teachers' attitudes toward digital competencies, had significantly predicted their use of ICT during the COVID-19 pandemic. Higher self-evaluated levels of importance of possessing digital competencies predicted more effective use of ICT during the COVID-19 pandemic (explained about 18% of variance). In contrast, attitude about the usefulness of digital technology was not a significant predictor. Understanding how important digital competencies are for distance learning is a basic prerequisite for the effective use of ICT. As in previous studies, a positive attitude toward the use of ICT has proven to be significant in predicting preschool teachers' effective use of ICT [28-31]. However, it seems that attitude about usefulness of digital technology was not important for preschool teachers' effective use of ICT. It is possible that preschool teachers do not believe that ICT is much helpful in their everyday work (only necessary during the period of COVID-19 pandemic) due to the developmental stage of young children and the lack of online platforms and digital educational technology optimized for children under six years old [18, 34].

Preschool teachers' effective use of ICT during the COVID-19 pandemic was best predicted by the perceived ease of use. An additional 18% variance was explained, whereas both predictors (use of ICT during the regular educational process and ICT competencies necessary for work during the COVID-19 pandemic) were significant. Preschool teachers who self-evaluated higher levels of previous use of ICT during the regular educational process (specific prior knowledge) and with a higher level of ICT competencies necessary for work during the COVID-19 pandemic used ICT during the COVID-19 pandemic more effectively. These results confirmed our expectations with modified TAM. Previous experience with ICT and specific ICT pandemic make using ICT effortless and hassle-free for preschool teachers, which has proven to be crucial for the effective use of ICT [35, 36].

One-way ANOVA (Table 3) was performed to check for the main effects of previous work experience on ICT use and digital competencies of preschool teachers. Previous work experience was also an indirect measure of participants' age, i.e., older preschool teachers had more than 20 years of work experience and were more than 50 years old. Previous work experience had shown only a significant effect on ICT competencies necessary for work during the COVID-19 pandemic; for other subscales of the Use of ICT Questionnaire and Digital Competencies Questionnaire, there were no statistically significant differences. Post hoc analysis using Gabriel's correction (for non-equal group size) showed that preschool teachers with the least work experience self-evaluated the highest

level of ICT competencies necessary for work during the COVID-19 pandemic. In contrast, preschool teachers with the most work experience self-evaluated the lowest level of ICT competencies necessary for work during the COVID-19 pandemic which could be the effect of their age and practical experience in everyday life with ICT. These results were in accordance with previous studies, which have shown a lack of knowledge and ICT competencies in senior preschool teachers [25, 35, 41].

TABLE III. RESULTS OF ONE-WAY ANOVA: TESTING MAIN EFFECTS OF WORK EXPERIENCE OF ICT AND DIGITAL COMPETENCIES

subscales	Work experience	N	M	SD	$F_{(2,102)}$
effective use of ICT during the COVID-19 pandemic	1-10 years	36	3.23	0.93	F = 0.241 p = 0.786
	11-20 years	41	3.36	0.98	
	> 20 years	27	3.21	1.13	
use of ICT during the regular educational process	1-10 years	36	3.54	0.73	F = 0.532 p = 0.589
	11-20 years	41	3.71	0.67	
	> 20 years	27	3.68	0.87	
ICT competencies necessary for work during the COVID-19 pandemic	1-10 years	34	3.52	0.58	F = 24.10 p = 0.000
	11-20 years	38	2.93	0.66	
	> 20 years	27	2.40	0.66	
importance of possessing digital competences	1-10 years	36	3.87	0.72	F = 0.273 p = 0.762
	11-20 years	41	3.98	0.71	
	> 20 years	27	3.87	0.63	
need for additional ICT education	1-10 years	36	3.14	0.91	F = 1.62 p = 0.202
	11-20 years	41	3.31	0.85	
	> 20 years	27	3.53	0.77	
usefulness of digital technology	1-10 years	36	3.75	0.71	F = 0.372 p = 0.690
	11-20 years	41	3.83	0.95	
	> 20 years	27	3.94	0.99	

IV. CONCLUSION

The results of the research are in accordance with previous worldwide studies about emergency distance learning, i.e., the use of ICT in early education [4-7]. The proposed modified TAM was partially confirmed:

- Extended TAM, which included external variables such as users' characteristics that shape experience in educational settings, was not confirmed in our study (years of work experience and need for additional ICT education were not significant predictors of effective use of ICT during the COVID-19 pandemic);
- Basic TAM, with main variables that predict effective use of ICT during the COVID-19 pandemic, i.e., Perceived Usefulness and Perceived Ease of Use, was confirmed;
- Perceived Usefulness as a measure of users' attitudes (i.e., the measure of indirect influence on ICT behavior) has shown to be a significant predictor (positive attitude about the importance of possessing digital competencies predicted more effective use of ICT during the COVID-19 pandemic);
- Perceived Ease of Use which measures recent experience during the COVID-19 pandemic with ICT use and represents a direct measure of effective preschool teachers' use of ICT during the COVID-19 pandemic, has shown to be the most important and significant predictor (use of ICT during the regular educational process and possession of ICT competencies necessary for work during the COVID-19 pandemic predicted more effective use of ICT during COVID-19 pandemic);
- At last, it was shown the main effect of previous years of work experience on ICT competencies necessary for work during the COVID-19 pandemic as was found in recent studies (preschool teachers with the least work experience self-evaluated the highest level of ICT competencies necessary for work during the COVID-19 pandemic, while preschool teachers with the most work experience self-evaluated the lowest level of ICT competencies necessary for work during the COVID-19 pandemic).

The methodological limitations of the research are reflected in the preschool teachers' subjectivity in the self-assessment of their digital competencies and ICT knowledge. These limitations may be overcome in future research with objective assessments by the users of the educational service (e.g., young children and their parents), objective tests of ICT knowledge, or observations by a co-teacher as a critical friend, i.e., by the researcher applying a non-participatory monitoring methodology.

Earlier findings are confirmed in the obtained study, further emphasizing the importance of giving ICT skills to preschool teachers, especially in conditions of a CO0VID-19 pandemic. After all, it can be concluded that previous experiences with ICT in everyday work and a specific set of ICT competencies were crucial for effective emergency distance education of preschool teachers, i.e., for the successful transformation of activities into a virtual environment during the COVID-19 pandemic.

REFERENCES

[1] H. Alomyan and A. Alelaimat, A. "Employing ICTs in kindergartens in remote areas of Jordan: Teachers' perspectives on uses, importance and challenges", European Journal of Educational Research, Vol. 10, No. 4, 2021, pp. 2145–2157.

[2] A. Crawford, K. A. Vaughn, C. L. Guttentag, C. Varghese, Y. Oh and T. A. Zucker "Doing hat I can, but I got no magic wand: a snapshot of early childhood educator experiences and efforts to ensure quality during the COVID-19 pandemic", Early Childhood Education Journal, Vol. 49, No. 5, 2021, pp. 829–840.

[3] T. G. Ford, K. A. Kwon and J. D. Tsotsoros "Early childhood distance learning in the US during the COVID pandemic: Challenges and opportunities", Children and Youth Services Review, No. 131, 2021, pp. 106-297.

[4] M. J. A. Dias, M. Almodóvar, J. T. Atiles, A. C. Vargas and I. M. Zúñiga León "Rising to the Challenge: Innovative early childhood teachers adapt to the COVID-19 era", Childhood Education, Vol. 96, No. 6, 2020, pp. 38–45.

[5] C. Dong, S, Cao and H. Li "Young children's online learning during COVID-19 pandemic: Chinese parents' beliefs and attitudes", Children and Youth Services Review, 118, 2020, article 105440.

[6] M. Gayatri "The implementation of early childhood education in the time of Covıd-19 pandemic: A systematic review", Humanities & Social Sciences Reviews, Vol. 8, No. 6, 2020, pp. 46–54.

[7] M. Nagasawa and K. Tarrant. Who will care for the early care and education workforce? COVID-19 and the Need to Support Early

Childhood Educators' Emotional Well-being. New York Early Childhood Professional Development Institute; NY, USA: 2020.

[8] J.T. Atiles, M. Almodovar, A. Chavarria Vargas, M. J. Dias and I. M. Zuniga Leon "International responses to COVID-19: Challenges faced by early childhood professional", European Early Childhood Education Research Journal, Vol. 29, No.1, 2021, pp. 66–78.

[9] H. C. Dayal and L. Tiko "When are we going to have the real school? A case study of early childhood education and care teachers' experiences surrounding education during the COVID-19 pandemic", Australasian Journal of Early Childhood, Vol. 45, No.4, 2020, pp. 336–347.

[10] X. Hong, M. Zhang, Q. Liu, B. Zhong, and H. Li "Preschool education demand in China by 2035: A prediction analysis", Early Education and Development, Vol. 33, No. 3, 2022, pp. 391–406.

[11] E. Munastiwi "Colorful online learning problem of early childhood education during the COVID-19 pandemic", Al-Ta Lim Journal, Vol. 27, No. 3, 2020, pp. 227–235.

[12] A. Railienė, O. Merfeldaite, and R. Prakapas "COVID-19 lessons: Experience in organization of distance preschool education", Digital Education Review, Vol. 40, 2021, pp. 141–153.

[13] J. Szente "Live virtual sessions with toddlers and preschoolers amid COVID-19: Implications for early childhood teacher education", Journal of Technology and Teacher Education, Vol. 28, No.2, 2020, pp. 373–380.

[14] C.S. Ugwuanyi and C. I. Okeke "Determınants of preschool teachers' competence ın the use of onlıne teaching space durıng and after the covıd-19 pandemıc", Proceedings of EDULEARN21 Conference, Vol. 5, article 6, July 2021.

[15] Yildiz, S., Kilic, G. N., & Acar, I. H. (2022). Early childhood education during the covıd-19 outbreak: the perceived changing roles of preschool administrators, teachers, and parents. Early Childhood Education Journal, 2022, 1–11.

[16] P. Foti "Research in distance learning in Greek kindergarten schools during the pandemic of COVID-19: Possibilities, dilemmas, limitations", European Journal of Open Education and E-Learning Studies, Vol. 5, No. 1, 2020, pp. 19-40.

[17] N. Bigras, L. Lemay, J. Lehrer, A. Charron, S. Duval, C. Robert-Mazaye and I. Laurin "Early Childhood Educators' Perceptions of Their Emotional State, Relationships with Parents, Challenges, and Opportunities During the Early Stage of the Pandemic", Early Childhood Education Journal, Vol. 49, No. 5, 2021, pp. 775–787.

[18] A. Kruszewska, S. Nazaruk and K. Szewczyk "Polish teachers of early education in the face of distance learning during the COVID-19 pandemic – the difficulties experienced and suggestions for the future", Education, Vol. 50, No 3., 2020, pp. 304-315.

[19] D. Bassok, M. Michie, D. M. Cubides-Mateus, J. B. Doromal and S. Kiscaden "The divergent experiences of early educators in schools and child care centers during COVID-19: Findings from Virginia" [Panel presented 2020 at the 42nd Annual Association of Public Policy Analysis and Management]

[20] N. Brown, K. Te Riele, B. Shelley, and J. Woodroffe. Learning at home during COVID-19:Effects on vulnerable young Australians. Hobart: University of Tasmania, Peter Underwood Centre for Educational Attainment, Australia: 2020.

[21] V. G. Gayvoronskiy "The teachers' opınıon on effıcıency of dıstance educatıon", Bulletin of Nizhnevartovsk State University, Vol. 3, 2020, pp. 11–17.

[22] Y. J. Joo, S. Park and E. Lim "Factors influencing preservice teachers' intention to use technology: TPACK, teacher self-efficacy, and technology acceptance model", Educational Technology & Society, Vol. 21, 2018, pp. 48–59.

[23] H. Rafique, A. O. Almagrabi, A. Shamim, F. Anwar and A. K. Bashir "Investigating the acceptance of mobile library applications with an extended technology acceptance model (TAM)", Computer and Education, Vol. 145, No. c, 2020, article 103732.

[24] R. Bagozzi "The legacy of the technology acceptance model and proposal for a paradigm shift", Journal of the Association of Information Systems, Vol. 8, No. 4, 2007, pp. 244–254.

[25] O. O. Durodolu "Technology Acceptance Model as a predictor of using information system' to acquire information literacy skills", Library Philosophy and Practice (e-journal), 2016, article 1450.

[26] R. K. Yeh and T. C. Teng "Extended conceptualisation of perceived usefulness: empirical test in the context of information system use continuance", Behaviour & Information Technology, Vol. 31, No. 5, 2012, pp. 1097-1115.

[27] S. M. Muinde and P. Mbataru "Determinants of implementation of public sector projects in Kenya: a case of laptop project in public primary schools in Kangundo sub-County, Machakos County", International Academic Journal of Law and Society, Vol. 1, No. 2, 2019, pp. 328–352.

[28] J. Kim "Learning and teaching online during COVID-19: Experiences of student teachers in an early childhood education practicum", International Journal of Early Childhood, Vol. 52, No. 2, 2020, pp. 145-158.

[29] N. Kara and K. Cagiltay "In-service Preschool Teachers' Thoughts about Technology and Technology Use in Early Educational Settings", Contemporary Educational Technology, Vol. 8, No. 2, 2017, pp. 119-141.

[30] P. Mertala "Teachers' beliefs about technology integration in early childhood education: A meta-ethnographical synthesis of qualitative research", Computers in Human Behavior, Vol. 101, 2020, pp. 334–349.

[31] H. P. Miranda and M. Russell "Understanding factors associated with teacher directed student use of technology in elementary classrooms: A structural equation modeling approach", British Journal of Educational Technology, Vol. 43, 2021, pp. 652–666.

[32] N. Zaranis, V. Oikonomidis, and M. Linardakis, M. "Factors Affecting Greek Kindergarten Teachers to Support or Oppose ICT in Education", in Research on e-Learning and ICT in Education: Technological, Pedagogical and Instructional Perspectives, P. Anastasiades and N. Zaranis, Eds. Berlin: Heidelberg Springer 2017, pp. 203-2015.

[33] K. Nikolopoulou, V. Gialamas, K. Lavidas and V. Komis "Teachers' Readiness to Adopt Mobile Learning in Classrooms: A Study in Greece", Technology, Knowledge and Learning, Vol. 26, No. 1, 2021, pp. 53–77.

[34] C. Dreifuss-Serrano and P. C. Herrera "Early childhood online education in the COVID-19 context. Behavioral Patterns for User Interface Design" Proceedings of 2020 IEEE Learning With MOOCS (LWMOOCS), pp. 90-95, October 2020.

[35] V. Venkatesh, M. G. Morris, G. B. Davis and F. D. Davis "User acceptance of information technology: toward a unified view", Management Information Systems Quarterly, Vol. 27, 2003, pp. 425–478.

[36] D. Zhu, C. T Linb and Y. Hsu "Using the technology acceptance model to evaluate user attitude and intention of use for online games", Total Quality Management, Vol. 23, No. 8, 2012, pp. 965-980.

[37] M. Alkhawaldeh, M. Hyassat, M., E. Al-Zboon and J. Ahmad "The role of computer technology in supporting children's learning in Jordanian early years education", Journal of Research in Childhood Education, Vol. 31, No. 3, 2017, pp. 419- 429.

[38] Ü. Alan "Distance Education During the COVID-19 Pandemic in Turkey: Identifying the Needs of Early Childhood Educators", Early Childhood Education Journal, Vol. 49, No. 5, 2021, pp. 987-994.

[39] S.A. Winarto. Analysis Effect of External Variables on System Usage and User Satisfaction Using Technology Acceptance Model (Empirical Study on Bank Pekreditan Rakyat in Semarang City Region). Undergraduate thesis, Faculty of Economics, Diponegoro University Semarang, 2011.

[40] S. J. Yousafzai, G. R. Foxall and J. G. Pallister "Technology acceptance: a meta-analysis of the TAM", Journal of modelling in management, Vol. 2, No. 3, 2007, pp. 251–280.

[41] J. L. Woei, A.. Ali, M. Abdul Malik and N. Abdul Malek "The Private Preschool Teachers' Perception of ICT Integration Usefulness in Teaching and Learning During MCO", Proceedings International Conference on Education (ICE 2021, virtual conference), pp. 330-337, September 2021.

[42] M. Gayatri "The implementation of early childhood education in the time of Covıd-19 pandemic: A systematic review", Humanities & Social Sciences Reviews, Vol. 8, No. 6, 2020, pp. 46–54.

[43] K. Timmons, A. Cooper, E. Bozek and H. Braund "The Impacts of COVID-19 on Early Childhood Education: Capturing the Unique Challenges Associated with Remote Teaching and Learning in K-2", Early Childhood Education Journal, Vol. 49, No. 5, 2021, pp. 887-901.

Data Acquisition and Corpus Creation for Phishing Detection

I. Dunđer*, S. Seljan* and M. Odak**
* Faculty of Humanities and Social Sciences, University of Zagreb, Department of Information and Communication Sciences, Zagreb, Croatia
** Faculty of Humanities and Social Sciences, University of Mostar, Department of Information Sciences, Mostar, Bosnia and Herzegovina
ivandunder@gmail.com, sanja.seljan@ffzg.hr, marko.odak@sum.ba

***Abstract* - Detecting phishing attacks is not straightforward, since there are many obstacles that derive from language complexity and technical aspects. Studying phishing attacks and other related issues heavily relies on computer datasets, i.e. digital corpora that reflect these linguistic and technical intricacies. Diverse studies using phishing datasets have been performed, but mainly for the English language. Research for other languages is scarce, and especially for not widely spoken languages. For the Croatian language there is an evident lack of corpora that are essential for diverse analyses and for constructing models that are capable of recognizing phishing attacks and protecting users. These datasets are necessary for natural language processing and building machine learning workflows, where results largely depend on corpora that must be specifically crafted for this purpose. Therefore, creating high-quality domain-specific corpora is of great importance in the domain of information security. Such corpora can be employed for teaching purposes in various courses in higher education, and could be analyzed in numerous ways in order to understand the underlying principles of phishing attack strategies. The aim of this paper is to demonstrate the entire process of data acquisition and corpus creation for the phishing detection domain. In addition, an analysis of the corpus is presented with regard to different aspects, such as descriptive attributes, terminology characteristics, metadata and language.**

Keywords* - *data acquisition; digital corpus creation; computational data analysis; natural language processing; phishing; information privacy; information security

I. INTRODUCTION

A phishing attack is a type of social engineering in which an attacker poses as a reliable entity in order to deceive a victim into disclosing sensitive information [1], including login passwords, personal data or financial information [2]. Phishing detection, on the other hand, is the effort to recognize and stop phishing attacks on people and business enterprises, and as such represents a crucial endeavor in the field of information and cyber security [3].

Nowadays, phishing attacks can be assessed in a variety of ways, as the idea is to gauge how well phishing attacks and phishing detection methods work [4]. However, detecting phishing attacks is not an easy task, since there are many difficulties that arise not only from linguistic complexity and diversity, but also from technical and technological aspects [5].

In the process of analyzing and evaluating the effectiveness of phishing attacks and phishing countermeasures, computer datasets that contain a large number of phishing messages play a significant role [6]. These datasets, also known as digital corpora, are purposefully crafted and prepared in order to account for these linguistic and technical nuances.

The importance of digital corpora in the fields of natural language processing (NLP), data science, artificial intelligence (AI) and especially machine learning (ML), cannot be overstated, as they enable computers (machines) to comprehend, interpret and manipulate data, but also to interact with human languages. All these fields offer various techniques and methods that are suitable for detecting phishing attacks [7].

Phishing datasets include electronic messages, such as e-mails that have been previously marked as malicious or suspicious [8]. They can be fed into security systems that use machine learning so that, based on prior observations, they can learn to differentiate between valid and illegitimate messages [9]. In that way, digital corpora help to determine the level of effectiveness of phishing detection technologies and approaches, which can then be utilized to implement any necessary adjustments in security systems.

II. MOTIVATION

Timely analysis and comprehension of incoming data, the discovery of trends and changes in data content and communication volume can support the development of security campaigns in business enterprises.

Processing enormous amounts of data from many sources, such as e-mails, social media platforms, internet portals, product reviews, web feeds or news articles, is one of the main advantages of computers and modern technology, such as natural language processing and machine learning [10].

Digital corpora are crucial for the study of phenomena in human languages, as they enable researchers access to a vast and varied collection of textual data, which can be employed for building machine learning models that are

applicable to several tasks [11], like text filtering, classification, clustering, summarization, sentiment analysis, language translation etc. These models would be significantly less accurate and practical without access to a wide range of digital corpora.

Implementing phishing detection requires extensive training and education on how to spot phishing attacks and how to react to them [12, 13]. This knowledge, however, derives from several disciplines, such as computer science, information science, computational linguistics, statistics etc. Therefore, such specially crafted corpora can also be used for teaching purposes in various courses in higher education [14], where they could be examined in numerous ways in order to comprehend the underlying fundamentals of phishing attack methods and strategies.

III. Related Work

Diverse studies have been carried out using phishing datasets, but primarily for the English language. A recent paper has presented two dataset variations that consist of almost 150,000 websites labeled as legitimate or phishing, which allow training of classification models, building phishing detection systems, and mining association rules [15].

Another research explained key dimensions of data quality relevant for security, illustrated them with several popular datasets for phishing, intrusion detection and malware, and presented operational methods for assuring data quality in datasets for security challenges [16].

One paper addressed the absence of benchmark datasets for phishing detection. The authors stated that this is due to the fact that phishing websites are short-time living and dead URLs cannot be used in content-based analysis. A dataset was constructed by following a set of proposed guidelines, and afterwards it was utilized for assessing the level of effectiveness in systems based on the random forest classifier algorithm [17].

Another paper presented datasets and tools that are helpful for researching phishing. The authors describe the problem of creating high-quality, diverse and representative phishing datasets. They also discuss the problems of datasets that already exist on the market. Then a benchmarking framework is proposed that automates the extraction of more than 200 features, that implements more than 30 classifiers and a dozen evaluation metrics for the detection of phishing [18].

Ref. [19] suggested a technique that analyzes ULRs, extracts several features, reduces the problem dimensionality, and that when integrated with the support vector machine classifier showed high efficiency in phishing detection. The proposed approach exhibited good results when benchmarking with a range of standard phishing datasets.

Research for other languages is performed less, particularly for those that are not commonly spoken. For the Croatian language there is an evident lack of recent research in the field of phishing detection, and especially in research of corpora that are essential for conducting analyses and for building models that are capable of identifying phishing messages, stopping attacks and protecting users.

One Croatian research has focused on examining the familiarity of users in Croatia with threats in form of social engineering and phishing attacks. Here a practical assessment of users' capabilities to identify phishing attacks was conducted. Also, the research investigated the potential features that are problematic for users when trying to detect phishing attacks [20].

Another Croatian research examined the various types of phishing attacks, such as e-mail phishing, instant messaging phishing, smishing, bulk phishing, spear phishing, whaling, vishing and pharming. The paper also proposed machine learning algorithms suitable for phishing detection, such as decision trees, random forest, support vector machine, k-nearest neighbors etc. [9].

One approach that could be used for discerning phishing messages was presented in a recent Croatian study. The paper focused on the suitability of online services for machine learning of models for predicting classification outcomes [21].

IV. Research, Results and Discussion

Since phishing messages are written in human languages, it is not surprising that creating high-quality and domain-specific textual corpora is of great importance for various analyses in security-related domains and protective strategies.

The aim of this section is to demonstrate the entire process of acquiring data and creating a digital corpus for the task of phishing detection, followed by content analysis using text mining techniques [22] for topic detection. In addition, the resulting corpus is presented and discussed with regard to descriptive attributes, terminology characteristics, metadata and language.

A. Dataset Acquisition and Corpus Creation

The initial dataset was acquired and crafted by extracting data that derived from several personal official e-mail accounts. It consists of 520 phishing e-mails written in Croatian and English dating from the year 2022. In addition to the 520 phishing e-mails, there were some phishing e-mails that were written in other languages, such as French, German and Spanish, but they were less represented and therefore discarded.

The dataset was divided into two subsets containing e-mails that were manually checked and verified as phishing:

- **Set A** which contains 260 e-mails originally written in Croatian (not necessarily grammatically correct), and
- **Set B** which contains 260 e-mails originally written in English and translated automatically into Croatian with Google Translate and without any post-editing.

Each subset contains some duplicates in order to reflect the frequency of certain topics of phishing messages. The initial dataset was pre-processed which

included operations, such as lowercasing, removing accents and numbers, and text filtering.

A stop word list of 290 words was created, consisting of exclamations, conjunctions, pronouns, auxiliary verbs, particles (so-called function words) etc. in order to obtain a corpus of semantically full words, and to filter out all the redundant words.

The resulting phishing corpus consists exclusively of e-mails with the corresponding information and text body, whereas attached files, such as executables, PDF files and textual documents were discarded. Descriptive attributes of the phishing corpus and metadata are presented in Table I.

TABLE I. DESCRIPTIVE ATTRIBUTES AND METADATA

Attributes	Set A	Set B
Sentence number	1,471	939
Token number	37,886	21,427
Word number	31,532	17,660
Mentioning e-mail address(es)	68	51
Containing web link(s)	48	95
Containing attached file(s)	12	6
Containing image(s)	2	5
To field: *personal*	45	85
To field: *undisclosed-recipients*	167	48
To field: *other or not available*	48	67

When comparing the two subsets, Table I presents information on the number of phishing e-mails sent directly to personal e-mail accounts (more than 40 in Set A and more than 80 in Set B), undisclosed recipients (almost 170 in Set A and almost 50 in Set B) and when the recipient was obfuscated (almost 50 in Set A and almost 70 in Set B).

It also shows the number of phishing e-mails that mentioned other e-mail addresses in the text itself (more present in Set A), that contained web links in the body, attached files or images. It is evident that Set A contained more attachments and is larger in terms of sentence, token and word number, whereas Set B contains more web links (95) and images (5). Overall, Set A and Set B consist of almost 60,000 tokens, divided into 2410 sentences, i.e. 1471 in Set A and 939 in Set B.

B. Frequency of tokens

Word clouds as a method for analyzing frequencies of tokens are very helpful, as they offer a visual depiction of the most commonly used terms in a given text or group of texts, which makes it simpler to comprehend and swiftly evaluate important themes and subjects in large datasets [23].

Fig. 1 presents a word cloud with the 50 most frequent tokens from Set A (after pre-processing). The word cloud shows semantic words related to financial fraud, personal data, credit cards, lottery wins, business offers, fundings and bank payments.



Figure 1. Word cloud for Set A

Fig. 2 presents a word cloud with the 50 most frequent tokens from Set B (after pre-processing). The word cloud reveals tokens and semantic words that are related to obscene and indelicate content, e.g. girls, craigslist, service, teenagers, hot, cute, dating, escort, secret, night etc.



Figure 2. Word cloud for Set B

C. Frequent Terminology

The identification of frequent terminology enables more fine-grained evaluation with the distinct aim of differentiating the distribution of words and concepts from the field of phishing attacks. Applying such a straight-forward method allows researchers to interpret the structure of phishing messages with their unique style, vocabulary and specific terminology. This can provide a deeper, objective, unbiased and consistent insight into the ways of crafting phishing messages.

However, such analyses are altogether based on word occurrences and this approach yields problems related to the qualitative spectrum of data analysis [24]. Therefore, it is important to warrant an investigation of concordances that can explain the diverse contexts of such words in a phishing message.

Two word lists containing the most frequent meaningful lemmas of nouns and verbs, reflecting the content of e-mails from Set A and Set B are given in Table II.

TABLE II. MOST FREQUENT TERMINOLOGY

Set A		Set B	
Nouns	**Verbs**	**Nouns**	**Verbs**
banka	kontaktirati	djevojka	željeti
fond	moći	e-pošta	tražiti
e-pošta	poslati	stranica	razgovarati
ime	moliti	račun	podijeliti
račun	dobiti	pratnja	poslati
adresa	željeti	usluga	kontakirati
sredstvo	primiti	pozdrav	moliti
broj	odgovoriti	transakcija	pogledati
podatak	pomoći	url	pozivati
pozdrav	dati	mreža	pomoći
kartica	potvrditi	trošak	koristiti
iznos	odlučiti	aplikacija	kliknuti
plaćanje	obavještavati	datoteka	dati
dolar	učiniti	poziv	javiti
zemlja	kliknuti	zajednica	upoznati
milijun	prijaviti	veza	čekati
novac	donirati	novac	znati
informacija	dijeliti	dolar	primiti
poruka	osvojiti	nsa	morati

The extracted word lists presented in Table II display the content of phishing e-mails. Word list from Set A contains mostly nouns and vocabulary related to financial fraud (*banka – bank*, *fond – fund*), valuable data (*račun – account*, *adresa – address*, *broj – number*), verbs reflecting communication of receiving and giving, sending, helping, replying, and mentions e-mails as a means of communication (*e-pošta – e-mail*, *podatak – data*, *sredstvo – medium*/*way*).

Word list from Set B contains mostly obscene vocabulary (*djevojka – girl*, *usluga – service*, *pratnja – escort*), mentions websites and e-mails as means of communication (*e-pošta – e-mail*, URL, *mreža – network*), transaction-related data (*račun – account, transakcija – transaction*), and greetings at the end of messages (*pozdrav – greetings*). Verbs reflect more a polite tone of communication, using words, such as *željeti – wish*, *razgovarati – talk*, *moliti – beg*, *pomoći – help*, *tražiti – ask*, *podijeliti – share*, *poslati – send*.

D. N-gram analysis

In order to detect how and what words are frequently combined in a language, collocations can be used. They are commonly applied in various natural language processing tasks, text mining and linguistic analyses.

Collocations can be examined through n-grams, which typically represent a sequence of *n* words in a text. They can help to capture the context and understand word semantics.

In this research, the focus is on bigrams and trigrams. Bigrams are sequences of two words that appear together in a text, and can be employed for identifying common word combinations and text patterns. A trigram is a sequence of three words and can be used to find more intricate structures and patterns in a text.

When it comes to Set A, the most frequent bigrams are related to financial fraud, personal data and personal communication, such as *poštovani korisniče – dear user*, *bankovni račun – bank account*, *banka africa – bank Africa*, *afrička unije – African Union*, *broj telefona – phone number*, *kartice adresu – card address*, *atm kartica – ATM card*, *pošaljite informacije – send information*, *godine spol – years gender*, *milijun dolara – million dollars*, *pošaljite podatke – send data*, *visa kartice – Visa cards*, *prijenos sredstava – money transfer*, *najbliža rodbina – close relatives*, *odmah odgovorite – answer immediately* etc. A few expressions in Set A are also written in English, e.g. *united bank – united bank*.

Trigrams reveal the same content as bigrams, but provide a broader context, such as *united bank africa – united bank Africa*, *atm kartice adresa – ATM card address*, *telefon godine spol – phone years gender*, *pošaljite podatke visa – send Visa data*, *iznos milijun tisuća – amount million thousands*, *poštovani korisniče računa/pošte – dear account/mail user*, *odgovorite united bank – answer united bank*, *međunarodni monetarni fond – International Monetary Fund* etc.

Most frequent bigrams in Set B are primarily related to obscene content, such as *zgodna pratilja – handsome companion*, *povremeni spojevi – occasional dates*, *craigslist zakona – craigslist law*, *tajna zajednica – secret community*, *nabaviti zgodne – get a pretty*, *usluga pratnje – escort service*, *noć povremena – occasional night* etc.

Trigrams in Set B are, e.g. *noć povremene spojeve – night occasional dates*, *odjeljku craigslist zakona – craigslist law section*, *rastuća tajna zajednica – growing secret community*, *nabavite zgodne djevojke – find hot girls*, *djevojka na poziv – call girl*, *tajna zajednica nsa – secret NSA community* etc.

E. Context analysis

Concordances are very useful when data analysis requires examining and comprehending context in which a certain word, n-gram or phrase is used in a text. This helps to spot patterns and trends in a text, such as how a term is used repeatedly or in distinct contexts, how word collocations are formed, or when and what words frequently follow each other in a text.

In this research a few keywords in context (KWIC) [24] were chosen. Table III shows concordances from Set A and Set B with regard to four keywords (KWIC) with responding contexts: odmah – immediately, *sada – now*, *razgovarati – talk*, *podijeliti – share*. The integer in parentheses indicates the number of occurrences of a particular word.

TABLE III. CONCORDANCE ANALYSIS

Set A	Set B
odmah (44) – immediately	**odmah (5) – immediately**
savjetujemo da ***odmah*** pošaljete svoje ime, kontakt adresu i broj mobilnog telefona.	ljubazno mi odgovorite ***odmah***. Hvala.
ako budem plaćen, ***odmah*** ću uništiti video	Odgovorite mi i nazovite me da ***odmah*** dobijem besplatnu ponudu
(MMF) Afričke regije ***odmah*** unutar sljedećih 168 sati.	Dopustite mi da počnem ***odmah*** od točke
Napomena: ***odmah*** po zaključku transakcije imate pravo na 45%	kliknite ovdje za nadogradnju i ***odmah*** ponovno potvrdite svoj račun
kontaktirajte me ***odmah*** za daljnju komunikaciju	***Odmah*** potvrdite svoj interes
sada (60) – now	**sada (41) – now**
Sada kontaktirajte Službu za korisnike UBA banke	Chatajte ***sada*** (samo unesite svoju e-poštu)
Sada vaše novo plaćanje, broj odobrenja	Možemo li ***sada*** uspostaviti video poziv
Sada želim da hitno kontaktirate	Provjerite ***sada*** gdje možete
Sada se obratite generalnom direktoru	**Sada** sam spreman za poziv
Sada sam u skrovištu samo da zaštitim svoj život i	***Sada*** dostupno WhatsApp Facebook Instagram
razgovarati (10) – talk	**razgovarati (44) – talk**
Mogu li ***razgovarati*** s vama putem ove e-pošte?	pregledajte našu stranicu i **razgovarajte** -> https:
kako bismo ***razgovarali*** o mojoj hitnoj potrebi za nasljednikom	privatne fotografije ***Razgovarajte*** sada (samo unesite svoju e-poštu)
investicijskog projekta o kojem smo ***razgovarali***	želite li **razgovarati** ?? Ako se slažete, pošaljite mi svoje ime, dob
podijeliti (10) – share	**podijeliti (44) – share**
kako bi mi pomogli ***podijeliti*** ovih 5,5 milijuna dolara	***podijelila*** je datoteku s vama
spreman sam ***podijeliti*** novac s vama.	***Podijelit*** ću slike i više detalja o sebi čim mi se javite
novac vraćen od prevaranata ***podijeli*** među 100 sretnih ljudi	povjeriti vama i ***podijeliti*** s vama ovaj povjerljivi posao.
ovaj će se posao ***podijeliti*** u ovom omjeru	bih želio ***podijeliti*** s tobom ako samo možeš odgovoriti na moju poštu

Many phishing e-mails ask for an urgent response, and therefore oftentimes contain a call for an emergency action. Such messages include the words *odmah – immediately* and *sada – now*. This is especially noticeable in Set A (originally written in Croatian), where these words appear 44 and 60 times, respectively.

However, there are some differences in the use of such words between the two subsets. Urgency is more present in Set A, which is more focused on financial fraud. On the other hand, most phishing messages from Set B (phishing e-mails originally written in English and translated into Croatian) are in the style of begging for help or asking for a conversation, hence the of use of verbs, such as *razgovarati – talk* and *podijeliti – share* (both used more than 40 times). This was accompanied by a call to click on links that were integrated into the text body of e-mails (95 times out of 260 e-mails).

V. CONCLUSION

This paper presents the process of data acquisition and corpus creation for the phishing detection domain. Collected phishing e-mails were divided into two subsets of data: e-mails originally written in Croatian (Set A), and e-mails translated from English into Croatian (Set B). Both subsets were evaluated in a number of ways, e.g. by analyzing n-grams, collocations, and frequencies of e-mails and web links that were integrated into phishing e-mails.

In this dataset, phishing e-mails written in Croatian were predominantly sent to undisclosed recipients, whereas those originally written in English were mostly sent to personal e-mails. The authors assume that these e-mail addresses were probably extracted from public research papers, presentations or speeches.

Phishing e-mails originally written in the Croatian language refer more to other e-mail addresses as a malicious means of connecting with victims, whereas those originally written in English contain more web links, enticing potential victims to click on them and instructing victims to disclose private or sensitive information.

Set A reveals content that is related to finance, personal data, bank details, business offers, funding and payments, while Set B contains more mature and obscene content.

Bigrams and trigrams show collocations belonging to the same domains. Phishing e-mails that are related to financial transactions and credit cards are often marked as urgent, and ask for an immediate response (frequent use of terms *odmah – immediately* and *sada – now*). Phishing messages that are, however, related to indecent content are more in a kind and begging tone, not focusing on urgency. Here the intention is to lure victims into sharing data, either by clicking on web links that are integrated into the text body, or by inviting them to friendly conversations.

For future research, the authors of this paper plan to increase the volume of the phishing dataset, and then to categorize messages into different topic classes, to examine the sentiment that is present in phishing messages, and to apply machine learning algorithms for predicting various classes of phishing e-mails for the Croatian language.

REFERENCES

[1] Australian Cyber Security Centre, "ACSC Annual Cyber Threat Report: July 2019 to June 2020",

https://www.cyber.gov.au/sites/default/files/2020-09/ACSC-Annual-Cyber-Threat-Report-2019-20.pdf

[2] M. Dadkhah, S. Shamshirband, and A. Wahab, "A hybrid approach for phishing web site detection", The Electronic Library, vol. 34, no. 6, pp. 927–944, 2016.

[3] M. Rosenthal, "Must-Know Phishing Statistics: Updated 2022", Tessian, https://www.tessian.com/blog/phishing-statistics-2020/

[4] Z. Alkhalil, C. Hewage, L. F. Nawaf, and I. Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy", Frontiers in Computer Science, vol. 3, article 563060, 2021.

[5] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges", Multimedia Tools Applications, vol. 82, pp. 3713–3744, DOI: https://doi.org/10.1007/s11042-022-13428-4, 2023.

[6] A. ALmomani, T.-C. Wan, A. Manasrah, A. Altaher, M. Baklizi, and S. Ramadass, "An enhanced online phishing e-mail detection framework based on "Evolving connectionist system"", International Journal of Innovative Computing, Information and Control (IJICIC), vol. 9, no. 3, pp. 1065–1086, 2013.

[7] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs", Expert Systems with Applications, vol. 117, pp. 345–357, 2019.

[8] W. A. Awad, and S. M. ELseuofi, "Machine Learning methods for E-mail Classification", International Journal of Computer Applications, vol. 16, no. 1, pp. 39–45, 2011.

[9] A. Kovač, I. Dunđer, and S. Seljan, "An overview of machine learning algorithms for detecting phishing attacks on electronic messaging services", 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Croatia, 2022, pp. 954–961, DOI: 10.23919/MIPRO55190.2022.9803517.

[10] R. H. Hariri, E. M. Fredericks, and K. M. Bowers, "Uncertainty in big data analytics: survey, opportunities, and challenges", Journal of Big Data, vol. 6, no. 44, DOI: https://doi.org/10.1186/s40537-019-0206-3, 2019.

[11] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text Classification Algorithms: A Survey", Information, vol. 10, no. 4, DOI: https://doi.org/10.3390/info10040150, 2019.

[12] M. Adil, R. Khan, and M. A. Nawaz Ul Ghani, "Preventive Techniques of Phishing Attacks in Networks", 3rd International Conference on Advancements in Computational Sciences (ICACS), Pakistan, 2020, pp. 1–8, DOI: 10.1109/ICACS47775.2020.9055943.

[13] A. Sumner, and X. Yuan, "Mitigating Phishing Attacks: An Overview", ACM Southeast Conference (ACM SE '19). Association for Computing Machinery, USA, 2019, pp. 72–77, DOI: https://doi.org/10.1145/3299815.3314437.

[14] R. Jaworski, S. Seljan, and I. Dunđer, "Towards educating and motivating the crowd – a crowdsourcing platform for harvesting the fruits of NLP students' labour", 8th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, vol. 1, 2017, pp. 332–336.

[15] G. Vrbančič, I. Jr. Fister, and V. Podgorelec, "Datasets for phishing websites detection", Data in Brief, vol. 33, ISSN 2352-3409, DOI: 10.1016/j.dib.2020.106438, 2020.

[16] R. M. Verma, V. Zeng, and H. Faridi, "Data Quality for Security Challenges: Case Studies of Phishing, Malware and Intrusion Detection Datasets", ACM SIGSAC Conference on Computer and Communications Security (CCS '19). Association for Computing Machinery, USA, 2019, pp. 2605–2607, DOI: https://doi.org/10.1145/3319535.3363267.

[17] A. Hannousse, and S. Yahiouche, "Towards benchmark datasets for machine learning based website phishing detection: An experimental study", Engineering Applications of Artificial Intelligence, vol. 104, ISSN 0952-1976, DOI: https://doi.org/10.1016/j.engappai.2021.104347, 2021.

[18] V. Zeng, S. Baki, A. El Aassal, R. Verma, L. F. Teixeira De Moraes, and A. Das. "Diverse Datasets and a Customizable Benchmarking Framework for Phishing", Sixth International Workshop on Security and Privacy Analytics (IWSPA '20). Association for Computing Machinery, USA, pp. 35–41, DOI: https://doi.org/10.1145/3375708.3380313.

[19] J. Rashid, T. Mahmood, M. W. Nisar, and T. Nazir, "Phishing Detection Using Machine Learning Technique", First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Saudi Arabia, 2020, pp. 43–46, DOI: 10.1109/SMART-TECH49988.2020.00026.

[20] J. Andrić, D. Oreški, and T. Kišasondi, "Analysis of phishing attacks against students", 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Croatia, 2016, pp. 1423–1429, DOI: 10.1109/MIPRO.2016.7522363.

[21] I. Zatezalo, and I. Dunđer, "Online service for accessible machine learning of prediction models", Zbornik radova Međimurskog veleučilišta u Čakovcu, vol. 12, no. 2, p. 10, 2021.

[22] M. Pejić Bach, Ž. Krstić, S. Seljan, and L. Turulja, "Text Mining for Big Data Analysis in Financial Sector: A Literature Review", Sustainability, vol. 11, no. 5, pp. 1-27, DOI: 10.3390/su11051277, 2019.

[23] M. A. Hearst, E. Pedersen, L. Patil, E. Lee, P. Laskowski, and S. Franconeri, "An Evaluation of Semantically Grouped Word Cloud Designs", in IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 9, pp. 2748–2761, 2020, DOI: 10.1109/TVCG.2019.2904683, 2020.

[24] M. Pavlovski, and I. Dunđer, "Is Big Brother watching you? A Computational Analysis of Frequencies of Dystopian Terminology in George Orwell's 1984", 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Croatia, 2018, pp. 638–643, DOI: 10.23919/MIPRO.2018.8400120.

Information Extraction from Security-Related Datasets

S. Seljan*, N. Tolj* and I. Dunđer*
* Faculty of Humanities and Social Sciences, University of Zagreb, Department of Information and Communication Sciences, Zagreb, Croatia
sanja.seljan@ffzg.hr, ntolj@ffzg.hr, ivandunder@gmail.com

***Abstract* - There are various approaches to executing security breaches which are nowadays massively occurring in electronic communication environments, and phishing attacks are one of the most applied ones. A vast majority of phishing attacks are initiated using electronic messages, which attackers utilize to direct users to harmful or fake websites, to infect computers or to obtain personal or sensitive data for malicious purposes. Consequently, it is necessary to identify phishing messages in order to provide suitable user protection. Research and numerous studies have included machine learning algorithms and techniques from the field of artificial intelligence which predominantly depend on language-specific datasets and characteristics of phishing messages, and which have demonstrated to be effective for extracting critical information and for data-driven decision making. However, phishing datasets exist mainly for the English language. The aim of this paper is to present an information extraction pipeline that encompasses phases, such as corpus pre-processing, generating predictions of phishing messages using selected machine learning algorithms, along with a basic analysis, confusion matrices and evaluation scores for Croatian phishing messages. This type of key information can be used for teaching in higher education, e.g. in security-related courses or subjects that deal with artificial intelligence, machine learning, big data analysis, computational linguistics etc. This is essential as it can provide deeper insights into phishing attack strategies and potential countermeasures.**

Keywords - information extraction; machine learning; corpus analysis; security datasets; information security; information and communication sciences

I. INTRODUCTION

In a phishing attack, which is a form of social engineering [1], an attacker assumes the identity of a trustworthy party in order to trick a victim into exposing sensitive data, private or financial information [2].

The effort to identify and prevent phishing attacks on individuals and companies is known as phishing detection. Phishing detection is a vital component of information security, and without adequate identification and prevention procedures, a phishing attack might result in data breaches, financial loss and reputational harm [3].

Protecting against these attacks requires setting up strong security measures and training people in how to spot and deal with phishing. However, identifying phishing attacks is not easy, since there are numerous challenges brought on by linguistic complexity and diversity, as well as technical and technological issues.

Moreover, to make phishing messages seem very authentic, attackers also known as "phishers" [4] frequently use sophisticated terminology and techniques, such as misusing official branding and logos of reputable companies and other organizations [5]. In order to further boost the persuasiveness of their messages, attackers also tend to imitate the language, style and tone of a known sender.

This makes it harder for automated security systems to identify attackers and to discern between phishing and legitimate messages [6], especially if they lack the linguistic knowledge or contextual awareness needed in order to properly examine the messages' content.

When it comes to countermeasures, machine learning can be an effective tool for detecting phishing messages, as it enables automatic identification and classification [7]. Machine learning algorithms can accurately determine the difference between phishing and authentic messages [4], especially when trained with domain-specific data and after applying model fine-tuning.

In order to assess and prevent phishing attacks, special datasets, so-called digital corpora, are needed for training machine learning models [8]. The quality and volume of datasets have a significant impact on the performance of security systems that are based on such models.

Furthermore, a dataset that truly reflects the variety of popular phishing methods and strategies is essential for phishing detection [9]. That means that such datasets should contain large amounts of phishing e-mails, texts, and web pages, along with examples of authentic messages to serve as a comparison. It is important that the dataset accurately reflects real examples, so that the model has enough information and context in which it will be applied [10].

Nevertheless, machine learning is not an infallible method for detecting phishing. Machine learning models are susceptible to errors [11], particularly when the training dataset is not diverse enough or when the model is not adjusted correctly.

Therefore, it is crucial to have a dataset that is fairly large for the model to generalize well to observations that were not seen during model training. In order to prevent or reduce bias toward one output class (also known as label)

during model training, the dataset should be balanced, meaning that it should contain an equal number of phishing and non-phishing messages. Additionally, it should be tagged by domain specialists or security experts who are knowledgeable about phishing strategies. This will increase the validity of the labels and, consequently, the quality of the corpus.

The paper is organized in the following way. In the Introduction section the authors present a vital topic in information security, i.e. phishing and potential countermeasures. The second section states the motivation for conducting this research, whereas the third section presents related work and relevant research on datasets and machine learning algorithms for detecting phishing messages. The fourth section deals with ensemble learning methods, which is followed by the research section which presents all results and discusses important findings. Finally, in the last section conclusions are given with regard to research outcomes.

II. MOTIVATION

Information extraction deals with applying natural language processing methods to automatically extract essential details from text documents [12]. These methods are language dependent and focused on the domain of the text. Machine learning, on the other hand, is fast and language independent, and can been used in order to facilitate the process of information extraction. Machine learning algorithms have proven to be effective for extracting critical information and making data-driven decisions.

Nevertheless, they primarily depend on language-specific and high-quality datasets, and characteristics of phishing messages that need to be representative, diverse and labeled in a correct way. Model performance directly depends on the features and properties of the dataset.

However, Croatian-English phishing datasets and recent analyses are lacking, and this very absence of studies was the main motivation behind this research. New insights could be incorporated into various courses in higher education, such as natural language processing, machine learning, artificial intelligence, data analysis, information security, computational linguistics, statistics etc. This is crucial since it can offer more in-depth information about phishing attack techniques and potential responses.

Depending on the perspective of the problem, students could be taught how to evaluate attackers and assess their profiles, how to distinguish different phishing strategies, and how to deal with the principles of phishing techniques and potential countermeasures.

III. RELATED WORK

One research dealt with detecting phishing websites by employing a novel dataset, which contains 5,000 legitimate and 5,000 phishing websites. In order to obtain the best results, various machine learning algorithms, such as Random Forest algorithm and Multilayer Perceptron were explored [13].

Ref. [14] proposed a system that uses the Support Vector Machine algorithm and Naïve Bayes to train predictive models on a 15-dimensional feature set. The experiments were based on datasets consisting of 2541 phishing instances and 2500 benign instances. Using 10-fold cross-validation, experimental results showed 0.04% false positives and 99.96% accuracy for both predictive models.

Another paper also used Random Forest classification and the Support Vector Machine algorithm in order to achieve a 99.87% accuracy score on a dataset that was comprised of 1605 e-mails (1191 phishing and 414 safe e-mails) [15].

In a recent research various types of phishing attacks, such as e-mail phishing, instant messaging phishing, smishing, bulk phishing, spear phishing, whaling, vishing and pharming were examined [4]. For instance, phishing in general is considered as a type of cyberattack in which a hacker tries to trick people into disclosing private or important information by impersonating a trustworthy organization or user, whereas, e.g. spear phishing is a highly specialized form of phishing in which a malicious e-mail is specifically intended for a particular person, business, or group of people. This paper also proposed machine learning algorithms suitable for phishing detection, such as Decision Trees, Random Forest, Support Vector Machine, k-Nearest Neighbors etc.

Another study concentrated on the appropriateness of online services for machine learning of models for predicting classification outcomes, and provided a method that could be utilized for identifying phishing messages [16].

One paper discussed the absence of relevant datasets for phishing detection. Therefore, a specific dataset was constructed by following a set of proposed guidelines, and was used afterwards to assess the level of effectiveness in phishing detection systems based on the Random Forest classifier algorithm [17].

Ref. [18] proposed a method that concentrates on ULRs, extracts features, lowers the dimensionality of the problem, and, when combined with a Support Vector Machine classifier, showed high efficacy in the detection of phishing.

IV. ENSEMBLE LEARNING METHODS

Ensemble learning methods are used in machine learning to combine predictions from multiple models in order to improve predictive performance. There are three main types of ensemble learning methods: Bagging (Bootstrap Aggregation), Stacking and Boosting [19, 20].

Bagging or Bootstrap Aggregation uses multiple Decision Trees on different samples of the same dataset, and then averages predictions [21]. Bagging employs different members of an ensemble group by using variations of training data. It uses replacement, meaning that once the instance (row) is selected, it is returned, and can be selected again. This technique of bootstrap resampling is often used in statistics on small datasets, where many training datasets can be prepared to achieve an overall better estimation [21].

An example of Bagging is the popular Random Forest algorithm, which is used for solving classification and regression problems. This algorithm contains numerous Decision Trees on different subsets of a dataset, and therefore "takes the average to increase the predictive accuracy of that dataset" [22]. It makes predictions depending on each tree, and makes a final decision based on the maximum votes of predictions, which is especially useful when individual trees are not correlated among themselves. By using multiple Decision Trees, it overcomes the problem of overfitting. On the other hand, the accuracy of Random Forest depends on the larger variety of trees [23].

According to [24], advantages of the Random Forest algorithm are automatization of lost values in data and, overcoming overfitting issues with efficiency in handling large datasets. However, disadvantages are in the context of more computing and more resources that are needed for efficient results. Random Forest requires more time for training as well since it relies on many Decision Trees.

Stacking uses different models (e.g. Linear Regression for regression tasks, or Logistic Regression for binary classification) on the same dataset and another model to learn how to make predictions. Ensemble members are referred to as level-0 model, and the model that is used to combine predictions as level-1 model [19].

Boosting is an ensemble learning method that aims to change the training data in order to focus on examples that were erroneously predicted during previous fitting of models on the training dataset [19]. Is focused on instances that were misclassified by previous classifiers. It is called Adaptive Boosting since weights are re-assigned to each instance, i.e. each row, having higher weights assigned to incorrectly classified instances. Instances are weighted to indicate the needed amount of focus on the dataset during training. The output is given as a weighted average of predictions.

Popular examples of the boosting method are Adaptive Boosting (AdaBoost), Stochastic Gradient Boosting (XGBoost and similar) and Gradient Boosting Machines [19]. AdaBoost is an adaptive machine learning algorithm where weights are reassigned to each instance, i.e. each row which is misclassified. Incorrectly classified instances have higher weights, and are widely used for imbalanced data. While Decision Trees in the Random Forest algorithm have equal contributions, in AdaBoost they have different contributions.

V. Research, Results and Discussion

The objective of this paper is to present a pipeline for information extraction that includes steps like corpus pre-processing, generating predictions for phishing messages using specific machine learning algorithms, performing an accuracy-based analysis, creating confusion matrices and evaluating scores for Croatian phishing messages.

A. Dataset Characteristics

The dataset used in this research consists of 550 phishing e-mails that arrived at private e-mail addresses during 2022. The entire dataset was pre-processed with special focus on lowercasing, removing accents and numbers, text filtering with help of a specially crafted stop word list, discarding e-mail attachments etc. All e-mails were manually checked and annotated. The dataset was then divided into two subsets:

- **Set A**: 275 e-mails originally written in Croatian,
- **Set B**: 275 e-mails originally written in English that were automatically translated into Croatian by the web service Google Translate.

Examples of phishing e-mails from both sets are given below:

Set A:

- "Obavijest: Vaša kompenzacijska bankovna kartica na bankomatu u vrijednosti od 1.500.000,00 dolara registrirana je kod voditelja kurira DHL-a g. Marka Adjovija. Radi trenutne isporuke kontaktirajte ga putem e-pošte (dhlcourierexpressbj1@outlook.com) za više informacija o tome kako ćete je zatražiti."
- „Draga moja, jesi li primila poruku koju sam ti poslao? Pozdrav, Jerry Ngessan"

Set B:

- "Poštovani, Vaša verzija e-pošte je zastarjela, nenadogradnja na najnoviju verziju ffzg.hr 7.1 sada će dovesti do trajnog zatvaranja računa. Sukladno odredbi 17.9 Uvjeta, ffzg.hr može u bilo kojem trenutku prekinuti svoje usluge za račune. Za nadogradnju kliknite ovdje i odmah ponovno potvrdite svoj račun. Hvala, ffzg.hr Mail Team"
- „Sada sam spreman za poziv, mogu li vas nazvati? Odgovorite na -> angelodicosta@gmail.com. Možemo li se naći u kafiću?"

Each phishing e-mail was manually marked with one of the following categories (also known as labels or classes):

- Finances – commercial phishing e-mails with the aim to scam users for monetary benefit,
- Health – offers and promises life changing products that do not exist, in order to acquire valuable personal information from victims,
- Adult content – e-mails containing erotic and raunchy content, or having allusions to pornographic topics, nudity, explicit sexual material or violence,
- Short communication – short e-mails containing brief and generic greetings to encourage victims to continue a conversation in which they provide sensitive information to the attacker.

Distribution of categories in the dataset is presented in Table I, showing the proportions of different classes (labels) of e-mails. In both subsets, the test set equals to 10% of the training set. The most represented categories in the training and test sets are "Adult content" and "Finances". In almost all cases, the least represented category is "Health".

Although divided into four different categories, all phishing e-mails are related to financial matters, which might contribute to later-stage misclassification. E-mails that do not deal with financial topics are aimed at collecting data through a friendly conversation, with the goal of obtaining financial benefit through fraud.

TABLE I. DATASET CHARACTERISTICS

	Set A		**Set B**	
	Training set	**Test set**	**Training set**	**Test set**
No. e-mails	250	25	250	25
Finances	89	11	90	10
Health	8	2	17	5
Adult content	94	8	103	9
Short communication	59	4	40	1

A large amount of phishing e-mails focuses on the financial aspect of victims, asks for valuable data or directly for money, employs blackmailing strategies, presents inappropriate services, asks for charity donations or promises lottery wins. Such e-mails frequently offer victims money in exchange for their data.

B. Prediction of Categories

In order to perform prediction of all categories, both subsets were split in the ratio of 90:10 (proportion of training and test set). Training was performed using the following classifier algorithms [4, 19]: Multivariable Logistic Regression (LR), which is suitable for multiple variables (categories), Random Forest (RF), k-Nearest Neighbors (kNN), Naïve Bayes (NB) and AdaBoost (AB).

These algorithms were chosen on purpose in order to detect the algorithm and approach with the highest efficacy, i.e. percentage of correct predictions for this task.

Table II shows predictions achieved by all algorithms, and across all four categories. When comparing the two subsets, Set A, originally written in Croatian, has slightly better predicted scores than Set B, which was automatically translated from English into Croatian.

TABLE II. PREDICTED ACCURACY RESULTS ACROSS ALL CATEGORIES

	Set A		**Set B**	
Algorithm	Accuracy	F1	Accuracy	F1
LR	0.684	0.686	0.680	0.675
RF	**0.776**	**0.767**	**0.732**	**0.711**
kNN	0.704	0.691	0.648	0.613
NB	0.460	0.454	0.448	0.453
AB	**0.748**	**0.743**	**0.760**	**0.753**

Remark: more is better.

When comparing prediction outcomes, accuracy and F1 measures achieved similar results. The highest predictions were achieved by the algorithms Random Forest and AdaBoost, ranging from ca. 71% up to ca. 77% for accuracy and F1 scores.

The F1 score is based on the balance between the metrics of precision and recall, i.e. it is the harmonic mean of precision and recall [25], usually used for imbalanced datasets.

Precision identifies all correctly identified cases, divided by all positives (true and false positives). Recall is the measure that counts true positives divided by true positives and false negatives.

F1 and classification accuracy as prediction measures obtained by Random Forest and AdaBoost are compared with manual testing results.

C. Confusion Matrix

The confusion matrix is a predictive analytics tool used widely in machine learning and in various classification tasks [26]. It is a summarized *NxN* table that presents the number of correct and incorrect predictions in a classification task, and therefore estimates the performance of a classification algorithm.

The confusion matrices determine the differences between the actual and the predicted accuracy across all four categories. The confusion matrix takes the total number of all training datasets in all categories and predicts accuracy.

Fig. 1 and Fig 2. show the confusion matrices for the Random Forest and AdaBoost algorithms for each of the four categories. Both algorithms predicted similar order of categories in both datasets.

Interestingly, both algorithms predicted that “Health” in Set A will gain a 100% score, followed by close scores for the categories “Adult content” and “Finances”. For Set B, both algorithms predicted the same ranking of categories: “Adult content”, followed by “Finances”, then “Short communication”, and finally “Health”. This ranking is contrary to the rankings in Set A.

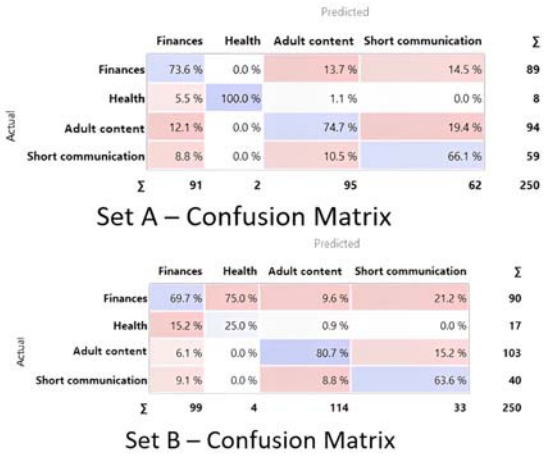


Figure 1. Confusion Matrices per Category for Random Forest

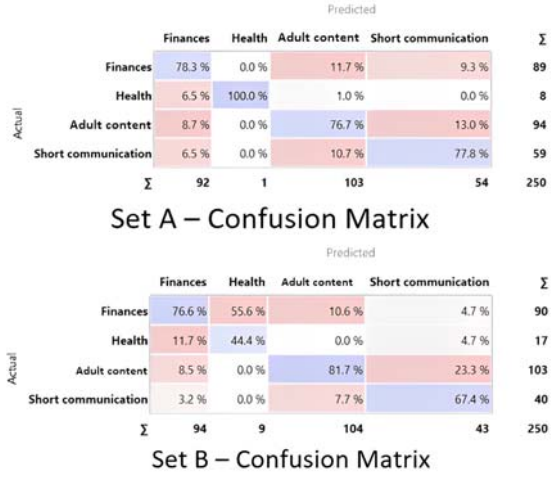


Figure 2. Confusion Matrices per Category for AdaBoost

Most represented categories in datasets are predicted more consistently: "Adult content" and "Finances" with predictions ranging from 70-80% for both datasets and both algorithms, whereas less represented categories demonstrated significant variations.

D. Results of Predicted and Tested Accuracies

Predicted scores are based on the complete training set, and were obtained by each algorithm across all four categories. Tested scores were obtained through test sets (25 phishing e-mails) which are used to test the accuracy for prediction of categories by each algorithm.

Fig. 3 presents results of tested and predicted scores across all four categories. Accuracy of tested scores is lower than accuracy of predicted scores, generally by 2-6%.

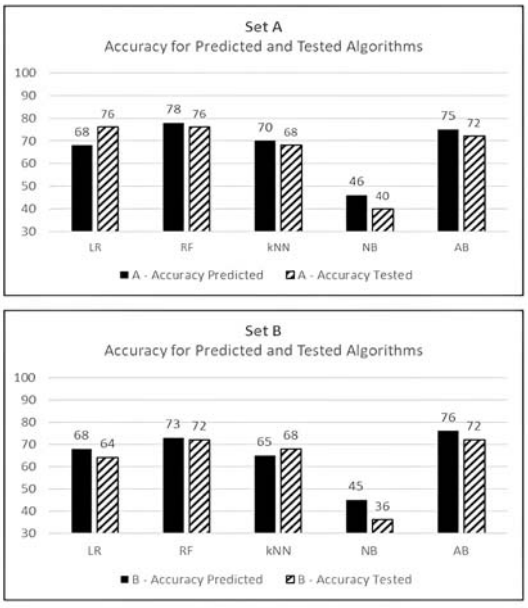


Figure 3. Tested and predicted accuracies by algorithm

Fig. 4 presents results for predicted and tested average scores per category, showing significant variations among categories.

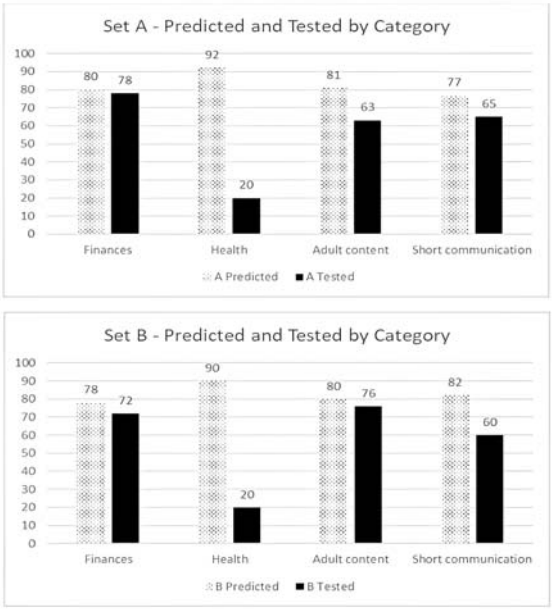


Figure 4. Predicted and tested accuracies by category

Results show a decline of tested accuracy scores in all categories, but the smallest decrease is in the "Finances" category, followed by the category "Adult content", which are the most represented categories in the training sets. The most significant decline was in the "Health" category (72% in Set A and 70% in Set B), which had the smallest amount of training data.

VI. CONCLUSION

The main goal of this research was to demonstrate a corpus-based information extraction pipeline that is language independent and built on machine learning techniques for predicting categories of phishing e-mails.

The research was performed on two sets, differentiated by language originality (Croatian or automatically translated from English into Croatian). Training and test sets were imbalanced in terms of categories. The most represented categories in training and test sets were "Adult content" and "Finances". The category "Health" was the least represented category, and achieved the lowest accuracy scores. However, a significant number of phishing e-mails contain mixed content, i.e. all of them deal more or less with financial issues, and therefore can belong to multiple categories. This was confirmed to be the main limitation of this research.

Results are compared between predicted accuracy scores and tested accuracy scores, with regard to the chosen machine learning algorithm and with regard to the predicted category. When comparing algorithms with manual testing results, the best predictions were obtained by Random Forest and AdaBoost, while Naïve Bayes performed worst.

Predicted results obtained by classification accuracy by applying Random Forest and AdaBoost are compared with manual testing results for each of the four categories. Both algorithms predicted most consistently categories of "Adult content" and "Finances", which were the most represented categories in the training datasets, with scores of 70-80% for both subsets (A and B). This result shows the immense importance of data quantity in the training set. The biggest variations in prediction scores and when comparing predicted and tested score values are for the category "Health", which is the least represented category in the training dataset, exposing the weakness of machine learning algorithms on small datasets.

Results in this paper confirm that predicted accuracy increases with data quantity. The best predictions are obtained for categories that were most represented in datasets, whereas worst results were obtained for the least represented category "Health". Overall, algorithms that exhibited best prediction results were ensemble machine learning algorithms – Random Forest and AdaBoost.

In order to improve accuracy results, the authors plan to increase sample size equalize the number of phishing e-mails that are used for in training and test sets, and harmonize the length of phishing messages.

ACKNOWLEDGEMENT

This research is funded by institutional research projects "61-920-2553" and "11-933-1053" at the Faculty of Humanities and Social Sciences, University of Zagreb.

REFERENCES

[1] M. Dadkhah, S. Shamshirband, and A. Wahab, "A hybrid approach for phishing web site detection", The Electronic Library, vol. 34, no. 6, pp. 927–944, 2016.

[2] Australian Cyber Security Centre, "ACSC Annual Cyber Threat Report: July 2019 to June 2020", https://www.cyber.gov.au/sites/default/files/2020-09/ACSC-Annual-Cyber-Threat-Report-2019-20.pdf

[3] M. Rosenthal, "Must-Know Phishing Statistics: Updated 2022", Tessian, https://www.tessian.com/blog/phishing-statistics-2020/

[4] A. Kovač, I. Dunđer, and S. Seljan, "An overview of machine learning algorithms for detecting phishing attacks on electronic messaging services", 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 954–961, DOI: 10.23919/MIPRO55190.2022.9803517.

[5] D. Aljeaid, A. Alzhrani, M. Alrougi, and O. Almalki, "Assessment of End-User Susceptibility to Cybersecurity Threats in Saudi Arabia by Simulating Phishing Attacks", Information, vol. 11, no. 12, DOI: https://doi.org/10.3390/info11120547, 2020.

[6] F. Carroll, J. A. Adejobi, and R. Montasari, "How Good Are We at Detecting a Phishing Attack? Investigating the Evolving Phishing Attack Email and Why It Continues to Successfully Deceive Society", SN Computer science, vol. 3, DOI: https://doi.org/10.1007/s42979-022-01069-1, 2022.

[7] S. Pandiyan, P. Selvaraj, V. K. Burugari, J. P. Benadit, and P. Kanmani, "Phishing attack detection using Machine Learning, Measurement", Sensors, vol. 24, ISSN: 2665-9174, DOI: https://doi.org/10.1016/j.measen.2022.100476, 2022.

[8] O. Yavanoglu, and M. Aydos, "A review on cyber security datasets for machine learning algorithms", IEEE International Conference on Big Data, Boston, USA, 2017, pp. 2186–2193, DOI: 10.1109/BigData.2017.8258167.

[9] H. F. Atlam, and O. Oluwatimilehin, "Business Email Compromise Phishing Detection Based on Machine Learning: A Systematic Literature Review", Electronics, vol. 12, no. 42, DOI: https://doi.org/10.3390/electronics12010042, 2023.

[10] H. Hindy et al., "A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems", IEEE Access, vol. 8, pp. 104650–104675, DOI: 10.1109/ACCESS.2020.3000179, 2020.

[11] S. Doroudi, "The Bias-Variance Tradeoff: How Data Science Can Inform Educational Debates", AERA Open, vol. 6, no. 4, DOI: https://doi.org/10.1177/2332858420977208, 2020.

[12] A. Téllez-Valero, M. Montes, and L. Villaseñor-Pineda, "A Machine Learning Approach to Information Extraction", Lecture Notes in Computer Science, vol. 3406, pp. 539–547. DOI: 10.1007/978-3-540-30586-6_58, 2005.

[13] M. Almseidin, A. Abu Zuraiq, M. Al-kasassbeh Mouhammd, and N. Alnidami, "Phishing Detection Based on Machine Learning and Feature Selection Methods", International Journal of Interactive Mobile Technologies (iJIM), vol. 13, no. 12, DOI: doi:10.3991/ijim.v13i12.11411, 2019.

[14] A. A. Orunsolu, A. S. Sodiya, and A. T. Akinwale, "A predictive model for phishing detection", Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 2, pp. 232–247, DOI: https://doi.org/10.1016/j.jksuci.2019.12.005, 2022.

[15] S. Rawal, B. Rawal, A. Shaheen, and M. Shubham, "Phishing Detection in E-mails using Machine Learning", International Journal of Applied Information Systems, vol. 12, no. 7, pp. 21–24, 2017.

[16] I. Zatezalo, and I. Dunđer, "Online service for accessible machine learning of prediction models", Zbornik radova Međimurskog veleučilišta u Čakovcu, vol. 12, no. 2, 2021.

[17] A. Hannousse, and S. Yahiouche, "Towards benchmark datasets for machine learning based website phishing detection: An experimental study", Engineering Applications of Artificial Intelligence, vol. 104, ISSN: 0952-1976, DOI: https://doi.org/10.1016/j.engappai.2021.104347, 2021.

[18] J. Rashid, T. Mahmood, M. W. Nisar, and T. Nazir, "Phishing Detection Using Machine Learning Technique", First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 2020, pp. 43–46, DOI: 10.1109/SMART-TECH49988.2020.00026.

[19] J. Brownlee, "A Gentle Introduction to Ensemble Learning Algorithms", 2021, https://machinelearningmastery.com/tour-of-ensemble-learning-algorithms/

[20] Y. Zhang, J. Liu, and W. Shen, "A Review of Ensemble Learning Algorithms Used in Remote Sensing Applications", Applied Sciences, vol. 12, DOI: 10.3390/app12178654, 2022.

[21] R. Odegua, "An Empirical Study of Ensemble Techniques (Bagging, Boosting and Stacking)", Deep Learning IndabaX Conference, Ibadan, Nigeria, 2019, DOI: 10.13140/RG.2.2.35180.10882.

[22] K. Nikhil, D. S. Rajesh, and D. Raghavan, "Phishing Website Detection Using ML", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), vol. 7, no. 4, pp. 194–198, 2021.

[23] H. Kumar, A. Prasad, N. Rane, N. Tamane, and A. Yeole, "Dr. Phish: Phishing Website Detector", International Advanced Research Journal in Science, Engineering and Technology (IARJSET), vol. 8, no. 1, pp. 176–182, 2021.

[24] M. Gupta, and S. D. Pandya, "A Comparative Study on Supervised Machine Learning Algorithm", International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 10, no. 1, pp. 1023–1028, 2022.

[25] B. E. Boukari, A. Ravi, and M. Msahli, "Machine Learning Detection for SMiShing Frauds", IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, USA, 2021, DOI: 10.1109/CCNC49032.2021.9369640.

[26] K. M. Ting, "Confusion Matrix", in: C. Sammut, and G. I. Webb (eds), Encyclopedia of Machine Learning. Springer, DOI: https://doi.org/10.1007/978-0-387-30164-8_157

A Gamified Learning Scenario for Introducing Pupils to the Scientific Method through Experimentation and Programming

Dalia Kager*, Alexander Marinšek**
* Eugen Kvaternik Elementary school, Velika Gorica, Croatia
** KU Leuven, ESAT-WaveCore, Ghent, Belgium
dalia.kager@skole.hr
alexander.marinsek@kuleuven.be

***Abstract* - Elementary and high school students learn about numerous physical phenomena that make up our everyday life. However, physics is most often perceived as boring and only a handful of students remember what they learned in the classroom. Schools nowadays have the necessary equipment, such as micro:bit single board computers (SBCs), to digitalize physical experiments and make them more attractive to students. Due to limited teacher training and learning material availability, these modern tools are unfortunately often left out of the physics class. In this paper, we describe Galileo XR (Extended Reality) – a learning activity, ranked among the best three proposals at the Science is Wonderful! 2023 competition, wherein students' study everyday physical phenomena in a fun and inclusive way. The gamified scenario guides them through 6 workstations, corresponding to 6 physical phenomena, where they are embraced to make certain observations and record measurement data using different sensors on the micro:bit SBC, for which they are rewarded with encrypted clues regarding the final solution. Students are motivated to apply the scientific method throughout their quest to find the forgotten scientist, Galileo Galilei. Additionally, we evaluate the observations and student questionnaire results, obtained during Galileo XR's prototype launch.**

Keywords - microcomputer, physics, extended reality, gamified learning, deciphering

I. INTRODUCTION

When we talk about teaching science, technology, engineering, and mathematics (STEM) in schools, the emphasis is on the upper secondary level of science teaching. The results of the SECURE project [1], which deals with encouraging and preparing children for future careers in MST (Math, Science, Technology) from an early age, demonstrate that positive attitude towards all MST subjects decreases with age, and the biggest drops take place between age 8 and 11. Also, in all three MST subjects, activities in all age groups had the greatest impact on the students' positive attitude towards the subject, except for science at the age of 13 when 'themes' predominate. Although the number of people with higher education degrees is increasing, this increase has not occurred in the field of MST, and among students who graduate in these fields, women are extremely underrepresented. This is supported by the PISA 2018 Science results, in which only 6 EU countries achieved an average score of more than 500 [2]. In 2018, the underachievement rate stood at 21.7% in reading, 22.4% in mathematics and 21.6% in science. Over the 2009-2018 period, performance in science and reading deteriorated at the EU level [3]. Hence, children around the age of 10 require additional attention to nurture their interest in STEM during these pivotal years. In addition, the digital literacy of students should also be mentioned. Contrary to common belief, findings from the first two cycles of the International Computer and Information Literacy Study (ICILS) in 2018 [4] show that today's youth is not a generation of digital natives. Instead, according to the results, young people do not develop sophisticated digital skills just by growing up with digital devices. In 9 of the 14 EU Member States participating in ICILS, more than one third of students scored below level 2 on the ICILS CIL (Computer and Information Literacy) scale, which can be considered the threshold for failure in digital competence. This emphasizes the importance of focusing on the development of digital skills through formal education and enabling the use of digital tools as productive resources for learning and participation in the digital society.

There is general agreement among practitioners [5] that the way science is taught at the primary school level affects students' perception and attitude towards science and the subsequent inclination towards STEM subjects and careers. Elementary school teachers may have a central role; however, they often feel underprepared to approach STEM subjects in their classrooms. Their lack of self-confidence and ability to teach science is a major issue of concern in primary science education [6]. Accordingly, key strategies concerning teacher training have been devised to provide more effective STEM training in schools and foster the STEM culture [7]. A large-scale approach to STEM promotion has been rolled out by the European Commission (EC) in 2015 under the name Science is Wonderful (SiW), which bring together students, teachers, and researchers to discuss STEM topics on a more abstract level [9]. For the 2023 edition, the respective subject area researchers and teachers were given a rare opportunity to team up and to co-create learning material (experiment, quiz, game, or any other activity) that must have a real-life application [10]. Among all SiW proposals, three are chosen by the EC for translation into multiple languages, spoken within the EU, and for further professional development. The work at hand describes our submission to SiW 2023 – ranked by the EC among the best three proposals – *Galileo XR escape room: application of the scientific method in the discovery of physical phenomena with micro:bit programming, experiments and deciphering*

or shortly *Galileo XR: finding the forgotten scientist*. Our contributions are threefold:

- Describing the STEM learning scenario (LS), where students aged 7 – 15 improve their programming, physics, and auxiliary knowledge through the application of the scientific method.
- Evaluating the proposed activity's impact on students, participating in a prototype implementation, highlighting the positive aspects and the shortcomings that need to be addressed.
- Providing public access to the developed material through an online repository [11], hence, empowering educators to teach science at the primary education level by the International Standard Classification of Education (ISCED 1) and the lower secondary education level (ISCED 2) [12] from a more modern perspective.

The paper is structured as follows: Section II. describes the proposed Galileo XR LS, student questionnaire results from a prototype classroom implementation are presented in Section III., and Section IV. summarizes the work at hand.

II. GALILEO XR: FINDING THE FORGOTTEN SCIENTIST LEARNING SCENARIO AND ITS APPLICATION IN ELEMENTARY SCHOOL

A. About the learning scenario

Affordable single-board computers (SBCs), like the micro:bit, are equipped with several sensors, and are easy to operate [13], allowing students to increase their digital literacy through a multidisciplinary and interdisciplinary approach. It only takes a few visual blocks or lines of code to add an extra dimension to (often outdated) physics experiments. However, measurements alone do not add value in the absence of critical thinking. The glue between the two is the scientific method, which includes thorough observation, review of existing sources, formation of hypotheses, evaluation through measurement, and dissemination of results. By applying the so-called Inquiry-Based Science Education (IBSE) 5-step procedure, students gain an understanding of the studied physical phenomena from a hands-on perspective and are offered insight into the fundamentals of each research activity [14].

The *Galileo XR: finding the forgotten scientist* LS proposes a game-based educational process for students aged from 7 to 15 years that motivates them to critically evaluate basic physical phenomena (e. g. light), formulate a hypothesis, evaluate their opinion by programing an SBC to capture relevant measurement data, and disseminate the lessons learned among their peers.

B. Work process

The LS is named after Galileo Galilei, the father of the scientific method, who is the "forgotten" scientist that students are looking for. Teachers can choose one of the two proposed workflows – a *group activity* or a *competition activity* for students divided into a total of 6 teams, depending on how much time they have available for the implementation. Both proposals include the same activities: students participate in a game like a treasure hunt, face physics-based challenges, decipher clues and answer questions through the XR application, with the aim of finding the forgotten scientist (based on collected clues (ciphered facts) from his life), and prepare a presentation on the studied physical phenomena, the corresponding micro:bit sensor together with cases of its use in everyday life and clues from the life of Galileo Galilei. In the end, the teams disseminate their learning results and conclusions by presenting them in front of their peers.

Fig. 1 shows the group activity workflow in which one team of students works on one workstation, and together they uncover the forgotten scientist. Alternatively, the 6 teams can partake in a competition where each team must locate all the micro:bit beacons and attend all workstations with the goal of finding the forgotten scientist first, hence, repeating steps 2 – 6 for each workstation.

C. Treasure hunt

The treasure hunt activity takes advantage of the localization capabilities of the Bluetooth RF (Radio Frequency) transceivers found on the micro:bits. Each of the six beacons consists of a micro:bit and a QR code. The latter acts as a gateway to the XR quiz application, while the former continuously transmits a signal on a predefined radio group index (unique channel). The micro:bits first given to students are pre-programmed to measure the RSS (Received Signal Strength) of a particular beacon, allowing students to find its location as if they were participating in a treasure hunt. The beacons scattered throughout the school (or even the school yard, if plausible) will make the activity more fun and give the students much needed mobility.

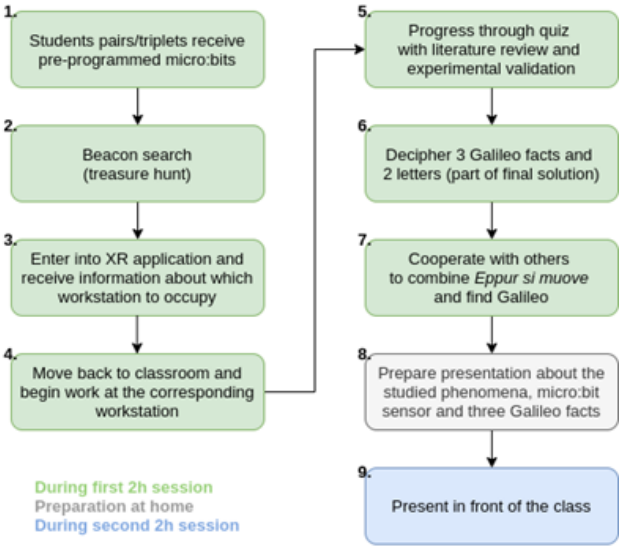


Figure 1. *Group activity* workflow

D. XR quiz application and measurement tables

An XR quiz application, pictured in Fig. 2, is designed for each workstation. It features 3 quiz questions about the corresponding physical phenomena and the micro:bit sensor's applications. Upon completion, the students are granted 3 encrypted clues about Galileo Galilei's life and 2 letters (part of the famous Galileo quote *Eppur si muove*) that make up the final key. The quiz is designed as an XR application, since this has proven as a particularly effective way of motivating students in the 7 – 15 age group [15].

In addition, each workstation comes with pre-designed measurement tables, shown in Fig. 3, to provide the students with more guidance on which measurements to conduct, and to give them a tangible resource that they can use for orientation when designing the final presentation.

E. Deciphering task

The students receive three encrypted clues from Galileo Galilei's life upon completion of the XR quiz at one workstation. These are encrypted using Caesar's cipher, a form of encryption where the letters in a word are substituted by their counterparts in a shifted alphabet. A straightforward, yet very mysterious encryption the first-time students encounter it. The encryption offset is fixed for all deciphering tasks, and students are provided with conversion tables, in addition to a deciphering procedure example. With this task, students are offered insight into the basics of information theory, useful for password and data encryption, as well as encoding in communication systems.

F. Workstation overview

Fig. 4 shows the 6 workstations that comprise Galileo XR, further detailed in the following paragraphs. Examples of students experimenting at workstations 1 and 6 are provided in Fig. 5 and Fig. 6, respectively.

1) Nikola who? (Magnetometer) The students learn about the Earth's magnetic field, how strong it is, and that not only cars are named after Nikola Tesla. They confirm that magnets have a magnetic field and discover that an electrical current will produce its own magnetic field (optional); hence, they understand what is meant with the word electromagnet. Hopefully, they will remember the lessons learned and not follow the compass (analog or on a smartphone) in the wrong direction the next time they go on an outdoor adventure.

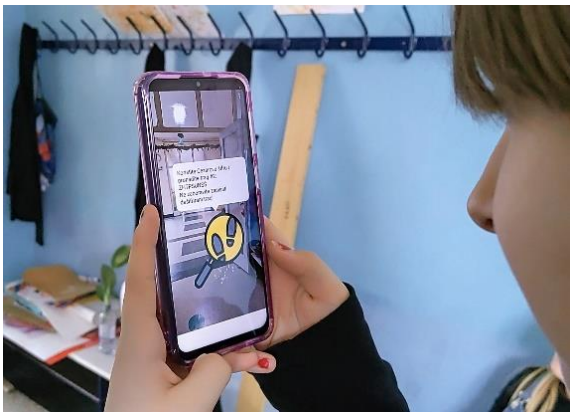


Figure 2. XR app for engaging students in solving quiz questions and deciphering clues from life of Galileo Galilei

Workstation: *You can run, but you can't hide - I still see you (radio communication)*

Remember to take notes from the material (micro:bit sensor and physical phenomena) at each station. You will need it later to create a team presentation. Notes write below.

Think, discuss with your peers and mark the answer - which material will block the radio signal best?

a) Normal paper
b) Baking paper
c) Aluminum foil

Now, use the measuring tables to note the measuring results and to check your assumption.

1. Wireless signal attenuation by different materials

	Normal paper	Baking paper	Aluminium foil
Received strength when obstructed			

Figure 3. Pre-designed measurement tables at each workstation

2) The laws of attraction (Accelerometer) The students grasp the concept of physical force through acceleration. Starting from gravitational acceleration, they evaluate the forces exerted on different parts of their bodies during physical activity and discover the shock absorbing role that joints play in our bodies. E.g., that's why we don't see blurry while running.

3) You are as cool as deodorant. Wait, what? (Temperature sensor) The students learn about the connection between gas temperature and density. They get to experience how warm air rises upwards, making way for colder air to sink towards the ground. Hence, they become acquainted with convection. They also get acquainted with an everyday phenomenon – compressed gas expansion (deodorant) – to get an idea of how gas expansion / compression can be used for heat conversion. For example, in refrigerators and air conditioning. Finally, they are introduced to the concept of efficiency by comparing the light intensity of a projector (through heat conversion) to the temperature of the projector's outlet air.

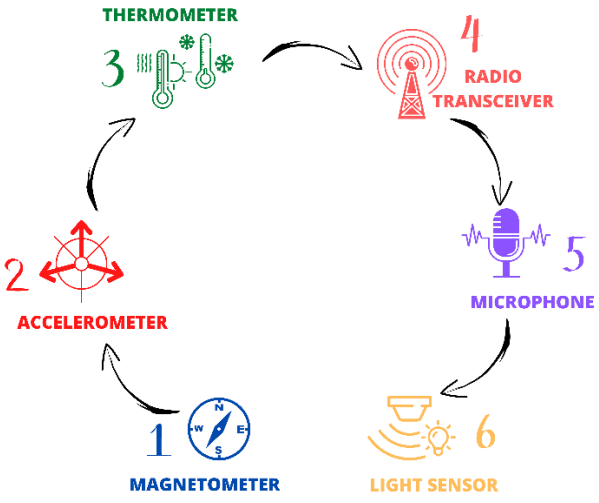


Figure 4. Workstation overview



Figure 5. Example of a classroom workstation setup

4) You can run, but you can't hide - I still see you (Radio transceiver) The students become acquainted with the propagation characteristics of electromagnetic waves – mainly their obstruction and amplification using different materials. They also grasp the concept of radar, letting them to understand how their ears enable them to locate sound sources, in addition to realizing how everyday technology, such as cruise control in cars or aircraft / vessel tracking works.

5) You are too loud! (Sound sensor – microphone) The students are introduced to machine learning (ML) through process characterization and knowledge-based inference. They use the micro:bit's microphone to determine the dependency of the measured noise in relation to the number of people clapping, which they can then use for occupancy estimation. Moreover, the students are introduced to sound level measurements and embraced to look further into more complex measurement techniques (frequency, tempo).

6) Roses are red, violets are blue, don't be so sad cause I'll enlighten you (Light sensor) The students understand that the perceived color of objects is the result of light reflection. For example, a leaf is green since chlorophyll reflects green light and not the other way around. Furthermore, they get a glimpse into the color spectrum by dissecting white light into red, green, and blue (RGB) color components, elaborating on sunlight's spectrum, and evaluating the micro:bit's frequency response at all three color components.

G. Finding the final key

Successful quiz completion also grants students with 2 letters of the final, most well-known, hint connected to the life of Galileo Galilei. The 6 letter pairs are: *EP PU RS IM UO VE*. This is Galileo Galilei's most widespread statement ("and yet it moves"), showing his determination in supporting scientific fact despite the lack of public support. A statement that the students should keep in mind during their educational journey and in their later careers. The students are provided with an online shuffler application where they can rearrange the letter pairs until they obtain the correct phrase *Eppur si muove*. Once the students start to assume that the forgotten scientist is Galileo Galilei, they can verify their assumption by entering his name as the passphrase in an online form.



Figure 6. Evaluating the micro:bit's frequency response at all three components (RGB)

H. Work process instructions

Student instructions are arranged in four tiers in accordance with the level of guidance they provide, giving teachers full liberty in how much support they wish to give, e.g., age groups or even particular teams based on the ambitions of the students. The teacher can, appropriate with the age and programming skills of their students, choose the materials according to which the students will program the micro:bits (in the visual MakeCode editor or in the Python editor) for performing experiments. The four tiers offer incrementally less guidance: tier 1 provides the students with full instructions on how to design the program using MakeCode or Python, including screenshots of the final solution, while tier 4 only describes the studied physical phenomenon and employed sensor.

Student instructions are designed using MS Word (printed version) and MS Sway (web version), shown in Fig. 7, both of which feature built-in *text translation*, allowing students of all nationalities to access the material in their native language. Moreover, Sway's adaptability – adjustable contrast levels and the screen reader functionality – makes it applicable to student groups with special personality traits and visual impairments.

III. CLASSROOM IMPLEMENTATION RESULTS

The Galileo XR escape room prototype was tested at the Eugen Kvaternik Elementary School in Velika Gorica during the *22nd Winter School of Informatics* from January 2 to 5, 2023, organized by the Association of Technical Culture and Informatics Pedagogues, Velika Gorica, for a

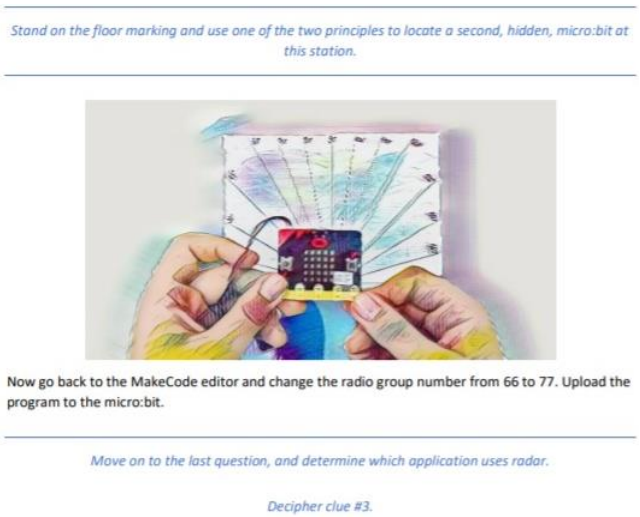


Figure 7. Student instructions are available online and on paper

total duration of 6 full hours (6 x 60 minutes). Workshops for the Winter School of Informatics are usually designed and conducted by ICT or technical culture teachers to enable students to spend quality free time during the winter holidays. The total number of students who participated in testing the learning scenario prototype (contest version) over 4 days was 18 (6 girls and 12 boys) aged 7 to 14 (ISCED 1 and ISCED 2 level). 15 students (4 girls and 11 boys) participated in the evaluation [16] of the activities carried out during the learning scenario testing. Fig. 8 shows the students' evaluation responses regarding the activities carried out and the content of the learning scenario itself. The overall conclusion obtained through the evaluation form, including other questions and statistical information (e.g., age), is as follows:

- 60% of students found the tasks at the workstation's were always fun and interesting, 26.7% found the tasks often fun and interesting, while 13.3% of them found that the tasks were difficult (ISCED level 1 students),
- 20% of students occasionally needed the teacher's help during work, and 6.7% of them needed help all the time (ISCED level 1 students),
- 46.7% of students said that it was never difficult for them to work in a team, and for 40% of students it was rarely difficult, which is supported by the fact that 53% of respondents selected the sound workstation as their first choice as the most entertaining and most instructive, because they cooperated with each other through joint clapping, as they explained in the free form comments,
- 40% of students believe that all team members contributed equally to solving tasks.
- 60% of students felt comfortable and relaxed during work,
- in reference to the clarity and transparency of the written instructions for micro:bit programming, 73.3% of students indicated that the instructions were always clear, 13.3% thinks that they were often clear, and 13.3% decided that micro:bit programming instructions were occasionally clear,
- 60% of students declared that the instructions for conducting the experiments were always clear to them, 13.3% consider that the instructions were often clear, and 26.7% students declared that the instructions were occasionally clear to them,
- 80% of students think that deciphering with Caesar's code was fun and easy.

Students emphasized their own activity and engagement in work, experimenting and research, multidisciplinary application of micro:bit coding through physical computing in solving real-life problems, as well as team collaboration. Their impression of the whole workshop is that they felt comfortable and learned a lot while having fun.

After implementing the learning scenario prototype, analysing the evaluation results and other data collected by

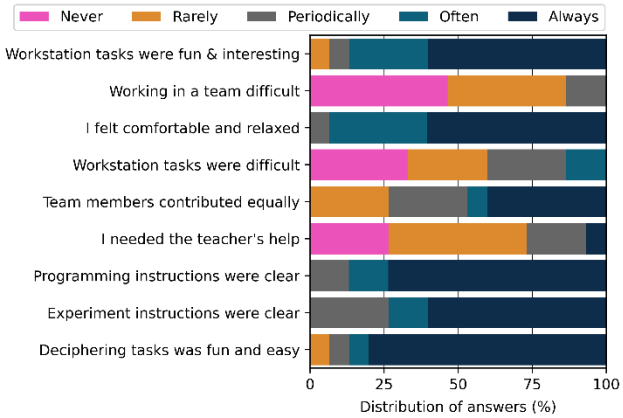


Figure 8. Students responses in evaluation of the Galileo XR activities, Likert scale

observing the work process, we concluded that the activities and instructions should be organized and adapted for the three student age groups, since they have different knowledge and concentration levels that will require different amounts of support. An exception are students with special needs (students with learning and / or developmental difficulties, i.e., gifted students). For such students, the teacher can choose materials adapted for a higher or lower age group, to meet their needs and ensure inclusion in the organized activities.

Age group 7 – 9:

- work in pairs or groups of 3 members and only the group activity is carried out (one team – one workstation),
- program the micro:bit using the visual programming language MakeCode,
- receive tier 1 or tier 2 instructions (full guidance or assembling pre-defined code blocks).

Age group 10 – 12:

- carry out activities in one of the two proposed versions (group or competition), according to the teacher's decision,
- create programs for micro:bit in the visual MakeCode editor according to the instructions for either of the 4 programming skills levels.

Age group 13 – 15:

- carry out activities in one of the two proposed versions (group or competition), according to the teacher's decision,
- create programs for micro:bit in the Python MakeCode editor according to the instructions for one of the 4 programming skills levels; optionally, the teacher can also choose a visual programming language for this age group, but only for programming skills levels 3 and 4.

IV. CONCLUSION AND FUTURE WORK

The work at hand proposes a multidisciplinary STEM learning scenario for improving the programming knowledge of students, aged 7 – 15, during the performance

of fundamental physical experiments, while also acquainting them with the scientific method. Gamification [17] is employed to motivate the students to actively participate, as well as to freshen up their minds through moderate physical activity. Using micro:bit SBCs as the principal learning platform, we enable a cost-efficient way for schools to incorporate the proposed STEM learning scenario into their curriculum. A prototype implementation in the classroom shows that the benefits of this STEM learning scenario are multiple: the engagement of students and their motivation are increased by applying the Inquiry Based Learning (IBL) method that increases student autonomy; students are active during work, which achieves a more significant pedagogical effect; the activities in which they participate relate to real situations and problems, so students build not just knowledge and understanding, but higher-order skills needed in everyday life - ability to synthesize and evaluate, creative thinking and critical thinking, i.e. competencies needed in the 21st century. Physical computing, which involves using specialized hardware to interact with the real world, i.e. writing programs to control input and output devices, such as in this case the buttons, sensors and LEDs on the micro:bit, promotes a broader perspective and deepens students' programming skills with theoretical knowledge of how hardware works. In this way, digital skills are linked to other subjects, such as science in this case, showing students how these skills can be applied to solve real-world problems. Furthermore, the use of microcomputers in the classroom helps bring learning to life through a holistic computing experience by combining hardware and software, linking it to subjects outside of computing, and evening out gender inequality by allowing students to be expressive and creative. Based on the results of the student evaluation and the data collected by observing the implementation process of the learning scenario prototype, we concluded that future work should address the limited inclusion of ISCED level 1 students by arranging the otherwise written instructions into short video excerpts or animations, that would better capture their full attention. Furthermore, ISCED level 2 students would benefit from a preliminary study of the addressed phenomena in physics class, preceding the learning scenario activities, thereby acquiring the necessary prior knowledge needed to formulate a well-thought-out hypothesis.

ACKNOWLEDGMENT

This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 861222 (MINTS).

REFERENCES

[1] European Commission, SECURE project report summary, 2013, accessed: 12.1.2023. [Online]. Available: https://cordis.europa.eu/project/id/266640/reporting

[2] A. Schleicher, *"PISA 2018 Insights and Interpretations,"* (Organization for Economic Cooperation and Development (OECD) report), 2019.

[3] European Comission, Key Competences and basic skills, accessed 12.1.2023. [Online]. Available: https://education.ec.europa.eu/education-levels/school-education/key-competences-and-basic-skills

[4] European Comission, The 2018 International Computer and Information Literacy Study (ICILS), 2018, accessed 12.1.2023. [Online]. Available: https://education.ec.europa.eu/sites/default/files/document-library-docs/icils-2018-policy-note.pdf

[5] S. Boon Ng *"Exploring STEM competences for the 21st century,"*, 2019, UNESCO International Bureau of Education, IBE/2019/WP/CD/30 REV, accessed 15.1.2023. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000368485

[6] C. Murphy et al. *"Primary Science Teacher Confidence Revisited: Ten Years on,"* (Educational Research), 2008, DOI: 10.1080/00131880701717289

[7] A. Nistor et al. *"STEM in primary education,"* (Scientix Observatory report), 2019.

[8] A. Nistor et al. *"Science, Technology, Engineering and Mathematics Education Practices in Europe,"* (Scientix Observatory report), 2018.

[9] *Marie Skłodowska-Curie Actions, Science is Wonderful!*, accessed 12.1.2023. [Online]. Available: https://marie-sklodowska-curie-actions.ec.europa.eu/science-is-wonderful

[10] *Science is Wonderful! Competition*, accessed 12.1.2023. [Online]. Available: https://event.scienceiswonderful.eu/event/competition

[11] D. Kager and A. Marinšek, "Galileo XR: finding the forgotten scientist, learning scenario material repository," [online] accessed 4.2.2023. Available: https://zenodo.org/record/7607988#.ZCiV2XvP1PY

[12] European Comission, *"The Structure of the European Education Systems 2018/19: Schematic Diagrams,"* (Eurydice Facts and Figures), Luxembourg: Publications Office of the EU, 2018.

[13] *Details of the latest micro:bit hardware revision*, accessed 15.1.2023. [Online]. Available: https://tech.microbit.org/hardware/

[14] AEMASE, "Inquiry-based Science Education: Primer to the international AEMASE conference report," Berlin: All European Academies, 2014.

[15] M. Meccawy *"Creating an Immersive XR Learning Experience: A Roadmap for Educators,"* (MDPI Electronics), 2022.

[16] D. Kager. "Galileo XR: finding the forgotten scientist escape room evaluation," accessed 24.1.2023. [Online]. Available: https://cutt.ly/GalileoXR_SiW_evalu

[17] M. Kalogiannakis et al. "*Gamification in Science Education. A Systematic Review of the Literature*," (Educ. Sci.), 2021, DOI: 10.3390/educsci11010022

The Reality of Digital Teaching of History in Croatia

M. Hajdarović
1st Elementary School Čakovec, Čakovec, Croatia
miljenko@hajdarovic.com

***Abstract* - The educational system has traditionally been slow to accept innovations and advances in technology used in everyday life. The high penetration of networking, digitalization, and the use of devices is ordinary daily life in the 21st century. They also have a severe impact on the everyday life of schools. Teaching history has not remained immune to the fruits of the IT revolution. In the Croatian framework, several attempts to reform history teaching must be added, rounded off by the School for Life project and pandemic attempts at online and combined teaching in recent years. The use of digital technologies and adapted methods of teaching history in that period went from indignation followed by individual utopian attempts to broader acceptance of the use of ICT. However, what does this wider acceptance mean for teaching history? In qualitative research using the structured interview method, 11 history teachers (advanced in the profession) were interviewed who, in the initial contact, expressed themselves positively towards the use of digital tools. The results point to possible future directions for the education of history teachers and specific strategies for using digital tools.**

Keywords* - the *digital revolution, use of ICT, teaching history

I. INTRODUCTION

The use of ICT in education and the so-called digital revolution are no longer novelties about which opposed utopian and dystopian works would be written. Fortunately, the first decades of the 21st century are increasingly filled not only with experimentation but also with quality research on the use of ICT in schools. As a result, educational policymakers have accepted ICT not only as a supporting part of the educational system but also as a necessary part of the education of new generations.

Teaching history has not remained immune to the fruits of the IT revolution. In publications and magazines on didactics and methods of teaching history, the use of ICT was shyly mentioned already in the late 1960s and early 1970s. Only at the turn of the century and thanks to the translation of foreign literature did Croatian history teachers receive professional literature that gave instructions and examples of the use of the then-current computer technology and the Internet. Scientific research on history teaching using digital media and tools is rare [1]. In the last twenty years, in Croatian scientific journals covering the topic of history teaching, the application of ICT has only been written about in rare cases. The most significant impetus for using ICT was the COVID-19 pandemic, during which almost all teachers used the available technology in different ways to teach. In the Croatian case and for elementary schools, the already implemented e-School (and School for Life) project, in which all elementary school students received tablets for use, certainly has great merit in the high use of ICT. According to the project plan, free Internet should be available to students in addition to tablets in schools. There is no research or report on the use of tablets, but in communication through teacher groups on social networks, it could be concluded that they are mostly not used or used adequately [2].

One positive development influencing the broader acceptance of digital history teaching is upgrading digital parts of history textbooks from 2019 until today. In that series, all history textbooks, in addition to the usual digitized textbook (usually in PDF format with links to external multimedia content), also received various additional forms of interactivity. Depending on the textbook publisher, additional interactivity is in digital educational objects or digital educational content. *Digital educational objects* vividly display a particular phenomenon in a digital format. *Digital educational content* is didactically-methodically designed units that can be part of a lesson or replace a regular lesson [3]. The new textbooks were issued after the publication of the new History curriculum, which only slightly encourages the use of digital technologies in the recommended teaching methods. It mentions that digital resources (websites, journals, tools, archives, collections, and repositories) can also be used. The new textbooks were issued after the publication of the new History curriculum in 2019, which only slightly encourages the use of digital technologies in the recommended teaching methods. It mentions that digital resources (websites, journals, tools, archives, collections, and repositories) can also be used. It can only be assumed that the authors of the history curriculum relied on the fact that students should acquire part of the ICT skills according to educational expectations Cross-curricular topic "Use of information and communication technology."[2]. The two pandemic years brought more online classes, whether synchronous or asynchronous teaching. Teachers have increased digital communication, and the number of materials passed on to students. Due to the increased Internet use among students, there is also a growing problem with information verification. The problem of spreading false and inaccurate information is certainly contributed by leading statesmen who use fabricated data from the past to strengthen their arguments. There is agreement amongst media commentators that the spread of so-called fake news poses a severe threat to a democratic society. Therefore, history

educators should help stem the rising tide of misinformation. [4] Although the issue of using artificial intelligence is not new in the academic community, school practitioners have only begun to open up the topic. Tools like ChatGPT and Midjourney and even the general issue of artificial intelligence have not yet been mentioned in the official training programs for Croatian history teachers. With remarks about possible abuses in the matter of plagiarism, Chat GPT is perceived mostly positively among early adopters in recent research.[5]

The status of history as a school subject in the countries of the European Union is not uniform, nor is the procedure for acquiring the qualifications of history teachers. Regarding the state of digital teaching in Croatia, examples from other countries can be presented, but they are not comparable. In discussions about the relationship of teachers to technology, a romanticized notion is still used, according to which the new generation of students are digital natives while teachers are digital immigrants. [6] The scientific community warns that the rapid spread of fake news and digital media that transmit historical falsifications shows that the so-called digital natives intuitively use digital technology but lack digital literacy. Recent literature calls this current period post-factual or the era of alternative theories. [7] This challenge of digital literacy also puts increasing pressure on Croatian history teachers. The use of digital technology supports today's prevailing paradigm of constructivist teaching, which, by placing students in the leading role, should contribute to developing critical and historical thinking while expanding digital literacy. [8]

II. RESEARCH

In qualitative research using the structured interview method, 11 history teachers were interviewed (working in elementary and high school). Questionnaire testing was avoided to give respondents more freedom to clarify their work and think about the use of digital technologies. With a questionnaire, the possible answers could be too conditioned by the limitations set by the researcher. Two criteria for selecting respondents were an initially positive attitude towards using digital technology in teaching and professional advancement. The criterion of professional advancement was chosen because it testifies to the professional activity of the respondent. The gender, age of the respondents, and their work experience are not relevant if we consider that advancement in the profession is the primary variable for selecting respondents.

The opening research question to starting the interviews with respondents was: How do you use digital technology in your teaching?

Emphasized follow-up questions were:

- what are the advantages of using digital technology,

- what are the limitations and/or disadvantages of using digital technology,

- do you think that today's generation of students are digital natives who use digital technologies independently and meaningfully,

- in what ways can teachers guide students for the meaningful use of digital technologies in learning history?

The questions were comprehensive to avoid contamination of the answers or influence on the respondents. For example, a possible area of inquiry could have mentioned learning management platforms, 3D models, or augmented or virtual reality. However, such questions could lead the respondents to give possible, desirable answers to show the situation better.

From the collected answers, the following common characteristics and a few outstanding developments can be distinguished.

When answering the initial question, respondents usually began to explain in which part of the teaching process they most often use digital technology. All respondents use computer presentations as a teaching tool for presenting content. This includes presenting new facts and using presentations to ask questions and tasks. A minor part (3) of the respondents expressed a more remarkable ability to use computer presentations with which they create interactive games (such as association, quizzes, puzzles, and gradually discovering graphics). Computer presentations are used in all parts of the lesson - from the introduction through the central part to repetition or even self-evaluation at the end. According to most respondents (9), using other applications, websites, multimedia, simulations, or 3D models depends primarily on the historical topics and content covered in the lesson. When choosing whether to use digital technology or "classical" textbook materials, most respondents prefer textbook materials or materials they have prepared themselves for a particular lesson. A large part of the respondents (7) leaves the use of digital technologies to the students in preparation for the presentation of certain content or during the repetition. In practice, this means that teachers give students the opportunity or task to present a part of the topic, some historical source (text or picture) for an upcoming class, or in another creative way, take over part of the creation of the teaching material. For this purpose, students use different presentation tools available on the market. Finally, students independently create quizzes or interactive tasks for repetition. Among the respondents, there is a consensus that for specific teaching topics, there is suitable digital material for teaching. At the same time, there is no acceptable digital content for a significant part of the topics. In doing so, they highlight the problem of foreign languages, the inappropriateness of content for a specific age of students, and, for example, the duration of certain content if video clips, documentaries, or feature films are used.

To the related question of why they give preference to classic teaching materials, the respondents indicated: uncertainty as to whether the technology available to students will be helpful (frequent tablet failures and insufficient quality of internet connection in classrooms), loss of time for technology preparation, more complex control of individual tablets in terms of are the students doing what they should or taking the opportunity to play or surf the Internet. Some respondents noted that it is extremely difficult or almost impossible to achieve equality, i.e., the availability of digital technology for

every student in the class. For example, some students cannot work with the help of digital technology. In that case, classic materials must be provided for them in the teaching process, leading to a division of opportunities in the class.

All respondents pointed out the possibility of better presentation of historical events, personalities, culture, and historical stories in general with the help of digital multimedia as the most crucial advantage of digital technology in teaching history. Less than half of respondents (5) mentioned that using well-designed and interactive digital educational content enables students to independently research and draw conclusions about the past. Only one respondent recognized the possibility for additional research of gifted or especially interested students in publishing houses' existing digital educational content.

Regarding the limitations of digital technology, the respondents pointed out the issue of available time as the most significant limitation - insufficient time for the quality preparation of teachers (who must familiarize themselves with some digital content, sufficiently master the use of some digital technology or materials) as well as the big problem of lack of time in the weekly number of hours as well as total time during the entire school year. A small number of respondents (4) cited that students are continuously exposed to screens and various digital tools as a disadvantage of digital technology. That is why they often and intentionally insist on using paper materials, plain text, or images. Searching the Internet for historical content presents a significant challenge to students because they need prior knowledge to assess the value or veracity of the information found. In the absence of time, teachers rarely decide to spend time gradually and studiously take students through the research process. Unfortunately, the problem of fake news was not explicitly mentioned among the respondents.

Respondents were divided on whether today's generation of students is genuine digital natives who skillfully use digital technologies. Most respondents agree that students regularly use digital technologies. At the same time, six respondents believe that this use is skillful and that students use technology correctly. Another part of the respondents believes that students, although they often use digital technology, do not know how to use it for functional educational purposes. At the same time, they point out that most students, for example, refrain from critically evaluating information found on the Internet during research.

Due to the inadequate use of digital technologies, the respondents agree that students should be taught how to use digital technology adequately. One of the interviewees highlighted the question of whether students should learn how to use digital technology only in history lessons meaningfully or should it be part of the basic knowledge that students would acquire from the very beginning of school education (the question of the need and expediency of computer science classes in schools).

Among the respondents, there is a consensus about how to use the possibilities of digital technology meaningfully. They emphasize project teaching that is carried out individually or in groups. Such teaching would be carried out during several teaching hours, and teachers should prepare enough tasks that are carried out in steps or in a specific order. Students would independently research most of the content, present it to others, and create a review. Only one respondent tentatively explained the flipped classroom model as the preferred method of using digital technologies.

III. CONCLUSION

The period of teacher insecurity in using ICT is a thing of the past. Teachers are familiar with the possibilities of digital technology, know that there are different tools, and use them depending on personal preferences. Digital technology is used throughout the teaching process, often not continuously but at appropriate teaching stages.

This research was conducted as a private initiative without the support of educational authorities and is therefore limited in the number of participants. The small number of respondents can only indicate the direction of thinking, but conclusions cannot be generalized for the entire population of history teachers in Croatia. The interview research method is time-consuming because it includes the time needed to organize the interview and the analysis time. Institutional support for conducting research of this type is necessary. The research showed that among teachers, there is a need to question the use of digital technologies. However, it cannot be concluded that teachers are skeptical about using digital technologies, but there is much uncertainty about acceptable use and the need to create adequate teaching materials. When listing digital tools used during history classes, the number of named tools leads to additional questions. For example, how does this amount of tools positively or negatively affect the attention of students who adapt to various tools through all school subjects?

It is interesting and encourages further research that no respondent mentioned the use or possibilities of LMS (such as Moodle). The teachers were exposed to LMS while preparing to introduce the new subject curriculum, i.e., during the recent education system reform. It is also important to note that none of the interviewees mentioned artificial intelligence during the conversation. However, we can only assume that this topic has yet to enter the teaching population, but whether the students will be the forerunners who will, for example, use possibilities like ChatGPT in their little school research or homework. During the research, respondents did not mention gamification, virtual and augmented reality, or active use of 3D models.

Educational authorities and creators of educational policies should systematically and continuously conduct research that would include the teacher population. In addition, there is an urgent need to examine the effects of all forms of teaching, not just teaching with digital technology.

REFERENCES

[1] Demantowsky, M., & Pallaske, C. (Ed.). Geschichte lernen im digitalen Wandel. De Gruyter Oldenbourg. 2015.

[2] Hajdarović, M. History teaching in Croatia in the jaws of the Internet. 1st International Online Scientific Conference ICT in Life - Conference proceedings, 184–193. 2022.

[3] Hajdarović, M., Glučina, M., & Turk-Presečki, V., Metodički paket za poučavanje povijesti (2. izd.), Profil Klett. 2021.

[4] Bekker-Nielsen, T. Fake News from the Past – Lessons For the Future. Public History Weekly, 2021(10). https://doi.org/10.1515/phw-2021-19199.

[5] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), 15. https://doi.org/10.1186/s40561-023-00237-x

[6] Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. On the Horizon, 9(5), 1–6. https://doi.org/10.1108/10748120110424816

[7] Peters, M. A., Rider, S., Hyvönen, M., & Besley, T. (2018). Post-truth, fake news. Springer.

[8] Tribukait, M. (2020). Digital learning in European history education: Political visions, the logics of schools and teaching practices. History Education Research Journal, 17(1). https://doi.org/10.18546/HERJ.17.1.02

Increasing Students' Motivation and Improving Outcome by Changing the Course Examination from Test to Project on a University Course

Daniel Bele, Danijel Kučak, Đani Pašić
University College Algebra
Gradišćanska 24, Zagreb, Croatia
daniel.bele@algebra.hr, danijel.kucak@algebra.hr, dani.pasic@racunarstvo.hr

***Abstract* – This paper measures the impact of changing the university course examination from test to project. It was witnessed that the student's motivation was low, as most students took the exam at the last deadline, and their grades were deficient. In order to increase students' motivation, grades, and earlier delivery, the paper authors changed the approach. Instead of forcing students to learn for the exam on particular dates, students were given a set of projects to deliver regularly. This paper is trying to answer the following questions: Does the usage of project-based course results in scoring better grades (Q1)? Does using a project-based course increase the student's motivation to deliver earlier (Q2)? The paper analyzes scores and examination dates for four generations of students – the first two generations were examined through the tests, and the other two were examined through individual projects. The course's learning objectives remained the same, as well as the learning objectives for the courses that precede the course being examined. The paper shows the following results: The average grade for the students rose from 3.31 to 3.74 (Q1), while the average value of exam date changed from 3.01 to 2.20 (Q2).**

Keywords – examination; tests; project-based learning; students' motivation

I. INTRODUCTION

Governed by the well-being of our students, as teachers, paper authors often feel obligated to create content and approach that can light the spark in the students. In order to improve students' achievements at the university, it is often required to use alternative methods and procedures of teaching [1]. The usual organization of the educational process and the philosophy oriented toward convergent thinking that there is "only one right answer" often inhibit creativity rather than develop it [2]. Therefore, it is vital for teaching to be modern and creative in order to increase the motivation for learning and participation of the students in the teaching process [3]. Modern education programs should start with the teachers, to take the role of reflective practitioners and action researchers to create changes to improve the educational system from the inside out [4].

The above presents a challenge and motivation factor to the paper authors and the students. The generations of students today should be governed in learning to solve problems by different research methods in controlled conditions with teacher supervision to become active participants and not passive observers [5]. For example, modern teaching methods like research and interactive teaching, which represent a departure from "ex-cathedra" and aim at the active participation of the students and independent application of the knowledge [6], are proven to create greater interest in students. Also, since the classical approach to studying creates much stress for students, "the teachers should support students' autonomy and facilitate students' academic flow in order to prevent students' burnout" [7].

One of the approaches to be taken is using case studies that are first shaped by teachers' view of the subject and later built upon by expectations, backgrounds, and the student's abilities which help to develop ability in analysis and thinking [8]. As another example, 'design thinking' introduces practical problem-solving that puts the student in the center of the problem, often leading to differentiation and even innovation [9]. Further, applying gamification principles in teaching has shown very good results [10] because "gamification utilizes game mechanics to transform the learning experience into a game, while game-based learning integrates games into the learning process" [11]. Finally, many studies have proven project-based teaching to encourage students' activity and satisfaction that, can increase the student's success [12]. It is even more true for programming courses because students tend to fear the abstract nature of programming. After all, it is tough to realize the purpose of the knowledge gained [13]. It is the situation where project-based teaching and examination shine, which this paper should also prove.

Based on the previously mentioned facts, the current education system requires significant changes to prepare students for future work challenges. Examining the knowledge gained should focus more on the student's creative expression and less on the reproduction of the content [14].

II. EXAMINATION METHODS

As there are many different approaches in programming teaching methods [15], the students should also be provided with a better way of their examination. This study should prove that students examined by tests often need to be more motivated and focused on the minimum of the content that ensures them a passing grade. That presents a considerable problem because classes and exams should reflect their future jobs; therefore, being unmotivated for the content

means being unmotivated about the work they will soon perform as programmers.

Traditionally, the course Accessing data from program code tested students on midterms, where students were obligated to answer the tests related to the classes' content. The problem with this approach is that students learn for the test only and, therefore, easily find patterns examined on the written tests and concentrate on the minimum to pass. This fact has been proven throughout the years because it was witnessed that most students took the test at the last deadline and scored poorly, just about enough to pass.

In order to increase students' motivation and improve the outcome, it is decided to change the course examination to project-based. For that matter, students should focus more on the content and on applying their understanding of the concept practically. The skills generated in this process become part of their programming knowledge and, therefore, must be remembered. Paper authors are aware of that fact from their approach to programming – every concept applied to a solution of a problem becomes part of our toolbox used to fix future problems. In that sense, the grade achieved by the student reflects their learned skill because this skill becomes practical and applicable. Project-based learning is already recognized in Croatia's current curricular reform [16].

Moreover, this paper tries to prove that when students engage individually, their interest rises. They are more motivated to deliver quality projects they can eventually put in their portfolios to show their future employers. In that sense, it simulates the natural world, which is not artificial but realistic.

While tests create much stress manifested in a particular moment [17], the project gives a lot more time to investigate the problem and research other solutions to the problem to apply them as their own. The project has to be well thought out, planned, and programmed in a lot more time, giving students a chance to refactor and change wrong ideas into new cognition.

Finally, throughout the years of teaching, authors concluded that programming is not a field to be studied in a conventional sense but to be constructed and built upon solving problems. The more students do programming, the better they understand the concepts and the reasoning behind them. It is impossible to understand the need for an approach or a concept if the student does not face a problem that is hardly solved without it. Because of that, it is of great importance to face as many problems as possible in lifelike projects to become more versatile and ready for the problems in real projects they should be prepared for [18]. On the other hand, tests make them memorize concepts without deepening the true nature of the problems the concepts should solve.

III. RESEARCH

This research aimed to evaluate the impact of changing the course examination from test to project on the university course Accessing Data from Program Code (ADfPC). For years, it was witnessed that the student's motivation for the course mentioned above is relatively low, which was proven by the fact that most students took the exam at the last deadline, and their grades were deficient. In order to increase students' motivation and grades, the paper authors decided to change the approach. Instead of forcing students to learn for the exam on particular dates, students were given a set of projects to deliver periodically, which should have encouraged them on a more individual approach and greater interest. In essence, this paper aims to answer the following questions.

- Does the usage of project-based course results in scoring better grades? (Q1)
- Does using a project-based course increase the student's motivation to deliver earlier? (Q2)

The study involved 419 students from 6 generations who attended the ADfPC course and took a course exam in the spring semesters starting from the academic year 2016/2017 to 2021/2022. Table I shows the number of students included in the research according to the academic year in which they studied and took the course. One hundred seventy-three students took the classic exam, and 246 took the project assignment. For this research, we will name the 173 students who took the classic exam as G1 and 246 students given by project assignment as G2. G1 and G2 were taught by the same lecturer, according to the same materials and methodology. The only difference was that G1 students took the classic exam, while G2 students were given the exam as a project assignment.

Table I. Distribution of students in the academic year

	16/17	17/18	18/19	19/20	20/21	21/22
Students	51	56	66	58	98	90
Is project-based	NO	NO	NO	YES	YES	YES

At the beginning of the study, the aim was to check whether there was a statistically significant difference in the prior knowledge of Object Oriented Programming (OOP) and Introduction to Databases (ItD) for both groups, students given the definitive test and students given by project assignment. It has to be determined whether to exclude the difference in prior knowledge for experimental and control groups as an essential factor in the exam results. Based on previous teaching experience, the knowledge from the second-year courses OOP and ItD is a crucial prerequisite for successful learning and adoption of course content ADfPC, which students enroll in the third year of study. For this purpose, data on students' grades from the control and experimental groups were processed. All students involved in the experiment passed the OOP and ItD courses. Grades at higher education institutions in Croatia are as follows:

- Grade 2 - Sufficient
- Grade 3 - Good
- Grade 4 - Very good
- Grade 5 - Excellent

Table II shows the distribution of students' grades in the OOP course.

Table II. Distribution of students' grades on the OOP course

	Grade 2	Grade 3	Grade 4	Grade 5	Total
G1	27 (15.6%)	59 (34.1%)	47 (27.2%)	40 (23.1%)	173
G2	42 (17.1%)	80 (32.5%)	76 (30.9%)	48 (19.5%)	246
Total	69 (16.5%)	139 (33.1%)	123 (29.4%)	88 (21.0%)	419

Thus, the total sample consists of N = 419 students, of which 173 students from G1 are in the control group and 246 from G2 are in the experimental group. The $\chi 2$ test was used for statistical data processing. A null hypothesis was set:

H_0 = the G1 and G2 groups do not differ statistically significantly according to the obtained grades from the OOP course distribution.

Assuming no significant difference exists between the control and experimental groups, the proportion of specific scores observed should be the same in both groups. Based on that, a table of expected frequencies was obtained. According to χ^2 test specification, the number of degrees of freedom (labeled v) is 3.

The limit value of χ^2 with 3 degrees of freedom at the significance level $p<0.05$ is 7.815. Since the calculated value of χ^2 in the experiment is smaller (1.3745), the H_0 was accepted, i.e., there is no statistically significant difference in the distribution of the obtained grades on the OOP course between G1 and G2 groups.

The same method also processed data on grades from the ItD course. Tables III shows students' grades' distribution on the ItD course.

Table III. Distribution of students' grades on the ItD course

	Grade 2	Grade 3	Grade 4	Grade 5	Total
G1	37 (21.4%)	49 (28.3%)	47 (27.2%)	40 (23.1%)	173
G2	50 (20.3%)	65 (26.4%)	66 (26.8%)	65 (26.4%)	246
Total	87 (20.8%)	114 (27.2%)	113 (27.0%)	105 (25.1%)	419

The limit value of χ^2 with 3 degrees of freedom at the significance level of $p<0.05$ is 7.815. Since the calculated value of χ^2 in the experiment is smaller (0.8413), the H_0 was accepted, i.e., there is no statistically significant difference in the distribution of the obtained grades on the ItD course between G1 and G2 groups.

In this research, we were interested in whether the change in how knowledge is tested will lead to statistically significant differences in students' final results and whether this change will motivate students to take the exam earlier.

Q1: By analyzing the grades of students from group G1, as well as from group G2, we concluded that the average grade of students from group G1 is 3.31, while the average grade of students from group G2 is 3.74, which made a significant difference in favor of the G2 group.

Table IV. t-test results

	G1	G2
Mean	3,312139	3,735772
Variance	0,97177	1,223776
Observations	173	246
Pooled Variance	1,119831	
Hypothesized Mean Difference	0	
df	417	
t Stat	-4,03457	
p(T<=t) two-tail	6,51E-05	
t Critical two-tail	1,965669	
H_0: The difference between G1 and G2 group is not statistically significant		
t = -4.03457, t critical two-tail = 1.9657, df = 417, $p = 6.5x10^{-5}$, $p < 0.05$		
We do not accept H_0 hypothesis		

Two-tailed t-test has been made to check if there is a statistically significant difference between those values. The results are shown in Table IV. H_0 hypothesis cannot be accepted based on the calculated p-value, so an alternative hypothesis has been accepted to show that the difference between values is statistically significant.

Q2: At Algebra University College, for each course, students have 5-time slots spread over approximately 12 months during which they can take the exam. We can mark these terms for research with 1, 2, 3, 4, and 5. We are interested in the average value of the term in which students from group G1 passed the course AdfPC, as well as students from group G2, and we are interested in whether the differences between these significantly different values.

The research was conducted after all 419 students had passed the exam on one of the five possible dates. Two-tailed t-test has been made to check if there is a statistically significant difference between those values. The results are shown in Table V. H_0 hypothesis cannot be accepted based on the calculated p-value, so an alternative hypothesis has been accepted to show that the difference between values is statistically significant.

Table V. t-test results

	G1	G2
Mean	3,00578	2,203252
Variance	1,982525	1,550357
Observations	173	246
Pooled Variance	1,728613	

Hypothesized Mean Difference	0	
df	417	
t Stat	6,151699	
p(T<=t) two-tail	1,8E-09	
t Critical two-tail	1,965669	
H_0: The difference between G1 and G2 group is not statistically significant		
t = 6.1517, t critical two-tail = 1.9657, df = 417, p = $1.8x10^{-9}$, p < 0.05		
We do not accept H_0 hypothesis		

Analyzing the dates for which the students took the exam, we concluded that the mean value of the exam date of students of group G1 was 3.01, while the mean value of the exam date of students of group G2 was 2.20. That is a visible change in the sense that students who worked on the project assignment, on average, took the exam earlier.

Based on many years of experience teaching ADfPC, the authors used grades from OOP and ItD courses to measure the student's prior knowledge as the prerequisite courses for the ADfPC. It would be interesting to show by statistical processing the correlation between the output results of these courses and the output result of the ADfPC course, which can be an exciting area to explore in the future. The plan for the future is to expand the scope of research to other courses.

IV. CONCLUSION

For years, it has been noticed that students' motivation is low when they collect points by taking classic exams. The idea behind this paper was to positively affect the observation by introducing project assignments instead of well-known exams.

The study targeted the course Accessing Data from Program Code (ADfPC), on which project assignments have been used for three consecutive years. The study involved 419 students from 6 generations. The first three generations (G1 group) have 173 students that took the classic exam, and the remaining generations (G2 group) have 246 students that took the project assignment. It was shown that there is no statistically significant difference between the groups concerning their level of prior knowledge. The authors of this research were interested in two questions:

- Does the usage of project-based course results in scoring better grades? (Q1)
- Does using a project-based course increase the student's motivation to deliver earlier? (Q2)

By processing the results of the groups, the obtained results showed the following:

- The average grade of students from group G1 is 3.31, while the average grade of students from group G2 is 3.74, which made a significant difference in favor of group G2.
- Students from group G1 passed the course in an average of 3.01 exam periods, while group G2 in 2.20 exam periods, which is a visible change in the sense that students who worked on the project assignment, on average, took the exam earlier.

The research showed a significant correlation between introducing project assignments in the ADfPC course, scoring better grades, and increasing students' motivation to deliver projects earlier.

The study's results could encourage other researchers to investigate further the impact of introducing project assignments to different university courses, primarily to motivate students to practice more.

REFERENCES

[1] Močinić, S.N. (2012) 'Active teaching strategies in higher education', Metodički obzori, 7(2012)2(15), pp. 97-105.

[2] Bognar, L. (2012) Kreativnost u nastavi. Napredak (Zagreb), 153, pp 9-20.

[3] Debeljak De Martini, A., Ristić, R. & Potočnjak, J. (2020) 'Nastavne metode u edukaciji zdravstvenih djelatnika', *Journal of Applied Health Sciences = Časopis za primijenjene zdravstvene znanosti*, 6(2), 297-304

[4] Jukić, T. & Nemeth Jajić, J. (2022) 'Mixed-Methods Approach in Action Research: Challenges and Possibilities in Higher Education' in: Blažević, I., Bulić, M. & Tomaš, S. (ed.)ICFE 2022 – 14th International Conference in the Field of Education Education Reforms: The Role of Policy and Educational Research - Book of Abstracts.

[5] Vinduška, I. (2016) 'Inovativne edukacijske metode i pristupi u nastavi stručnih predmeta u srednjim medicinskim školama', Master's thesis, University of Zagreb, School of Medicine

[6] Cigan, V. & Šlogar, H. (2012) 'Istraživanje stavova studenata o metodama poučavanja na visokim čilištima u cilju razvoja poduzetničke kompetencije', Učenje za poduzetništvo, 2(1), 179-190

[7] Ljubin Golub, T., Rijavec, M. & Olčar, D. (2020) Student flow and burnout: The role of teacher autonomy support and student autonomous motivation. Psychological Studies, 65, 145-156

[8] Ozretić Došen, Đ. (2002) The use of the case method for effective education and business training. U: Vukonić, B. & Čavlek, N. (ur.)Rethinking of Education and Training for Tourism.

[9] Pap, M., Vdović, R. & Baletić, B. (2019) Design Thinking metoda u znanstvenom istraživanju, edukaciji i poslovnoj praksi. *Prostor : znanstveni časopis za arhitekturu i urbanizam*, 27 (2 (58)), 334-347.

[10] Kučak, D., Bele, D., Pašić, Đ.: Climbing up the leaderboard: an empirical study of improving student outcome by applying Gamification principles to an object-oriented programming course on a university level. In: 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 527–531.

[11] Kučak, D., Pašić, Đ., Mršić, L. (2023). An Empirical Study of Long-Term Effects of Applying Gamification Principles to an Introductory Programming Courses on a University Level. In: Vasant, P., Weber, GW., Marmolejo-Saucedo, J.A., Munapo, E., Thomas, J.J. (eds) Intelligent Computing & Optimization. ICO 2022. Lecture Notes in Networks and Systems, vol 569. Springer, Cham

[12] Lukša, Ž., Kerić, E., Radanović, I. & Garašić, D. (2009) Utjecaj aktivnog učenja projektima i terenskom nastavom na uspješnost, motivaciju i stavove učenika. U: Vijtiuk, Nada, Šiljković, Željka (ed.)Zbornik sažetaka, 3. međunarodna konferencija o naprednim i sustavnim istraživanjima ECNSI-2009.

[13] Konecki M., Lovrenčić S. & Kaniški M., "Using real projects as motivators in programming education," 2016 39th International Convention on Information and Communication Technology,

Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2016, pp. 883-886

[14] Dubovicki, S. (2019) 'Professors' Views on the Relationship Between the Curriculum of the Teacher Education and the Development of Students' Creativity Based on the Delphi Method – Longitudinal Research' in: Dziurzyński, K. & Duda, E. (ur.) What is new in the Field of Education?. Warsaw, Publishing House of The Maria Grzegorzewska University, 61-81.

[15] Mohorovičić, S. & Strčić, V. (2011) 'An Overview of Computer Programming Teaching Methods' in: Hunjak, T., Lovrenčić, S. & Tomičić, I. (ed.)Proceedings of the 22nd Central European Conference on Information and Intelligent Systems.

[16] Ilić, Š., Šenjug Krleža, A. & Knežević, Ž. (2019) 'Project-Based Learning in German Language Teaching – Teachers' Opinions and Perceptions' in: Kolar Billege, M. (ur.)International Scientific and Art Conference Contemporary Themes in Education - CTE. Book of abstracts.

[17] Šimić, N. (2003) 'Neke psihofiziološke reakcije na ispitni stres', doktorska disertacija, Filozofski fakultet, Zagreb, Hrvatska.

[18] Letina, A. & Diković, M. (2012) Problem-Based Teaching vs. Programmed Teaching: Challenges for the Future of Education. U: Pixel (ed.)International Conference: The Future of Education, Conference Proceedings.

Model for Assessing e-Learning Courses Considering Multiple Visual and Technical Indicators

M. Bankovska*, D. Borissova**, K. Rasheva-Yordanova*
* University of Library Studies and Information Technologies, Sofia, Bulgaria
** Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences, Department of Information Processes and Decision Support Systems, 1113 Sofia, Bulgaria
m.bankovska@unibit.bg, daniela.borissova@iict.bas.bg, k.rasheva@unibit.bg

***Abstract* – The rapid progress of information and communication technologies and their penetration into the world of education has led to a gradual transition from conventional learning to modern e-learning. More and more universities and businesses have switched to web-based learning environments. Integrating electronic media into learning improves productivity by influencing both the effectiveness and efficiency of the learning process and the motivation of learners. This article aims to systematize the main indicators for evaluating e-learning media, considering multiple indicators. The identified important indicators are used to formulate a mathematical model for the overall assessment of the e-learning courses' performance. The proposed model could be applied to every e-learning course and is independent of the field of its application and the used learning environment. The obtained results demonstrate the applicability of the formulated mathematical model.**

Keywords – e-Learning; indicators for evaluation; ICT; mathematical model.

I. INTRODUCTION

The rapid progress of information and communication technologies and their entering the world of education leads to a gradual transition from conventional face-to-face learning to learning accessible through the use of digital media. In recent years, we have witnessed a trend towards an expanded use of learning web-based solutions by both schools, universities and businesses worldwide [15, 16, 24, 32, 37]. According to Guri-Rosenblit, e-learning is related to the use of electronic media for different learning proposals that can happen in the classroom environment and even replace the face-to face virtual meetings [32].

Using innovative approaches [8] and media as a communication environment for learning [22, 23] raises the interest of students in the learning process [26] creates a fruitful learning atmosphere, and stimulates students to acquire more knowledge in a shorter period of time. This on its part leads to raising the efficiency of the learning process. Consequently, according to some researchers [5, 25, 31] the use of instructional media is the most effective way to be used so that all learning objectives can be achieved.

In its broadest sense, e-learning media refers to anything that can be used to stimulate the thoughts, feelings, attention, abilities or skills of students so that it can promote the learning process in an electronic environment. According to Munawar, the application of e-learning media also leads to the improvement of learning competence [25]. A study conducted by Wahid revealed that the use of e-learning based learning media caused feelings of pleasure, quickly understood, and is more attractive than conventional learning media [3]. In addition, it is worth mentioning that an essential part of e-learning is represented by e-tests, which are also part of the educational process [9, 10, 17]. This means that when evaluating the overall performance of e-courses, e-tests have to be taken into account as an integral part of them [18].

The main aspiration of any educational institution offering e-learning should be aimed at automating learning and transferring knowledge to learners, while ensuring the effectiveness and efficiency of the educational process. Building an e-learning course, however, is an intersection of various competencies that a course designer should possess. In this regard, Michael Allen argues that design and development of good e-learning is a complex undertaking, and it requires content knowledge and expertise in a wide range of areas, including text composition, illustration, testing, instruction, interactivity design, user interface design, authoring or programming, and graphic design [21]. Not knowing or neglecting any of the mentioned elements can lead to deviations from the main goal of teaching and learning.

In order to be adequate to the needs of students, e-learning media should meet specific requirements. This, in turn, will ensure readiness to ensure the three characteristics: effectiveness, efficiency, and satisfaction with the learning process. It is extremely important to develop a mechanism for evaluating e-learning courses. Factors relevant to the above are: (1) technical and functional aspects of the educational platform used, (2) substantive aspects of the learning content, and (3) design approaches for its presentation. To assess all of these specific e-learning course requirements, a mathematical model based on multiple indicators needs to be formulated. This model should be applicable to any e-course and be independent of the area and field of application of e-learning media. Taking into account the formulated

problem, the current article aims to systematize the leading indicators for evaluating e-learning media and derive a generally applicable model for assessing e-learning courses.

The rest of the article is organized into several self-contained sections. Section II is focused builds a framework of metrics for evaluating e-learning systems, including various functional and technological requirements as well as constraints imposed by the design of the environment. In Section III a mathematical model based on the identified indicators for the assessment of e-courses is formulated. Section IV describes the results and analysis of conducted numerical testing of the proposed model, and the conclusions are drawn in Section V.

II. METRICS FOR ASSESSING E-LEARNING SYSTEMS

In our research, the most significant two groups of criteria for evaluating e-learning courses will be discussed namely: (1) visual performance and (2) technical performance. In the following lines, these two groups of criteria will be described along with as well as the evaluation indicators belonging to each of them.

A. Visual Indicators for Assessing e-Learning Courses

Evaluating the visual design in our research focuses on the evaluation of graphic design applied in e-learning media. Graphic design is important for conveying information in a clear and concise way, helping students learn more effectively and remember information better in a more engaging way. Imbalance and inconsistent use of colors, graphics, and typography, have all been found to reduce learning [2].

Based on Wynand et al. [4] we can claim that designing e-learning media is directly linked to the effective information exchange between the student and the educational platform, which directly affects the user's perceived experience of the e-learning environment [4]. A general rule is that the design should be tailored to the age group and educational needs of the learners, be as clean and simple as possible. It is often evaluated on user experience and usability.

A number of researchers have focused their research on the impact of graphic design on users, but few have addressed the evaluation of visual design and its impact on learning. According to Anglin et al. [14] the study of the effects of visual elements on learning is becoming important due to the increasing number of graphics used in educational materials. According to FAO [13] when conducting e-learning, the use of graphics (including illustrations, photos, diagrams and icons) is appropriate, as their use leads to the facilitation of learners.

Along with the use of graphics, by visual indicators we refer to the indicators of color, typography, graphics, contrast, as well as the compositional relations between them, creating harmony. Evans and Thomas [28] theorize that elements of visual design are line, texture, shape, space, which is why we add them to the analysis and evaluation factors. The arguments for the selection of these factors in our model can be found in the following:

- **Color** bring a clearer structure to the text, but also can be used as an aid that helps facilitate learning. Color can direct and hold the attention of learners.
- **Graphics** enhance many aspects of learning [30] including recognition, recall, comprehension and problem-solving. Graphics help to organize, interpret or transform textual information [11], which is of utmost importance for the learning process.
- **Lines** are used to add structure to a composition, to frame, and to separate information. In combination with text, lines can be used to highlight and emphasize a word, phrase or paragraph. In the context of e-learning media and the presentation of learning content, lines are particularly significant.
- **Visual texture** concerns some graphic techniques that draw attention to an element on the page or serve as a background to it. Texture is an essential element for any design, which is why we don't rule it out as a metric to be evaluated.
- **Shapes** are the basis of design, which on the other hand are the basis of every drawing, painting and graphic design [1]. Forms are three basic types – geometric, natural (or organic) and abstract ones. Regardless of which type they fall into, it is important to specify that shapes in e-learning media help separate and connect learning material in a visually appealing and organized way.
- **Space** in graphic design is the distance around the area between the different design elements involved in the composition. Space can be used to both separate and connect elements in visual design. Wider spaces separate elements from each other, and narrower spaces connect elements to reveal the connections between them. Overlapping elements maximizes their connection. The use of spaces in e-learning media will allow easier following of elements in the course, give visual space and create emphasis when needed.
- **Typography** is the art of arranging letters and text in a way that makes e-learning media legible, clear and visually appealing to learners. This includes elements such as font style, appearance, and structure.
- **Harmony** represents the balance between design decisions and design elements, and also the balance between typography and layout. The goal is that all elements of e-learning media design complement each other and form a harmonious relationship. According to Norman [7], designers must consider both cognitive and emotional aspects of design, as human cognition and emotion are closely related. All design elements have an influence in this direction at the same time.
- **Contrast** occurs in the juxtaposition of opposite elements – large versus small or dark versus light. Using contrast can highlight important design elements. Contrast is easily achieved with color, but can also be achieved with texture, type, and

graphic elements. An element brought to the fore and defined as primary by means of contrast is easier to find in the particular environment. This, in turn, facilitates the perception of information and its learning.

- **Composition** is used to direct the viewers [29]. Properly constructed composition affects the easy perception of visual design4 in e-learning environments. It refers to the overall organization of the fields – comparison of small to large elements, those grouped by common features and relating to a common theme, bringing to the fore the most important and significant objects for learning.

We can argue that the correct use of visual design indicators can lead to higher learning outcomes in the following several directions:

a) Attracting the attention of learners to a specific element of the content;

b) Creating analogies between new and existing content, respectively between information based on old and new knowledge;

c) Supporting the understanding of the basic concepts of the training course;

d) Motivating learners by working with more interesting and interactive materials.

Visual design evaluation indicators are directly related to technical indicators and will be discussed in the next section.

B. Technical Indicators for Assessing e-Learning Courses

Based on the fact that a web-based learning media platform is subject to evaluation, in the category of technical indicators, indicators such as: 1) code quality; 2) security; 3) speed; 4) usability; and 5) accessibility.

- When evaluating the **code quality**, it is necessary to carry out an analysis for the presence of errors or technical problems that interfere with the work with e-learning media. In short, an assessment of the internal quality of e-learning media is needed and in particular their adherence to good design principles. Quality can be represented as "totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs" [19]. Adapting the results from the research by Kanellopoulos et al. [38] to this paper needs, when evaluating the code quality the following will be taken into account a) functionality, concerned with what the e-learning media does to fulfil students' needs; b) reliability, evaluating e-learning media's capability to maintain a specified level of performance; c) usability, assessing how understandable and usable the e-learning media is; d) efficiency, evaluating the capability of the e-learning media to exhibit the required performance with regards to the amount of learning resources needed; e) maintainability, concerned with the e-learning platform's capability to be modified; and f) portability, measuring the e-learning media's capability to be transferred across environments.

- **Security** indicator is essential with a view to securing personal data on the one hand and carrying out unauthorized breaches in the media on the other. An e-learning media vulnerability can be represented as a security flaw, problem or weakness found in software code that can be exploited by an attacker [20] to cause harm to the stakeholders of a e-learning platform [34]. It has been proven that for an application that is not reviewed for security flaws, there is a 100% probability of problems related to unregulated accesses and other breaches [27]. Therefore, we consider it of utmost importance that this indicator participates in our evaluation model. Availability of a suitable SSL certificate, use of secure access to site administration, use of an up-to-date CMS version (if available), etc. are evaluated.

- **Speed** is another indicator influencing the work with the system, which in turn directly affects the learning process. In terms of usability level, speed is one of the most fundamental characteristics [6, 12]. As a general rule, media and the related resources should work equally well at different Internet speeds, with requests being executed quickly and accurately. Factors such as having a CDN to serve the content or other type of optimization implementation are evaluated, as well as the overall loading speed of the resources.

- **Accessibility** indicator is an important indicator required by international web standards, referring to correctly developing and coding web-based platforms so disabled people can use them. According to W3C [36] accessible design improves overall user experience and satisfaction, in a variety of situations, across different devices. Accessibility ensures equal access to knowledge through educational media, which is why we include it in the list of indicators to be evaluated.

- **Usability** is another important indicator related to the extent to which e-learning media can be used by users to achieve educational goals with effectiveness, efficiency, and satisfaction. Availability for responsive design, availability of error visualization, search functionality and resource review are subject to evaluation; resource loading speed and last but not least – presence of intuitive structure and navigation [35]. Elements such as adaptability of content, changing navigation methods, readability of information and usability on touch devices should be monitored and evaluated here.

In conclusion, it can be argued that code inspection ensures the detection of vulnerabilities and flaws leading to delays in the execution of requests to and from the server side. Accessibility and usability are partially linked and depend both on other technical indicators and on some indicators of visual design. All of them are significant for the preparation of the summary evaluation of the technical indicator.

III. MODEL FOR ASSESSING E-LEARNING COURSES CONSIDERING MULTIPLE INDICATORS

The described above indicators could be used to assess different e-learning courses realized in various environments. To considers simultaneously all of these indicators, the following mathematical model is proposed:

$$C_q = \max\{\alpha \sum_i^v w_i^v e_{iv} + \beta \sum_j^t w_j^t e_{jt}\} \quad (1)$$

$$\alpha + \beta = 1 \quad (2)$$

$$\sum_i w_i^v = 1 \quad (3)$$

$$\sum_j w_j^t = 1 \quad (4)$$

where $q = \{1,2,\dots,Q\}$ represents set of e-learning courses, coefficient α expresses the importance of criterion related to visual performance while coefficient β expresses importance of criterion related to technical performance, the coefficients w_i^v and w_j^t express relative importance between indicators related to the visual and technical performance, e_{iv} and e_{jt} represents evaluations of *i-th* course about *v-th* criterion related to the visual performance, and corresponding *i-th* course about *t-th* criterion related to the technical performance.

IV. RESULTS ANALYSIS AND DISCUSSION

The proposed groups of criteria about visual and technical indicators along with mathematical model (1) – (4) could be used to assess different e-learning courses. In order to demonstrate the applicability of this approach, some preliminary tests have been done. A case study of two e-learning courses of basic of web programming is used to show the core idea of the described approach. In the assessment of these courses, four visual indicators and four technical were used as shown in Table I.

TABLE I. GROUPS OF CRITERIA AND INDICATORS FOR ASSESSING OF E-LEARNING COURSES

No	Criteria Groups		Indicators
1.	Visual performance	1.1	Composition and harmony
		1.2	Color
		1.3	Typography
		1.4	Contrast
2.	Technical performance	2.1	Security
		2.2	Usability
		2.3	Accessibility
		2.4	Speed

To be able to assess the overall performance of these e-learning courses based on the indicators in Table I, it is needed to determine coefficients for the relative importance between visual indicators and respectively between technical indicators. For the particular testing experiment, both visual and technical indicators are considered with equal importance with values $w_i^v = 0.25$ and $w_j^t = 0.25$. According to the formulated mathematical model (1) – (4), along with these coefficients, it is necessary to evaluate each e-learning course on all visual and technical indicators. These estimates and coefficients are summarized in Table II.

TABLE II. EVALUATIONS OF THE VISUAL AND TECHNICAL CHARACTERISTICS OF THE E-LEARNING COURSE

Groups of Criteria	Relative importance between indicators	Course-1	Course-2
Visual performance	$w_1^v = 0.25$	0.75	0.90
	$w_2^v = 0.25$	0.88	0.90
	$w_3^v = 0.25$	0.75	0.85
	$w_4^v = 0.25$	0.90	0.80
Technical performance	$w_1^t = 0.25$	0.90	0.88
	$w_2^t = 0.25$	0.85	0.87
	$w_3^t = 0.25$	0.95	0.85
	$w_4^t = 0.25$	0.88	0.85

The visual and technical performance evaluations are done using a scale from 0 to 1 where the bigger value means better performance of the evaluated criterion. Here should be mentioned that the proposed model could be integrated with other measurement approaches. For example, the Absolute Category Rating (ACR) method is used for quality tests. This method can be implemented using a five-point overall quality rating scale (*Bad, Poor, Fair, Good, Excellent*), which is the most widely used scale, or if required using a higher nine-point scale. Regardless of whether the five-point or the nine-point scale will be used, it is necessary to normalize the scores obtained in this way to the interval from 0 to 1. This is mandatory to provide dimensions similar to the other coefficients of the proposed model in order to obtain reliable results.

Last but not least, two additional coefficients related to visual and technical performance need to be determined. Two different cases are investigated using different values for the importance of visual and technical performance. Case-1 considers visual and technical performance with equal importance, while Case-2 gives more importance to technical performance indicators over visual ones. These two cases are shown in Table III.

TABLE III. VALUES OF COEFFICIENTS FOR VISUAL AND TECHNICAL CHARACTERISTICS IMPORTANCE

	Visual performance	Technical performance
Case-1	$\alpha = 0.5$	$\beta = 0.5$
Case-2	$\alpha = 0.4$	$\beta = 0.6$

The input data from Table II and Table III are used to verify the applicability of the proposed modeling approach (1) – (4) to assess the e-learning courses' performance according to the used visual and technical indicators. Obtained preliminary results of the conducted assessment of two similar e-learning courses are visualized in Fig. 1.

According to the results of Case-1 where equal values are used for the coefficients for visual and technical performance expressed by α and β ($\alpha = \beta = 0.5$) along with the equal importance between indicators of these two groups of indicators show that Course-1 has a better performance than Course-2. This is due to the obtained overall assessment equal to 0.8625 for Course-2 compared the value 0.8575 for overall assessment of Course-1 (see Fig. 1).

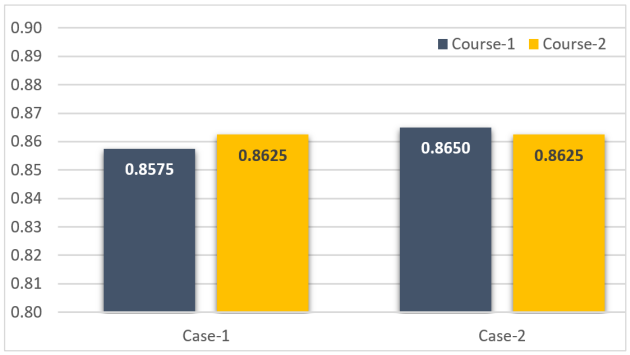


Figure 1. Assessment of e-learning courses using different importance values of visual and technical performance

The Case-2 simulates the situation where more attention is given to the technical indicators as a whole (see Table III). In this case, better performance is obtained for Course-1 with the overall assessment value equal to 0.8650 compared to the value of 0.8625 for Course-2 performance.

It is interesting to see how these courses are evaluated separately on visual and technical performance alone. This scenario is visually illustrated in Fig. 2.

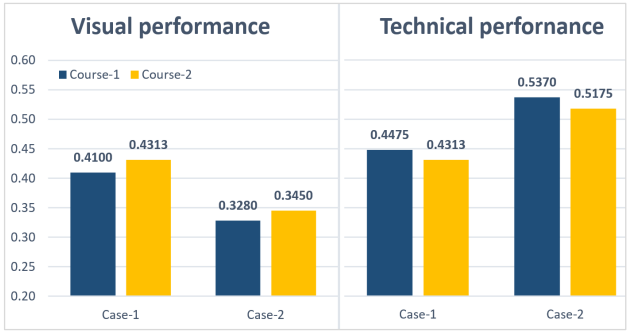


Figure 2. Cources assessment in respect to: a) visual performance; b) technical performance

Taking into account visual assessment separately from the technical one, it could be seen from Fig. 2 that the Course-2 has better visual performance compared to Course-1. Considering only the technical evaluation, it is easy to find that Course-1 has a better performance in both case: assessment of 0.4475 compared to Course-2 with value 0.4313 for case-1 and 0.5370 vs. 0.5175 for case-2 (see Fig. 2).

Although the visual parameters are better on Course-2 compared to Course-1 in both cases, and the technical parameters of Course-1 are better in both cases, the overall evaluation using the introduced coefficients α and β show different ranking for both cases (see Fig. 1).

It should be noted that the conducted preliminary tests do not consider situations where the coefficients w_i^v and w_j^t for relative importance between indicators for visual and technical performance differ. It is interesting to show how these similar sources will be performed when all identified indicators within this article are involved. These activities are planned as future investigations along with extending the proposed model for assessing e-learning courses by using additional indicators.

V. CONCLUSION

Digitization and related processes affect all areas of our lives, having a significant impact on e-learning. Along with these processes, the Covid-19 crisis has stimulated the transformation of various learning materials into e-learning courses. Therefore, the need to evaluate the offered electronic courses naturally arises. That is why the core idea of the current article aims to formulate a generally applicable mathematical model for the assessment of e-learning courses. The distinguishing characteristic of the proposed modeling approach is its independence of the area application of e-learning courses. The formulated model is capable to evaluate e-learning courses' performance taking into account simultaneously two types of criteria related to visual and technical indicators. It should be noted, that the proposed model uses of two types of coefficients to express: 1) the relative importance between visual and technical criteria, and 2) the relative importance between indicators that are part of visual and technical criteria. The conducted testing demonstrates the applicability of the proposed approach. The formulated mathematical model can be easily modified using additional indicators or a group of indicators. In this regard, it is interesting to involve an additional point of view namely the users' point of view. This evaluation composed of three groups of criteria is planned as future activities.

ACKNOWLEDGMENT

This work is supported by the Bulgarian National Science Fund through the project “Mathematical models, methods and algorithms for solving hard optimization problems to achieve high security in communications and better economic sustainability”, KP-06-H52/7/19-11-2021.

REFERENCES

[1] A. C. Bradshaw. Effects of Presentation Interference in Learning with Visuals. Journal of Visual Literacy, 23 (1), 41–68, 2003.

[2] A. E. Arntson. Graphic Design Basics. University of Wisconsin – Whitewater. Thomson Wadsworth, 2007.

[3] A. Wahid, Najiburrahman, K. Rahman et al. Effectiveness of Android-Based Mathematics Learning Media Application on Student Learning Achievement. Journal of Physics: Conference Series, 1–7, 2020. https://doi.org/10.1088/1742-6596/1594/1/012047.

[4] B. Wynand; H. Gelderblom, and A. van der Merwe. Why designers responsible for websites of large organisations disregard basic web design principles. CONF-IRM 2016 Proceedings. 4. http://aisel.aisnet.org/confirm2016/4.

[5] C. Muali, S. Minarti, M. Taufik Qurohman, Haimah. Analysis of metacognitive capability and student learning achievement through edmodo social network. Journal of Physics: Conference Series, 1175(1), 2019. https://doi.org/10.1088/1742-6596/1175/1/012150.

[6] C. Xilogianni, F. R. Doukas, I.C. Drivas, D. Kouis. Speed Matters: What to Prioritize in Optimization for Faster Websites. Analytics, 1, 175–192, 2022. https:// doi.org/10.3390/analytics1020012.

[7] D. A. Norman. Emotion & Design: Attractive things work better. Interactions Magazine, ix (4), 36–42, 2002.

[8] D. Borissova, D. Keremedchiev, G. Tuparov. Multi-criteria model for questions selection in generating e-education tests involving gamification. TEM J. Technol. Educ. Manag. Inform. 9(2), 779–785, 2020. https://doi.org/10.18421/TEM92-47.

[9] D. Borissova, N. Buhtiyarov, R. Yoshinov, M. Garvanova, I. Garvanov. Integrated Models-Driven Framework to Generate Various Online and Print Tests. In: Saeed, K., Dvorsky, J. (eds), Computer Information Systems and Industrial Management, LNCS, vol. 13293, pp. 316–329, (2022).

[10] D. Borissova, T. Ivanova, N. Buhtiarov, N. Naidenov, K. Rasheva-Yordanova, R. Yoshinov, M. Garvanova, I. Garvanov. Application of information technology in the teaching of mathematics when study of 2D geometric shapes. In: 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO 2022), pp. 638-643, 2022.

[11] D. Guo, S. Zhang, K. Landau Wright, and E. M. McTigue. Do You Get the Picture? A Meta-Analysis of the Effect of Graphics on Reading Comprehension. AERA Open, 2020.

[12] E. Fundingsland, J. Fike, J. Calvano, A. Raja, D. Lai, S. Silacci, S. He. Website usability analysis of United States emergency medicine residencies. AEM Educ. Train., 5, e10604, 2021.

[13] FAO. 2021. E-learning methodologies and good practices: A guide for designing and delivering e-learning solutions from the FAO elearning Academy, second edition. Rome. https://doi.org/10.4060/i2516e, ISBN 978-92-5-134401-9.

[14] G. J. Anglin, H. Vaez, and K. L. Cunningham. Visual Representations and Learning: The Role of Static and Animated Graphics. In D. H. Jonassen (Ed.), Handbook of research on educational communications and technology (pp. 865–916). Lawrence Erlbaum Associates Publishers, 2004.

[15] H. Amin, M. Shehata, S. Ahmed. Step-by-step guide to create competency-based assignments as an alternative for traditional summative assessment. MedEdPublish., 9(1):120, 2020.

[16] H. Djidu, S. Mashuri, N. Nasruddin, A. E. Sejati, Rasmuin, R., Ugi, L. E., and Arua, A. L. Online learning in the post-Covid-19 pandemic era: Is our higher education ready for it? Jurnal Penelitian Dan Pengkajian Ilmu Pendidikan: E-Saintika, 5(2), 139–151, 2021. https://doi.org/10.36312/esaintika.v5i2.479.

[17] I. Kostadinova, K. Rasheva-Yordanova, I. Ivanov, P. Petrova. Automated system for generating and validation a learning tests. In: ICERI2017 Proceedings, pp. 414–424, 2017.

[18] I. Mustakerov, and D. Borissova. A conceptual approach for development of educational Web-based e-testing system. Expert Systems with Applications, 38(11), 14060–14064, 2011. https://doi.org/10.1016/j.eswa.2011.04.214.

[19] ISO/IEC 9126. Software Engineering – Product Quality Int'l Standard, 2003.

[20] K. Dempsey, P. Eavy, and G. Moore. Automation Support for Security Control Assessments. Technical Report. Technical Report NISTIR 8011, National Institute of Standards and Technology, 2017.

[21] M. Allen. Michael Allen's guide to e-Learning (second edition): Building Interactive, Fun, and Effective Learning Programs for Any Company. Hoboken, USA, Wiley, 2016.

[22] M. Bali, Musrifah. The Problems of Application of Online Learning in the Affective and Psychomotor Domains During the Covid-19 Pandemic. Jurnal Pendidikan Agama Islam, 17(2), 137–154, 2020. https://doi.org/https://doi.org/10.14421/jpai.2020.172-03.

[23] M. H. Shehata, A.P. Kumar, M. R. Arekat, M. Alsenbesy, M. al Ansari A, Atwa H, et al. A toolbox for conducting an online OSCE. The Clinical Teacher, 18(3), 236–242, 2020. https://doi.org/10.1111/tct.13285.

[24] M. Muhammad Mushfi El Iq Bali, H. Baharun, A. Madanibillah, C. Muali et al. Innovative Learning Media Based on e-Learning in the New Normal Era. Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore, March 7-11, 2021.

[25] M. Munawar. Supervisi Akademik: Mengurai Problematika Profesionalisme Guru di Sekolah. Al-Tanzim: Jurnal Manajemen Pendidikan Islam, 3(1), 135–155, 2019. https://doi.org/10.33650/al-tanzim.v3i1.522.

[26] N. Sudarwati, Rukminigsih. Evaluating E-Learning as a Learning Media: A Case of Entrepreneurship E-Learning Using Schoology as Media. International Journal of Emerging Technologies in Learning (iJET), 13(09), 269–279, 2018. https://doi.org/10.3991/ijet.v13i09.7783.

[27] OWASP Project. OWASP Code Review Guide 2.0 https://owasp.org/www-pdf-archive/OWASP_Code_Review_Guide_v2.pdf, 2017.

[28] P. Evans, M. Thomas. Exploring the elements of design (3rd ed.). Delmar, NY: Cengage Learning, 2013.

[29] R. E. Wileman. Visual Communicating. Englewood Cliffs, NJ: Educational Technology Publications, 1993.

[30] R. N. Carney, J. R. Levin. Pictorial illustrations still improve students' learning from text. Educational Psychology Review, 14(1), 5–26, 2002. https://doi.org/10.1023/A:1013176309260.

[31] S. Ahmed, M. Shehata, M. Hassanien. Emerging faculty needs for enhancing student engagement on a virtual platform. MedEdPublish, 9(1):75, 2020.

[32] S. Guri-Rosenblit 'Distance education' and 'e-learning': Not the same thing. Higher Education, 49(4), 467–493, 2005.

[33] S. Islam, H. Baharun, C. Muali, M. Ghufron, M. Bali, M. Wijaya, I. Marzuki. To Boost Students' Motivation and Achievement through Blended Learning. Journal of Physics: Conference Series, 1114(1), 1–11, 2018. https://doi.org/10.1088/1742-6596/1114/1/012046.

[34] The OWASP Foundation. OWASP Foundation. https: //owasp.org/, Last accessed December 2022.

[35] V. S. Moustakis, C. Litos, A. Dalivigas, L. Tsironis. Website Quality Assessment Criteria. Proceedings of the Ninth International Conference on Information Quality (ICIQ-04), pp. 59–73, 2004.

[36] W3C. Introduction to Web Accessibility. https://www.w3.org/WAI/fundamentals/accessibility-intro/, 2022.

[37] Y. B. Hermanto, and V. A. Srimulyani. The Challenges of Online Learning During the Covid-19 Pandemic. Jurnal Pendidikan Dan Pengajaran, 54(1), 46–57, 2021. https://doi.org/10.23887/jpp.v54i1.29703.

[38] Y. Kanellopoulos, P. Antonellis, D. Antoniou, C. Makris, E. Theodoridis, C. Tjortjis, N. Tsirakis. Code Quality Evaluation Methodology Using The ISO/IEC 9126 Standard. International Journal of Software Engineering & Applications (IJSEA), 1(3), 2010. http://dx.doi.org/10.5121/ijsea.2010.1302.

.

Teaching Computer Security in a Secondary School

J. Mottl
Faculty of Science, University of Hradec Králové/Department of Applied Cybernetics,
Hradec Králové, Czech Republic
jan.mottl@uhk.cz

***Abstract* - The article focuses on computer security, especially on the teaching of this topic in secondary schools. From the broad topic of security, the article focuses mainly on authorization and authentication and the related tools used for these processes in web development. The main objective of the paper is to determine the extent of education on these topics in secondary schools. The paper provides an insight into the issue by means of a questionnaire survey conducted in a secondary vocational school with a focus on computer science. The results of this survey were compared with the results of the same questionnaire given to students of the Bachelor's degree programme in Applied Informatics at the Faculty of Informatics and Management at the University of Hradec Kralove. The paper summarizes the results of the analysis of these questionnaires and provides insight into the importance of teaching computer security in secondary schools.**

Keywords - computer security; authorization; authentication; secondary school

I. INTRODUCTION

Computer security is a very broad topic that has been addressed in many research papers. This paper focuses on teaching cyber security and provides insight into teaching this topic in a secondary school.

The paper also attempts to provide a basic theoretical insight into the topic, so the first chapter focuses on computer security itself, specifically the principles of authorization and authentication, as well as basic methods of data encryption. In addition, a preview of the basic tools used to secure websites is provided.

Subsequently, the paper discusses the teaching of these topics, and, in particular, the articles focused on the topic of security itself.

The main contribution is the practical part of the paper, which builds on the topic and provides insight into the current situation of teaching the topic at Delta High School in Pardubice, Czech Republic. The level of knowledge and familiarity with computer security tools was determined using a questionnaire survey conducted at the school. The paper compares the results with a smaller sample of students at a higher education institution, specifically the University of Hradec Kralove, Faculty of Informatics and Management.

II. THEORETICAL BASIS

Cybersecurity is a broad topic; this chapter is intended to provide a basic insight into the subject at a theoretical level so that it can be linked to education.

Security can be summarised in two key points: defending against threats and dealing with the consequences. For the latter, it is necessary to include not only dealing with the damage caused but also the skill of immediate response during an attack [1].

Author in [2] summarizes cybersecurity into six points: confidentiality, integrity, authorization, availability, authentication, and accountability. In this paper, the important point is the authorization process, which ensures the authenticated user has access to secured data, the transmission of which is protected by encryption.

Authorization is defined as determining which data is accessible. Individual operations on these data are subject to authorized control, which is provided by the authorized owner or server. Similarly, an authentication system works by continuously verifying that the entity connected to the data is the authorized owner. Authentication serves as the primary authentication property that ensures the integrity of data manipulation [3].

There are two types of attacks on computer systems: one targeting users and known as social engineering (e.g., phishing, pharming). The second type of attack is aimed directly at the system, most often attacking the operating systems, servers, applications, or data stores. Often the attacker gets access to the data through a breach of the authentication or authorization process [1].

The most common identity attacks are phishing, spear phishing, credential stuffing, password spraying, and man-in-the-middle.

Phishing is a part of social engineering; it is an attempt to obtain login credentials by using a deceptive email or message on a social network.

Spear phishing is a targeted phishing attack on an organization or individual.

Password spraying is a brute force attack where an attacker, ignoring the basic rules of a given server, tries to guess the password in a crude manner. The defense against this type of attack is two-factor identity verification.

Man-in-the-middle is a targeted attack that tries to get between the user's communication and the server. An example is attacking communication through a public Wi-Fi network, where the server and the user's computer are secure, but the network that mediates the communication is insecure [4].

A. Data encryption

The security of all communication through the network is ensured through encryption methods. These methods focus on encrypting the content of messages and authenticating the users, which is done through substitution and transposition.

Substitution is the replacement of characters in a message with other characters, and transposition is changing the order of characters in a message. The science that deals with this is cryptography [5].

The encryption can be symmetric in asymmetric; in either case, a encryption key is important. In symmetric encryption, a one (private) key is shared between the communicating parties. In asymmetric encryption, two keys are used, public and private. The private key is not shared within the communication [3].

Symmetric encryption is typical for dealing with large amounts of data. This type of encryption emphasizes the transmission speed [6]. An example of such encryption is the DES (Data Encryption Standard) algorithm, developed by IBM in 1967. The encryption blocks that determine the length of the key are typical of this type of encryption. Its successor was the AES (Advanced Encryption Standard), also known as the Rijndael algorithm. It uses keys 128 to 256 bits long and is used in web communication as well as file encryption [7].

Asymmetric encryption, as mentioned earlier, uses a pair of keys (private and public), with one being important for encryption and the other for decryption. In the process of decryption, each party generates its own private key and is responsible for keeping it secure. Compared to symmetric encryption, this method is slower.

The Diffie-Hellman algorithm is one example of asymmetric encryption. However, this algorithm has a major weakness and cannot defend against a man-in-the-middle attack, where one party can be presented with a fraudulent key.

The Rivest-Shamir-Adleman algorithm is often used for electronic communications, digital signatures, and building private networks. This algorithm uses keys over 100 decimal digits long and is therefore considered to be strongly secure. Asymmetric encryption is also used in certification and signature authority [6].

B. Identity management (authentication and authorization)

Identity management, which includes authentication and authorization, uses various tools and protocols. An example of this is the SAML 2.0 protocol, which is difficult to implement. OpenID Connect layer on top of the OAuth 2.0 protocol and SAML 2.0 are significant standards in authorization and authentication [8].

Identity management is often described as a cycle, at the beginning of which is registration, where a user account receives its unique identification number/symbol. Subsequently, the new account is assigned the available actions it can perform (there is a difference between an administrator account and a regular user within the system privileges). At the same time, the user's identity must be continuously verified (by the authentication process).

The most common authentication is by password or by sending an authentication email. Two-factor authentication, which helps to ensure account security, is standard. After the user has been authenticated, the process of enforcing the content that the user is entitled to, depending on the type of account, takes place. Next, the length of the login (session) is limited; at the end of this time, the user is re-authenticated or directly logged out. That ensures the secure length of the connection. Subsequently, either at the user's request or at the expiration of the session, the user is logged out.

The identity management cycle ends only when an account is deleted, which is the ultimate termination of access to the server [8].

OAuth 2.0 is an authorization protocol that allows third-party Application Programming Interface (API) calls via tokens without the protocol sharing sensitive user information. In addition, the protocol makes the content available for a limited time. The use of tokens is crucial, as their interception does not compromise the client or the server [8].

Authentication is handled by the OpenID Connect protocol. In particular, protocol ensures only limited user authority as well as limited token accesses and validity [8].

SAML (Security Assertion Markup Language) is based on an XML framework. It allows single sign-on from multiple domains. The key factors are identity providers and service providers who communicate with each other to verify the user's identity and possible credentials [8].

Currently, there are several modern open-source identity managers. These include the Ory and Gluu projects and, most importantly, the Keycloak project.

This project offers an identity and access manager, as well as single sign-on services, federated identities, and REST APIs. Keycloak was originally developed by Red Hat, which later abandoned the project. It has been further developed by the community and now offers the ability to style the login page, set up strong authentication, and a large range of ready-made programming flows. In its basic toolset, it offers the possibility to reset passwords and encourages users to update them. Keycloak includes a single sign-on system and management of the length of that sign-on [9].

Compared to Keycloak, Gluu offers the possibility to log a user in without a password, using, for example, a mobile phone, but otherwise using a similar set of tools. Ory, on the other hand, focuses mainly on working with cookies in a web environment.

III. TEACHING COMPUTER SECURITY

Awareness of computer security is being raised outside the school environment in the Czech Republic. The primary organizations dealing with the topic of prevention are the Police of the Czech Republic and the Primary Prevention Centre. The largest organization concerned with cyberspace is the National Cyber and Information Security Agency, which also serves as an educational institute for cybersecurity. Other examples of raising awareness of cyber security include various courses, competitions, and other educational events.

One of the larger projects that educate about cybersecurity is Counties for a Safe Internet. It is an e-learning course established in 2013 and supported by the Ministry of the Interior and Microsoft. A similar e-safety project called Czech Children and Facebook 2015 focused on safe behavior within saidsocial network. It was a part of a broad research project of the Faculty of Education of Palacký University in Olomouc.

A. Related works

In this subsection, several research studies concerning teaching cybersecurity at different levels of education in different countries are presented.

In the Netherlands, 140 primary school students and 96 secondary school students were asked (through a form) about their behavior in terms of password handling and cyber security. The questionnaires revealed that secondary school students were much better able to recognize a phishing site or email than elementary school students. Both groups were shown to at least check the security of a given website when accessing it. Both groups handled downloaded files similarly, as they did not check if they are secure. Overall, cybersecurity appeared to be slightly more discussed among high school students, while it was completely absent among elementary school students [10].

Similar research has been conducted at the University of the Western Cape in South Africa, where the primary focus was on the cybersecurity of university students, specifically smartphone usage behavior. This research spanned the years 2016 and 2017, while a total of 252 students were interviewed. The results were positive; students demonstrated knowledge of safe smartphone usage [11].

The quantitative study was conducted in Bartin province, Turkey. Several secondary schools and a total of 8 299 students participated in this study, thus, it provides comprehensive results. The data were collected using a voluntary questionnaire offered to the students. The research primarily dealt with model situations and investigated how students behave in them. These situations were related to cybersecurity, whether it was sharing login credentials or logging out of the system. Similarly, the research was interested in sharing personal information within social networks. The results were primarily positive as they showed that students were aware of the various threats. However, the research showed a weakness in recognition of social engineering [12].

In 2017 and 2018, extensive research was conducted on secondary school students, focusing on students' social media behavior. As part of the research, the social network Fakesbook was created. This platform allowed the authors to conduct simulations of student behavior within this network. Within these simulations, it was found that students share sensitive data and information within social networks. Students have little awareness of the interconnectedness of accounts and share content across networks. In follow-up work with students and using graphs to show account and data interconnectedness, the authors sought to educate and try to make students aware of the threats in disseminating information within such networks [13].

Research conducted in China had a similarly alternative approach to teaching computer security. In research realized between 2014 and 2016, a series of simulations were conducted to teach students computer network security. This study revealed wide gaps in understanding basic terms and principles [14].

These papers provided insight into how security is taught across countries as well as levels of education. All of these works have become essential sources for the development of the research for this paper.

IV. RESEARCH

The research of this paper aimed to test the knowledge of theoretical and practical computer security. This research took place in the fall of 2022 and was conducted at Delta secondary school in Pardubice, Czech Republic. Delta is a private school that focuses on teaching computer science and management.

The results from the secondary school were compared to a smaller sample of college students studying Applied Computer Science at the University of Hradec Kralove, Faculty of Informatics and Management.

A. Method

This quantitative research was conducted through a voluntary questionnaire administered to secondary school students. The same questionnaire was administered to college students in order to compare the results.

The questionnaire was constructed with an emphasis on the topics described in the theoretical part of the paper an emphasis on data encryption and identity management, specifically, specifically the level of knowledge on these topics. In the questionnaire, respondents answered fully anonymously; neither gender nor the age of the student was asked, only the year in which the student was. Similarly, the undergraduate students were asked only about the year of study.

The questionnaire was divided into a section dealing with cryptography and a section dealing with identity management.

The first section focused primarily on the difference between asymmetric and symmetric encryption, specific encryption algorithms (DES, AES, etc.), and their categorization.

The second section dealt with identity management, specifically the processes of authorization and authentication. Here, the focus was particularly on the use of multi-

TABLE I. ANSWERS TO ATTITUDE QUESTIONS

Statement	Secondary school				University			
	Agreed	*Rather agree*	*Rather disagree*	*Disagree*	*Agreed*	*Rather agree*	*Rather disagree*	*Disagree*
I am familiar with cryptography and have been involved in this field in school or self-study.	15	30	42	36	0	6	6	2
I am familiar with the concepts of authorization and authentication and have been involved in this field in school or self-study.	27	55	30	11	6	5	3	0
I used one of the well-known authorization protocols when creating the website.	9	10	22	82	3	3	4	4

factor password authentication and password management through software developed for this purpose.

As both schools are focused on teaching computer science, the final part of the questionnaire focused on questions related to the use of authorization and authentication tools in web development.

The questionnaire includes attitude questions, where students comment on individual statements and whether they agree or disagree with them (or to what extent). They are further divided into factual questions, where the aim is to determine the degree of realistic insight, especially regarding the different types of encryption. The questionnaire then contains questions directed towards specific written responses, mainly in terms of how students use multi-factor identity authentication and, where applicable, for which services or tools. For such questions, multiple choice was often used, as well as the option to enter one's own answer if none of the selected ones were suitable.

B. Results

123 secondary school students and 14 college students participated in the survey. Specifically, for the secondary school, there were 51 first-year students, 15 second-year students, 26 third-year students, and 31 fourth-year students. The questionnaire was voluntary. The large number of first-year students represented is due to the larger number of students in these years, while the upper years have a smaller number of students per class. For the college students, there were 14 students third-year undergraduate students.

I have decided to analyze the results within individual categories, with cryptography being the first to be looked at in depth. The results of the attitudinal question of cryptography knowledge are shown in Table I. It can be seen that the majority of the secondary school respondents (82/123), as well as college respondents (8/14) are not familiar with the topic of cryptography. In the case of the college, it should be noted that none of the respondents fully agreed they were knowledgeable about this question. 28 of the negative responses come from first-year secondary school students who have not yet discussed the topic in their lessons and so have knowledge from primary school.

The questionnaire also addressed the issue of cryptography by asking whether students were able to distinguish between symmetric and asymmetric encryption. Students were asked to identify which of these encryptions uses one and which uses two keys, as well as to distinguish which of these encryptions is faster. In the second series of questions, students were asked to match specific encryption algorithms to the correct category, whether the encryption type was symmetric or asymmetric.

The results of these knowledge questions are presented in Table II. Students were given the option of not answering the question to see the difference between those who had poorly grounded knowledge and those who had never encountered the terms and thus were unable to answer the question at all.

As can be seen in Table II., both groups of students have positive results in the basic distinction of encryption types by key and speed. In both cases, more than half of the respondents answered correctly. In contrast, knowledge of specific algorithms is very low. It is interesting to note that the results achieved by high school students were better than those achieved by university students. That, however, may be caused by the small number of university respondents, so it cannot be considered conclusive..

The second category studied was related to identity management, specifically authorization and authentication processes. This category was divided into topics related to identity management itself and a final section that dealt with the development of websites and the tools used in these processes as part of their creation.

Regarding the attitude question, the results are shown in Table I., which clearly indicates that the respondents' attitude is more positive towards this topic than cryptography. There are 82 positive responses for the secondary school, which is over half of the respondents. The results are similar for college, with 9 positive responses; it is also worth noting not a single respondent disagreed with the statement they were unfamiliar with the topic.

The next question in the questionnaire was directed at

TABLE II. RESULTS OF CRYPTOGRAPHY KNOWLEDGE QUESTIONS

Type of question	Secondary school			University		
	Correct answer	*Incorrect answer*	*No answer*	*Correct answer*	*Incorrect answer*	*No answer*
Numbers encryption keys	47%	16%	37%	64%	7%	29%
Encryption speed	49%	13%	38%	64%	11%	25%
Algorithm assignment to encryption type	17%	16%	67%	10%	16%	74%

the use of two-step identity verification. Here, I did not differentiate between secondary school and university students, as their responses did not differ significantly. Fig. 1 shows the results of where students use two-factor authentication; it demonstrates that two-factor authentication is now the standard. The students use it not only for online banking, where two-factor authentication is usually mandatory, but also for other online services, whether it is a mailbox, social networking or discussion forums, or gaming clients (such as Steam). On the contrary, there were 3 respondents who do not use two-factor authentication at all.

In terms of the specific method of two-factor authentication, the majority of respondents (133) chose to authenticate using a phone device, whether it is via SMS or an application. Another common verification method (108 respondents) was a second verification through sending an email with a check password. Other methods, such as USB keys or specific identity verification programs, were also present in units of responses. Again, it confirms how widely used two-factor authentication is today.

Another part of the questionnaire was devoted to password managers. It started with a simple question about whether students use identity managers. For secondary school, 55.3% (or 68 respondents) do use a password manager, while for university, only 28% (4 respondents) do. Here, it can be seen password managers are more likely to be used in secondary school, where students often manage various accesses to many systems and use software tools to remember passwords.

The most commonly used tool was Google Password Manager. Similarly, the corresponding Apple tool often appeared among the responses. In addition, software such as Bitwarden, LastPass and, KeePass was represented among the respondents. The results can be seen in Fig. 2.

The last part of the questionnaire was devoted to web development and the use of authorization programs. Here, the attitudinal type responses, which can be seen in Table I., show that vast majority of secondary school students are not familiar and even less of them are experienced with web development. However, results are somewhat better for university students who generally have more experience in web development. When asked about their knowledge of specific tools, the secondary school respondents appeared to have better knowledge, with 18 respondents knowing and sometimes using KeyCloak, which was mentioned in the theoretical section, followed by the secondary school respondents mentioning their knowledge of Glu and Ora (4 and 3 responses, respectively). Amongst college students, only one respondent knew one of these tools.

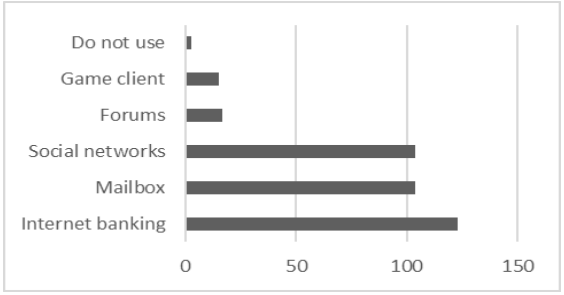


Figure 1. Using two-step identity verification

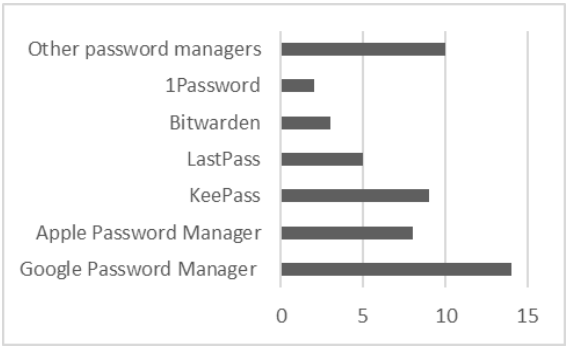


Figure 2. Password Manager

V. Conclusion

The aim of the paper was to introduce cybersecurity, especially data encryption and identity verification processes, as well as provide insight into teaching these topics within topics in secondary schools.

The main contribution is the research conducted in a secondary school with a focus on Computer Science and a university in Applied Computer Science. The results of this questionnaire showed that there is still little knowledge among students of symmetric and asymmetric encryption as fundamental knowledge in understanding how to send data across a network. However, even if students were lacking in theoretical knowledge about cryptography, they demonstrated very good knowledge of practical topics such as identity management, students from both schools appeared to have very good knowledge, as two-factor authentication of that identity has proven to be standard among students. In password management, the situation was similar; it became apparent that more education is needed in web development and using tools for user authorization and authentication processes.

This paper is just a light insight into the issue and I believe it can help with a deeper investigation in this regard.

Acknowledgment

I would like to thank Mgr. Josef Horálek, Ph.D., who helped me with the selection of the topic and cooperated as a supervisor of the article itself. I would also like to thank Mgr. Long Do, whose work provided me with a good insight into the topic and provided an overview of inspiration for the research.

I would like to thank the Department of Applied Cybernetics, Faculty of Science, University of Hradec Králové for financial support.

References

[1] J. Kolouch and P. Bašta, CyberSecurity. Praha: CZ.NIC, 2019.

[2] P. C. Van Oorschot, Computer security and the internet: tools and jewels from malware to bitcoin. Switzerland: Springer, 2021

[3] R. Pužmová, TCP/IP v kostce, 2. upravené a rozšířené vydání. České Budějovice: Kopp, 2009.

[4] “Identity is key to stopping these 5 cyber security attacks,” Okta. [Online], 2001. Available: https://www.okta.com/resources/whitepaper/identity-is-key-to-stopping-these-5-cyber-security-attacks/. [Accessed: 03-Feb-2023]

[5] S. Singh, The code book : the science of secrecy from Ancient Egypt to quantum cryptography. New York: Anchor Books, 2000.

[6] K. Burda, Kryptografie okolo nás. Praha: CZ.NIC, 2019.

[7] L. Dostálek, Velký průvodce protokoly TCP/IP: Bezpečnost, Praha: Computer Press, 2003.

[8] Y. Wilson and A. Hingnikar, Solving Identity Management in Modern Applications: Demystifying OAuth 2.0, OpenID Connect, and SAML 2.0,. Berkeley, CA: Apress, 2019.

[9] S. Thorgersen and P. I. Silva, Keycloak - identity and access management for modern applications: harness the power of Keycloak, OpenID Connect, and OAuth 2.0 protocols to secure applications, Birmingham: Packt, 2021.

[10] J. W. A. Witsenboer, K. Sijtsma and F. Scheele, “Measuring cyber secure behavior of elementary and high school students in the Netherlands,” Computers & Education, vol. 186, pp. 104536, Sep 2022.

[11] I. M. Venter, R. J. Blignaut, K. Renaud and M. A. Venter, “Cyber security education is as essential as “the three R's”,” Heliyon, vol. 5, issue 12, e02855, Dec 2019.

[12] R. Yilmaz, F. G. Karaoğlan Yilmaz, H. T. Öztrük and T. Karademir, “Examining secondary school students' safe computer and internet usage awareness: An example from Bartın province,” Pegem Journal of Education and Instruction, vol. 7, issue 1, pp. 83–114, Feb 2017.

[13] M. Zikus, O. Curry, M. Moore, Z. Peterson and Z. J. Wood, “Solving A social networking platform for teaching security and privacy concepts to secondary school students,” in Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19), pp. 892–898, New York, NY, ACM, 2019.

[14] X. Dai and S. Tsai, “Experimental Teaching of Information Security Based on Virtual Simulation,” Mobile Information Systems, vol 2021, 8062065, Nov 2021.

What Do We Know about Learning – Conversations with chatGPT

J.Henno*, H.Jaakkola** and J.Mäkelä***
* Tallinn University of Technology, Tallinn, Estonia
** Tampere University, Pori, Finland
*** University of Lapland, Rovaniemi, Finland
jaak@cc.ttu.ee

***Abstract* – We are witnessing a major achievement of machine learning – appearance of several chatbots which are able to produce human-quality conversation. Appearance of such 'fluid conversationalists' has aroused big interest especially among teachers, since they have already passed several exams designed for humans and have been used by students in winter fall 2022 exams. Their performance is not based on linguistic research, but is an achievement of data science. This performance cannot yet be rationally explained (it is based on interplay of billions of variables), thus in the following are used conversations with the currently most visible member of this family – the chatGPT to clear some myths around these programs.**

Keywords – learning, statistics, language, neural net, chatGPT, WordNet

I. INTRODUCTION

We are living in an era of catastrophes – climate catastrophe, exhaustion of natural resources, population explosion, the 'Knowledge Explosion' [1] (predicted already in Bible, Daniel 12:4 [2]).

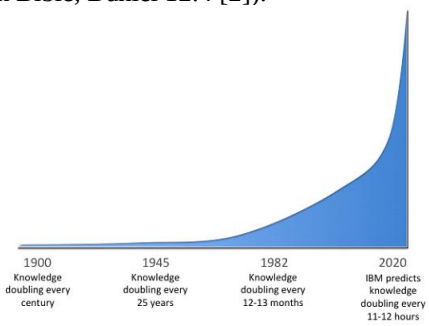


Figure 1. The 'Knowledge-Explosion' graphically

And at the end of the last year begin the explosion of chatbots capable to produce fluent text in natural languages (Croatian, Estonian, Finnish - not only on English!): on Nov.30, 2022 OpenAi announced chatGPT (General Purpose Transformer – transforms text to text) [3]; as a response to this Google announced of Feb 6, 2023 a rival – Bard [4] and on Feb 7, 2023 Microsoft announced new, AI-based (Artificial Intelligence) search engine and homepage Bing [5].

These 'super-polyglots' are based on research in neural networks and presented as 'deep machine learning'; in recent years have appeared umpteen papers proposing new methods of deep machine learning.

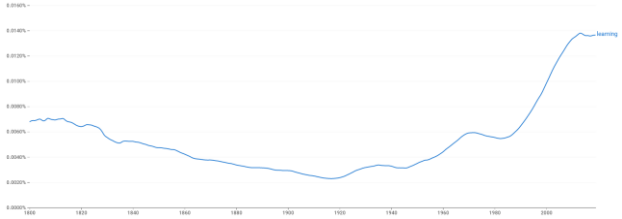


Figure 2. The world-wide interest in learning has in recent decade grown more than twice [6]:

With such explosion of new learning theories the old ones becomes quickly obsolete [7] and many teachers may be confused – should I now start working entirely differently – in a new, 'innovative', 'deep' way?

In the following is presented brief overview our understanding of learning – the 'classical' theories and the new, 'deep' ones..

II. CLASSICAL THEORIES OF LEARNING

Human brain contains ca 10^{11} neurons and each neuron is directly connected via synapses with ca 10^3 other neurons [8]. When learning something, brain constantly rearranges this network. Although this structure has been extensively studied by neurobiologists, teachers, computer scientists, mathematicians, physicists, psychologists, linguists etc. We still do not have workable theories and models of brain, since we can follow brain only from 'outside', follow processes (neuron activations), assuming inner workings of brain, what current methods of neurophysiology cannot yet explain.

Many theories are trying to explain human learning mechanisms [9], [10]. All these theories: the behaviorist learning theory (good students receive positive reinforcement - a prize or a gold star - primitive gamification), Cognitivism (learning is based on changes in the brain), Constructivism (learners construct their understanding of the world through experiences and interactions), Social Learning theory (emphasizes the role of society), the transfer theory (learning occurs when applying already existing knowledge in new situations) etc. are all trying to explain our experience, century's old practice of teaching-learning. They are not 'rules for action' – 'do this, then do that' – they only allow to reflect on our practice. Every teacher is a jazz artist who has

momentarily respond to current situation in class – who did what, who asked what. Learning and teaching both are hard work based mainly on improvising.

III. BOTTOM-UP VS TOP-DOWN

As every process/action, learning can be described from several perspectives. For humans is important the main direction of knowledge acquisition – top-down (from description of external features of the process, describing 'how?') or form bottom-up – from description of inner mechanisms, forces, structures what drive the whole process, describing 'why?'.

The classical learning/teaching is considered in bottom-up manner. You have first learn the basics – learn to read and write, math, laws of physics. Nobody can learn to drive a bicycle by listening lectures on bicycle construction and angular momentums (the top-down approach); you have to try to run a bicycle (bottom-up approach).

With overall advance of computers – the humanity's first 'Tool for Intellect' - many basic intellectual skills are now questioned – could they be performed by computers? In several countries handwriting (calligraphy) has been replaced with drawing 'block letters' (e.g. in USA handwriting become not compulsory in many states already in 2010), ability to calculate is questioned – "everyone has a mobile device which can do all these calculations!" and Google is trying to nullify even ability to read, replacing everything with video - in search results the YouTube videos are always placed first. The 'how it looks' is considered more important than 'how human does it'.

Currently we are witnessing the next step in 'top-down' revolution in considering human thinking – the neural-net revolution of handling natural (English, Croatian, Estonian, Finnish etc.) languages.

We still do not understand our natural languages enough to use them with computers. With computers we have to use mysterious sequences of strange symbols with lot of hidden meaning; what makes programming difficult to learn and understand. For example one line (in Python) [11], [12]

$$a_2, a_1, a_0 = \text{np.polyfit}(x, y, 2)$$

is sufficient to calculate second-order approximation

$$y = a_0 + a_1 x + a_2 x^2$$

This one line hides (in libraries) 18378 other lines of Python code [11], thus from the code is visible only 0.0000979431-th part and it is very difficult to understand. Everything were much simpler, if we could address computers with human languages.

But "If the Mountain won't go to Mohammed, then Mohammed must come to the Mountain" – while we still fight computer languages, computers learned to speak human languages.

IV. THE MACHINE LEARNING REVOLUTION

The 'machine learning' approach to human learning is totally top-down – computers are mimicking the order of tokens (characters, syllables, words) what humans have used in their texts, they do not understand their meaning. Linguists are not involved, the computer scientist Frederick Jelinek (studied automatic speech recognition) boasted: "Every time I fire a linguist, the performance of the speech recognizer goes up".

The beginning of machine learning was discovery of neurons in 1990-s. Neurons are discrete entities, conducting their impulses in one direction only, down their axon and receiving information through dendrites. American neurophysiologist Warren McCulloch and mathematician Walter Pitts presented in 1943 a model of neuron functioning.

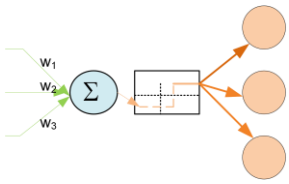


Figure 3. Neuron fires (sens signal to next neurons) if sum of its weighted inputs becomes sufficient.

Donald Hebb took the idea further and proposed that information is processed in brain step-by-step: neurons activate others (fire) by layers and neural pathways strengthen over each successive use (connectionism).

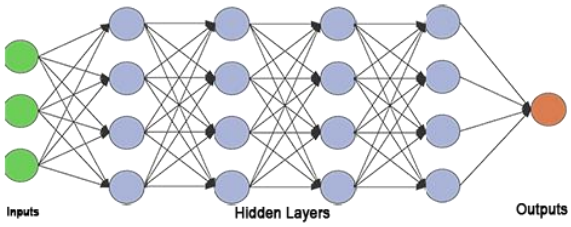


Figure 4. Neural net

Neural nets are devices for classification, compressing large amounts of data into small set of classes, e.g. recognizing an object in a visual scene or understanding the meaning of a sentence or phrase. In statistics these tasks were called clustering, but neural net researchers started to use term 'learning' and since modern neural nets use in the process large number of layers and parameters (weights), the (machine) learning is stated as 'deep'. At the end of the last century were proved universality results – neural nets can approximate any continuous function [13],[14],[15]. Unfortunately this result is often cited not correctly (even by Google), omitting the essential word 'continuous'. Neurons are helpless with e.g. the XOR function [16].

In NLP (Natural Language Processing) and machine learning are used Recurrent Neural Nets (RNN), which have also backward links to implement short-term memory. In text producing the main problem is prediction of the next word (or character) [17]. For prediction is used information - which token follows which one – derived from earlier (human-generated) texts, thus every machine

learning project begins with collecting a large corpus of human-produced text, e.g. from WWW.

Everything is based on statistics about 'what follows what' and only on this, the program does not use linguistic information – word types, sentence parts etc. This information is distilled by a neural net from the corpus, stored in a large number of parameters and called 'text model', e.g. the text model GPT-3 which was used for creating chatGPT has 175 billion parameters.

Text model creation (inferring tokens order information) is based on simple statistics. Suppose we already have a sequence of words:

$$w_1, w_2, ..., w_{i-2}, w_{i-1}$$

The next word w_i could not be arbitrary, it depends on preceding words. The next word could be guessed maximizing the relative probability [18]:

$$\max(w_i \in V) P(w_i \mid w_1, ..., w_{i-1})$$

Here $P(w_i \mid w_1, ..., w_{i-1})$ is the conditional probability that after words $w_1, ..., w_{i-1}$ follows the word w_i. In practice probabilities are estimated from frequencies in the corpus, i.e. the relative probability of a word following previous words is calculated from the frequencies of use of these words in the corpus:

$$P(w_i \mid w_1, ..., w_{i-1}) \approx \frac{Fr(w_1, ..., w_{i-1}, w_i)}{Fr(w_1, ..., w_i)}$$

The probability of the whole phrase is the product of probabilities, i.e. probability of the beginning phrase $P(w_1, ..., w_{i-1})$ and the conditional probability that it is followed by the word w_i:

$$P(w_1, ..., w_i) = P(w_1, ..., w_{i-1}) P(w_i \mid w_1, ..., w_{i-1})$$

Human language is often considered as a process with limited memory (a Markov process), assuming that the meaning of the next word depends only on a limited number of preceding words. This is generally not true, we expect often that the reader/listener already knows the meaning of many words which have been use. But applying the Markov process assumption '*probability of word depends only on few number of previous words*' simplifies programs and is used in NLP (Natural Language Processing) everywhere. Thus for prediction of the next word is used only a sequence of fixed length *k* (the Bayes assumption) and the search goal is

$$\arg\max(w_i \in V) P(w_i \mid w_{i-k}, ..., w_{i-1}) \qquad (1)$$

After replacement $i - k \to 1$ we are looking for

$$P(w_1, ..., w_i) \approx P(w_1, ..., w_{i-1}) P(w_i \mid w_1, ..., w_{i-1})$$

Probability of the first word/phrase could be expressed the same way or those words are given as a seed.

In practice (to speed up calculations) the last formula is simplified even more. By the Bayes conditional independence assumption the probabilities $P(w_j \mid w_{j-1})$ are independent. So in language models are often used only binary probabilities (a very rough assumption, but gives huge speed-up of calculations), thus

$$P(w_1, ..., w_n) = P(w_1) \prod_{i=2}^{n} P(w_{i-1} \mid w_i) \qquad (2)$$

The formula (2) allows to create new texts – give some words $w_1, ..., w_{i-1}$ as a seed and find a word w_i which maximizes probability $P(w_i \mid w_1, ..., w_{i-1})$, then shift the 'action window' one step to right and repeat the process with sequence $w_2, ..., w_{i-1} w_i$. The value of the parameter *n* is big, e.g. in chatGPT n=2048 (see below).

Chaining words which in some context did follow each other could produce expressions which have opposite meaning to the meaning which was expressed in original sentences. For instance, with (2) from sentences:

John loves Mary. Mary loves Marc

would follow (using the underlined pairs of words):

John loves Marc

In the corpus could be also a statement

John hates Marc

But the program would not "see" it – it does not have any pairs of words which could be used for linking with the words from the first two sentences, thus the program would produce a wrong statement. For instance from a corpus containing texts of presentations from the "Mipro.CE" conference was produced nonsence [12]:

Students are not included in the process of e- learning in the context of the process of teaching and learning.

For arithmetical expressions does not hold the Universality Theorem - they do not implement a continuous function except for unary representation, thus they cannot be learned by neural networks. Some websites [19] claim it is possible: "You'll get to 99+% validation accuracy after ~30 epochs". This holds only for expressions with similar structure and length, testing their program with expressions of different length gives rather different results:

```
Q 49+232  T 1    - 0
Q 134+19  T 1    - 0
Q 450+29  T 1    - 0
Q 0+364   T 0    OK 0
…
Correct: 56%
```

With small modifications: multiplying binary strings using variables gives already a very poor result:

```
x=101001  y=1101011010  z=x*y        A:  z=1000100101101010      -
z=1000100111110110
x=1000111 y=1001 z=x*y     A: z=1001111111    - z=1001100011
…
Correct = 7.0%
```

The probabilities $P(w_j \mid w)$ allow to consider the whole vocabulary as an *n*-dimensional space (*n* – the size of the vocabulary) with distances between words determined by these probabilities. For instance, the list of

word vectors of the popular GloVe model from Stanford [20] contains 19171496 lines (one for every word) of length ~ca 2587; the upper-left corner (first 8 elements from first 3 lines; the first word is comma ',') is:

, 0.18378 -0.12123 -0.11987 0.015227 -0.19121 -0.066074 -2.9876 ...
the -0.20838 -0.14932 -0.017528 -0.028432 -0.060104 -0.2646 -4.1445 ...
and -0.09611 -0.25788 -0.3586 -0.32887 0.5795 -0.51774 -4.1582 -....

Creating text model from a text corpus is not a "rocket science", but already a rather common adventure using many openly available programs [21]. Big models show also some semantic relationships, e.g. in the 'vocabulary cloud' created from Mipro.CE presentations 2016-2021 appeared several groups of closely related words:

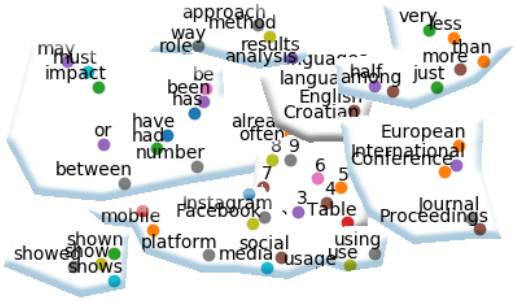


Figure 5. Some semantic groups from the 270021-dimensional space of the vocabulary of Mipro.Ce English-language presentations [22]

V. CHATCBT AND OTHERS

The models created from large text corpuses using openly available algorithms can produce amazingly good, 'human-like' texts. Nobody has not yet presented good explanation, why adding to a word another word which sometimes has followed it could produce 'human-like' text, but many people are over-exited. Professor of the USA university said that he “fell in love” with chatGPT after reading its answers to MBA exam questions [23].

In the following are considered some features of chatbots produced by neural nets. These features hold for all of them (the reasons are the same), but are here illustrated by answers produced on Feb.7-9, 2023 by the chatGPT program itself; its answers are not edited (copy-pasted from the program's window), only a bit reduced in order to comply with length limitations for this paper.

A. chatGPT does not have any understanding of truth

Authors of the chatGPT warned already in their first announcement [3]: "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers."

To check the truthfulness of chatGPT we asked chatGPT itself:

We: *"Can you say false statements ?"*

chatGPT: *" ...I can't guarantee that all the information is true or up-to-date."*

The program cannot distinguish between true and false statements.

We: *"Can You check truth of your statements ?"*

chatGPT: *"I do not have the capability to verify the truth of my own statements. It is ultimately up to the person receiving the information to critically evaluate and verify its accuracy. "*

Some prestigious Internet publications state "ChatGPT lies about scientific results" [24]. This is an over-reaction: lying is a deliberate action, but chatGPT does not know that it lies.

B. chatGPT can not learn

We: *"Are You learning ?"*

chatGPT: *"As an AI language model, I am not capable of learning ..."*

This is the second important difference between chatbots and humans – humans learn constantly, but the chatbot's 'learning' (usually called 'training') is one big computation in order to create the text model. Cost of computation of the chatGPT-s text model GPT-3 having 175 billion parameters was 46 million USD; cost of using GPT-3 (full model) is ca 14400 USD per month [25]; after training the model is fixed, cannot be re-trained.

C. Chatbots are circular

Chatbots are finite deterministic programs (if they were not deterministic, then the computer is severely broken); every finite deterministic system is circular [26].

We: *"What is the definition of set ?"*

chatGPT *"... a set is a well-defined collection..."*

We: *"What is ' a well-defined collection of distinct objects' ?"*

chatGPT: *"A well-defined collection of distinct objects refers to a set"*

Thus '*set*' = '*well-defined collection*' = '*set*', i.e '*set*' is a '*set*'.

D. Chatbot will eventually fall in loop

Every finite and deterministic dynamic system falls in loop, starts repeating itself:

We: *"As a finite system you start looping when producing very long statements ?"*

chatGPT: *"Yes, as a finite machine learning system, I can repeat patterns or get stuck in loops..."*

E. chatGPT (sometimes) does not know math:

The language model produced by neural net cannot guarantee correct math (the same problem was reported in [23]). In our first conversation with chatGPT it made several mistakes.

We: *"Calculate 17*7-26!"*

chatGPT: *"The answer is 69"*

Wrong, 17*7-26 = 89

We: *"Calculate 237.3+36.2*7-382.3!"*

chatGPT: *"The answer is 176.4"*

Not correct, should be 108.4.

After some days chatGPT calculated correctly and produced fluent explanations:

We: *"Please add 24/6 to 12*5!"*

chatGPT: *"To add 24/6 to 12 * 5, first we need to simplify 24/6 to a whole number. Dividing both the numerator and denominator by 6, we get:*

24/6 = 4

*Next, we can multiply 12 * 5 to find the value of the second term:*

*12 * 5 = 60*

Finally, we can add these two values together:

4 + 60 = 64

*So, the result of 24/6 + 12 * 5 is 64."*

F. chatGPT can explain itself

We: *"What is the size of the text corpus which was used to create your text model?"*

chatGPT: *"As an AI language model trained by OpenAI, I was trained on a large text corpus called the "WebText" corpus. This corpus consists of over 8 million web pages and contains a diverse range of content, including articles, books, and websites on a variety of topics.*

The specific version of the WebText corpus that I was trained on contains approximately 40 GB of uncompressed text data. However, it's worth noting that the exact size of the corpus and the training process can vary depending on the specific version of the model and the training configurations used."

We: *"How many tokens to you link to construct your answers ?"*

chatGPT: *"... the maximum number of tokens I can use in a single response is 2048, which is the maximum sequence length that my training allows.*

To give you an idea, an average sentence in English consists of around 15-20 tokens, so I can typically construct a response using anywhere from a few sentences to a few paragraphs, depending on the question."

Google announced on Feb 6, 2023 its own chatbot Bard; it made an error already in its very first demo what resulted in Google's parent company Alphabet loosing $100 billion in market value [27]. Microsoft demoed its new AI-powered search engine Bing on Feb 7, 2023 and placed all the responsibility to user: "Make sure to check the facts!" What is the sense of using expensive chatbot if every its statement should be checked for truth?

G. chatGPT can respond to several languages

Currently Internet is already not only in English – there are texts from many languages and if these texts were included in the text corpus used to train language model, the resulting model will also respond to other languages. This does not mean, that it understands anything – it just links together words/characters/syllables what in some text were following each other:

We: *"Objasnite mi - što je ljudsko učenje i po čemu se razlikuje od strojnog učenja?"*

chatGTP: *"Ljudsko učenje je proces učenja koji se odvija kroz iskustva i interakciju s okolinom. To može uključivati učenje putem iskustva, učenje od drugih, obrazovanje i mnogo drugih oblika učenja. Ljudsko učenje obično zahtijeva nekoliko puta da se nešto uči i da se stekne iskustvo kako bi se steklo znanje.*

Strojno učenje je oblik računalne tehnologije koja omogućuje računalima da nauče iz podataka bez eksplicitnog programiranja. Umjesto da se računalu eksplicitno kaže kako riješiti problem, strojno učenje koristi algoritme koji automatski identificiraju značajne značajke iz podataka i stvaraju modele koji se mogu koristiti za predviđanje ili klasifikaciju. Strojno učenje zasniva se na korištenju velikih količina podataka i računalne snage da bi model bio učinkovit.

Zaključno, razlika između ljudskog i strojnog učenja je u tome što ljudsko učenje uključuje interakciju s okolinom i stjecanje iskustva kroz mnoge učenje, dok strojno učenje zasniva se na automatskom identificiranju značajki i stvaranju modela na temelju velikih količina podataka."

VI. CONCLUSIONS – SHOULD TEACHERS USE CHATBOTS ?

Chatbots as many other inventions did not appear to answer a specific need, but as a result of human curiosity, so now the problem is "What should we do with this?" In terms of Gartner's hype cycle [28], chatbots are currently close to the 'Peak of Inflated Expectations'; it takes time until they move to 'Plateau of Productivity'.

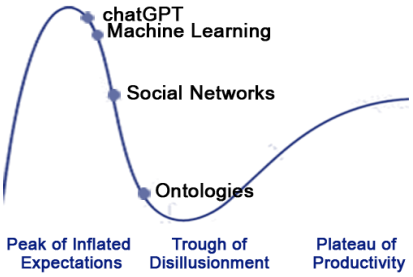


Figure 6. Some popular technologies in Gartner's 'hype cycle' scheme (according to results from Google Trends).

Although teachers did not participate in development of chatGPT (at least there is no information), chatbots will certainly affect our education, e.g. cheating has become incredibly easy. The program was able to pass MBA exam with result B..B- [23] and approach or exceed the score needed to pass three tests of the US Medical Licensing Exam (USMLE) [29] (this also says something about those exams); 17% of prestigious Stanford University students confessed that they used chatGPT to assist with their assignments and exams [30].

The 'demons of cheating' have forced several strong actions: the New York City's education department and growing number of other educational institutions blocked the site on its networks [31], a question and answer site

for programmers StackOverflow banned chatGPT after a week of test use [32].

But 'the troll is already out of bottle'. Universities are accepting the new technology and modified their academic integrity policies. Several possibilities for chatbots are investigated, e.g. if they could be used as automated training and grading systems [33], [34].

We: "*Should the AI tools (such as this program) be used in schools?"*

chatGPT: *"The use of AI tools in education is a complex issue with potential benefits and drawbacks.*

On the one hand, AI can help schools in many ways, such as by providing personalized learning experiences, assisting with grading and assessment, and streamlining administrative tasks. AI can also provide access to vast amounts of educational resources and can help to engage students in new and innovative ways.

On the other hand, the use of AI in education raises concerns about privacy, bias, and the potential for job loss in the education sector. It is also important to ensure that students develop critical thinking skills and are able to critically evaluate information, rather than becoming overly reliant on technology.

In conclusion, the use of AI in education can be beneficial in many ways, but it is important to carefully consider the potential drawbacks and to implement AI in a way that supports student learning and well-being. AI should be used as a tool to support and enhance education, rather than replacing human teachers and other educators."

REFERENCES

[1] J. H. Prince. The knowledge explosion -(or the exciting era). Sept. 1966, doi:10.1111/j.1444-0938.

[2] Daniel 12:4. https://www.biblegateway.com/passage/

[3] ChatGPT: Optimizing Language Models for Dialogue. https://openai.com/blog/chatgpt/

[4] CNBC Tech Drives. Google announces Bard A.I. in response to ChatGPT. https://www.cnbc.com/2023/02/06/google-announces-bard-ai-in-response-to-chatgpt.html

[5] CNBC. Microsoft announces new A.I.-powered Bing homepage. https://www.cnbc.com/2023/02/07/microsoft-open-ai-chatgpt-event-2023-live-updates.html

[6] https://books.google.com/ngrams/graph?content=learning&year_start=1800&year_end=2019&corpus=26&smoothing=3

[7] P. Singhai. 'You don't learn that at university': 40 per cent of degrees will soon be obsolete, report finds. https://www.smh.com.au/education/university-degrees-obsolete-report-ernst-young-20180501-p4zcn5.html

[8] Herculano-Houzel S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. . doi:10.1073/pnas.1201895109

[9] Guy R. Lefrançois. Theories of Human Learning. Cambridge University Press; 7th edition (Oct. 24, 2019), ISBN: 9781108484633, 566 pp

[10] E. Sørensen, The Materiality of Learning. Cambridge University Press; 7th edition (Oct. 24, 2019), ISBN: 780521182713, 226 pp

[11] J. Henno; H. Jaakkola; J. Mäkelä. Handling Software Icebergs. CEUR Workshop Proceedings, 2022, Vol.3237

[12] J. Henno; H. Jaakkola; J. Mäkelä. Machine Learning vs Human Learning. MIPRO 2021 - The 44th International Convention on Information, Communication and Electronic Technology Sept 27-Oct 1, 2021, Proceedings, pp 594-602, DOI: 10.23919/MIPRO52101.2021.9596771

[13] K. Hornik et al. Multilayer Feedforward Networks are Universal Approximators. Neural Networks. Vol. 2. pp. 359–366.

[14] Cybenko, G. Approximations by superpositions of sigmoidal functions. Mathematics of Control, Signals, and Systems, 2(4), (1989) pp 303–314. doi:10.1007/BF02551274

[15] Baker, M.R., Patil, R.B. Universal Approximation Theorem for Interval Neural Networks. Reliable Computing 4, 235–239 (1998). doi.org/10.1023/A:1009951412412

[16] M. Fabien. The Rosenblatt's Perceptron. https://maelfabien.github.io/deeplearning/Perceptron/

[17] Jaakkola, H., Thalheim, B., Henno. About the Essence of Intelligence - Will Artificial Intelligence Cover Human Intelligence? J.Frontiers in Artificial Intelligence and Applicationsthis link is disabled, 2021, 343, pp. 19–43

[18] D. Jurafsky, J. H. Martin. Speech and Language Processing. web.stanford.edu/jurafsky/slp3/4.pdf

[19] Smerity et al. Sequence to sequence learning for performing number addition. https://keras.io/examples/nlp/addition_rnn/

[20] GloVe: Global Vectors for Word Representation. https://nlp.stanford.edu/projects/glove/

[21] GPT text generation with KerasNLP. https://keras.io/examples/generative/text_generation_gpt/

[22] J. Henno; H. Jaakkola; J. Mäkelä. Machine Learning vs Human Learning. MIPRO 2021, Proceedings, pp 594-602, DOI: 10.23919/MIPRO52101.2021.9596771

[23] C. Terwiesch. Would Chat GPT Get a Wharton MBA? https://mackinstitute.wharton.upenn.edu/2023/would-chat-gpt3-get-a-wharton-mba-new-white-paper-by-christian-terwiesch/

[24] T.Ray. ChatGPT lies about scientific results, needs open-source alternatives, say researchers. ZD Net, Feb 16,2023, https://www.zdnet.com/article/chatgpt-lies-about-scientific-results-needs-open-source-alternatives-say-researchers/

[25] OpenAi. Contact our Sales. https://openai.com/contact-sales

[26] R. Sutton, A. Barto. Reinforcement Learning: An Introduction. A Bradford Book. The MIT Press. 2017, 356 pp

[27] M. Coulter, G. Bensinger. Alphabet shares dive after Google AI chatbot Bard flubs answer in ad. Reuters Feb.9, 2023, https://www.reuters.com/technology/google-ai-chatbot-bard-offers-inaccurate-information-company-ad-2023-02-08/

[28] T.Ray. ChatGPT lies about scientific results, needs open-source alternatives, say researchers. ZD Net, Feb.16, 2023, https://www.zdnet.com/article/chatgpt-lies-about-scientific-results-needs-open-source-alternatives-say-researchers/

[29] Gartner Hype Cycle. https://www.gartner.com/en/research/methodologies/gartner-hype-cycle

[30] T. Kung et all (incl. chatGTP). Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models. doi:10.1101/2022.12.19.22283643

[31] The Stanford Daily. Scores of Stanford students used ChatGPT on final exams, survey suggests. https://stanforddaily.com/2023/01/22/scores-of-stanford-students-used-chatgpt-on-final-exams-survey-suggests/

[32] A. Johnson. ChatGPT In Schools. https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students/

[33] stackoverflow. Temporary policy: ChatGPT is banned. https://meta.stackoverflow.com/questions/421831/temporary-policy-chatgpt-is-banned

[34] A. Kumar. AI's New Role In Education: Automated Grading. https://elearningindustry.com/artificial-intelligence-new-role-in-education-automated-paper-grading

[35] Xu Quian et al. The effectiveness of artificial intelligence-based automated grading and training system in education of manual detection of diabetic retinopathy. Front Public Health. 2022; 10: 1025271, doi: 10.3389/fpubh.2022.1025271

A Treasure Trove at Your Fingertips: Analysis of Contents from Europeana with Regard to Their Copyright Status

Mirko Duić
University of Zadar, Department of Information Sciences, Zadar, Croatia
miduic@unizd.hr

Abstract - Libraries, museums, archives and other institutions from more than 35 countries in Europe as well as several countries outside Europe, shared with Europeana the metadata about their heritage documents in digital form. This enabled the creation of huge integrated digital library which gives access to more than 55 million digital documents. Among these documents are books, newspapers, photographs, illustrations. Lots of documents from Europeana are freely available to use for various purposes, given that it is content in the public domain or content that has one of the Creative Commons licences or is otherwise free to use without having to pay compensation to rights holders. This kind of open access content was especially important during the Covid-19 pandemic when, for example, many teachers and school librarians could use it for the improvement of online teaching and learning. The goal of this study was to research the copyright status of contents from Europeana. The Europeana's contents of different copyright status were analyzed with regard to the following aspects: type of media, type of right statement, country in which the content is located, institution in which the content is located. Insights were acquired using the quantitative content analysis of the data from Europeana.

Keywords - Europeana; copyright status; digital library; open access; creative commons; public domain

I. INTRODUCTION

In 2005, several high-ranking politicians from different European countries sent a message to the European Commission to create a digital European library. This library should contain numerous works of European cultural heritage and provide free access to these works for all interested. In 2007, the Europeana Foundation was founded, and 'a year later the website www.europeana.eu goes online, while the European Digital Library (EDL) gets its new name: the Europeana' [1]. The Europeana is funded by the European Commission and its goal is to 'make Europe's cultural and scientific heritage accessible to the public' [2, p.85-86]. Since its inception, the Europeana has been building technological infrastructure and collaborative relationships with a number of libraries, archives, museums, and other institutions, to aggregate metadata about digital objects that are available in various collections across Europe and beyond. Today, the Europeana is one of the world's largest digital libraries through which you can search and access digital heritage content from all European and a few non-European countries. In 2012, the Europeana had more than 20 million objects in its catalog [2]. In 2016, it had 33 million objects from approximately 2500 participating institutions [3]. Today, it has more than 55 million objects from more than 3700 participating institutions [4]. This means that Europeana users have quite a wide selection of digital content at their disposal, content that they can use for various purposes. According to research from 2017, in which around 250 users of Europeana were surveyed, it was found that 30.4% of them were users from the academic field, there were 24.6% of cultural heritage enthusiasts, as well as 18.3% of cultural heritage professionals. Also, there were 13.3% of students and 4.6% of school teachers [5]. Insights from this research indicate that among the Europeana users there are many of those whose interests and positions are related to cultural heritage, research and education. It was also established that the majority of users 'were searching Europeana with the intention of using the information found to create a new work, e.g. to write a book, to prepare an exhibition, to use images for a presentation...' [5, p.216]. The possibilities for using Europeana contents in education and related fields are quite diverse. For example, biology teachers can use contents about the different types of plants that exist around the world. The Europeana has millions of records of this type and many of them are in the public domain, so students can use them to create a wide variety of projects: they can use the images and descriptions of plants they find on Europeana to write and publish essays with images of rare plants. Also, school librarians can find books at Europeana, which are not available in their library. They can inform their users about contents from Europeana that are available for use in various ways, including modification of original works from Europeana. Employees in the education sector, can organize workshops to present valuable sources at Europeana and explain in which ways specific Europeana contents can be used with regard to copyright status. Insights gained from research that will be described in this paper, as well as from topically similar research papers, could support and encourage employees in the education sector to get to know and use Europeana contents more fully and frequently. The topics that will be presented in the following part of the text are rights statements about the contents of Europeana, research goals, methodology and findings. Also, important findings will be discussed and further research opportunities related with copyright status of Europeana contents, will be proposed.

II. EUROPEANA, METADATA AND RIGHTS STATEMENTS

The Europeana collects metadata about digital objects from different institutions. Also, it formats and manages

this metadata. Based on the aggregated metadata, users from all over the world can search a large amount of digital content of European cultural and scientific heritage. The metadata, which is provided to Europeana is licensed with Creative Commons CC0 1.0 Universal Public Domain Dedication, which means that it could be freely re-used by other institutions and individuals. One of the main challenges for Europeana is the issue of copyright and licensing related to the digital objects described by this metadata. In order to ensure that digital cultural heritage contents described through collected metadata can be legally shared and used, content providers, such as galleries, libraries, archives and museums (the 'Europeana data providers'), must select the proper right statements for their materials [6]. They have to designate if their content is in the public domain, if it is in copyright or if it has some other type of copyright status. There are 14 rights statements, which provide information about re-use of contents for research, education, or other purposes. These statements designate if the contents described by metadata can be re-used and under what conditions. Among the statements are six Creative Commons licences, which are used for 'works that are in copyright and for which the rights holder wants to authorise some reuse': 'Creative Commons - Attribution (BY)'; 'Creative Commons - Attribution - ShareAlike (BY-SA)'; 'Creative Commons - Attribution - No Derivatives (BY-ND)'; 'Creative Commons - Attribution - Non-Commercial (BY-NC)'; 'Creative Commons - Attribution - Non-Commercial - ShareAlike (BY-NC-SA)'; 'Creative Commons - Attribution - Non-Commercial - No Derivatives (BY-NC-ND)'. Also, among the statements are two Creative Commons tools, which should be 'used on works that are in the public domain or that want to be dedicated to the public domain': 'The Creative Commons CC0 1.0 Universal Public Domain Dedication (CC0)' and 'Public Domain Mark'. Finally, there are six Rights Statements by the Rights Statements Consortium: 'No Copyright - non commercial re-use only (NoC-NC)'; 'No Copyright - Other Known Legal Restriction (NoC-OKLR)'; 'In Copyright (InC)'; 'In Copyright - Educational Use Permitted (InC-EDU)'; 'In Copyright - EU Orphan Work (InC-OW-EU)'; 'Copyright Not Evaluated (CNE)'. These Rights Statements 'were designed to complement Creative Commons Licences and Tools: so that only when these are not suitable for a specific situation, Rights Statements are used' [7]. The intention of Europeana is to encourage the use of Creative Commons licences and tools to facilitate the use and reuse of digital cultural heritage. But, there are cases when those rights statements are not suitable, for example, if the providing institution cannot legally apply Creative Commons licences, then Rights Statements should be used. It needs to be emphasized that Europeana most often does not receive metadata about rights, as well as other metadata, directly from the individual institutions that created them and that own the items described by metadata, but instead, Europeana receives metadata from the aggregators that collect data from libraries, museums or other institutions. Currently, the Europeana has 145 aggregators that collect metadata according to different criteria: there are national, regional, domain and thematic aggregators. While 'national and regional aggregators define their scope by specific country or region', 'domain and thematic aggregators define their scope by cultural sector (such as museum, archive or library) or by topic and theme (such as fashion)' [8]. Domain and thematic aggregators work with contributors from various countries, while national and regional aggregators work with contributors from a specific country or region. Most of the content available through Europeana is in the public domain. These are works that are not protected by intellectual property rights, therefore users can dispose with them in any way they want - 'the contents in the public domain can be freely combined, re-arranged, transformed, and reworked, enabling unlimited possibilities for creative play' [9, p.963]. Also, organizations that digitize public domain content do not need to search for the copyright owner, nor do they need to obtain an owner's permission to digitize and publish the work. On the other hand, if the work to be digitized is protected by copyright, this presents important challenges, such as, finding the copyright owner, obtaining permission for the digitization and publication of these works. Also, institutions that digitize, collect and publish content should take care to protect the moral rights of authors during and after digitization - for example, they should take care that photographs are not used in contexts that may be offensive to their authors or the persons in the picture. In particular, works in digital form are 'more vulnerable to alteration, addition of content or distortion' [9, p.969]. In the description of EuropeanaPhotography project it was pointed out that private contributors to archives want that the moral integrity of the works or documents they contribute is preserved: 'Many people would love to donate family photo heritage to their local archive, but fear that when it becomes published online as public domain material, it could be reused without any moral restraint. Photos of their family might be altered, defaced, ridiculed and republished' [10, p.414]. Regarding this issue, it is worth to mention that in some countries, including Croatia, there is legal protection of moral rights for photos and other types of creative works, regardless of their copyright status [10].

III. RESEARCH

The goal of this study was to research the copyright status of contents from Europeana. The Europeana's contents of different copyright status were analysed with regard to the following variables: type of media, type of right statement, country in which the content is located, institution in which the content is located. The quantitative content analysis method was used, with the goal to acquire insights about these aspects. Publicly available data regarding the explored variables were collected from January 13 to 15, 2023 from the Europeana digital library. The data was collected through the search interface of Europeana, which enables the search of all Europeana contents using different search facets. The following facets, i.e. search categories were used to acquire data : 'Type of media', 'Can I use this?', 'Providing country', 'Rights statements'. Acquiring detailed insights about the availability and copyright status of items in Europeana is potentially useful to various subjects - scientists dealing with similar research topics, employees of heritage institutions working to create and share digital content

collections, educators and artists who plan to use items from Europeana in their educational and artistic work, etc.

IV. FINDINGS

As of January 15, 2023, Europeana had metadata in its catalog on over 55 million items stored in institutions from a large number of countries. The exact number of item metadata available in Europeana was 55,859,454. Each of the Europeana items is marked with one of 14 right statements. Most of the items have the following rights statement: 'Public Domain Mark'. There are 21.23% of these items (see Table I). In second place, according to the quantity, are items with right statement 'In Copyright' - 18.55%. In third place are items with 'CCO' right statement - 12.79%, and in fourth place, according to the quantity, are items with 'CC BY' right statement - 10.68%. In fifth place are objects with 'CC BY-SA' right statement - 7.33%. Items with these five most numerous rights statements comprise about 39.4 million items or 70.58% of the total number of items. Among the items with these five rights statements, only for items with the 'In Copyright' right statement the permission to use them should be requested, while the items with the following two rights statements can be used without fulfilling any condition - 'Public Domain Mark', 'CC0'. Also, for items with the 'CC BY' right statement the name of the author of the work should be provided, when they or their derivatives are used and published. Finally, for items with the 'CC BY-SA' rights statement the name of the author of the work should be provided and used item should have the same licence.

TABLE I. RIGHTS STATEMENTS OF EUROPEANA ITEMS

Rights statement	Items	%
Public Domain Mark	11,858,735	21.23
In Copyright	10,360,359	18.55
CC0	7,142,342	12.79
CC BY	5,968,215	10.68
CC BY-SA	4,092,748	7.33
CC BY-NC-SA	3,711,793	6.64
CC BY-NC-ND	3,609,838	6.46
No Copyright - Other Known Legal Restrictions	3,321,452	5.95
No Copyright - Non-Commercial Use Only	1,681,341	3.01
In Copyright - Educational Use Permitted	1,571,979	2.81
CC BY-NC	1,287,387	2.30
Copyright Not Evaluated	923,788	1.65
CC BY-ND	280,837	0.50
In Copyright - OW-EU	30,179	0.05
Rights Reserved - Free Access	9,862	0.02
Out of Copyright - non commercial re-use	4,552	0.01
Total number of items	**55,859,454**	

The Europeana items belong to one of the following five types of media: image, text, sound, video, 3D. Of the total number of items, the largest percentage share have image items - there are 56.5% or about 31.4 million image items. In second place are text items - 41.94% or about 23.4 million text items. Regarding the media types, image and text items convincingly prevail in Europeana with 98.18%. The remaining 1.82% of the items are sound items (1.2%), video items (0.61%) and 3D items (0.01%).

Out of approximately 31.4 million image items in Europeana, the majority have the following rights statement: 'CC0'. There are 18.89% of these items, among all image items. In second place, according to the quantity, are image items with 'CC BY' rights statement - 17.43%. In third place are image items with 'In copyright' statement - 16.53%, and in fourth place are image items with 'CC BY-SA' rights statement - 11.99%. In fifth place are image items with 'Public Domain Mark' statement - 8.44%. Image items with these five most numerous rights statements comprise about 23 million or 73.28% of the total number of image items. Among these image items, only for items with the 'In Copyright' statement, the permission should be requested to use them. Out of approximately 23.4 million text items in Europeana, the most text items have the following rights statement: 'Public Domain Mark'. There are 29.09% of these items, among all image items. In second place, according to the quantity, are text items with 'In Copyright' right statement - 18.53%. In third place are text items with 'No Copyright - Other Known Legal Restriction' statement - 11.31%, and in fourth place are text items with 'No Copyright - non commercial re-use only' rights statement - 7.12%. In fifth place are text items with 'CC0' rights statement - 5.15%. Text items with these five most numerous rights statements comprise about 19 million or 81.2% of the total number of text items. Among these text items, the items with 'Public Domain Mark' and 'CC0' rights statements can be freely used for any purpose, while the text items with 'No Copyright - non commercial re-use only' statement can be used for any non-commercial purpose. Among approximately 0.67 million sound items in Europeana, the majority of them have the following rights statement: 'In Copyright'. There are about 0.52 million sound items with 'In Copyright' statement or 78.25% of the total number of sound items. Also, among approximately 0.34 million video items in Europeana, the majority of them have the following rights statement: 'In Copyright' (see Table II). There are about 0.3 million video items with 'In Copyright' statement or 88.64% of the total number of video items.

TABLE II. RIGHTS STATEMENTS OF EUROPEANA VIDEO ITEMS

Rights statement	Type of Media VIDEO	%
In Copyright	304,309	88.64
In Copyright - Educational Use Permitted	8,808	2.57
Public Domain Mark	7,880	2.30
CC BY-NC-ND	5,744	1.67
CC BY-NC-SA	2,309	0.67
CC BY-SA	1,485	0.43
CC BY	2,132	0.62
Total number of items	**343,309**	

In Table III, the terms of use of all Europeana items are presented. However, for the better understanding of data in Table III, as well as data in the succeeding tables, explanations are needed. In its faceted search interface Europeana has a following search category, 'Can I use this?'. When the user selects this search category, then three search subcategories i.e. answers to this question, are displayed: 1. 'Yes'; 2. 'Yes, with conditions'; 3. 'Maybe, seek permission'. If the user selects one of these subcategories, he/she is presented with search results containing all items in that subcategory. The subcategory

'Yes', contains all Europeana items which are marked with one of the following four rights statements: 'Public Domain Mark', 'CC0', 'CC BY', 'CC BY-SA'. Items with these rights statements can be used for various purposes without asking the rights holder's permission. Rights statements 'Public Domain Mark' and 'CC0' are particularly generous statements toward users, because they do not impose any conditions if users want to use items marked with these statements. Rights statements 'CC BY' and 'CC BY-SA' are also generous statements toward users, although they impose some minor conditions for using items marked with these statements. 'CC BY' is a Creative Commons rights statement that allows the use of item without asking the rights holder's permission, but user should attribute the author. Finally, 'CC BY-SA' is a Creative Commons rights statement that allows the use of item without asking the rights holder's permission, but the user should attribute the author and should license the adaptations of the work under the same terms. This means that all 'new works based on the original licensed work will carry the same license' [7]. After these explanations, we can turn our attention to Table III. In it we can see the number and percentage of Europeana items which belong to a subcategory 'Yes', the subcategory which contains items marked with the four previously mentioned rights statements. We can see that there are about 29 million items that belong to 'Yes' subcategory, which makes 51.88% of the total number of Europeana items. Also, there are about 15.4 million items that belong to 'Yes, with conditions' subcategory, which makes 27.52% of the total number of Europeana items. This subcategory contains all Europeana items which are marked with one of the following eight rights statements: 'CC BY-NC-SA', 'CC BY-NC-ND', 'No Copyright - Other Known Legal Restrictions', 'No Copyright - Non-Commercial Use Only', 'In Copyright - Educational Use Permitted', 'CC BY-NC', 'CC BY-ND', 'Out of Copyright - non commercial re-use'. Items with these rights statements are less available to use for different purposes than items in a subcategory 'Yes'. For example, the rights statement 'CC BY-NC-SA', which belongs to 'Yes, with conditions' subcategory, is a Creative Commons right statement that allows the non-commercial use of item without asking the rights holder's permission. Also, the user should attribute the author and license the adaptations of the work under the same terms. Another somewhat restrictive rights statement, which belongs to 'Yes, with conditions' subcategory, is 'CC BY-NC-ND' rights statement. It is a Creative Commons statement that allows the non-commercial use of item without asking the rights holder's permission. User cannot change the item in any way and, also, user should license the adaptations of the work under the same terms [7].

In Table III, we can see the number and percentage of Europeana items which belong to a subcategory 'Maybe, seek permission', the subcategory which contains items with the following rights statement: 'In Copyright'. There are about 11.2 million items that belong in this subcategory, which makes 20.06% of the total number of Europeana items. Although items with 'In Copyright' rights statement are available in Europeana and its partner institutions, any kind of their use requires permission from the rights holder [7]. Therefore, items with the 'In Copyright' rights statement, i.e. items in a subcategory 'Maybe, seek permission', are in the most restrictive subcategory for use, compared to items in subcategories 'Yes' and 'Yes, with conditions'.

TABLE III. TERMS OF USE OF ALL EUROPEANA ITEMS

Can I use this?	Items	%
Yes	28,980,652	51.88
Yes, with conditions	15,371,030	27.52
Maybe, seek permission	11,206,028	20.06
Total number of items	55,859,454	

In Table IV, we can see the number and percentage of Europeana items which can be used without any conditions or only with minor conditions. There are about 11.9 million items marked with 'Public Domain Mark' statement, which makes 40.87% of the total number of Europeana items in 'Yes' subcategory. Also, there are about 7.1 million items marked with CCO statement, which makes 24.65% of the total number of Europeana items in the same subcategory. Items marked with the two already mentioned rights statements - 'Public Domain Mark' and 'CC0' - can be used for various purposes without asking the rights holder's permission and without the need to fulfill any condition, when using them. There are about 19 million such items, which makes 65.52% of the total number of Europeana items in 'Yes' subcategory.

TABLE IV. RIGHTS STATEMENTS OF EUROPEANA ITEMS WHICH CAN BE USED WITHOUT ANY CONDITIONS OR WITH MINOR CONDITIONS

Can I use this?	Items	%
Yes		
Public Domain Mark	11,844,684	40.87
CC0	7,142,342	24.65
CC BY	5,900,827	20.36
CC BY-SA	4,092,799	14.12
Total number of items	28,980,652	

In Table V, the terms of use of Europeana items with regard to their types of media, are presented. Image items are the most numerous items in Europeana, regarding the type of media. Majority of image items are in 'Yes' subcategory - 56.53% out of the total number of image items, which is about 31.4 millions. Next, there are 26.43% of image items in 'Yes, with conditions' subcategory and 16.78% of image items in 'Maybe, seek permission' subcategory. Text items have the similar division by the mentioned subcategories. For example, almost half of text items are in 'Yes' subcategory. Sound, video and 3D items have very different division by the mentioned subcategories, in comparison with image and text items. There are more than 80% of sound and video items in 'Maybe, seek permission' subcategory, and there are almost 90% of 3D items in 'Yes, with conditions' subcategory. This means that, on average, sound and video items have the most restrictive rights statement among Europeana items. 3D items have less restrictive rights statements and image and text items have the least restrictive rights statements.

TABLE V. TERMS OF USE OF EUROPEANA ITEMS WITH REGARD TO THEIR TYPES OF MEDIA

Can I use this?	Image	Text	Sound
Yes	56.53%	47.73%	4.16%
Yes, with conditions	26.43%	29.67%	12.71%
Maybe, seek permission	16.78%	21.65%	83.13%
Total number of items	31,416,291	23,426,302	667,709

Can I use this?	Video	3D
Yes	3.41%	5.77%
Yes, with conditions	7.80%	89.41%
Maybe, seek permission	88.79%	4.83%
Total number of items	343,309	5,843

In Table VI, the terms of use of Europeana items with regard to the providing countries, are presented. For the better understanding of data in this table, some explanations are needed. In the second column ('All items (1,2,3)') - quantities of items 1, 2, and 3 from a specific country are listed: items 1 comprise all items in 'Yes' subcategory; items 2 comprise all items in 'Yes, with conditions' subcategory; and items 3 comprise all items in 'Maybe, seek permission' subcategory. In the third column ('Items 1 and 2') - percentages of items 1 and 2 from a specific country are listed. Items 1 and 2 are items which are marked with less restrictive rights statements, in comparison with items 3, which have the most restrictive rights statement: 'In Copyright'. Countries in Table VI are listed according to the percentages from the third column. The higher the percentage share of items 1 and 2 originating from a particular country, the higher that country is positioned in the list of countries in the second column.

TABLE VI. TERMS OF USE OF EUROPEANA ITEMS WITH REGARD TO THE PROVIDING COUNTRIES

Can I use this?			
Country	All items (1,2,3)	Items 1 and 2	Items 1
Norway	3,469,641	99.65%	85.89%
Finland	1,144,806	96.02%	91.25%
Denmark	1,010,867	95.17%	34.11%
Sweden	4,439,064	94.61%	70.20%
France	3,970,078	90.33%	15.78%
United Kingdom	4,816,412	83.91%	69.87%
Netherlands	9,175,688	83.47%	79.48%
Hungary	1,017,249	80.09%	11.20%
Spain	3,566,741	79.91%	32.02%
Germany	6,248,078	74.57%	20.64%
Belgium	2,506,863	71.55%	59.43%
Poland	3,649,817	67.30%	65.66%
Austria	2,636,522	64.25%	38.88%
Czech Republic	1,244,959	61.31%	37.35%
Italy	2,185,087	45.78%	9.19%
Slovenia	469,480	51.38%	44.20%
Serbia	80,107	91.24%	49.88%
Croatia	58,178	28.97%	21.49%

It can be seen from the Table VI that Norway has the largest percentage of items 1 and 2, in comparison with the percentage of items 3 from the same country. This means that Norway has the largest percentage of items that are marked with less restrictive rights statement (99.65%), in comparison with its items with 'In Copyright' statement (0.35%), which is a most restrictive rights statement. For example, Italy is positioned quite low in the second column, which means that it has a much lower percentage of items 1 and 2 (45.78%), and much higher percentage share of items 3 (54.22%), in comparison with Norway. In the fourth column ('Items 1'), it can be seen that Norway has 85.89% of items 2 (items in 'Yes, with conditions' subcategory). This means that it has only 13.76% of items 2 (items in 'Yes' subcategory). On the other hand, Italy has only 9.19% of items 2, which means that it has 36.59% of items 2 (items in 'Yes' subcategory'). The Europeana contains items from more than 40 countries and in this table the data for 15 countries that have the largest number of items in Europeana, are listed. Additionally, the table contains data related to Slovenia, Serbia, and Croatia, the countries which are not among those 15 countries. Slovenia is in the 18th position, Serbia is in the 28th position, and Croatia is in the 32nd position, with regard to the number of items in Europeana.

In Table VII, the terms of use of Europeana items with regard to the providing institutions, are presented. The Europeana contains items from more than 3700 institutions. In this table, data from 15 institutions that have the largest number of items in Europeana, are listed. It can be seen that that there are seven institutions with 100% of items 1 and 2. This means that these institutions have only provided items with less restrictive rights statement. On the other hand, Austrian National Library is in the last position in the list of countries, because it has provided 54.46% of items 1 and 2. It means that it has provided 45.54% of items 3, which are marked with 'In Copyright' rights statement.

TABLE VII. TERMS OF USE OF EUROPEANA ITEMS WITH REGARD TO THE PROVIDING INSTITUTIONS

Can I use this?			
Institution	All items (1,2,3)	Items 1 and 2	Items 1
Naturalis Biodiversity Center	4,512,513	100%	100%
The National Archives of Norway	2,995,810	100%	86.57%
Trustees of the Natural History Museum	2,453,996	100%	100%
Meise Botanic Garden	1,383,210	100%	100%
National Library of Spain	619,942	100%	0%
Cultural Heritage Agency of Netherlands	600,185	100%	99.90%
Royal Botanic Gardens Kew	574,117	100%	99.98%
Deutsche Fotothek	1,311,703	99.92%	10.68%
Swedish National Heritage Board	784,264	99.17%	100%
Bavarian State Library	1,239,897	98.68%	0%
Jagiellonian Digital Library	569,564	98.63%	100%
National Library of Poland	608,455	97.78%	100%
National Library of France	2,976,646	97.78%	0.19%
National Library of the Netherlands	1,308,024	64.07%	100%
Austrian National Library	1,481,153	54.46%	50.11%

V. CONCLUSION

According to research insights, it can be concluded that Europeana provides a huge amount of content that can be used for various purposes, including teaching and research. It is particularly significant that tens of millions of Europeana items have rights statements that are very generous to users, regarding the right to use and re-use these items. For example, it has about 19 million items with rights statements 'Public Domain Mark' or 'CC0', which makes up 34.02% of the items in Europeana. These are very valuable and extensive resources for users,

because these rights statements allow them to use the contents in the most free way possible: they can use the contents in their original form, they can rework them and create new works based on the original contents, they can use the contents in educational and any other contexts, non-commercial or commercial. Also, users can decide for themselves under which licence they will distribute the works created on the basis of the items from Europeana and partner institutions, which have 'Public Domain Mark' or 'CC0' rights statements. When publishing these Europeana items outside of Europeana or when publishing works created on the basis of these items, it is not necessary to mention the names of their authors. Additionally, it is important to point out that there is a large number of items in Europeana with an almost equally non-restrictive rights statements. Namely, Europeana has around 10 million items with rights statements 'CC BY' or 'CC BY-SA', which makes up 18.01% of Europeana items. These items are just as free to use as items with the rights statements 'Public Domain Marks' and 'CC0', except that when using 'CC BY' and 'CC BY-SA' items, the names of their authors should be mentioned. Also, when publishing Europeana's 'CC BY-SA' items outside of Europeana, or when publishing new, derivative works based on 'CC BY-SA' items, it is necessary to share those new works with the same rights statement: 'CC BY-SA'. But, regardless of these minor restrictions, those two statements still leave great freedom to use the items for various purposes. It should be noted that Europeana also has a large number of items with somewhat more restrictive terms of use, but which still leave plenty of room for users to use these items in various contexts. For example, there are items with rights statements 'CC BY-NC-SA' (6.64%) and 'CC BY-NC-ND' (6.46%). And besides, even the items that have the most restrictive rights statement regarding the possibility and conditions of further use - namely, works with 'In Copyright' rights statement - are available to read, view or listen via Europeana and its partnering institutions. Although for publication of 'In Copyright' items outside Europeana or for their modification, the permission of the rights holder should be obtained, nevertheless these items are accessible through Europeana. These items can, in their original form, be included as required reading literature in educational courses or as valuable visual materials in the graphic design or arts educational process. Considering all the contents offered by Europeana and its partners, it can truly be said that it is an open access treasure trove, which is at everybody's fingertips. But it should also be said that, according to the type of media, there are huge inequalities in terms of the amount of items in Europeana. Image and text items make up as much as 98.18% of Europeana items. The remaining 1.82% items are sound items (1.2%), video items (0.61%) and 3D items (0.1%). And while less than 20% of image items have the most restrictive 'In Copyright' rights statement, even 78.25% of sound items and 88.64% of video items have this rights statement. It means that Europeana is primarily a source of images and texts, while there are significantly fewer sounds and videos. Additionally, most image and text items have less restrictive rights statements than most sound and video items. With regard to the countries from which Europeana items originate, there are also pronounced inequalities regarding the amount of items and their rights statements. It is noticeable that a large number of Europeana items come from countries in the north of Europe - Norway, Finland, Denmark, Sweden. A very low percentage of items from these countries have the most restrictive 'In Copyright' rights statements (less than 5%). They are followed by France, UK, the Netherlands and Hungary, which have a slightly larger share of 'In Copyright' items - they have less than 20% of these items. This situation is stimulating for additional research. Interesting topic to explore could be specific reasons and contexts that are contributing to the fact that the majority of items from all these mentioned countries are licensed with very non-restrictive rights statements. For example, could it be that an extended collective licence model, which was established in the Nordic countries, is responsible for the favorable situation in the Nordic countries regarding the huge amount of digitized and available items and their non-restrictive types of rights statements. Also, other good practice examples of digitization, licensing and publication of digital content, could be explored and presented in order to provide quality information for members of institutions that have a lot of valuable content for digitization, as well as born-digital content, which they want to make publicly available on web with non-restrictive rights statements.

REFERENCES

[1] E. Buelinckx, "Digitizing photographic archives: Project-linked opportunities and pitfalls: The role of Europeana in an institutional digitization landscape," in 2013 Digital Heritage International Congress (DigitalHeritage). 28.10.-01.11.2013. IEEE, 2013, vol. 2, pp. 87-90.

[2] M. Damova, "Bulgariana: A Bulgarian Aggregator to Europeana," Digital Presentation and Preservation of Cultural and Scientific Heritage, (II), pp. 85-92, 2012.

[3] R. Matulionyte, "10 years for Google Books and Europeana: copyright law lessons that the EU could learn from the USA," International journal of law and information technology, vol 24 (1), pp. 44-71, 2016.

[4] Europeana. Search, Retrieved January 26, 2023. from https://www.europeana.eu/en/search?view=grid&query=

[5] P. Clough, T. Hill, M. Lestari Paramita and P. Goodale, "Europeana: What users search for and why," in Research and Advanced Technology for Digital Libraries: 21st International Conference on Theory and Practice of Digital Libraries, TPDL 2017, Thessaloniki, Greece, 18.-21.09.2017. Proceedings 21. Springer International Publishing, 2017, pp. 207-219.

[6] A. Malliari, I. Nitsos, S. Zapounidou and S. Doropoulos, "Copyright implications for the aggregation of audiovisual content in Greece,"Digital Library Perspectives, Vol. ahead-of-print No. ahead-of-print, 2022. Retrieved January 26, 2023. from https://doi.org/10.1108/DLP-04-2022-0033

[7] Europeana Pro. Understanding the rights statements used by Europeana, Retrieved January 26, 2023. from https://pro.europeana.eu/page/available-rights-statements

[8] Europeana Pro. Europeana Aggregators Forum, Retrieved January 26, 2023. from https://pro.europeana.eu/page/aggregators

[9] N. Klass, "Europeana, ARROW and orphan works: bringing Europe's cultural heritage online," in EU Copyright Law. Edward Elgar Publishing, 2014, pp. 946-987.

[10] F. Truyen and A. Fresa. "PHOTOCONSORTIUM: Digitizing Europe's Photographic Heritage," in Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 6th International Conference, EuroMed 2016, Nicosia, Cyprus, 31.10.–5.11.2016. Proceedings 6. Springer International Publishing, 2016, pp. 409-419.

QR Codes as an Educational Tool for Implementing the BYOD Approach in Physics Lessons

M. Drushlyak [*], Yu. Sabadosh [**], P. Mulesa [***], E. Diemientiev [*], A. Yurchenko [*] and O. Semenikhina [*]
[*] Sumy State Pedagogical University named after A. S. Makarenko, Sumy, Ukraine
[**] Vinnytsia national technical university, Vinnytsia, Ukraine
[***] Uzhhorod National University, Uzhhorod, Ukraine
marydru@fizmatsspu.sumy.ua, avataric87@gmail.com, mulesa@uzhnu.edu.ua, edementev8@gmail.com, a.yurchenko@fizmatsspu.sumy.ua, e.semenikhina@fizmatsspu.sumy.ua

***Abstract* - The integration of students' own computer means into the learning process can be solved within the implementation of the BYOD approach (Bring Your Own Device). One of the digital technologies used in education today that works in favor of the BYOD approach are the use of QR codes. The authors analyzed the practices of using the BYOD approach. Examples of the use of QR codes are given: the creation of information leaflets; organization of surveys, tests, web quests, and virtual exhibitions; expanding the content of textbooks; information support of objects of study, visualization and/or research of phenomena and processes. It is shown that the use of QR codes in physics lessons within the framework of the BYOD approach allows not only to solve educational goals but also to increase interest in learning, develop teamwork skills, and expand students' cultural horizons, that is, develop universal students' skills. Ukrainian resources for generating QR codes are provided. The results of an expert analysis of ways to use QR codes in teaching physics are presented. It was found that the most effective is the use of QR codes when organizing web quests and for the purpose of informational support of objects of physics study.**

***Keywords* - BYOD approach; QR code; digital technologies; teaching physics; ways of using QR codes; innovative educational approaches.**

I. INTRODUCTION

The need to involve IT in the organization and support of the educational process today is indisputable. A significant number of teachers in the world use digital technologies and tools to visualize educational material, organize surveys, keep electronic journals, etc. Many teachers use specialized software to conduct virtual experiments and model various processes or phenomena.

Ukrainian teachers, along with the use of multimedia, note that schools have problems with sufficient computers and students' access to them [1]. Only computer science lessons are held in computer classes. Other studies do not involve students' access to computer technology. This often leads to pupils using their mobile devices, which teachers do not like. Teachers note that young people actively use their digital means more for entertainment (social networks, games, surfing the Internet) and less for learning (finding the necessary materials for homework) [2].

After the pandemic, the usage of computers for distance learning became the norm. Teachers actively use Zoom, Google class, etc., but at the same time also note the active use of pupils' own mobile devices (and often not for educational purposes). The teachers believe that this reduces the cognitive activity of pupils.

This actualizes the contradiction between the spread of digital technologies among young people, the openness of young people to the use of mobile devices anywhere and anytime, and the limited ideas of teachers on how to use the activity of young people about their own mobile devices for educational purposes.

Traditional pedagogical approaches to the organization of the educational process increasingly do not satisfy the needs of both teachers and schoolchildren ineffective, intensive knowledge of the surrounding world. A traditional approach is focused on the formation of knowledge. This approach is authoritarian, as it lacks conditions for the manifestation of individual interests and creative abilities, and learning. The authors have analyzed the essence of the main approaches to teaching physics in Ukrainian schools. The authors focus on modern approaches, but to compare them with the traditional approach, they will also refer to the latter's principles (Table 1).

TABLE I. APPROACHES TO TEACHING PHYSICS IN SCHOOLS (SOURCE: MADE BY AUTHORS)

Approach	Characteristic
Traditional	Communication of ready-made knowledge to students. Learning by example. Inductive logic of presentation (from partial to general). Mechanical memorization. Survey methods on reproduction. External evaluation of results.
Interactive	Constant, active interaction of all participants in the educational process. The student and the teacher are equal educates. Simulation of life situations, use of role-playing games, and joint problem-solving based on the analysis of circumstances and the relevant situation.
Problem-centered	Creation of a problem situation and independent activity of students in solving it. Independent research work of students to solve a system of interrelated intra- and inter-subject

	educational problems.
Task-based	A list of general education skills and abilities is being formed. The system of problems should cover all the main phenomena, concepts, and laws, and the problems should be solved in order of increasing complexity so that each previous problem was a certain basis for solving the next one. Based on the teacher's analysis of the student's ability to solve the problem, a conclusion can be drawn about the understanding content of the relevant theoretical provisions.
BYOD approach	Involvement of students' IT and mobile devices (mobile phone, tablet, laptop) in the educational process for quick access to reference information, modeling processes, speeding up calculations, etc.

Learning with a traditional approach corresponds to a passive model of learning: the pupil listens, perceives, and remembers. Another model of learning, the active model, involves situations where the pupil himself is looking for a solution, solving a problem, analyzing the optimal solution, etc. (Figure 1), and therefore uses all the technologies, methods, and means that are available to him. Since young people are used to solving current problems using the Internet, they actively use their own mobile devices.

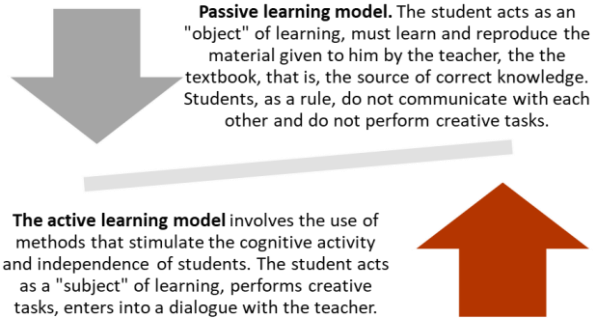


Figure 1. Comparison of passive and active learning models (source: made by authors)

This argues for the expediency of organizing the educational process based on the BYOD approach (Bring Your Own Device). Therefore, the goal of the study was to determine modern educational tools for the implementation of the BYOD approach in the example of teaching physics.

II. CHALLENGES OF USING THE BYOD APPROACHES IN THE TEACHING OF PHYSICS

There are many options for using mobile devices to teach physics. Smartphones are increasingly used in universities during the study of mechanics [3], acoustics [4], and thermodynamics [5] because they provide a unique way to conduct a simple scientific experiment.

A. Analogues of measuring devices

V. Sipiy believes that modern smartphones have a huge number of sensitive and accurate sensors, which allow measuring in real-time and storing various data about the external environment in the device's memory. For the most part, it is available to users to obtain data about movement, and the strength of light and sound flows, the smartphone can be used to replace such physical devices as a stopwatch, metronome, sound generator, stroboscope, etc. [6]-[7]. With the help of a smartphone, you can examine, for example, the volume of sound created by various sources [8].

B. Slow motion mode

But when using the BYOD approach in teaching physics, smartphones are not only analogs of measuring devices. The teacher can use such a function of the smartphone as a slow-motion video mode. For example, when measuring the uniformly accelerated motion of a falling body near the Earth's surface, an inclined plane is traditionally used to "reduce" the acceleration (so that it can be measured) – this approach was proposed by Galileo.

Now the free fall of the body in the vertical direction can be recorded by the student using the slow-motion function of the smartphone. This is somewhat equivalent to a decrease in the value of the free-fall acceleration g. A standard video consists of a sequence of images captured at 30 frames per second. The slow motion function, available in several smartphones, allows you to shoot videos at a speed of 120 to 960 frames per second. For example, the image scan time is approximately every 4.2 m/s for a video shot at 240 frames per second. This corresponds to the required accuracy for measuring the free-fall acceleration g [9].

The above examples of the use of mobile devices are not implemented by every smartphone. Also, not every smartphone of Ukrainian pupils has the opportunity to install specialized applications. This leads to the search for other areas of use of the mobile phone within the BYOD approach. One of these is the use of QR codes.

Analysis of the scientific results of the introduction of QR codes as an educational tool shows that QR codes are an effective tool that stimulates the interest and cognitive activity of lower secondary school pupils, and increases their motivation to learn [10]. N. Mileva and D. Stoyanova describe the use cases of QR codes: QR codes are printed on transparent films and placed on the corresponding pages of the textbook (educational content is deepened by additional resources); QR codes are printed on a separate sheet according to the topic of the lesson and have questions and tasks with multimedia materials related to QR codes; QR codes are printed on worksheets, but for collectively checking the answers and fixing errors.

In [11] T. Chicioreanu, E. Bilal, and M. Butnariu noted the use of QR codes by university professors to accompany professional activities to facilitate access to information that is given by a long internet address. The authors claim that professors are satisfied with such a digital tool for education.

Research [12] justifies that the use of QR codes in the classroom is simple. It does not depend on the discipline but allows you to combine real objects with any additional web content. J. Artal-Sevil, E. Romero, and J. Artacho

claim that QR codes allow you to generate virtual reality and augmented reality environments in the classroom, and also focus on the fact that the QR code can have access to text, videos, and photos. A QR code can be used in educational activities for reading and writing (educational material to deepen knowledge or a simple questionnaire where the student inserts information).

E. Artemova, Y. Shishalova, S. Melnikov, O. Orekhova, and G.Nikiporets-Takigawa [13] proved that the use of QR codes contributes to the development of professional skills, providing an increase in the quality indicators of the educational process. According to the authors, with the help of QR codes, effective educational activities for students are organized.

Another positive aspect of the use of QR codes is noted in [14]. S. Sharara and S. Radia suggest using QR codes to replace paper information leaflets. Patients then immediately access the necessary information on their own personal devices. The contactless way of transmitting information is more and more not dangerous, which became especially relevant during the spread of Covid-19.

There is also confirmation of the feasibility of using QR codes in teaching physics. I. Astra, L. Suryadi, and F. Bakri [15] explore the development of analytical skills in physics lessons and present the didactic development of a physical workshop equipped with 3D media based on QR-Code technology. The worksheet has been designed based on "using discovery learning stages".

Thus, the analysis testifies to the positive impact of the use of QR codes in the educational process and the commitment of professors and teachers to the use of such technologies in the world. This was the impetus for the study of the practices of using QR codes by Ukrainian physics teachers.

III. EXAMPLES OF USING QR CODES IN TEACHING PHYSICS

One of the digital technologies used in education today that works in favor of the BYOD approach are the use of QR codes. Virtually any mobile device can easily recognize and decode information encrypted using a QR code.

Today, the use of QR codes in education can be seen on posters and stands, and in libraries, teachers offer interactive lessons using QR codes and develop textbooks with QR codes.

Some possibilities for using QR codes in the educational process are given below.

- *Hyperlinks to multimedia sources and resources*. When accompanying the lesson with a presentation, students can be provided with handouts with QR codes for access to auxiliary applications (hyperlinks to multimedia sources and resources: video, audio applications, websites, drawings, animations, electronic educational publications, libraries, etc.). You can also place QR codes on the presentation slides. Instead of typing a URL into their phones, students can scan a code to get more information instantly.
- *Project activity*. During the organization of project activities, you can create link collections, information blocks, comments on project support website pages, and posters. Students can create their own portfolios or annotations on read books, educational and methodological literature on the researched topic and place them on the project website in the form of QR codes.
- *Survey and testing.* QR codes will allow you to organize quick surveys and conduct testing both in the classroom and outside it (web services ClassTools, Plickers, Mentimeler, etc.). For example, a printed QR code with the correct answers or a hint with an algorithm for solving the problem can be placed on each test ticket.
- *Game forms of activity*. QR codes can be used in-game quests to offer game tasks at one or more stages of the relevant events, in educational crossword puzzles.
- *Covers educational and methodical literature*. It is advisable to use QR codes to place reference material, information about the author, publisher, or any additional information on the covers of educational and methodological literature.
- *Information stands*. It is advisable to use QR codes for information saturation of standard information stands in classrooms, laboratories, recreation areas, libraries, museums of educational institutions; for placing the results of the educational process, etc.
- *Appendices to educational facilities*. QR codes can be placed on parts of mechanisms, electrical diagrams, and other physical objects. For example, QR codes placed on the periodic table of elements can contain physical and chemical properties of elements; QR codes placed on laboratory (demonstration) equipment may have a hyperlink to a virtual laboratory or control questions for self-study.

In physics lessons, QR codes can be used at any stage – from setting goals to homework.

QR codes can be utilized to set the algorithm of laboratory work or provide step-by-step instructions for its implementation. They can be used to expand the information space of the topic being studied when learning a new subject. By marking the elements or objects of a specific physical phenomenon with QR codes, educational videos or accompanying drawings can be provided. Self-check or mutual check can also be organized using QR codes. For example, students can solve a problem and then use a QR code to check their answers. The motivation, in this case, is to solve the problem faster and check it.

Fixing and repeating the studied material can be made more engaging with QR codes. Crossword puzzles, dominoes, lotto, and other games can be played using QR codes with coded physical terms, formulas, names of famous physicists, historical dates, and events.

The authors offer several other ways of using QR codes in physics lessons.

Information leaflets. In modern textbooks, little attention is paid to scientists, and if there is a mention of them, it is only the years of life and a few suggestions about the contribution to one or another branch of physics. With the use of the QR code, it becomes possible to expand the scope of such information through the creation and use of information leaflets that connect the QR code and larger information content for a more detailed study of the life and work of physicists. Leaflets with QR codes can be printed or sent to students, and they can familiarize themselves with the information on their own (Figure 2).



Figure 2. An example of using QR codes in information leaflets with physics tasks (source: made by authors)

Surveys, testing. Another way of using QR codes is online testing, which provides a test of students' knowledge, prevents writing off, and eliminates the subjectivity of the teacher, and the results are immediately displayed in the system.

A task in the form of a QR code can be offered during surveys at one of the stages of the lesson. The students can decipher the code to complete the task (Figure 3).

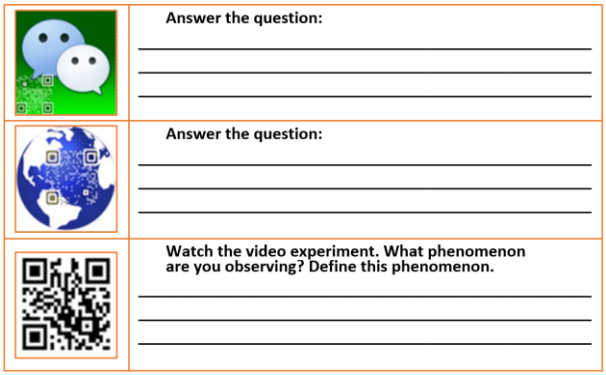


Figure 3. An example of a survey card (source: made by authors)

Web quests. Various web quests can be organized with the help of a QR code, which students tend to enjoy.

A web quest is a game activity of the type of a regular quest, in which the teams have to go through as many stations as possible in a limited time and answer questions or complete tasks that are encrypted using a QR code. Class students are grouped into teams. Each team receives:

- a smartphone with an installed application for reading QR codes and a presentation for photo tags;
- a map (printed or electronic) with stations indicated on it;
- a worksheet for recording the answers received at the stations.

The teams pass the points stations in the specified order. At each station, students have to decipher the QR code, answer the questions and enter the answer in the worksheet, make a photo tag, and place it on the appropriate slide of the presentation.

For each completed station, teams receive a certain number of points (provided they have a photo tag). The team is also awarded points for correct answers. If there is no answer (on the condition of passing this station), the team can be fined.

Physical web quest stations can be varied. For example, when studying the topic "Heat phenomena", the local stations can be the "Historian", "Metric", "Technique", and "Literature" stations. Here it is advisable to provide high-quality physics problems (it is better to choose short problems so that there are no difficulties with reading QR codes).

Textbooks. QR codes can also be used to display additional information in textbooks or special workbooks. For example, you can additionally place reference materials, links to videos of physics experiments, or simply link to online testing on the Internet in them via QR codes. QR codes can be used to organize a self-test of knowledge on the topic covered (Figure 4). Students receive a list of questions, the correct answers to which are posted in advance at various Internet addresses presented in the form of QR codes.

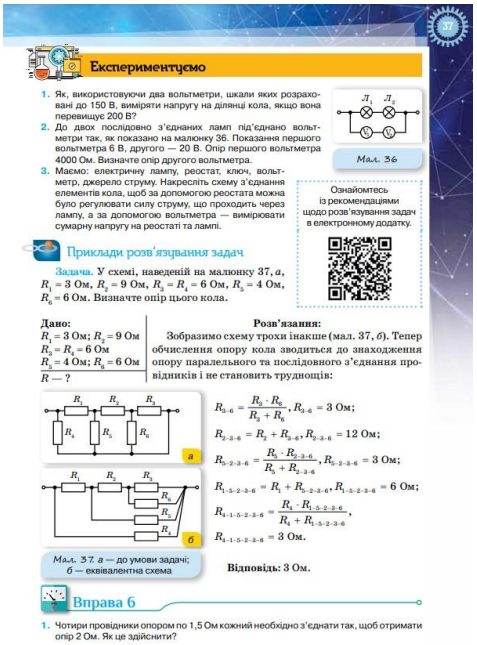


Figure 4. An example of using QR codes in a Ukrainian physics textbook (Source: [16])

Virtual exhibitions. The development of a virtual exhibition is necessary, for example, when studying the topic "Heat engines. The principle of operation of heat engines. Efficiency of a heat engine". The physics teacher needs to find images of heat engines on the school website or his own website, encrypt the hyperlinks to them in the form of QR codes, print them with signatures, and hang them on the walls of the classroom. Children can familiarize themselves with the exhibits of the exhibition during extracurricular hours.

Informational support of objects of study. QR codes are placed on laboratory equipment and allow visualization of operating rules, safety techniques, etc. Codes may provide a hyperlink to a virtual physics lab. In the office (laboratory) physics, you can use QR codes to provide information about the purpose of the equipment, conditions of use, history of creation, etc.

Study of physical phenomena and processes during laboratory work. For example, students can be offered to plot the Maxwell distribution function for O_2 (M=0.032kg/mol) at T=300K). Before the lesson, the physics teacher should code different options for model research tasks (Figure 5).



Figure 5. An example of a task (source: made by authors)

An expert assessment was carried out on the effectiveness of ways to use QR codes in a teacher's professional activity. Five service physics teachers who have at least 5 years of work experience and who implement the BYOD approach in their professional practice are involved in the examination.

Among the possible ways of using QR codes at physics lessons, the following were evaluated: 1) information leaflets; 2) surveys, 3) instructions; 4) web quests; 5) virtual tours, and exhibitions; 6) information support of physical objects of study. The experts had to rank the ways from the most effective (rank 1) to the most, in their opinion, ineffective (rank 6). Based on the results of the examination, we received data (Table 2).

TABLE II. THE RESULTS OF THE EXAMINATION (SOURCE: MADE BY AUTHORS)

Ways of using QR codes	Experts					Sum ranks	d	d2
	1	2	3	4	5			
1	4	1	3	6	3	17	-0,5	0,25
2	5	6	6	3	5	25	7,5	56,25
3	6	5	5	5	4	25	7,5	56,25
4	1	2	1	2	1	7	-10,5	110,25
5	3	4	4	4	6	21	3,5	12,25
6	2	3	2	1	2	10	-7,5	56,25
∑	21	21	21	21	21	105		291,5

The concordance coefficient W=0.66 indicates an average degree of agreement of opinions.

Pearson's test was used to assess the significance of the concordance coefficient: $\chi 2 = 5(6-1)\ 0.666 = 16.66$

For the significance level $\alpha=0.05$ calculated $\chi 2$ (16.66) is greater than the tabular one (11.07050), therefore W=0.666 – the value is not random.

Thus, the expert evaluation of the ways deserves trust: the most effective is the use of QR codes when organizing web quests and for the purpose of informational support of objects of physics study.

IV. CREATION OF QR CODES

A QR code can be created by using specialized online services that enable the encryption of necessary information in a single click. The most common software tools for generating and recognizing QR codes are QR-droid (https://play.google.com/store/apps/details?id=la.droid.qr), QR Reader (https://play.google.com/store/apps/details?id=uk.tapmedia.qreader), QR Barcode Scanner (https://play.google.com/store/apps/details?id=com.gamma.scan) and others. Often, scanners for reading QR codes are found as standard tools of many modern smartphones and are also found in messengers (in particular, Viber or Telegram).

Among Ukrainian-language online resources for creating QR codes, the following can be distinguished: qrcodes.com.ua, qr-code.com.ua, and ua.qr-code-generator.com.

There are many tools on the Internet that offer other, both free and commercial, services for converting various information into a QR code without downloading desktop software. Such online services can be effective, but the disadvantage can be the unauthorized use of the entered information in the public space. Therefore, it is not advisable to encrypt confidential information through online services.

Stand-alone, desktop, applications are more guaranteed to keep data confidential and that data will not be sent over the Internet. Such tools are called online or offline QR code generators.

When generating QR codes for the educational process, using different methods, you can customize the design of the code itself (the shape of the main part of the code, the shape of the square frame, the shape of the square itself, the color of the code, add a logo or image), the content of the code (you can add a small text, a link to the website, sound, video, e-mail address), the size of the generated code itself and the image format, etc.

V. CONCLUSION

Consequently, the spread of digital technologies and means affects both young people (use digital technologies and means actively) and the educational process, which acquires signs of active learning. The introduction of active teaching methods today is perceived as involving

young people in an independent search for answers to various educational tasks. Young people, in turn, use this purpose all available, including mobile ones, devices, and technologies that promote the BYOD approach to learning.

The tools for implementing the BYOD approach are, for example, mobile devices. Smartphones have sensitive and accurate sensors for measuring in real-time and can preserve data about the external environment (data and movement, the strength of light and sound fluxes, stopwatch, metronome, sound generator, stroboscope, etc.). Reachable for some smartphones is slow motion video mode.

The analysis of findings showed a significant potential for the use of QR codes for the BYOD approach implementation as an educational tool, which simplifies access to Internet resources, can build augmented reality or can provide multimedia support.

At the same time, in the context of Ukrainian schools, there are limitations on the use of certain smartphone functions, but QR codes are widely accepted for use.

QR-codes in the teaching of physics today are used in the following directions: the creation of information leaflets; organization of surveys, tests, web quests, and virtual exhibitions; improvement of textbooks; informational support for objects of physics study, visualization, and research of physical phenomena and processes.

The use of a QR code in physics lessons as part of the BYOD approach allows not only solves educational tasks in teaching physics but also provides an opportunity to involve students' personal devices in their educational activities, promotes increased interest in learning, develops teamwork skills, expands cultural horizons students, trains the skills of scanning QR codes and using relevant applications, that is, develops universal skills of students.

The QR code makes the lesson interesting, which additionally motivates students to study: it is convenient for students to read interesting content and, if necessary, quickly save it in the memory of smartphones. The approach, along with the visual channel of perception of initial information, allows you to use an additional (tactile) channel, which expands the information space of students.

The conducted expert analysis of the ways of using QR-codes in the teaching of physics showed that the most effective is the use of QR-codes when organizing web quests and for the purpose of informational support of objects of physics study.

The conducted research showed the expediency of using QR codes in the educational process in general and in teaching physics, in particular, and at the same time actualized other areas of research: the development of web quests to accompany the school physics course using QR codes, the research of the level of those educational results that are acquired using QR codes; analysis of mobile applications for teaching physics within the BYOD approach.

REFERENCES

[1] O. Lauta and A. Stelmashchuk, “The problem of informatization and computerization of the educational industry,” Bulletin of V.N. Karazin Kharkiv National University, series "Theory of Culture and Philosophy of Science", No. 56, pp. 26-30, 2017.

[2] “The virtual world and teenagers: what does the research show?”, Public Health Center of the Ministry of Health of Ukraine, 2021. Available at https://phc.org.ua/news/virtualniy-svit-ta-pidlitki-scho-pokazuyut-doslidzhennya, [accessed: 13.03.2023].

[3] P. Vogt and J. Kuhn, “Analyzing simple pendulum phenomena with a smartphone acceleration sensor,” The Physics Teacher, vol. 50, pp. 439-440, 2012. doi: 10.1119/1.4752056.

[4] M. Hirth, J. Kuhn and A. Müller, “Measurement of sound velocity made easy using harmonic resonant frequencies with everyday mobile technology,” The Physics Teacher, vol. 53(2), pp. 120-121, 2015. doi: 10.1119/1.4905819.

[5] J. Sweeney, “BYOD in Education (A report for Australia and New Zealand),” Intelligent Business Research Services, 2012.

[6] O. Slobodianyk, “Mobile applications at physics lessons. Physical and Mathematical Education,” Is. 4(14), pp. 293-298, 2017.

[7] O. Kolesnikova, N. Myslitska and D. Semeniuk, “Use Of BYOD Technologies For Formation Experimental Knowledge And Life Of Physics Masters,” Physical and Mathematical Education, Is. 2(20), pp. 48-53, 2019.

[8] V. Sepiy, “Use of the BYOD principle in the study of physics in high school at the professional level. Annotated results of research work of the Institute of Pedagogy of the National Academy of Sciences of Ukraine for 2017”, Institute of Pedagogy of the National Academy of Sciences of Ukraine. 2017.

[9] V. Zdeschyts and A. Zdeschyts, “The use of BYOD technology in the educational process in the conditions of distance learning of physics students: training. Manual,” Kryvyi Rih: Ed. Literia, 2022.

[10] N. Mileva and D. Stoyanova, “A methodological approach for using qr codes in education,” in 10th International conference of education, research and innovation (ICERI2017), pp. 8481-8486, 2017.

[11] T. Chicioreanu, E. Bilal and M. Butnariu, “QR codes in education – success or failure?” in 11th International Scientific Conference on eLearning and Software for Education (eLSE), pp. 180-187, 2015. doi 10.12753/2066-026 X-15-208.

[12] J. Artal-Sevil, E. Romero and J. Artacho, “QR-codes as a learning tool in advanced degrees,” in 12th International technology, education and development conference (INTED), pp. 7729-7739, 2018.

[13] E. Artemova, Y. Shishalova, S. Melnikov, O. Orekhova and G. Nikiporets-Takigawa, “The use of QR codes and their efficiency in the application of professional skills,” Apuntes Universitarios, vol. 12, Is. 1, pp.419-435, 2022. doi:10.17162/au.v12i1.978.

[14] S. Sharara and S. Radia, “Quick response (QR) codes for patient information delivery: A digital innovation during the coronavirus pandemic,” Journal of Orthodontics, 49(1), pp. 89-97, 2022. doi:10.1177/14653125211031568.

[15] I. Astra, L. Suryadi and F. Bakri, “Worksheets, Discovery Learning, and 3D Media based on QR-Code: The Ability to Analyze is Formed in Physics Practicum,” in 9th National physics seminar 2020, Vol.2320, 020011, 2020. doi10.1063/5.0037606.

[16] T. M. Zasekina and D. O. Zasekin, “Physics and Astronomy – 11,” Kyiv, UOVC “Orion”, 2019.

Windows Admin GUI Model for Learning PowerShell Commands

D. Tuličić*, D. Delija*, G. Sirovatka**, M. Mrkoci**
*Visoka škola za informacijske tehnologije, Zagreb, Hrvatska
**Tehničko veleučilište u Zagrebu, Zagreb, Hrvatska
dtulicic@vsite.hr, ddelija@tvz.hr, gsirovatka@tvz.hr, mmrkoci@tvz.hr

Abstract **This paper presents the development of a platform for the education of system engineers with the aim of insight into detailed data about computers under the MS Windows operating system. The application is based on the ideas of IBM's SMIT (System Management Interface Tool) interface for AIX systems, where standard commands are used through a pre-coded semi-graphical working environment. The developed application uses PowerShell commands and is primarily focused on the security aspects of system administration.**

Keywords Windows - PowerShell, educational application, information security;

I. INTRODUCTION

This paper presents the results of a master's thesis [1] in which an educational software tool was developed as a proof of concept. Even for systems like Microsoft Windows, modern system administration still relies on the skillful and mindful use of command line utilities (CLI). While Windows is based on a graphical user interface (GUI) for ordinary users, for any kind of professional system administration, a command line tool called PowerShell is used. However, from a learning point of view, transitioning from GUI to CLI can be an unusual and challenging path for an average student. To ease this situation and aid learning and task performance, a tool that acts as a safe wrapper around CLI commands with a GUI or semi-GUI interface can be used. If designed and built properly, such a tool has additional advantages. CLI commands are executed through scripts written by seasoned administrators and system programmers. These scripts act as a safe environment around CLI tools, managing arguments, options, inputs, outputs, results, and errors, and all actions are logged and provided with detailed help data and manual pages.

An operating system (OS) is a software program that manages all the other programs and resources on a computer. It acts as an interface between the hardware of the computer and the applications that run on it. In other words, it controls how the computer's hardware interacts with software and users. Examples of operating systems include Microsoft Windows, macOS, and Linux [6]. In the context of the operating system, a system administrator is someone responsible for managing and maintaining a computer system, including its hardware, software, and network infrastructure. They are typically responsible for tasks such as installing and configuring software, monitoring system performance and security, troubleshooting issues, and performing backups and data recovery.

The educational software tool described in the text is designed to help students learn how to perform some of these tasks using command line utilities, which are commonly used by professional system administrators.

Modern system administration still depends on the skillful and mindful use of command line utilities (CLI) even for such systems as Microsoft Windows. Windows are based on a graphical user interface (GUI) for ordinary users but for any kind of professional system administration a command line tool, PowerShell is used. From a learning point, it is an unusual path with a lot of obstacles for an average student. The students, the future system administrators start with GUI tools and then go into CLI with the growth of expertise and complexity of tasks to perform.

To ease that situation and help learn and perform tasks it is possible to use a tool that is a safe wrapper around CLI commands with GUI or semi-GUI interface. Such a tool has additional advantages if is designed and built properly, for example, CLI commands are executed through scripts written by seasoned administrators and system programmers.

The result of this approach is that students can perform serious system administration tasks with confidence, without damaging the system, and with the ability to see the correct syntax and procedures to use CLI. Such an approach to system administration already exists in some older commercial operating systems like IBM AIX, where there is a "System Management Interface Tool" (SMIT) [2] with the same functionality and purpose. As many security tasks are performed by system administrators, particularly data collection required for forensic analyses in incident response situations, it is sensible to create and test SMIT-like tools in the Windows environment, oriented mostly towards forensic data collection tasks.

Many ideas have already been explored using mainly PowerShell or a combination of Python and PowerShell. There are several notable works that are worth mentioning, such as "A Comparison of PowerShell and Python for Systems Administration" by Thomas Rayner and Michael Schneider. This paper compares the use of PowerShell and Python in systems administration and offers valuable background information. Additionally, Timothy Warner's "Designing a PowerShell-based System Administration Framework" investigates the design and implementation of a PowerShell-based framework for systems administration.

System administrators can also use the same tool to obtain information or supervise the system. The application is divided into several parts, from basic commands to more advanced commands such as getting information from networks, hardware, hard drives, and processes within the Windows environment. To implement this idea in the Windows philosophy and environment, CLI will run under the Windows PowerShell Integrated Scripting Environment (ISE), while an auxiliary application called AutoPlay Menu Builder (APMB) is used for the automated use of CLI commands and provides a basis for the command execution.

Although this project may seem simple in scope and purpose, it is too complex for one student to develop. Therefore, the first step was to build a working GUI prototype to prove the ability to interface scripting and GUI environment on Windows, which are both specific in their implementation and not Posix-like or typical business environments where such tools are usually implemented. This first step of implementation is described in [1].

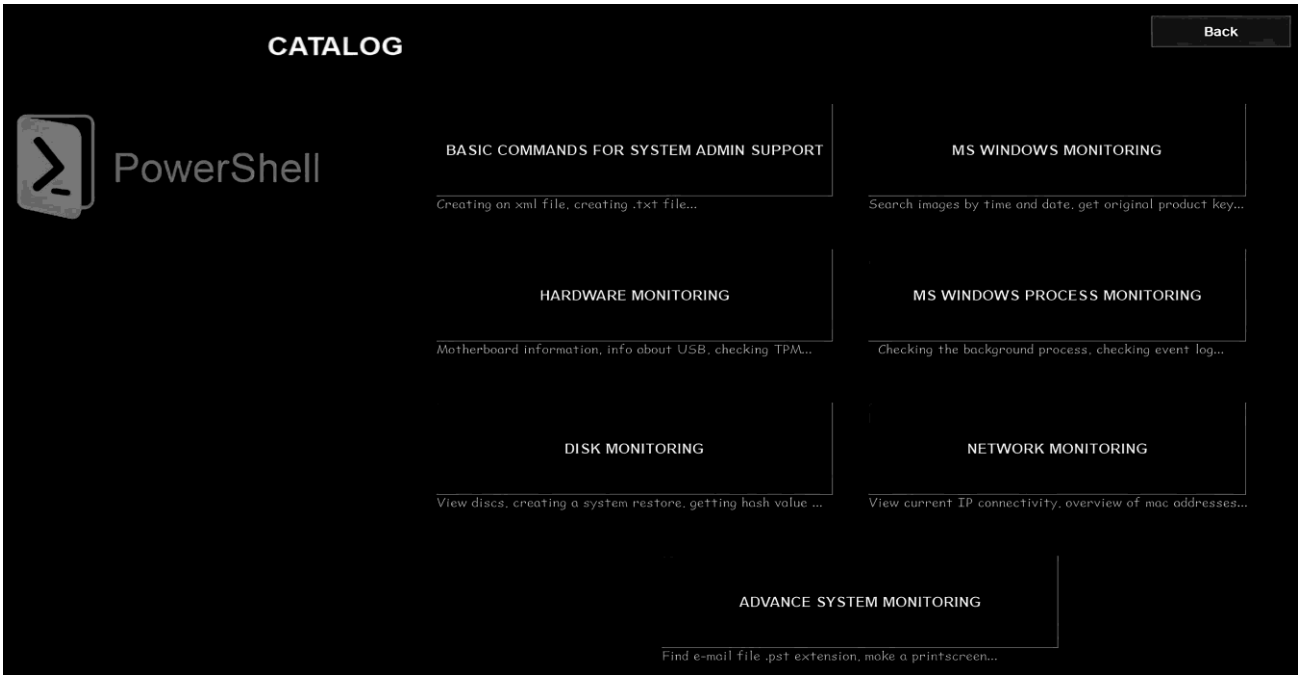


Figure 1. Main toolbar line in the application presetting the top menu, where an area of tasks can be chosen. (author's work).

This paper describes the process of developing an admin GUI application, including the tools and methods used to create the GUI. The GUI ADMIN modules and their functions will be described in detail, along with the underlying CLI commands in PowerShell. Finally, the admin GUI app will be presented, and a roadmap for its future development will be outlined.

II. TOOLS USED

A. PowerShell

PowerShell is a mixture of a command line, a functional programming language, and an object-oriented programming language [3]. It first appeared in 2006 as a replacement for various previous tools. As a tool, it is based on different languages like Python, C#, SQL, Tcl, and Puppet, with a strong resemblance to VMS CLI. The intention is to make PowerShell widely accepted, so it is ported on various other operating systems like macOS 10.12 and higher, Ubuntu 14.04, Debian 8.7+, Fedora 25, and CentOs 7. The first versions were written in the .NET Framework, while the new versions were written in the .NET Core. This intended portability makes it possibly a very useful forensic platform. Open-source software called PowerShell Core has also been released.

To implement proof-of-concept simple and complex cmdlet bar code commands representing the .Net class are used. It is important to note that cmdlets can use scripts, and the same scripts can be incorporated into modules.

With full implementation system administrators can perform tasks at remote locations but also locally because through enabled full access to Windows Management Instrumentation (WMI) and Component Object Model (COM) components tasks can be performed locally and remotely if needed.

B. Autoplay menu builder

The AutoPlay Menu Builder utility is used as a basis for application development. It enables the automatic execution of the PowerShell code in such a way that code is integrated into an executable script with the .exe extension. Implementing PowerShell code does not require programming experience, but all operations are done as drag and drop. There are security features to unauthorized prevent changes. The application supports Unicode standardization and various formats like JPG, PNG, BMP, and GIF, and it can implement video content, adobe reader, and other formats. It also supports a large font base.

C. *ISESteroids*

The environment (ISE) tool is built into the Windows operating system while The ISESteroids module is a commercially available addition that allows the encapsulation of PowerShell code into a self-contained executable. This module is not free but can be used for ten days without any restrictions. It provides many useful features in controlling the security and behavior of PowerShell code, like the application type in which code will be embedded, shell console behavior and user interaction, application signing, etc.

D. *The development process*

In the development and testing process, administrative tasks were chosen as a starting point, because of the importance of these tasks, if there is some serious misconception it is best to see them on the most important elements. Administrative tasks can be broadly separated into general system administration and operating system-specific tasks. This organization leads to three areas

1. basic commands for system admin support and advanced system monitoring;
2. MS Windows monitoring and MS Windows process monitoring;
3. hardware monitoring, disk monitoring, and network monitoring.

Because of the module functionality, each area is implemented as two or more subsets of commands. This organization is presented on the main screen of the tool "Catalog", where administrative tasks are grouped into seven different submenus, as it is shown in Figure 1.

For each chosen area tasks were defined, and appropriate commands were drafted. From this point, scripts were tested and prepared for integration.

After initial debugging and testing, there was an extended set of functional tests to explore the usability of the idea.

III. INTRODUCTION TO THE APPLICATION "EDUCATIONAL TOOL FOR SYSTEM ADMIN MONITORING "

The "Educational tool for system admin monitoring" application serves as a basis for system administrators to analyze, verify, and monitor operating systems. Also, it helps beginners in learning new commands and understands PowerShell's working environment. The application presents commands and information security aspects and a more transparent organization.

Each module presents the collection of information in a different way depending on the command that is predefined during application development (Figure 1). The modules are:

- Basic commands for system admin support,
- MS Windows Monitoring
- Hardware monitoring
- MS Windows process monitoring
- Disk monitoring
- Network monitoring
- Advance system monitoring

Each module or chapter is presented as a list of commands called "Command", followed by "Description," and an overview of commands or their execution.

The overview of commands or execution of commands is displayed in "CMD view" format, which is the basic standard text view, then "Tabular view" and "HTML View" which uses HTML elements to display the results, as it is presented in Figure 2..

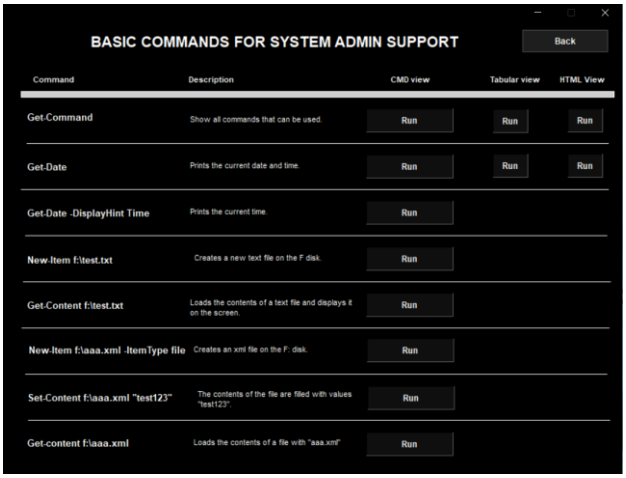


Figure 2 List of all commands on the system, cmdlet output. Each command is available as menu entry (author's work)

A. *Module "Basic commands for system admin support"*

This module will show some basic commands used for system manipulation and administration, how to display the date and time, create a new text record, display records from a text file, and retrieve and create an XML file.

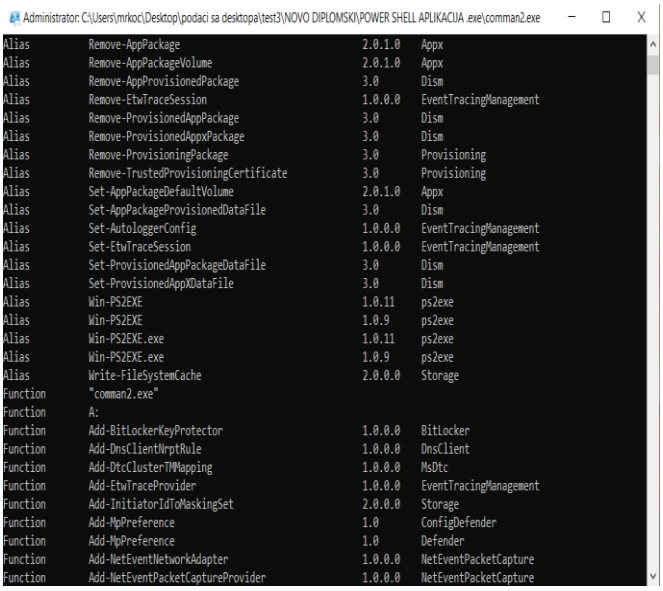


Figure 3 List off all basic commands for the system administrator support. (author's work)

As an example, Get-Command interface is presented.

1) Get-command module

Command: Get-Command

Synopsis: Displays all commands that are integrated on the local computer.

Description: Through this command, it is possible to get all command types as an alias, function, and Cmdlet commands. Results are presented in Figure 3.

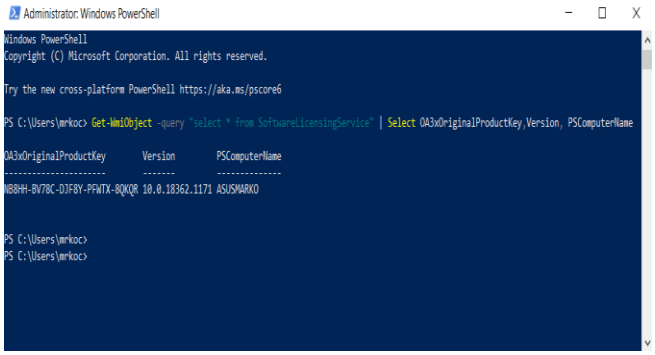


Figure 4 Windows product key as output (author's work)

Outputs: list of all available commands is printed.

In this module there are also other implemented commands:

- Get-Date – to get the system date
- Get-Content – to get file content
- Set-Content – set file content

2) Information about computer

Command: Get-WmiObject -query "select * from SoftwareLicensingService"|Select OA3xOriginalProductKey,Version, PSComputerCommand

Synopsis: This command finds the Windows key, computer name, and version.

Description: In this command, as we can see, it used the predefined commands "SoftwareLicensingService" and "OA3xOriginal ProductKey" to get the licensed key from the operating system.

Outputs: Windows product key as it is presented in Figure 4, it is in native blue background for PowerShell.

3) Hardware monitoring

Command: „Get-WmiObject Win32_Bios; Get-WmiObject win32_baseboard;Select-Object Banklabel, Manufacturer,Configuredclockspeed,Devicelocator,Capacity,Serialnumber|Get-WmiObjectwin32_physicalmemory; Get-ComputerInfo OsArchitecture,OSName“

Synopsis: Show information about the motherboard and the series and model.

Description: It can be seen that "Get-WmiObject" is used here to access digital information. With "Select-Object" we define which objects we want to display. Furthermore, the essential elements required can be seen, such as "GetComputerInfo OSName," and "OsArchitecture."

Outputs: are presented in Figure 5.

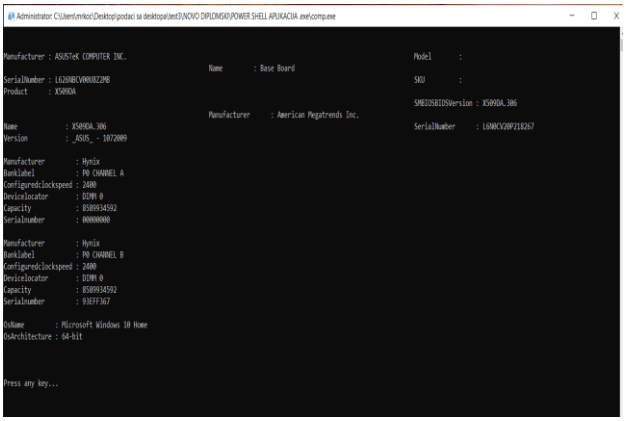


Figure ‘5 Information about the motherboard. All available info from components itselfs and registry data (author's work)

IV. FURTHER WORK

Since it is too complex to develop a whole application, it was decided to start with testing if the basic idea and needed elements are working appropriately and securely.

Examples listed here show that the concept is working and can be used for the intended purpose. There is a list of important functionalities to be implemented in the same test and proof fashion. As AIX SMIT [4] was used as a raw model additional elements needed are

- Database to store scripts, templates, and rules
- Templates for PowerShell code to be used
- Additional cmdlets for interfacing with OS and GUI
- Detailed logging facility
- Installation method
- Help subsystem
- Documentation is needed for development, usage, testing, and installation.

All these elements are essential for such a tool to be successfully and securely used, especially in incident response / digital forensic data acquisitions situation.

PowerShell as the interpreter is well suited to be sued as an interface and glue to keep all elements together. To avoid usual security problems with scripts with maximum defensive scripting and well-defined templates for the code it is possible to create secure scripts.

A. Database

The database is the backbone of such a tool, it stores code to be executed – commands, rules, and relations among this code, both with information about errors, and arguments [2]. Since the tool has to be as open as possible because of other OS support, the database should be one of the available lightweight databases with multiplatform support, also with a good interface to various programming languages [9]. The SQLite is very well suited for such a task and fulfills all these requirements

well, even though it is in extensive use on Windows from various applications and system tools [5].

B. Command templates

To keep reasonable system security and ease of maintenance command templates can be used in designing shell scripts, in this case, PowerShell commands. In such a way shape of the code, input, and command arguments can be standardized and tested before being consumed by the script itself. With this approach each script before execution is checked for validity and if approved it is executed.

C. Logging facility

Logging is essential for keeping records about events on the system and through this for general system security and reliability. As there are different logging approaches on the Windows operating system and other operating systems supported now by PowerShell, it is crucial to have a reliable and uniform interface to all logging subsystems included, to Windows logging g facility, and to syslog found on Unix-based systems.

Data that should be logged have to include:

- commands executed and their unique ids
- timestamps related to the start and stop of the commands
- arguments and inputs of the commands
- results and outputs of the commands
- debugging information if necessary
- account information under which commands were executed.

All this will provide an additional level of security and support for chain-of-evidence especially if security features like signing codes are used.

Because of the huge difference between Windows and syslog, it is possible that an additional log system should be used, where some raw data like outputs can be stored [2]. This is quite a common situation, especially on the Windows platform itself, where additional information about system events and history are stored in plain text files, while structured records are still logged into the windows logging system. The USB history is one such example. The abovementioned database can also be used as a facility to store logs, to avoid log data being too easily exposed to possible mishandling or change.

D. Utility and Interface cmdlets

To provide an efficient interface among PowerShell scripts executing tasks and the rest of the system and tool interface layer is needed, which in combination with script templates provides the required functionality. In essence, the current scripts presented in the examples in this paper require an additional wrapper layer that will provide an interface to the database and the system, while providing an environment for script execution, logging of results, presenting and preserving outputs, and all other tasks.

For example, while a task is selected through GUI, parameters, inputs, and environment data are generated for the code retrieved from the database, all this interaction is done by interface cmdlets and related utilities. The same is for handling script execution, controlling errors and results, and storing and logging generated data and script results. In essence, this layer behaves like a secure wrapper around scripts that change and modify the system keeping all tested and well-behaved.

It is important to notice that through such an approach, e.g., executable code stored in the database structures and then executed by another layer of shell code, it is possible to simplify maintenance and control of the code, resulting in more security. From a system security point of view, it adds a layer of security usually missing when system administration is done troughs scripts.

An additional benefit of such an approach is the ability to see code on request, which will run with all its arguments and variables set up, which gives huge insight into scripting practices and methods for novice system administrators, or anyone using such a system.

Development of such PowerShell code requires high programming skills and fluency in object programming, the Windows operating system, and secure coding, so it is obvious why such code was left to be implemented in further steps of development.

E. Installation Method

Installation and maintenance methods are usually overlooked in the case of system administration scripts, it is often done as simple archive extraction because of expected simplicity and “private” tools being distributed among machines. This is a huge security risk often overlooked. To avoid such a situation strict packaging into standard installation formats should be done. Since PowerShell is now available on Linux and on native Windows systems both ranges of installation packages have to be created. Again, this is a complex task same as developing the utility and interface cmdlets layer and that is why it was not implemented in the proof-of-concept tool. The current implementation is the simplest form for installation and executable which is transportable among systems.

The full implementation of the installation method will be chosen at the end of tool development, based on how other parts of the system will be implemented and how that will influence the whole tool [8]. To clarify this, the whole thing depends on the database where code and many other components are stored, so installation of the tool must incorporate the reliable installation of the database, its content, and interface layer as a set of packages. The same situation is with maintenance and upgrades where it is not yet clear which methods of updating and maintenance should be used, this heavily depends on the method of installation.

F. Help and Documentation

The help and documentation subsystem is an important part, which should be developed into a resilient and platform-independent format. Since there is a direct link between commands and help information and documentation, scripts should be developed with unique ids to ease connecting them with help files and documentation [8]. Scripts id can be based on various

identification systems, important is that each script has a unique id in each identification system, while the script still can have more than one id.

Relations among various parts of help and documentation should also be kept in the database, related to the code itself. The utility layer and GUI should be able to access help data and documentation straight from the tool itself during work, also a script code with its arguments and parameters should be visible through the GUI as part of the help system. This will also facilitate learning because the user can see the code which runs and how it interacts with the system [7].

V. CONCLUSION

As can be seen, using PowerShell and its scripts can significantly automate system processes to obtain timely information or information about a business information system. Also, the application is designed to enable the use of PowerShell commands quickly and easily, both in learning and education and in forensic research and analysis, which allows faster case resolution.

The application showed more straightforward commands such as calling only one command where an insight into the basics was obtained to complex scripts with multiple lines of code. In the following, the application is ready for possible upgrades and extensions to add more complex scripts to obtain quality information and design in the display of the same.

REFERENCES

[1] M. Mrkoci: “An educational tool for system admin monitoring using the PowerShell in the Windows environment”, Zagreb University of Applied Sciences, 2022.

[2] S. Segura: System Management Interface Tool (SMIT), IBM Redbooks, IBM, 30.11.2000

[3] C. Dent: “Mastering Windows PowerShell Scripting”, Packt Publising, 2019.

[4] IBM AIX documentation

[5] G. Allen, M. Owens: The Definitive Guide to SQLite, Apress, 2010.

[6] G.Silberschatz,,Galvin, P.,Gagne, G.: “Operating System Concepts” (10th ed.). John Wiley & Sons, 2018

[7] E.Holscher: “The Art of Documentation for Software Engineers”, Pragmatic Bookshelf, 2020

[8] W.Horton: “Developing Online Help for Computer Applications”, Wiley Year: 1991 (revised in 2003)

[9] J.Kreibich: "Using SQLite: Small. Fast. Reliable. Choose Any Three.", O'Reilly Media, 2010

Education of Teachers at Higher Education towards Changing Education Expectations

D. Paľová*, I. Korobaničová**
* Department of Applied Mathematics and Business Informatics
Faculty of Economics, Technical University of Košice, Slovakia
dana.palova@tuke.sk
** Department of Regional Sciences and Management
Faculty of Economics, Technical University of Košice, Slovakia
iveta.korobanicova@tuke.sk

***Abstract* - In the context of current technological development, its impact on the life of mankind and dramatic changes in society during the COVID-19 pandemic, the importance and its appropriate use in the education process has just arisen. The changes in education styles had to take place at all levels of the education system, including Higher education. Returning to the face-to-face form of education as it was before the pandemic is not possible anymore. Both students and teachers have discovered the wide options of technology used to further improve the educational process. However, in addition to the development of skills with the use of technology, it is necessary to develop other pedagogical skills and personal qualities of teachers, so that they can prepare their students for the continuously changing labor market. One of the projects that focuses on the development of teachers in all aspects of the educational process is "Holistic Education and Training of University teachers in the field of Economics". In this paper, the project is briefly described. Furthermore, the expected and already obtained outcomes and benefits for the entire pedagogical community and the educational process are showcased.**

Keywords - education platform, education tools, innovative teaching methods, teachers' skills development

I. INTRODUCTION

The 4th Industrial Revolution shifted the society and professional world into a more technology-supported space. And in the world of education, educators must remain abreast of these changes. The teacher's role in the education process never be the same anymore – he lost his/her role as a living "encyclopedia" that provide their student with important information. Nowadays the best teachers can help students to take ownership of their learning. That is why he/she needs to understand and be able to implement new technologies providing a variety of media and learning-support tools, that help students to receive education via the Internet, to increase their attention (for example using videos, visualization, storytelling, etc.) and approaches build the soft-skills and life-long learning abilities of the students.

Due to the emergence of advanced technologies and internationalization and globalization trends, higher education continuously changes [1]. Knowledge society, students' learning preferences and technological development lead to the process of curricula updating, and pedagogic methods innovation [2]. A lot of attention is paid especially to researching how to enhance students' learning outcomes in a more scalable, flexible environment secured by technology. Digital innovation provides wide-world access to educational opportunities and advances inclusion, enhances the relevance and quality of learning, builds ICT-enhanced lifelong learning pathways, strengthens education and learning management systems, and monitors learning processes. What is slightly omitted is the role of technological infrastructure in addressing individual and organizational advancement [3], [4].

The implementation of new technologies goes together with education transformation for almost decades, but the COVID-19 pandemic accelerated this process. Rich Henderson, Director of Global Education Solutions at Lenovo also confirmed the given situation by his *"Digital transformation has been underway in schools worldwide for the last several years. But it's been a slow transition. As a society, we had not yet crossed that threshold of full tech integration into our in-person instruction. When COVID-19 happened, it forced the digital transition on schools and teachers in a matter of weeks.*"

The available technologies like massive online courses, interactive whiteboards, virtual classrooms, augmented or virtual reality, and 3D printing support different education methods, like blended learning, etc. This transformation does not cover just the way of delivering education but promotes innovations and outside-the-box thinking [5].

However, as George Couros [6] mentions, in addition to the development of skills with the use of technology, it is necessary to develop other pedagogical skills and personal qualities of teachers, so that they can prepare their students for the continuously changing labor market and support the development of those skills that will enable them, not only the first successful entry to the labor market but also react dynamically to changes that are still taking place thanks to the development of society. One of the projects that focuses on the development of teachers in all aspects of the educational process is "Holistic Education and Training of University teachers in the field

of Economics". In the contribution, we will briefly present, the project itself, the expected and already obtained outcomes and benefits for the entire pedagogical community and the educational process. Finally, we will identify challenges to further developing higher education in Slovakia.

II. EDUCATION TRANSFORMATION

Education transformation is a wide topic, especially thanks to rapidly developing and available technologies. Today we have a lot of different technical tools and technology-based approaches, that could help teachers create more creative and experiencing learning environments for the students. Among the trendiest educational technologies, it is possible to mention virtual and augmented reality [7], [8], gamification [9], [10] and cloud technology [11], [12]. The advantages that teachers can get are:

- *Better interactive experience* – by implementing new technologies the education process becomes more dynamic by using various tools like videos, simulations digital books, etc. The learning can benefit from the deeper interest of the student followed by a higher level of his activity during the learning [13].
- *Enhanced retention of knowledge* – because of students' enhanced activity and interest in learning while he is using the technology, the retention of knowledge increases too [14], [15].
- *Support of every student's own pace of learning* – by using different tools and approaches the teacher can stimulate faster but also slower students to achieve the same educational goals [16].
- *Supported collaboration* – thanks to cloud solutions today students and teachers are now not limited to a classroom, but they can create learning groups and co-work on the activities through the school or wider space. Cooperative learning methods are tools that help to develop the hard and soft skills of the learners more effectively than in a traditional classroom [16], [17].

As stated in [18] among the main trends in education is soft skills training. The report of the World Economic Forum [19] represented the main technologies as advanced robotics, autonomous transport, artificial intelligence and machine learning, advanced materials, biotechnology, and genomics. These technologies change our lives dramatically and result in the disappearance of today's jobs and grow or create jobs that are not common yet. That is why the future workforce, our present and future students, need to be able to dynamically align its skillset to keep pace. The OECD [20] described that besides essential cognitive skills, metacognitive, social, and emotional skills become essential too. Already in 2013 Autor and Price [21] presented changes in the workplace. The Forbes [22] in 2022 ranked among the top 10 in-demand skills for the next 10 years skills Digital Literacy, Data Literacy, Critical Thinking, Emotional Intelligence, Creativity, Collaboration, Flexibility, Leadership Skills, Time Management and finally Curiosity and Continuous Learning. Among the important skills in the workplace for the future are critical thinking, problem-solving, people management, creativity, the ability to make hard decisions and leadership. Skills such as active listening, collaboration, and presenting ideas are highly valued in the modern workplace and increase the competitiveness of the person in the business environment [23]. While in the year 2000, the importance of soft skills was presented in 53% of job position offers, in 2015 it presented already 59% and it tends to reach almost 63% in 2030 (Fig. 1).

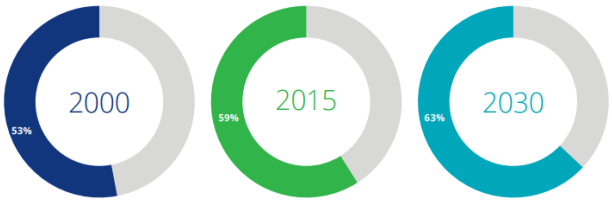


Figure 1. Soft skills intensive employment (% from all occupations); Source: [24]

The importance of soft skills doesn't decrease during the entire working life of a person; it does not matter if he is a graduate or a person who has been working for a long time. The changes in the workplace and society, evoke a huge need for education transformation. Available different technical tools enable change in the role of the teacher. Teachers are more facilitators now. Their job is to help students understand how to learn (and not in school while earning the degree but life-long), and how to uncover and understand the information they find.

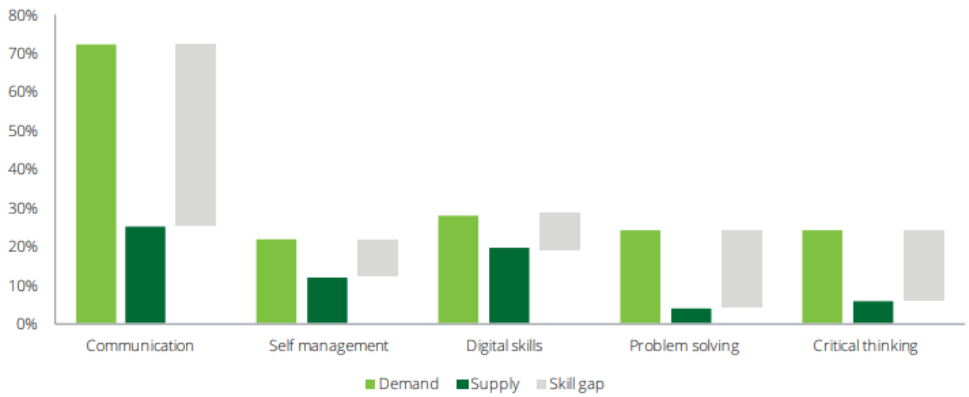


Figure 2.The soft skills gap; Source: [24]

Despite the fact, that there is a known need for more activities including the training of soft skills within the education process there is still an evident gap between the demand and supplied skills of the school graduates (Fig. 2).

Research by Majit et. al [25] identified that soft skills are considered important not only by employers but also by students. However, they found differences also in specific soft skills between these two groups. While students consider as important positive attitude, oral communication, self-motivation and self-direction, and problem-solving, employers put more emphasis on a skill like problem-solving (a skill that helps determine the source of a problem and find an effective solution. This skill at the same time includes also active listening, analysis, research, creativity, etc.), critical thinking (the purposeful, reasoned and goal-directed thinking involved in solving problems, formulating inferences, making decisions, etc.), effective communication (special kind of communication – communication in a clear, effective and efficient way), collaboration (one of the top employers' skills, that enables an employer to work well with others), and teamwork skills (qualities and abilities that allow an employer to work with others during conversations, projects, meetings or other collaborations) [26], [27]. The survey also shows that employers feel insufficient soft skills on the part of graduates and suggest the direct integration of soft skills development into the curriculum at all levels of education [28], [29]. The way how it will be done is important because students are fully aware of the importance of soft skills when entering the labor market and also in further career growth, but unfortunately, they do not make special efforts to acquire and develop these skills [30], [31]. Some authors realized that academic institutions are concerned more with hard skills development [32], [33], [34], but other ones present the possibilities of how to include this topic in education [30], [35], [36]. There are different ways how to solve this problem, i.e. to offer credit or non-credit dedicated modules for developing soft skills or to encourage teachers to include these skills in their courses' syllabi and provide an opportunity to practice soft skills in a given situation (for example by employing the group projects into education process).

The above-mentioned facts lead us to design the project "Holistic Education and Training of University Teachers in the Field of Economics", which will be shortly described following.

III. Project "Holistic Education and Training of University teachers in the Field of Economics"

Empirical studies have confirmed the critical impact of human capital on the development of innovation [37]. Analyses of the relationship between the innovation ecosystem and human capital are taking place at several levels of intellectual capital, e.g. individual, company, industry, region and/or country [38]. The formal education system has created an environment to prepare human capital that should be able to generate innovation in the labor market. Based on this knowledge the project "*Holistic Education and Training of University Teachers in the Field of Economics*" is oriented on the research of training needs and development of university teachers. The main objective of the project is to develop and evaluate an educational platform for university teachers, and lecturers. It has been focused on the research results: qualitative and quantitative research applied between university teachers and students.

During the first year of the project, the participants attended a few workshops about

- how the collaboration, creativity and idea exchange could help better to understand the end-user needs,
- learning design, i.e. how to design the training plans activities according to the needs of the end-users and eliminate areas of danger,
- education gamification principles, when and how it is possible to include it in the education process at HE.

However, to target the education of our teachers more precisely in the context of introducing innovation into education and better flexibility of the provided education regarding the development of the labor market, we decided to map the needs of our teachers in detail. The research question focuses on what the main requirements and needs are for the development of the pedagogical skills of teachers at the faculty. The questionnaire was prepared based on the previous research studies and the Teacher Questionnaire used by OECD (OECD Teaching and Learning International Survey – TALIS). The adapted questionnaire used in this study was based on the competence model: science and research competencies, teaching and learning, intra-personal qualities (self-esteem, self-confidence, coping with stressful situations, etc.), inter-personal (presentation skills, effective communication, social perception, etc.), motivational and technological (innovation in the teaching process) [39], [40], [41], [42].

In this article, the results of the questionnaire survey will be presented. The survey was carried out at the Faculty of Economics, Technical University in Košice in the online questionnaire in the academic year 2021/2022. The questionnaire was divided into a few parts: the first part of the questionnaire was consisting of identification questions. In the second part of the questionnaire were questions about personal development. The third part of the questionnaire focused on pedagogical development, and the next part was on the scientific research development of teachers. The last part of this questionnaire was devoted to the self-evaluation of the teaching process in terms of innovative learning methods. Based on the results from pilot testing on a sample of teachers from the Faculty of Economics in the year 2021, the final version of the questionnaire was prepared and implemented in the year 2022. The questionnaire contained 8 identification questions, 7 open-ended questions and 8 multiple choice or objective questions. The study was designed as a descriptive and cross-sectional study of the current state of the project mentioned above [41].

The respondents of this survey were teachers on the faculty. All teachers (61 faculty academic staff) were invited via e-mail two times during the period (from May to July 2022). A total of 28 teachers (23 teachers and 5

TABLE I. TYPE OF NEEDS IN PEDAGOGICAL DEVELOPMENT (%) – PART A

Level	Learning Design	Forms of student evaluation	Design Thinking	Team learning	Peer learning	Problem-solving method
1	25	17.9	17.9	17.9	21.4	32.1
2	42.9	42.9	50	50	32.1	32.1
3	10.7	17.9	3.6	3.6	10.7	7.1
4	7.1	7.1	3.6	14.3	21.4	10.7
5	7.1	10.7	14.3	10.7	10.7	10.7
I do not know	7.1	3.6	10.7	3.6	3.6	7.1

Source: Own Contribution

TABLE II. TYPE OF NEEDS IN PEDAGOGICAL DEVELOPMENT (%) – PART B

Level	Flipped classroom	Project-based teaching	Blended learning	Language skills	Development of digital skills
1	14.3	25	17.9	39.3	28.6
2	39.3	28.6	28.6	35.7	39.3
3	14.3	7.1	14.3	10.7	7.1
4	10.7	14.3	7.1	3.6	14.3
5	14,3	14,3	14.3	10.7	10.7
I do not know	7.1	10,7	17,9		0

Source: Own Contribution

PhD. students) participated in the online survey, 12 respondents were male (42,9%) and 16 of them (57,1%) were female in different occupations (Fig. 3).

Only some of the most interesting results will be presented in this article. Respondents were asked to indicate their agreements with their needs for personal development, pedagogical development, or science and research development on the five points scale: 1 = strongly agree; 2 = agree; 3 = Neither agree nor disagree; 4 = disagree; 5 = strongly disagree; I do not know.

According to the results of the survey (table I and II), the most suitable forms for participants' pedagogical development are courses to improve language skills (75% of respondents), digital skills (67,9%), learning design (67,9%), design thinking (67,9%) and team learning (67,9%). The academic staff would like to be educated in the research methods (85,7% of respondents would like to improve their knowledge of quantitative research and qualitative research) and would like to improve their skills with project writing (82,2%) and project and financial management (71,4%). Based on the findings from the questionnaire, the respondents in the age group 23-35 years old would like especially to improve their language skills (28,57% of all respondents) and digital skills (25%); in the age group 36-45 years would like especially to improve design thinking (28,57%) and team learning, problem-solving, digital skills and language skills (25%); and respondents in the age group 46-60 years would like especially improve team learning (21,43%) and language skills (21,42%).

The other question was about the methods or forms of personal development (training type). The most of academic staff indicated that they preferred to be educated through online webinars (71,4% of respondents), workshops (71,4% of respondents), courses to improve qualifications (67,9% of respondents), reading professional books (64,3%) and non-formal conversations with other university teachers about how to innovate teaching and learning process for students (peer sharing – 32,1%). 60.7% of respondents indicated that summer school is the least suitable method to improve their personal development. 60,7% of respondents felt that effective communication (active listening, development of personal characteristics, etc.) was another field of training for personal development. There were identified other fields for personal development: creative problem-solving methods (57,1%), self-motivation and self-management (57,1%), and coping strategies for stressful situations. For respondents, some of the following forms were not interesting for personal development: training in meeting facilitation (35, %) and emotional intelligence and self-awareness (35,7%).

IV. CHALLENGES FOR THE FUTURE HE AT THE FACULTY OF ECONOMICS

In the near history, the higher education curricula universities were more concerned with enriching the hard skills of their attendants. But, in connection with workplace requirements and the gap between the skills achieved by the students (as future employees) and the expectation of the companies, the higher education institution needs to undergo a transformation in a set of educational objectives to get more attention to soft skills development of the students. As is presented in Figure 3, the different kinds of jobs were replaced by computers (i.e. routine work such as calculating, typing, and production tasks), while the demand for non-routine interpersonal and analytical skills increased dramatically [20], [21], [43].

Including the abilities of self-regulation of learning, creativity, critical thinking, and other soft skills in higher education curricula is a challenge in itself [44].

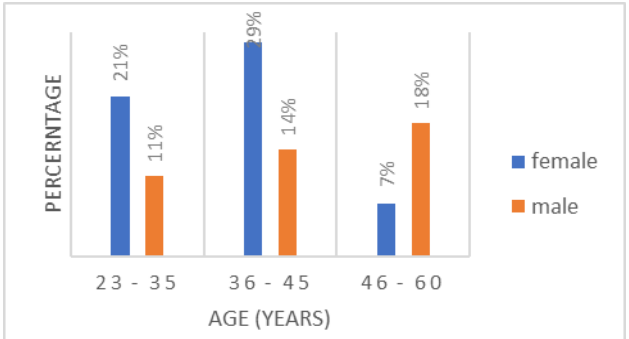


Figure 3. Representation of respondents based on gender; Source: own contribution

Based on the results of the questionnaire survey, we identified some of the challenges:

- *hard-skills development* – a crisis that arose during the pandemic clearly showed us the weak sides of digital literacy and the skills of the teachers, that need to be improved,
- *soft-skills development* – the marketplace requirements and development of society strongly demand that university graduates are expected to have not only know theories and have well-developed hard-skills, but also the soft-skills that are very useful in work-life,
- *identification and spread of innovative education methods and tools* that can be used in the education process at the faculty.

The mentioned challenges inspired us to update and enrich planned activities by following:

- workshops:
  - *presenting software* usable for qualitative research, text data management (Atlas.ti), learning design (Learning Design application), work with education platforms (LMS Moodle), teamwork (MS Teams),
  - oriented on *different teachers' soft-skills development* with practical exercises that can be implemented into the education process,
  - *presenting agile project management techniques – SCRUM methods.* Participants build the Lego robots based on the building instructions, and programming missions and underwent project management sprint planning. In this way, teachers achieve experience in how this project management method widely used in companies can be implemented into the HE educational process.
- *mentoring activities* for the establishment of professional support for teachers at the faculty.

V. CONCLUSION

The rapid changes in society go hand-in-hand with significant changes in the education field. Anymore, it is not possible to concentrate just on students' hard skills development, but more importance needs to be taken also on soft skills development. Digital innovation expands access to education and the quality of learning. As stated in [45], to transform education it is necessary to design, integrate and implement effective national policies and masterplans on learning making sure activities on the ground answer the needs of each country and community. All mentioned institutions, organizations and associations work on frameworks and different guidelines for the teachers working at different levels of education and different forms of education. At the same time, all worldwide institutions emphasize the importance of designing national frameworks and action plans for education innovation and transformation. Besides official documents, the best practice seems to be the community knowledge and experience sharing that help teachers to do better and up-to-date jobs. In this paper, we presented our activities concerned with the teachers at higher education institutions. Thanks to participation in the above-mentioned project, we identify the obstacles that slow down the education innovation process and propose possibilities for how to improve the situation at the faculty. The outcomes will be generalized and published to all Slovak educators interested in the transformation of education.

ACKNOWLEDGEMENT

This project "Holistic Education and Training of University Teachers in Field of Economics" (Project no. 053TUKE-4/2021) is funded by the grant scheme KEGA of the Ministry of Education, Science, Research and Sports of the Slovak Republic.

REFERENCES

[1] J. Khalid, B.R. Ram, M. Soliman, A.J. Ali, M. Khaleel, and M.S. Islam. "Promising digital university: a pivotal need for higher education transformation", in Management in Education, Vol. 12, No.3, 2018, pp.264–275. https://doi.org/10.1504/IJMIE.2018.092868

[2] R. Bridgstock, "Educating for digital futures: what the learning strategies of digital media professionals can teach higher education", in Innovations in Education and Teaching International, Vol. 53, No. 3, 2014, pp.306–315. https://doi.org/10.1080/14703297.2014.956779

[3] N. Dabbagh, and A. Kitsantas. "Personal learning environments, social media, and self-regulated learning: a natural formula for connecting formal and informal learning", in The Internet and Higher Education, Vol. 15, No. 1, 2012, pp.3–8. https://doi.org/10.1016/j.iheduc.2011.06.002

[4] M. Henderson, N. Selwyn, and R. Aston. "What works and why? Student perceptions of 'useful' digital technology in university teaching and learning", in Studies in Higher Education, Vol. 42, No.8, 2017, pp.1567–1579. https://doi.org/10.1080/03075079.2015.1007946

[5] D. Barker. "How the Education Industry Can 'Level Up' with Technology?", 2022, Available at: https://www.quatrrobss.com/articles-blogs/how-the-education-industry-can-level-up-with-technology/ [Accessed 15-Jan-2023]

[6] G. Couros. "Technology is Only Transformational if We Choose to Use It That Way", 2023, Available at: https://georgecouros.ca/blog/archives/9046 [Accessed 20-Mar-2023]

[7] J. Garzón, J. Pavón, and S. Baldiris. "Systematic review and meta-analysis of augmented reality in educational settings." In Virtual Reality, Vol. 23, 2019, pp.447–459. https://doi.org/10.1007/s10055-019-00379-9

[8] J. Garzón, Kinshuk, S. Baldiris, J. Gutiérrez, and J. Pavón. "How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis". In Educational Research Review, Vol. 31, 2020, https://doi.org/10.1016/j.edurev.2020.100334

[9] A. Manzano-León, P. Camacho-Lazarraga, M.A. Guerrero, L. Guerrero-Puerta, J.M. Aguilar-Parra, R. Trigueros, and A. Alias. "Between Level Up and Game Over: A Systematic Literature Review of Gamification in Education". In Sustainability, Vol. 13, No. 4, 2021. https://doi.org/10.3390/su13042247

[10] K. Ofosu-Ampong. "The Shift to Gamification in Education: A Review on Dominant Issues". In Journal of Educational Technology Systems, Vol.49, No.1, 2020, pp.113–137. https://doi.org/10.1177/0047239520917629

[11] O. Nooh Almotiry, M. Sha, M. Parveen Rahamathulla and O. S. Dawood, "Hybrid cloud architecture for the higher education system," Computer Systems Science and Engineering, Vol.36, No.1, pp.1–12, 2021. https://doi.org/10.32604/csse.2021.014267

[12] N.V. Valko, N.O. Kushnir, and V.V. Osadchyi. “Cloud technologies for STEM education”. In CTE Workshop Proceedings [Online], Vol.7, pp.435–447, 2020. https://doi.org/10.55056/cte.384

[13] I. Bouchrika. “10 Top Interactive Learning Trends: 2023 Data, Insights & Predictions”. 2022, Available at: https://research.com/education/interactive-learning-trends [Accessed 20-Mar-2023]

[14] M. Marco-Fondevila, M. Rueda-Tomás, and M. P. Latorre-Martínez. “Active Participation and Interaction, Key Performance Factors of Face-to-Face Learning”. In Education Sciences, Vol. 12, No. 7, pp. 429, 2022. https://doi.org/10.3390/educsci12070429.

[15] M. Tight. “Student retention and engagement in higher education”. In Journal of Further and Higher Education, Vol.44, No.5, pp.689–704, 2020. https://doi.org/10.1080/0309877X.2019.1576860.

[16] H. Mora, M.T. Signes-Pont, A. Fuster-Guilló, and M. L. Pertegal-Felices. “A collaborative working model for enhancing the learning process of science & engineering students”. In Computers in Human Behavior, Vol. 103, pp.140-150, 2020. https://doi.org/10.1016/j.chb.2019.09.008.

[17] F. Arató. “Towards a Complex Model of Cooperative Learning.” In Da Investigaçãoàs Práticas, Vol. 3, No. 1, 2013, pp. 57-79.

[18] E. Debétaz. “5 Trends in Education that continue in 2022”, 2022, Available at: https://hospitalityinsights.ehl.edu/education-trends-2022 [Accessed: 15-Jan-2023]

[19] A. Gray. “The 10 skills you need to thrive in the Fourth Industrial Revolution”, 2016, Available at: https://www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution/ [Accessed: 15-Jan-2023]

[20] OECD. “Future of Education and Skills 2030”, Concept Note, 2019, Available at: https://www.oecd.org/education/2030-project/teaching-and-learning/learning/skills/Skills_for_2030_concept_note.pdf [Accessed 2-Feb-2023]

[21] D. H. Autor, and B. Price, “The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003),” MIT Mimeograph, 2013, Available at: https://app.getsphere.com/david-autor/articles/the-changing-task-composition-of-the-us-labor-market-an-update-of-autor-levy-and-murnane-2003 [Accessed 15-Jan-2023]

[22] B. Marr. “The Top 10 Most In-Demand Skills For The Next 10 Years”, 2022, Available at: https://www.forbes.com/sites/bernardmarr/2022/08/22/the-top-10-most-in-demand-skills-for-the-next-10-years/?sh=62cf2d0a17be [Accessed 2-Feb-2023]

[23] J.J. Heckman, and T. Kautz. “Hard evidence on soft skills”. In Labour Economics, Vol. 19, No. 4, 2012, pp. 451-464. https://doi.org/10.1016/j.labeco.2012.05.014.

[24] Deloitte. “Soft skills for business success”, 2017, Available at: https://www2.deloitte.com/au/en/pages/economics/articles/soft-skills-business-success.html [Accessed 15-Jan-2023]

[25] S. Majid, C. M. Eapen, E. M. Aung, K. T. Oo. “The importance of soft skills for employability and career development: Students and employers’ perspectives.” IUP Journal of Soft Skills, Vol. 13, No. 4, 2019, pp.7-39. Available at: https://www.proquest.com/scholarly-journals/importance-soft-skills-employability-career/docview/2330760547/se-2 [Accessed 16-Mar-2023]

[26] J. Herrity. “What Are Problem-Solving Skills? Definitions and examples”, 2023, Available at: https://www.indeed.com/career-advice/resumes-cover-letters/problem-solving-skills [Accessed 20-Mar-2023]

[27] A. N. Kaminske. “Is Critical Thinking a Soft Skill?”, 2020, Available at: https://www.learningscientists.org/blog/2021/3/25-1 [Accessed 20-Mar-2023]

[28] C. Stewart, A. Wall and S. Marciniec. “Mixed Signals: Do College Graduates Have the Soft Skills That Employers Want?”. Competition Forum, Vol. 14, No. 2, 2016, pp. 276-281.

[29] A. Singh and P. Jaykumar. “On the Road to Consensus: Key Soft Skills Required for Youth Employment in the Service Sector”, Worldwide Hospitality and Tourism Themes, Vol. 11, No. 1, 2019 pp. 10-24. http://dx.doi.org/10.1108/WHATT-10-2018-0066

[30] M. Itani and I. Srour (2016). “Engineering Students’ Perceptions of Soft Skills, Industry Expectations, and Career Aspirations”, Journal of Professional Issues in Engineering Education and Practice, Vol. 142, No. 1, 2016, pp. 1-12. http://dx.doi.org/10.1061/(ASCE)EI.1943-5541.0000247

[31] L. Laari and B. M. Dube. “Nursing Students’ Perceptions of Soft Skills Training in Ghana”, 2017, Available at: http://www.scielo.org.za/pdf/cura/v40n1/16.pdf [Accessed 20-Mar-2023]

[32] J. Balcar. “Is It Better to Invest in Hard or Soft Skills?”, Economic and Labour Relations Review, Vol. 27, No. 4, 2016, pp. 453-470. http://dx.doi.org/10.1177/1035304616674613

[33] A. H. Cooley. “Piggybacking on Business Communication Through Interdisciplinarity: Developing Student Communication Skills in Legal Environment of Business Courses”, Business Communication Quarterly, Vol. 72, No. 4, 2009, pp. 431-438. http://dx.doi.org/10.1177/1080569909349525

[34] F. Hairi, M. N. A. Toee and W. Razzaly. “Employers’ Perception on Soft Skills of Graduate: A Study of Intel Elite Soft Skills Training”, Paper presented at the International Conference on Teaching & Learning in Higher Education (ICTLHE 2011), pp. 1-8, Melaka.

[35] M. Wats and R. K. Wats. “Developing Soft Skills in Students”, The International Journal of Learning, Vol. 15, No. 12, 2009, pp. 1-10. http://dx.doi.org/10.18848/1447-9494/CGP/v15i12/46032

[36] K. N. Tang. “Beyond Employability: Embedding Soft Skills in Higher Education”, Turkish Online Journal of Educational Technology, Vol. 18, No. 2, 2019, pp. 1-9.

[37] A. Kianto, J. Sáenz, and N. Aramburu, “Knowledge-based human resource management practices,” Intellectual capital and innovation. 2017(81), pp. 11-20. https://doi.org/10.1016/j.jbusres.2017.07.018.

[38] S. Kundu and S. Munjal, “Human Capital and Innovation. Examining the Role of Globalization,” 2015, Palgrave Studies in Global Human Capital ed. S.l.: Palgrave Macmillan.

[39] M. Blašková et al. “Key personality competencies of university teacher: comparison of requirements defined by teachers and/versus defined by students”, in 4th World Conference on Psychology, Counseling and Guidance WCPCG-2013, Procedia – Social and Behavioral Sciences 114, 2014, pp. 466-475.

[40] G. Kravčáková, I. Jeleňová. “Kľúčové kompetentntosti vysokoškolského učiteľa. In Verejná správa a spoločnosť, Vol. 11, 2010, pp. 40-56. [In Slovak] Available at: https://www.researchgate.net/publication/295548747_Klucove_kompetentnosti_vysokoskolskeho_ucitela [Acessed 24-Mar-2023]

[41] A.K. Yousif, O.Y. Ahmed, W.N. Osman. „Training needs assessment of academic teaching staff in Faculty of Dentistry“, in Education in Medicine Journal. Vol.11, No. 1, pp. 31–41, 2019. https://doi.org/10.21315/eimj2019.11.1.4

[42] OECD. „Teaching and Learning International Survey (TALIS). Available at: https://www.oecd.org/education/talis/ [Accessed 25-Mar-2023]

[43] OECD. “Trends Analysis Future Shocks and Shifts: Challenges for the Global Workforce and Skills Development”, 2017. Available at: https://www.oecd.org/education/2030-project/about/documents/Future-Shocks-and-Shifts-Challenges-for-the-Global-Workforce-and-Skills-Development.pdf [Accessed 25-Jan-2023]

[44] T. Plomp. “Preparing education for the information society: the need for new knowledge and skills”, in International Journal of Social Media and Interactive Learning Environments, Vol. 1, No. 1, 2013, pp.3–18. https://doi.org/10.1504/IJSMILE.2013.051651

[45] UNESCO. “What you need to know about digital learning and transformation of education”, 2023, Available at: https://www.unesco.org/en/digital-education/need-know [Accessed 25-Jan-2023]

Simulation of Corporate Work-life as a Part of Higher Education

D.Paľová*, M. Vejačka*
* Department of Applied Mathematics and Business Informatics
Faculty of Economics, Technical University of Košice, Slovakia
dana.palova@tuke.sk, martin.vejacka@tuke.sk

***Abstract* - The phrase "*Data is the new oil*" is becoming truer thanks to rapid technological development and our growing dependence on data. It is therefore very important that students' skills also include the ability to work with data (to search, understand, process, visualize, analyze, etc.). Today, data is used in decision-making and management in all areas of our lives. Therefore, also of Economics need to have an overview and experience in this field before entering the labor market. The long-term cooperation with the Deutsche Telekom IT Solutions company resulted in the idea of realizing a course where students would work with real corporate data during their studies, similarly to the employees in their normal working life. The course was created so that the leading employees of this company played an active part in its educational process. The paper presents the content and method of implementation of this subject, the results from its pilot run, and feedback from both, students and employees. Furthermore, the challenges that the pilot implementation of the subject posed to us in the coming years also are presented.**

Keywords - Data analysis; Higher education; Educational process; Innovative teaching methods;

I. INTRODUCTION

The importance of employability of higher education institution graduates is an often-discussed topic not only on the political scene and at the universities, but also in the business environment [1]. We live in a world where technology plays a larger role in many aspects of our day-to-day lives and education is no exception. Advancements in technology have changed nearly every facet of the educational system; it has not just transformed the way education is delivered and received but also promotes outside-the-box thinking and innovation [2]. Massive open online courses, interactive whiteboards, virtual classrooms, augmented reality, and 3D printing, all of which support blended learning methods, are testaments to these changing dynamics of education [3].

Novel educational technology trends can help develop interesting learning experiences for students (e.g. virtual reality, augmented reality, gamification, blockchain technology, virtual learning, and cloud technology). Thanks to them we could change and innovate the learning process in different ways to be very close to our students [4].

As trends like big data, machine learning, and the Internet of Things (IoT) have come into the picture, many educational institutions have been focusing on how they can adopt them. According to Holon IQ, in 20the digital spending on education was $227 billion, and it is forecasted to be $404 billion by 2025 [5]. Problem-based learning (PBL) emerges from discussions on the centrality of problem-solving in all aspects of everyday life and the need for education to go beyond the transmission of knowledge which often graduates find difficult to apply to new contexts and at the workplace [6]. In PBL contexts, learners solve problems in collaboration with peers as part of a self-directed learning approach. In that respect, responsibility for learning shifts toward the learner and/or the group of learners [7]. In PBL, knowledge and skills are organized around authentic problems which are often ill-structured rather than being organized around a hierarchical list of topics. As such, knowledge building in PBL approaches is stimulated by the problem, applied back to it, and structured around it [8], [9].

Simulation-based learning places students in situations where they can put their classroom knowledge into practice. Educational simulation is a teaching method that tests participants' knowledge and skill levels by placing them in scenarios where they must actively solve problems. The instructor defines the parameters to create a safe environment for hands-on learning experiences. A simulation is a form of experiential learning [10]. It is a strategy that fits well with the principles of student-centered and constructivist learning and teaching; that is, learning and teaching that gives students power over what they learn and how they learn it, and that allows students to build their understanding of what they're learning through their experiences and interactions, rather than just passively receiving information [11].

By promoting direct participation of students in their knowledge building, simulation-based education (SE) incentives intrinsic motivation in the hands-on learning activities designed by the instructors and developed under their supervision and guidance. Learning practices based on the use of SE software, tools, and games are receiving the attention of academic and industrial partners, who see SE as a way to improve the training of their students and future employees [12]. When participating in a scenario, students must quickly evaluate the situation, decide on the best course of action, and perform the correct procedural

steps. Educators can then assess whether the students understand the material and are translating their learned knowledge into skills.

II. DATA IS NEW OIL

The most useful resource of the 21st century is data. Data mean to the digital economy the same as oil meant to the industrial economy. But what is important to take into consideration, just data itself has no value. It needs to be mined and analyzed and just after this, it creates value [13]. Various types of information can be extracted from data because it promises an endless number of new uses, i.e. diagnosis of diseases, the direction of traffic patterns, estimation of customer behavior, economic forecasts, etc. But to achieve these outcomes, a lot of work in data science, engineering, and data interpretation is needed. The five most valuable companies (Apple, Amazon, Facebook, Microsoft, and Google) dominated their sector mainly thanks to the sale of access to data. At the same time, they collect information about every aspect of our lives and this data gives them more profit and power. Thanks to a huge mass of data every metric of business can be tested, measured, and improved [13]. The Global Sentiment Survey 2023 explored that due to the launching of the ChatGPT the AI and data analysis trends are back [14]. This leads to a constant need in the labor market for people who are oriented and have experience in this field.

III. IMPLEMENTING DATA ANALYSIS TOPICS INTO EDUCATION

The Faculty of Economics tries to provide up-to-date education to follow the marketplace expectations on higher education graduates. That is why that list of mandatory courses is continuously supplemented by the courses that cover the newest trends in the real world. One of the topics which we implement wider is data analysis and data visualization. That was the purpose of designing new optional courses (i.e. Introduction to R Statistics, Data Analysis and Reporting, Introduction to Datamining, Selected models and analyses in banking, Data Analysis and visualization) that create data analysis learning path across the whole study at the faculty. These courses develop students' skills and follow on from compulsory subjects such as Informatics I, Informatics II, Statistics and Probability, and Statistical Methods in Economics. Our objective is not only to acquaint students with the basic terminology in the field of data analysis but also through practical tasks to ensure the acquisition of skills in the use of data preprocessing methods as well as the subsequent use of methods of analysis and self-paced learning, data visualization on specific problems created by everyday practice. Following we describe in more detail our experience with the brand new course Data Analysis and Reporting.

This course is taught in the bachelor's degree of study branch Finance, Banking and Investment and can be chosen by students who have completed the compulsory subjects Informatics I and Informatics II, where they were introduced to programming in the environment of the R language, basic functionalities of MS Excel application, introduction into information systems and databases. Previously listed form the basis for the optional course, where students are gradually introduced to various business intelligence tools and methods that are used in practice in the analysis and subsequent interpretation of data. The course is focused on solving problems through various models and methods of data analysis on real business data.

The first run of the course was realized during the winter semester of the academic year 2022/2023. Because the course is optional, it is usual that a new course arouses little interest among students, especially if it is a subject connected with information technologies. So this first run was attended by 8 students who are interested in data science. The content covers the topics like introduction to data analysis, obtaining data from external sources, data pre-processing, detection of basic characteristics, and detection of anomalies and outliers.

These topics were divided into 13 face-to-face lectures, supported by an online course in LMS Moodle, where students can find a lot of additional information sources, best practices connected to the discussed topic, and more exercises to become more skilled in the field of data analysis and reporting, design, and development of simple or more complex visualization in MS Power BI application and interpretation of data. During the face-to-face lessons, we practically get across the whole data analysis process using anonymized real enterprise data, that are prepared for us by the company Deutsche Telekom IT Solutions Slovakia (DT). Cooperation with this company already started about five years ago, when employees of the Finance and Controlling Department participated in lectures concerned on the MS Excel application used in the DT company. This cooperation inspired us to involve more in the practical education of our students. Within the course, the DT employees actively participated in the design of the course, had lecturers with presenting the real connections to data saved in DT information systems, and following operations usually done in their daily routine, prepared and evaluated the semestral project.

The semestral project was fully covered by colleagues from the DT company and it was designed like an assignment of a project with which employees in the Financial Controlling department meet daily. The students had at their disposal dates, a short description of expected outcomes, delivery dates and a request in which way and in what form the performances should be submitted. The main idea was to prepare a report with financial results for the last 2 years. Students are in the role of Chief Business Analyst and they need to prepare a report for the CFO of the DT company to better understand the situation with the realized projects for the clients. The final solution was expected to be an application prepared for summary & detailed data analysis of your data set which needs to be visually attractive, cover all mentioned topics, and be ready to present to the CFO of the company.

The main objective of the implementation of the semestral project was to develop the practical skills and competencies of students and also enhance students to think and act more independently without intervention or instructional guidance by the teacher. For the students it was a new experience also of the amount of data that they

have to face, the dataset contained about 50 000 rows and 10 attributes for the years 2020 and 2021 and to solve the task students need also the search for missing information on the Internet. There was no right or accurate solution set in advance, so they need to be also creative in the design of visualization (Fig. 1) and following presentation of the achieved results (Fig. 2).

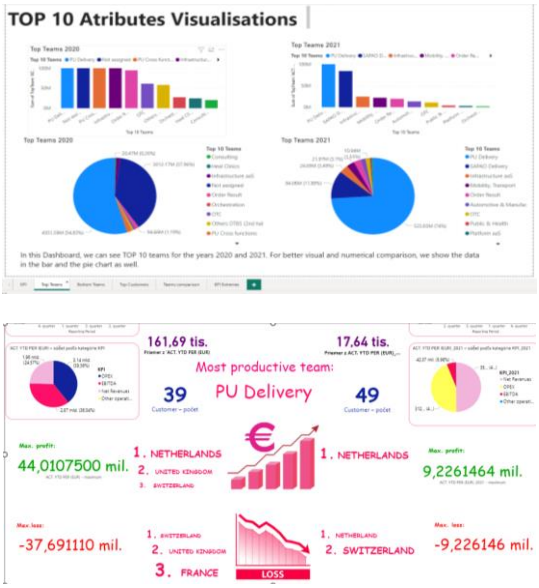


Figure 1. Examples of students' results; Source: own contribution



Figure 2. Final presentations of semestral projects; source: own contribution

The course evaluation is divided into two parts:

- practical credit exam (40 points of final available 100 points) that is aimed more at checking practical skills of working with data, their pre-processing, and subsequent visualization through the MS Power BI application.
- semestral project (60 points of final available 100 points) that is aimed more at the individual work of students and is described in more detail above.

Participating students surprised us. Over 75% of students achieved results better than 81 points in the final evaluation (82 points on average). They were able to prove their knowledge in working with the application used for pre-processing and working with data. The average result achieved for the credit exam was 32.63 points. The positive information is that our students achieved pretty good evaluations for their work on the semestral project which was evaluated by the DT employees according to the same criteria as the success of the delivery of the expected product would be evaluated in practice. They achieved 49.38 points from a maximum of 60 points. Particular results in more detail are presented in Table 1.

TABLE I. ACHIEVED RESULTS OF STUDENTS (IN THE PERCENTAGE OF THE GROUP OF ALL STUDENTS)

Success interval	Evaluation parts		
	Credit exam	***Semestral project***	***Final result***
91 – 100pt	25.00%	12.5%	25.00%
81 – 90pt	0.00%	62.5%	50.00%
71 – 80pt	62.5%	0.00%	12.50%
61 – 70pt	12.50%	25.00%	12.50%
51 – 61pt	0.00%	0.00%	0.00%

Source: own contribution

In general, based on the achieved final point results we can state, that course is well-designed and the main objectives are fulfilled. Besides the points, education has other associate effects and benefits:

- achievement of hard and soft skills that are required by the future employers of the students. This represents a significant benefit when looking for a job.
- getting an idea of how employees receive assignments in practice, what conditions they have to meet, and how they are evaluated in the end. This kind of achievement is slightly different from those they already have in their past; they can fulfil different criteria that are set usually by the teacher that follows the accreditation rules.
- get experience to present achieved results in to present and defend the method of data interpretation for people knowledgeable in the given field, who followed not only the "colors" of the presentation but also the complexity and meaning of the offered solution.
- and finally, the best three semestral projects get the DT gift thing and an offer to spend one day in the so-called shadow in their company, where they will have the opportunity to see how it works in practice, which can be followed by a part-time work offer.

The final evaluation points are just one side of the education process. Because the course realization was first run (and for the people from DT also first experience in cooperation such kind of course) of the course like this we were curious about the students' opinion. After passing the course we asked students to answer feedback to reach a better quality of data analysis education, to enhance

students' interest and motivation to study this topic. The feedback contained 77 questions and can be divided into the following parts:

- the reasons why they choose this course,
- course organization and published information,
- semestral project,
- course content and content of particular chapters,
- the way of evaluation of course
- and their opinion on the DT employees' participation in the education process.

As was mentioned above, the course was attended by 8 students, so the achieved results cannot be generalized, but we achieved the opinion of every student, which is why we decided to take it into account and prepare the next run better. After summarizing the answers, it is possible to mention the most interesting and important outcomes connected to specific parts of the feedback.

A. The reasons why they choose this course

The most frequently occurring reasons were the possibility of acquiring additional skills that students could use in the future in practice, as well as their awareness of the importance of knowledge in the field of data processing. An interesting reason, that was also mentioned, was a teacher with whom they already had positive experiences in the past.

B. Course organization and published information

The subject as a whole was evaluated with an average grade of 4.25 out of 5. Table 2 represents students' answers with assessments concerning the subject as a whole. Students would appreciate more practical exercises, which also supports the opinion that the number of hours per week in the given subject should be increased, with 62.5% of students proposing to increase the time span of the subject by at least 1-2 hours per week.

TABLE II. STUDENTS' SATISFACTION WITH THE COURSE ORGANIZATION AND PUBLISHED INFORMATION

Criteria	**Percentage of chosen answers**				
	Very satisfied	***Satisfied***	***Neither satisfied nor dissatisfied***	***Dissatisfied***	***Very dissatisfied***
Clarity of course organization	37.5	62.5	0.0	0.0	0.0
Amount of published information	50.0	37.5	12.5	0.0	0.0
The way of course realization	37.5	62.5	0.0	0.0	0.0
The proportion of theory vs. practice	25.0	37.5	0.0	37.5	0.0
Hourly subsidy for the subject per week	12.5	12.5	12.5	37.5	25.0

Source: own contribution

C. Course content and content of particular chapters

Over 87.5% of students stated that the published information inspired them to continue working with data in the future. As part of the feedback, we analyzed all parts of the content of the subject in detail (i.e. individual chapters of the subject, however, due to the scope of the contribution, it is not possible to present the obtained results to a broader extent). Each chapter of the subject was divided into the theoretical background, practical tasks, and additional study material. Students evaluated each of these components separately from several aspects (see Table 3).

TABLE III. STUDENTS' SATISFACTION WITH THE SEMESTRAL PROJECT

Criteria	**Percentage of chosen answers**				
	Very satisfied	***Satisfied***	***Neither satisfied nor dissatisfied***	***Dissatisfied***	***Very dissatisfied***
Clarity of SP organization	37.5	25.0	25.0	12.5	0.0
Amount of published information	50.0	25.0	12.5	12.5	0.0
The difficulty of SP (the project was not so difficult)	12.5	25.0	25.0	37.5	0.0
Way of evaluation	37.5	25.0	25.0	12.5	0.0
Way of realization	12.5	75.0	12.5	0.0	0.0

Source: own contribution

In the context of practical preparation and strengthening of the student's independence in the study process, we were also interested in whether they tried to do the tasks independently at home, although this activity was not evaluated. Almost 87.5% of students confirmed to us that the topic interested them and they tried to do the assignment independently, and 62.5% of students did not have a problem with the given tasks during homework. If there was a problem, all respondents confirmed that they solved it as part of a team with colleagues.

D. Semestral project

As part of the semester project, students could try out how projects are handled in an IT company. They could also evaluate this time of the subject, and we were interested in several aspects since this assignment was fully organized by the employees of DT. The results are presented in Table 4.

E. The way of evaluation of course

As already mentioned, the evaluation of the subject was divided into two parts - a report card and a semester project. The first thing we were interested in was whether they considered the distribution of points in the ratio of 40/60 to be fair. From the point of view of the students, this division is surprisingly unfair for some, they would accept more points for the credit report (25%), some students (25%) cannot decide and 50% are satisfied with the current setting.

TABLE IV. STUDENTS' EVALUATION OF THE COURSE CONTENT

Content criteria	Percentage of chosen answers				
	Very good	***Good***	***Acceptable***	***Poor***	***Very poor***
Quality of theory	75.0	12.5	12.5	0.0	0.0
Quality of supplementary study material	75.0	12.5	12.5	0.0	0.0
The proportion of theory vs. practice	50.0	25.0	25.0	0.0	0.0
Method of providing information	75.0	12.5	12.5	0.0	0.0
Difficulty in understanding the presented information	50.0	25.0	12.5	12.5	0.0
The difficulty of practical exercises	25.0	50.0	12.5	12.5	0.0
Supporting understanding of the topic through practical tasks	62.5	25.0	12.5	0.0	0.0

Source: own contribution

Most of the students rated credit written test as undemanding (Fig. 3).

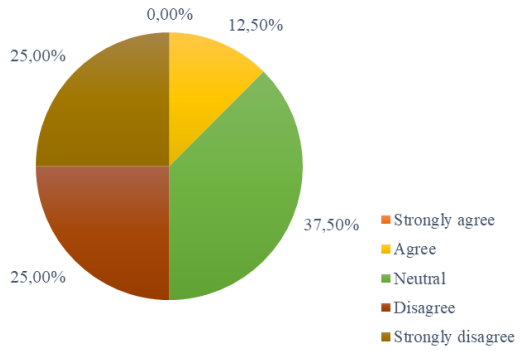


Figure 3. Rate of agreement with statement: "The credit test was demanding"

F. DT employees' participation in the education process

An important part of the course was the cooperation with external colleagues from practice. That is why we and they were interested in how the students perceived this kind of participation in the subject. The students evaluated several aspects of cooperation, including the level and implementation of presentations as well as the impact of such participation on their motivation to continue studying. Table 5 presents the resulting findings.

After completing the subject, they also identified the added value of the given subject as an insight into the environment of data analysts, they improved in working with applications such as MS Excel, MS Power BI, and MS PowerQuery, cooperation with a real company, sharing knowledge with practice and the possibility to solve problems completely independently a problem.

IV. CHALLENGES FOR THE FUTURE

Whenever we implement a new education method or realize a new course, we analyze the course from different points of view: the content, way of organization and realization, and level of achieved learning objectives. The course is analyzed and evaluated by the teacher, external partners, and by students. The evaluation process helps us identify the challenges for further improvement and set of actions how to fulfil them. Based on our current experience, the results of feedback from the students, and discussions with the employees of DT Slovakia, we identified several challenges for the future implementation of the given subject:

- *lack of time allocated for face-to-face learning.* In the current academic year, we incorrectly set the hourly subsidy for the subject. Many times we also discussed the topic with the students during the lectures, but there was not enough time for practical training or checking our conclusions from the discussion. It is, therefore, necessary to at least double the weekly one-hour subsidy for the subject.
- *The ratio of lectures and exercises*. This challenge is connected to the first one. Students would like to have more practical exercises with instructions for working with data since they have never worked with such large data before. This challenge could be fulfilled together with the first challenge.
- *semestral project realization and evaluation*. SP was evaluated by both parties as profitable, but in the future, it will not be appropriate to motivate students to solve SP points for individual partial tasks, because they often immersed themselves only in solving these tasks and underestimated the more complex interpretation of data in a wider context. In addition, the point evaluation weighed the students' hands in their creativity. In the future, it would be more appropriate to set up a friendlier environment, so that the student does not fear that for the incorrect presentation of data, they may lose points, which are transferred to the final evaluation. In addition, to improve the project and to test the students'

TABLE V. STUDENTS' OPINION ON THE PARTICIPATION OF DT EMPLOYEES IN THE EDUCATION PROCESS

Criteria/Statement	Percentage of chosen answers				
	Absolutely agree	***Agree***	***Neutral***	***Disagree***	***Absolutely disagree***
Participation in DT is beneficial for the education	62.5	37.5	0.0	0.0	0.0
Employees' approach was sufficient	62.5	37.5	0.0	0.0	0.0
Employees' approach was motivating	50.0	25.0	25.0	0.0	0.0
Lectures were captivating	75.0	25.0	0.0	0.0	0.0
Lectures had accurate proportion theory and practice	62.5	12.5	12.5	12.5	0.0

Source: own contribution

performance in practical life, it is possible to use the SCRUM methodology during the semester project, where students can try different roles, at the same time they will have to work regularly and report the results and thanks to the meetings after the sprints they will also get instant feedback on their work in the time when solving individual tasks.

- *lack of students in the course*. During the first run, the course was attended just by 8 students. That is why it is necessary to think of some way to promote the optional Data Analysis learning path in the compulsory courses and thus attract more students to the field. We have some ideas on how to motivate younger students, for example presenting the performances of colleagues from the course during lectures of their younger colleagues and possible future attendants; pointing out that this course is practical for their future work or development in this direction, as unlike other optional courses, where they could work with something that is used in practice; preparation of promotion activities in cooperation with DT; short promotion video prepared by students that they passed the course.

V. CONCLUSION

After processing the feedback on the course from the students, we got several interesting statements, such as the fact that 75% of the students agreed that better participation in the lessons would be helpful if the students would explain things to each other. All students confirmed that they consider the course realized in this way to be very beneficial for their future life. Thanks to passing the course, they could try out the way of working in a real company in a safe environment and the wide possibilities of working with data. The interesting outcome was that even though educational theory states that when using simulation-based learning it is important to have clearly defined assessment criteria, our students were in a sense less creative and more focused on the assessment they could get for completing individual tasks.

In the paper, we presented a positive confirmation of the fact that students are more interested in education if we continuously ask for their opinion if we add a huge part of the practice. Using an accurate mixture of it, we will reach a win-win situation. Thanks to the cooperation with practice and the simulation of real life in the company, we get motivated students who have the skills necessary for practice with an interest in dealing with the given field in the future. This was also confirmed by the students' interest in further cooperation within the framework of bachelor's theses focused on the field of data analysis. And last but not least, companies that cooperate in the education process also benefit, because, in the framework of cooperation, they can point out their requirements for future employees.

ACKNOWLEDGMENT

This project "Holistic Education and Training of University Teachers in Field of Economics" (Project no. 053TUKE-4/2021) is funded by the grant scheme KEGA of the Ministry of Education, Science, Research and Sports of the Slovak Republic.

REFERENCES

[1] N. Arranz, M. F. Arroyabe, V. Sena, C. F. A. Arranz, and J. C. Fernandez de Arroyabe, "University-enterprise cooperation for the employability of higher education graduates: a social capital approach," Studies in Higher Education, vol. 47, no. 5, pp. 990-999, 2022. https://doi.org/10.1080/03075079.2022.2055323.

[2] S. J. Teixeira, P. M. Veiga, and C. A. Fernandes, "The knowledge transfer and cooperation between universities and enterprises," Knowledge Management Research & Practice, vol. 17, no. 4, pp. 449-460, 2019. https://doi.org/10.1080/14778238.2018.1561166.

[3] QBSS, "How the Education Industry Can 'Level Up' with Technology?" 2022. Available at: https://www.quatrrobss.com/articles-blogs/how-the-education-industry-can-level-up-with-technology/.

[4] A. A. Pyrkin, A. A. Bobtsov, A. A. Vedyakov, S. V. Shavetov, Y. S. Andreev, and O. I. Borisov, "Advanced Technologies in High Education in Cooperation with High-Tech Companies," IFAC-PapersOnLine, Vol. 52, Iss. 9, 2019, pp. 312-317. https://doi.org/10.1016/j.ifacol.2019.08.226.

[5] Holon IQ, "Education Technology in 10 Charts," 2022. Available at: https://www.holoniq.com/edtech-in-10-charts.

[6] K. Mohd-Yusof, S. S. Hassan, M., Jamaludin, and N. Harun, "Motivation and Engagement of Learning in the Cooperative Problem-based Learning (CPBL) Framework," ASEE Annual Conference Exposition Proceedings, 2011. https://doi.org/10.18260/1-2--18387.

[7] M. Zenios, "Educational theory in technology enhanced learning revisited: A model for simulation-based learning in higher education," Studies in Technology Enhanced Learning, vol. 1, iss.1, 2020. https://doi.org/10.21428/8c225f6e.1cf4dde8.

[8] C. I. Fernandes, and J. J. M. Ferreira, "Knowledge spillovers: cooperation between universities and KIBS," R&D Management, vol. 43, iss. 5, pp. 461-472, 2013. https://doi.org/10.1111/radm.12024.

[9] W. Hung, D. H. Jonassen, and R. Liu, "Problem-based learning," In Handbook of research on educational communications and technology 3rd ed., Mahwah, NJ: Erlbaum, pp. 485-506, 2008.

[10] J. P. Hertel, B.J. Millis, "Using Simulations to Promote Learning in Higher Education: An Introduction," Stylus Publishing, LLC., 182 p., 2002.

[11] UNSW, "Simulations," 2022. Available at: https://www.teaching.unsw.edu.au/simulations.

[12] N. Campos, M. Nogal, C. Caliz, and A. A. Juan, "Simulation-based education involving online and on-campus models in different European universities," Int J Educ Technol High Educ vol. 17, no. 8, 2020. https://doi.org/10.1186/s41239-020-0181-y

[13] A. Adesina, "Data is the new oil," 2018. Available at: https://medium.com/@adeolaadesina/data-is-the-new-oil-2947ed8804f6

[14] D.H. Taylor, "L&D Global Sentiment Survey 2023: The Return to Data," 2023. Available at: https://donaldhtaylor.co.uk/insight/gss2023-results/

Potential Implementation of Augmented Reality Technology in Education

M. Brizar, D. Kažović
University of Applied Sciences Velika Gorica, Croatia
mato.brizar@vvg.hr, dkazovic@vvg.hr

***Abstract* - In today's age of development of science and technology, which is developing extremely fast, the way in which technology is implemented is also changing. One of the important technologies that is used in many areas of business, as well as education, is certainly AR technology, i.e. augmented reality technology. Numerous studies have shown great benefits of applying AR technology in all segments of education, from preschool education, elementary education, secondary education, higher education, as well as within the field of lifelong education. By using AR technology through augmented reality displays and 3D formats, it is possible to facilitate learning processes and experience of the material learned in all areas and segments of education. Scientific research will be carried out for the purposes of drafting this paper. The objectives of the research are to research the attitude towards augmented reality technology in learning and knowledge acquisition and to research the attitude towards the potential for improving educational processes using AR technology. The paper will present the results of research that will show the influence of previous experience with AR technology on the adoption of more positive attitudes towards the implementation of AR technology in learning and acquiring knowledge and will present the potential for the implementation of AR technology within the Croatian educational curriculum in the coming 5 years.**

Keywords: augmented reality, AR technology, education

I. INTRODUCTION

Augmented reality (AR) is a technology that integrates digital information, such as text, images, videos and 3D objects, into the real world [1]. The term "augmented reality" was first proposed by Boeing employees in 1990 [1]. AR technology also supports seamless interaction between real and virtual environments and allows tangible interfaces to be used for object manipulation [2]. Augmented reality integrates digital information with the user's real-time environment through various electronic devices, such as head-up displays, holographic displays and hand-held devices such as smartphones and smart glasses [3]. In augmented reality, a computer uses sensors and algorithms to determine the position and orientation of the camera, and AR technology then displays 3D graphics of the object as it would appear in the camera's view, superimposing the computer-generated images over the user's view of the real world [4].

The application of augmented reality in education is significant in the field of robotics, animation and programming basics since it enables the visualization of a complex 3D object, device scheme, technological process, etc. [5]. Although the real world is three-dimensional, we prefer to use two-dimensional media in education, which is very convenient, flexible, portable and cheap, however, such media is static and does not offer dynamic content, while the computer-generated three-dimensional virtual environment offered by augmented reality can be effectively used in education and make it more dynamic [6,7]. There are significant challenges in the potential implementation of AR technology in education for educators due to facing the increasing demands of students, institutions, industry and the expectations of society in general [8,9]. Therefore, educators also need to understand how to create educational content using augmented reality, and how to structure AR content at multiple points within the curriculum [8,10].

The usefulness and advantages of using AR technology in education have been recognized in many research studies [11-13]. AR technology as a tool for improving the educational process is applicable in all fields of education, from preschool education to higher education [13]. The motivation for introducing this technology varies from one application to another, but mostly it provides the user with additional information that the user cannot obtain using only his/her senses [12]. There are two principal ways of introducing AR technology in education - through smart glasses or AR headsets and through portable devices, such as mobile phones, tablets or laptops [10]. Augmented reality provides more enthusiasm for active learning, better understanding of object's appearance and features, as well as improving the ability to visualize an object in space [14-17].

The strategic framework for the digital maturity of schools and the education system in the Republic of Croatia (2030) contains a strategic area of support for learning and teaching using information and communication technology, in which the emphasis is on encouraging the development of digital educational content and teaching scenarios and supporting the culture of student creation using digital technology [18]. ICT technology is important in education and harmonizing educational materials and learning outcomes in accordance with constant changes in society [19]. The largest share of the use of modern technologies relates to the use of presentations in classes [20].

II. RESEARCH METHODOLOGY

The usefulness of applying AR technology in education has been recognized in scientific research. Considering insufficient research and insufficient implementation of AR technology in Croatia, any research effort related to the field of AR technology is quite significant. The research that was carried out for the purposes of writing this paper is focused on researching the extent to which AR technology can affect learning and knowledge acquisition, and what are the potentials of using AR technology to improve educational processes in Croatia.

At the beginning of the research, the research problem was defined, and scientific literature was extensively analysed. The research objectives were defined through literature analysis. A survey questionnaire was drafted for research implementation. Google Forms survey tool was used for drafting the questionnaire. The survey questionnaire could be completed between 16 January and 25 January 2023. The respondents voluntarily completed the questionnaire. The participation was anonymous.

Research participants: The participants were undergraduate students in Computer Systems Maintenance and graduate students in Information Systems at the University of Applied Sciences Velika Gorica. Participation in the research was voluntary. A total of 74 respondents participated, 87.8% of which were men and 12.2% were women. The age structure of the respondents was as follows: 74.3% between 18 and 25 years of age, 17.6% between 26 and 35 years of age, 5.4% between 36 and 45 years of age, 1.4% between 46 and 55 years of age and 1.4% over the age of 55. The type of completed secondary education by respondents is the following: 73% finished technical school, 10.8% finished grammar school and 16.2% finished other secondary schools.

Defining the problem: The research problem is to determine how much the application of AR technology in education can affect learning and knowledge acquisition and how much the implementation of AR technology could help improve educational processes. The scientific research problem is the potential of applying AR technology in education in Croatia.

Research goals: The research set two basic research goals: 1. Research the attitude on the application of AR technology in learning and knowledge acquisition and 2. Research the attitude on the potential for improving educational processes using AR technology.

Hypothesis 1 (H1). Respondents who have previous experience with augmented reality technology have more positive attitudes towards the implementation of AR technology in learning and knowledge acquisition.

Hypothesis (H2). Respondents who have a more positive attitude towards the implementation of AR technology in learning and knowledge acquisition will state a higher potential for the implementation of AR technology into education curricula in the coming 5 years.

Measuring instruments: The research used survey method. The Likert scale was used to determine the attitudes of respondents in the survey questionnaire. The statements used in the research were obtained based on the study of the literature on the implementation of AR technology in education.

III. RESEARCH RESULTS

Descriptive statistics showed great optimism among respondents in all areas, and it can be concluded that the respondents believe that the application of AR technology in classes can help them obtain more information on a subject or a field of study, that the application of AR technology can help to obtain information on a subject faster, that the implementation of AR technology can increase interest in subjects or a field of study, that the application of AR technology can increase knowledge acquisition efficiency, that the application of AR technology can improve the level of understanding. The majority of respondents believe that AR technology is easy to use and that applying AR technology in the educational process can be fun. Respondents do not feel nervous about using AR technology in class and are not insecure in their ability to use AR technology.

Principal component analysis with Oblimin rotation was performed to analyse the scale score. The results show that the two-component solution is optimal. The first component refers to the evaluation of usefulness, while the second component refers to the sense of concern. The analysis shows that variables: “The use of augmented reality (AR technology) can help me acquire more information on a subject or a field of study”; “The use of augmented reality (AR technology) can help me get information on a subject or a field of study faster”; “The use of augmented reality (AR technology) could increase my interest in subjects or fields of study”; “The use of augmented reality (AR technology) could increase knowledge acquisition efficiency”; “The use of augmented reality (AR technology) could improve my level of understanding” and “The use of augmented reality (AR technology) in education can be fun” fall under the first component. While the variables “I feel nervous (concerned) about the use of augmented reality (AR technology) within education curricula” and “I feel insecure about my ability to use augmented reality (AR technology)” fall under the second component. The analysis shows that the variable “I think augmented reality (AR technology) is easy to use” does not belong to any component.

Hypothesis 1 examines whether the respondents who have previous experience with augmented reality technology have more positive attitudes towards the application of AR technology in learning and knowledge acquisition. Pearson's correlation coefficient was used to analyse this hypothesis. At the level of scale results, there is no correlation between previous experience and the attitude towards the use of AR technology in education. It can be assumed that the participants show high trust in AR technologies regardless of previous experience, considering that, the assessment of usefulness is high and concern is low on average. Table I presents the details of the analysis. For a more detailed analysis, an analysis of

the correlations of individual variables with previous experience was performed. The correlation exists for the variable: “I think augmented reality (AR technology) is easy to use.” The results show that only the variable “I think augmented reality (AR technology) is easy to use” is positively related to previous experience, i.e. that the respondents who have more experience with the use of AR technology, simultaneously consider the technology easier to use. Table II presents the analysis details.

Hypothesis 2 examines whether the respondents who have a more positive attitude towards the application of AR technology in learning and knowledge acquisition will estimate a greater potential for the implementation of AR technology within the education curriculum in the coming 5 years. The results show that the respondents who gave a higher grade to the usefulness of AR technology also gave a higher grade to the potential for implementing AR technology in education in the coming five years. Table III presents the analysis details. For a more detailed analysis, an analysis of the correlations of individual variables with previous experience was performed. Correlation exists for variables: “The use of augmented reality (AR technology) can help me acquire more information on a subject or a field of study”; “The use of augmented reality (AR technology) can help me get information on a subject or a field of study faster”; “The use of augmented reality (AR technology) could increase my interest in subjects or fields of study” and “I feel insecure about my ability to use augmented reality (AR technology). The results show that the respondents who gave a higher grade to the potential for implementing AR technology in education also believe that AR technology can help to obtain more information on a subject, that information can be obtained faster, and that AR technology can increase interest in a subject. At the same time, an unexpectedly positive correlation between the potential for implementation and the sense of uncertainty in one's own ability to use AR technology was discovered. It can be assumed that considering that most of the respondents are students up to 25 years of age and lack experience, that the previous reflects their uncertainty in the knowledge of creating content using AR technology and is expected of them. Table IV presents the analysis details.

TABLE I. HYPOTHESIS 1 ANALYSIS

No.	Variable	1	2
1	Do you have previous experience with Augmented Reality (AR) technology, either through personal or professional experience?	1	.032
2	The usefulness of AR technology in education	.032	1
3	Concerns regarding AR technology	- .071	- .066

TABLE II. ANALYSIS OF CORRELATIONS BETWEEN INDIVIDUAL VARIABLES WITH PREVIOUS EXPERIENCE

Variable	Correlation
The use of augmented reality (AR technology) can help me get more information on a subject or a field of study.	.122
The use of augmented reality (AR technology) can help me get information on a subject or a field of study faster.	- .113
The use of augmented reality (AR technology) could increase my interest in subjects or fields of study.	.130
The use of augmented reality (AR technology) could increase knowledge acquisition efficiency.	.088
The use of augmented reality (AR technology) could improve my level of understanding.	- .019
I think augmented reality (AR technology) is easy to use.	.252*
The use of augmented reality (AR technology) in education can be fun.	- .070
I feel nervous (concerned) about the use of augmented reality (AR technology) within education curricula.	- .070
I feel insecure about my ability to use augmented reality (AR technology).	- .055

* $p < 0.05$

TABLE III. HYPOTHESIS 2 ANALYSIS

No.	Variable	1	2
1	How realistic do you think is it that the educational system in the Republic of Croatia will implement AR technologies within education curriculum in the coming 5 years?	1	.256
2	Usefulness of AR technology in education	.256*	1
3	Concerns about AR technology	.214	- .066

* $p < 0.05$

TABLE IV. ANALYSIS OF CORRELATIONS BETWEEN INDIVIDUAL VARIABLES WITH A MORE POSITIVE ATTITUDE

Variable	Correlation
The use of augmented reality (AR technology) can help me get more information on a subject or a field of study.	.262*
The use of augmented reality (AR technology) can help me get information on a subject or a field of study faster.	.240*
The use of augmented reality (AR technology) could increase my interest in subjects or fields of study.	.261*
The use of augmented reality (AR technology) could increase knowledge acquisition efficiency.	.184
The use of augmented reality (AR technology) could improve my level of understanding.	.195
I think augmented reality (AR technology) is easy to use.	- .012
The use of augmented reality (AR technology) in education can be fun.	.035
I feel nervous (concerned) about the use of augmented reality (AR technology) within education curricula.	.127
I feel insecure about my ability to use augmented reality (AR technology).	.262*

* $p < 0.05$

IV. CONCLUSION

Through strategic documents, the Ministry of Science and Education emphasizes raising the quality of education and the achieved learning outcomes, mainly by developing new approaches and contents [18]. The entire society and everyday life are changing every day mainly due to the accelerated development of technology and the introduction of the use of technology [21]. Given the changes in society and life in general, changes in education are also inevitable, which also change the learning methods [22,23].

Digitization of education in the Republic of Croatia started a long time ago, today we increasingly use smart boards and digital content in classes, students increasingly use computers and tablets, the use of e-learning tools has become a common practice at faculties, especially the Learning Management System for e-learning. All digitization of teaching thus far has been highly necessary, but it certainly cannot be final and must continue to develop rapidly. The above also applies to the technical equipment of schools and faculties, since the lack of equipment can often be a problem in the full implementation of IT technology. Today's modern technologies offer new possibilities that can shape new modern teaching aids and materials, among which AR technology is certainly very interesting. AR technology has proven its usefulness a long time ago and is today used successfully in all areas of business, from medicine, marketing, biology, mathematics, geography, engineering, etc. AR technology is still in development and offers new possibilities of use in different areas of society each day. AR technology provides significant help in teaching and better mastering the material and ultimately achieving better learning outcomes. For example, students can use AR technology to see information on the content on the screen in front of them, they can experience remote locations and different practical situations.

Through research, we aimed at finding out the students' attitudes towards AR technology and the implementation of AR technology in education. The research results certainly showed the great optimism of the respondents, and it can be concluded that the introduction of AR technology into everyday classes would be extremely useful for learning, knowledge acquisition and the final results achieved by the students. In addition to the usefulness of AR technology, a concern regarding the use of AR technology was recognized as a second factor, i.e. the respondents are confident in their knowledge and application of AR technology and do not feel worried about using AR technology.

The optimistic view on the application of AR technology can also be connected to the respondents' structure. Given that more than 70% of the respondents come from a technical profession, the assumption is that people with technical knowledge have a greater preference for modern IT technologies and are confident in the daily use of the AR technology.

It is interesting to point out the variable that shows that AR technology can be fun, which gives the conclusion that in today's era, in addition to teaching, an educator should also work on motivating students for classes, class content, and all of the aforementioned can be achieved by providing a certain dose of entertainment during learning and imparting knowledge.

Although the usefulness of AR technology has been confirmed, what could pose a problem in its application in the development of education is certainly the lack of optimism regarding the application of AR technology in the education system in the Republic of Croatia in the coming 5 years. The results show that the application of AR technology in the coming 5 years is unlikely. The aforementioned is very worrying considering that changes happen every day and that monitoring new modern technologies and exploiting their possibilities in creating quality educational content and experiences should represent one of our strategic goals in the near future. Certain participants pointed out the problems of the slowness of changes in the education system and the problem of obsolescence of many teaching contents and methods. Therefore, it is certainly necessary to monitor the development of education at all levels, from the government, the Ministry, the economy, educational institutions to educators and students. It is necessary to find ways for efficient implementation of AR technology as an educational technology that should be used by both educators and students.

In addition to formal education, the application of AR technology should also find a place in informal education. The research participants pointed out that they would rather choose a course featuring AR technology compared to a course that does not use AR technology even if the content of the course is identical.

This paper is an introduction to the field of AR technology research in education. Therefore, it is planned to continue the research in this field, and the next step is to investigate the possibilities and attitudes on the application of AR technology by educators, education holders, their attitudes, their capabilities, actions that would motivate them to apply AR technology in everyday teaching and determine whether classrooms are equipped with amenities necessary for the introduction of AR technology.

REFERENCES

[1] Y. Tan, W. Xu, S. Li, K. Chen, Augmented and Virtual Reality (AR/VR) for Education and Training in the AEC Industry: A Systematic Review of Research and Applications, Buildings, 2022. https:// doi.org/10.3390/buildings12101529

[2] N. F. Saidin, N. D. A. Halim, N. Yahaya, A Review of Research on Augmented Reality in Education: Advantages and Applications, International Education Studies, Vol. 8, No. 13, 2015. http://dx.doi.org/10.5539/ies.v8n13p1

[3] J. Belda-Medina, V. Marrahi-Gomez, The Impact of Augmented Reality (AR) on Vocabulary Acquisition and Student Motivation, Electronics, 2023. https://doi.org/10.3390/electronics 12030749

[4] J. B. Cowin, Digital Worlds and Transformative Learning: Google Expeditions, Google Arts and Culture, and the Merge Cube, International Research and Review, Journal of Phi Beta Delta Honor Society for International Scholars, Vol. 10, Number 1, Fall, 2020.

[5] D. Kozlitin, L. Kozak, T. Krystopchuk, D. Kochmar, The Application of Augmented Reality in Education and Development

of Students Cognitive Activity, Vol I: Main Conference, PhD Symposium, Posters and Demonstrations, Kherson, Ukraine, 2021.
[6] K. Mehmet, Y. Ozarslan, Augmented Reality in Education: Current Technologies and the Potential for Education, Article in Procedia - Social and Behavioral Sciences. December, 2012.
[7] D. Holley, M. Hobbs, Augmented Reality For Education, Bournemouth University, Bournemouth, Dorset, 2017.
[8] K. Dutta, Augmented Reality for E-Learning, RWTH Aachen University, Conference Paper, February, 2015.
[9] J.A. Sicaru, C. G. Ciocianu, C.A. Boiangiu, A survey on augmented reality, Polytechnic University of Bucharest, Article, December, 2017.
[10] S. Ribadiya, A Survey: Virtual, Augmented and Mixed Reality in Education. International Journal of Engineering Research & Technology (IJERT), Vol. 9, Issue 05, May-2020. Published by: http://www.ijert.org
[11] Augmented and virtual reality survey report, Industry insights into the future od immersive technology, Vol. 3, March, 2019. https://www.perkinscoie.com/images/content/2/1/v4/218679/2019-VR-AR-Survey-Digital-v1.pdf
[12] J. C. G. Vargas , R. Fabregat , A. Carrillo-Ramos, T. Jové, Survey: Using Augmented Reality to Improve Learning Motivation in Cultural Heritage Studies, Appl. Sci. 2020. doi:10.3390/app10030897
[13] M. A. Mundy, J. Hernandez, M. Green, Perceptions of the Effects of Augmented Reality in the Classroom, Journal of Instructional Pedagogies, Vol. 22, https://www.semanticscholar.org/paper/Perceptions-of-the-Effects-of-Augmented-Reality-in-Mundy-Hern%C3%A1ndez/bdf9ba0417f2a0ae962f86d3d41501020a52d1f8
[14] O. Huerta, A. Kus, E. Unver, R. Arslan, M. Dawood, M. Kofoğlu, V. Ivanov, A Design-based Approach to Enhancing Technical Drawing Skills in Design and Engineering Education using VR and AR Tools, Conference Paper, February, 2019.
[15] O. Çolak, L. Yünlü, A rewiev on augmented reality and virtual reality in enginnering education, Article, February, 2018. https://www.researchgate.net/publication/327793245
[16] L. Ding, V. Chau, Construction of Interactive Higher Education Model Based on "VR+Internet", Wireless Communications and Mobile Computing, 2022. https://doi.org/10.1155/2022/6444864
[17] H. Batool, Augmented reality applications as a digital learning innovation in response to the pandemic, doi: 10.3389/feduc.2022.937074
[18] Ministry of science and education, Strateški okvir za digitalno sazrijevanje škola i školskog sustava u Republici Hrvatskoj (2030), Zagreb, 2020.
[19] I. Smiljcic, I. Livaja, J. Acalin, ICT u obrazovanju, Zbornik Veleučilišta u Šibeniku, 3-4, Šibenik, 2017, pp 157-170
[20] T. Povic, K. Veleglavac, M. Carapina, T. Jagust, I. Boticki, Primjena informacijsko-komunikacijske tehnologije u osnovnim i srednjim školama u Republici Hrvatskoj // Proceedings of the Carnet User Conference 2015 (CUC2015), Zagreb: Carnet, 2015. 19, 26.
[21] KA. Abutayeh, OM Kraishan, EQ. Kraishan, The use of virtual and augmented reality in science and math education in Arab countries: A survey of previous research studies, doi: 10.3389/feduc.2022.979291
[22] I. Cicek, A. Bernik, I. Tomicic, Student Thoughts on Virtual Reality in Higher Education—A Survey Questionnaire. Information 2021. https:// doi.org/10.3390/info12040151
[23] M. E. C. Santos, A. Chen, T. Taketomi, G. Yamamoto, J. Miyazaki, H. Kato, Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation, IEEE transactions on learning technologies, Vol. 0, 2000.

Challenges of Adoption of Cloud Computing Solutions in Higher Education: Case Study Republic of Kosovo

L. Shala Riza, J. Ajdari and M. Hamiti
South East European University, Faculty of Contemporary Sciences and Technologies, Tetovo, North Macedonia
ls31402@seeu.edu.mk
j.ajdari@seeu.edu.mk
m.hamiti@seeu.edu.mk

***Abstract* - Cloud Computing has begun to be widely used in higher education institutions in the last few years. The transition from on-premises systems to cloud solutions in education is a smart initiative with its own set of rewards and obstacles and affects various stakeholders willing to accept this change. The research tries to highlight the challenges that could affect the Cloud Computing (CC) solution in higher education institutions in the Republic of Kosovo (RKS) as a developing country, regarding security, infrastructure, and performance. Potential challenges affecting the implementation of CC in education will be assessed by different stakeholders, such as students, academic and IT staff of public and private universities, who are the target participants of this research. Also, this paper tries to highlight the level of user awareness of CC technology, the advantages of this platform, as well as the willingness of shareholders to apply it in educational institutions. Expected results foresee the identification of challenges and benefits that can be encountered in the case of CC implementation in HEIs, with the case study of the RKS. The findings of the research can help institutions to assess the challenges during the integration of a CC solution in their HEIs.**

Keywords - cloud computing; challenges; higher education; e-Learning; Republic of Kosovo

I. INTRODUCTION

The Republic of Kosovo, as a country emerging from the war more than 20 years ago, is in its efforts to consolidate the education system in general, aiming to adopt effective educational techniques from examples of developed countries in the region and beyond. With all the tendency to apply numerous reforms, once again the education system is encountering various obstacles starting from the lack of infrastructure to the lack of sufficient budget for investments in this sector. On the other hand, the ICT sector has also experienced rapid development in the Republic of Kosovo since 1999. Various companies offer qualitative IT services, while public and private educational institutions in the field of IT, through certified teaching curricula, offer education for thousands of Kosovar youths, while the demand for various trainings is still growing [1].

Cloud computing plays an important role in boosting the quality of education by providing several benefits such as low-cost infrastructure, collaboration, scalability, and simplicity of use. [2]. In recent decades, there has been an increase in interest in e-learning, particularly among students enrolled in higher education. As stated in [3], "It is very important for governments to invest in high-quality universities and provide them the access to the most effective learning tools such as high standard IT infrastructure which can be used for educational software or easy data sharing within the classroom." Actually, today's students are continually connected to the internet through various programs such as Gmail, Google Classroom, Facebook, and Twitter, therefore they are a part of the cloud infrastructure. As a result, HEIs administrators are keen to embrace CC techniques, driving the trend of growing CC adoption in the higher education sector [4].

There are numerous papers and publications regarding cloud computing, but it is a need for continuous research work in this field in the sense of identifying and evaluating the prominent factors adopting CC implementation in HEIs. The purpose of this research work is to highlight numerous barriers or challenges that may influence CC implementation in the public and private universities of the Republic of Kosovo.

II. RELATED WORKS

A. The Concept of Cloud Computing

Though there is not a certain definition that has been accepted universally for cloud computing technology, in the literature we can find several interpretations and approaches that tend to elaborate on this concept. As stated in [5], "Cloud computing is the model for managing the effective access to the shared number of configurable resources like servers, network, application, and storage. It can minimize the management effort and the interaction of service providers. Cloud computing is used to minimize the IT infrastructure costs and maintenance". There are several service delivery and deployment models of CCT. The above-mentioned resources can be shared over the internet effortlessly. CCT has three types of services, infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). Through SaaS, the end users need only an interface to access needed information such as Gmail or Dropbox while the other architecture resources are hidden from the consumer. Through IaaS, users can deploy and execute software (such as Amazon Web Services or Microsoft Azure)

through the provision of computing resources. Lastly, PaaS provides a platform for end users to design, modify, and deploy applications utilizing libraries or languages established earlier by cloud providers.

There are four types of cloud deployment models: private, public, hybrid, and community. The private model is often utilized only within an organization, institution, or firm, and it is used when a higher level of safety is sought and critical for the organization. The most well- known cloud is the public cloud, which is administered by service providers such as Amazon, Google Drive, and Microsoft One Drive. The hybrid cloud model is a mixture of two or more patterns that is often used to maximize the benefits of each model combined. The community models are mostly employed by organizations whose infrastructure is hosted by a third party or another organization that is a member of the community [6].

B. Cloud Computing integration in Higher Educational Institutions

Cloud computing technology is already being used at higher education institutions in various industrialized countries throughout the world, including the United States, Japan, and Europe. The cloud-based program Google Educator, as well as the Google suite for education - GAFE, are probably one of the most essential applications that Cloud computing provides. Students and instructors may improve cooperation and productivity by having access to their data anywhere, at any time, with these tools. On the other hand, there are also many cloud-based platforms developed for teachers such as Microsoft Office 365, iGoogle, Book Search, News, Google Note-book, etc. [7].

According to the proposed model in Fig. 1, the university's system, and application within the hybrid cloud, has been divided into two parts: Private Cloud which is more secure and has on-premises servers and includes SIS, PMS, and UMS. The second part is Public Cloud where are hosted the less sensitive data and includes LMS, Video on Demand, VCR, and Mailing Systems. This model takes advantage of the hybrid cloud to achieve speed, security, and availability and it also reduces the burden of IT sectors by moving apps to the public cloud. As seen, it's a flexible model since it allows educational systems to balance internal/external loads based on the organizational needs or technical requirements [8] Despite the fact that developing nations have inadequate infrastructure, there is an increasing need to compare current e-learning approaches to conventional

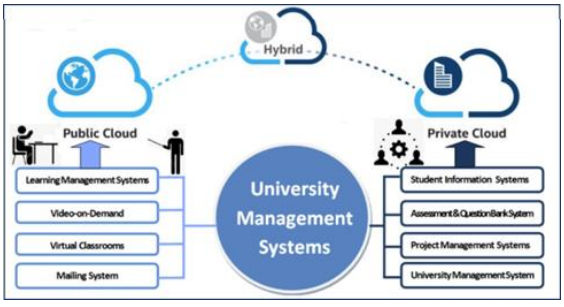


Figure 1. Cloud Computing proposed model [8]

ones by utilizing cloud services. This technical method may host Learning Management Systems and other university resources, although the hybrid model, in which the university's systems are spread between the private cloud on-premises and the public cloud, is viewed as the most ideal model for CC implementation [9].

According to [10], organizational capabilities are the key factors for adopting ICT in developing countries. This survey proposed a model to assess the real likelihood of CC or last technology adoption which allows organizations to make appropriate decisions. Service providers can implement this facility as a web-based DSS (Decision Support System) tool in order to measure customers' readiness, thus organizations can avoid unsuccessful adoption.

C. Key challenges and benefits when adopting Cloud Computing Technology in Higher Educational Institutions experiences from other countries

In the publication of [11], is found that factors such as reliability, relative advantage, security, compatibility, technology readiness, regulatory, top management support, policy, and competitive pressure have positive impacts on cloud computing adoption while tribalism culture has a negative impact. The paper of [12], highlights the key benefits and barriers to cloud computing adoption. According to this study, the main benefits of CCT adoption include greater operational efficiency, cost savings, simplicity of application deployment, flexibility, personnel redeployment, and sustainability, whereas key hurdles include security, availability, performance, and trust difficulties.

Further, the study of [13], identified two important fac- tors that were decision-making in adopting cloud solutions in the Philippines, such as internet access speed and lack of knowledge of cloud technologies. In order to overcome the aforementioned barriers, the study recommends that universities host training courses or workshops for aca- demic or IT staff. Similarly as in the above papers, in [14], authors recommend a model for adopting CC from HEIs in developing countries which consisted of five steps: Knowledge (increased awareness of CC technology), Per- suasion (find the key benefits of the solution aimed), Decision (adopt or reject the technology), Implementation (is proceeded in case of acceptance decision on the previous step) and Confirmation (continue using the solution for long term period).

In their research work, [14] proposed DOI theory to help organizations in accepting or rejecting innovations. They recommend a model for adopting CC from HEIs in developing countries which consisted of five steps: Knowledge (increased awareness of benefits and limitation of CC technology), Persuasion (find the key benefits of the solution aimed), Decision (this step may lead to adoption or rejection of the technology), Implementation (is proceeded in case of acceptance decision on the previous

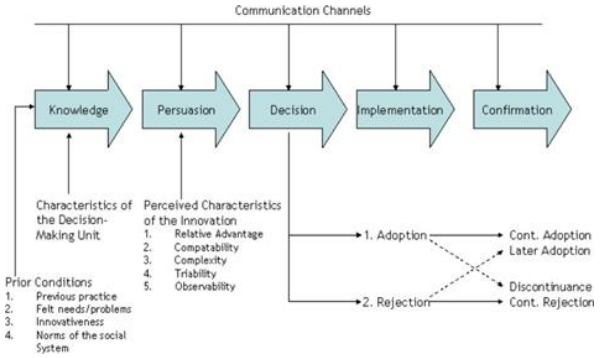


Figure 2. Innovation–decisions process model [14]

step) and Confirmation (continue using the solution for long term period).

Another study conducted in Jordan, scrutinized the main factors affecting CC adoption in HEIs of this country. Based on the findings, the following key elements were identified: cost savings, disaster recovery, scalability, compatibility, knowledge exchange, ICT utilization level, and HEIs size. The study also provided an implementation structure with four stages of CC adoption: CCT awareness, persuasion phase, decision phase, and implementation phase [15].

III. Research Methodology

The purpose of this study was to identify the main challenges that affect the adoption of cloud technology in higher education institutions in the Republic of Kosovo. The data were collected from secondary sources from different publications on the internet, later the variables were identified based on which the questionnaire was designed for the purpose of collecting primary data. Three phases were engaged in the questionnaire adaption. (1) First, the questionnaire instrument was translated from Albanian (source language) to English (target language), (2) the data acquired from respondents was administered and structured, and (3) the data was processed and analyzed graphically using Excel and the SPSS software. [16] The Google Forms questionnaire is built on the principle of confidentiality and includes both multiple-choice questions and questions with a 5-point Likert scale. To select the sample, we used convenience sampling methods in order to target the desired subjects. The targeted respondents are students or graduates, IT staff, and academic staff of public and private universities of the Republic of Kosovo. The total number of respondents that participated in the survey is 57, of which 21 are students/graduates, 13 are IT staff, and 23 belong to the academic community who actually work in either public or private universities. Because this study is descriptive in nature, we examined the data acquired from the research work using the quantitative approach and the Microsoft Excel application program.

A. Research questions

Based on the data collected systematically from the literature review, we could identify the independent variables such as security, privacy, lack of proper infrastructure, service provider support, reliability, and skills of IT personnel, management, performance, the price for accessing resources, internet speed, lack of electricity and institutional support. On the other hand, cloud computing adoption in HEIs in the RKS is considered a dependent variable. Thus, the research questions are:

R1: What are the most significant challenges that would affect cloud computing adoption in HEIs of RKS?

R2: What are considered to be the key benefits when adopting Cloud technology in higher education institutions in RKS?

B. Data analysis

1) *Questionnaire data analysis:* For data collection, we used a Google Forms questionnaire distributed online. The questionnaire targeted academic staff, IT staff as well as students or graduates from public and private universities of the Republic of Kosovo. This online questionnaire contains nine closed questions. In total, 57 respondents answered where the gender distribution of the respondents is almost equal between the two genders (51% female and 49% male). From Fig. 3, we can see that from the total number of participants, 23 are academic staff, 21 students or graduates and 13 are IT staff. From both public and private universities of RKS. From this total number, 40.4% have studied or are currently part of a private university, while 59.6% of respondents have chosen a public university as their institution. In embracing new technology, we assumed that the level of knowledge we have about that technology and awareness about the benefits it offers also play a significant role.

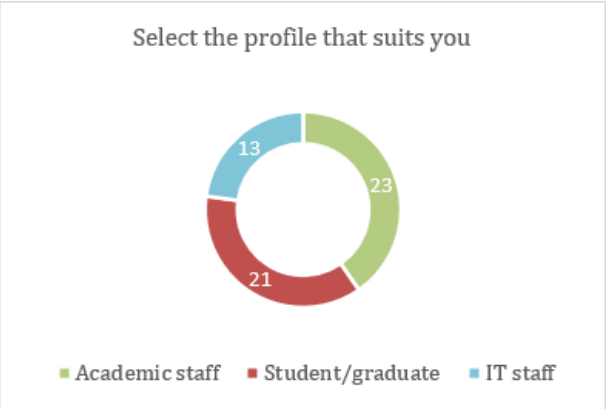


Figure 3. Respondents profiles

When asked about their knowledge of the Cloud Computing solution, 28% of the sample responded they are very well informed, while 17.5% confirmed they know only a little aware about this platform. The results of this question are visually illustrated in Fig. 4. As is shown in the graph presented in Fig. 5, in the question on the benefits of CC solution when used in HEIs, 43 respondents of which 22 were from public universities and 21 from private ones responded that they are aware of the benefits. Further, 9 of the participants answered with "Somehow", 4 answered they are not aware of CC benefits and only one respondent chose the "I have no answer" option.

Further, when asked if the lack of knowledge about Cloud Computing has an impact on the adoption of this solution in higher education institutions, most of the respondents agreed upon this fact, of which 17 were academic staff, 12 were IT staff and 18 were students or graduates (refer to Fig. 6).

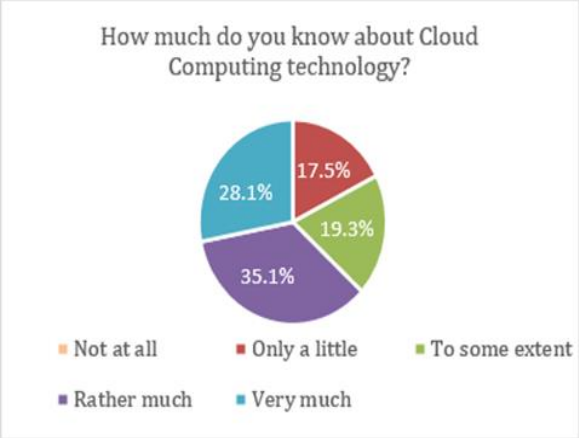


Figure 4. Respondents awareness over Cloud Computing

A smaller number of participants stated that the lack of knowledge or awareness would impact "Somehow" or has no impact on the adoption of this technology. The question results are shown in Fig. 6.

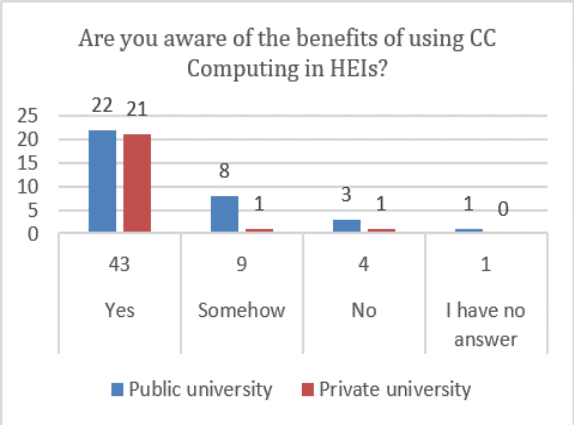


Figure 5. Respondents awareness of the benefits of using CC in HEIs

In terms of ethical standards, the respondents were informed about the goal and rationality of the questionnaire, as well as the destination of the gathered findings, throughout the study process, respectively during data collection via the online form. Respondent data is kept anonymous, and the obtained information is utilized for academic reasons.

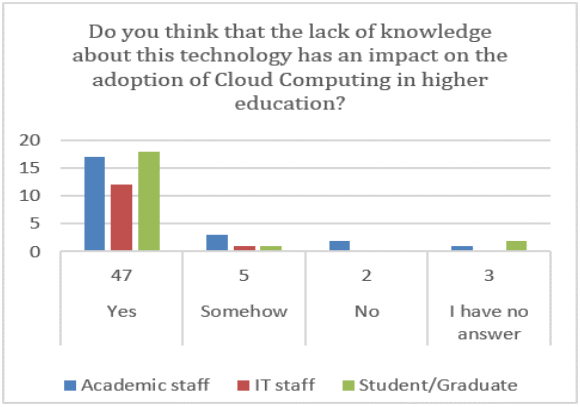


Figure 6. Respondents' thoughts over the of lack of knowledge impact of CC on the adoption of this technology in HEIs

2) *One-Sample Test: One-Sample T-test* is used to test the relationship between independent variables (security, privacy, reliability, SP support, lack of proper infrastructure for integration, IT skills and knowledge, management, performance, price for accessing resources, internet speed, lack of electricity, institutional support) on the dependent variable (Cloud Computing adoption in HEIs).

From the output results in the Sig. (2-tailed) column in Figure 6, we can see that p-value is < .05. Also, the findings indicate that these variables are all accepted as the mean value is (>0.3) and the standard deviation is (> 0.7). This indicates that the above-mentioned independent

TABLE I. ONE-SAMPLE TEST ON KEY FACTORS AFFECTING CLOUD COMPUTING ADOPTION IN HEIS

	Test Value = 0					
	t	df	Sig.(2-tailed)	Mean Diff.	95% Conf. Interval of the Diff.	
					Lower	Upper
Security	26.63	56	.000	4.00	3.69	4.30
Privacy	25.42	56	.000	3.82	3.52	4.12
Reliability	27.91	56	.000	3.98	3.69	4.26
SP support	24.84	56	.000	3.71	3.41	4.01
Lack of infrast.	22.94	56	.000	3.68	3.36	4.00
IT skills	30.36	56	.000	4.05	3.78	4.32
Management	27.02	56	.000	3.73	3.46	4.01
Performance	21.82	56	.000	3.54	3.21	3.86
Price for resource	29.77	56	.000	3.87	3.61	4.13
Internet speed	23.43	56	.000	3.78	3.46	4.11
Institutional support	32.26	56	.000	4.03	3.78	4.28
Lack of electricity	24.03	56	.000	3.84	3.52	4.16

variables have a significant impact on CC adoption in HEIs in the Republic of Kosova. At the same time, we answered the first research question: **R1: What are the most significant challenges that would affect CC adoption in HEIs of RKS?**

Consequently, we can conclude that all above mentioned independent variables such as security, privacy, reliability, Service Provider support, lack of proper infrastructure for integration, IT skills and knowledge, management, performance, the price for accessing resources, internet speed, institutional support and lack of electricity are key factors or challenges affecting Cloud Computing adoption in public and private universities in the Republic of Kosovo.

3) *Cronbach's Alpha and KMO and Bartlett's Test:* Since the questions are Likert-scale type, it is worth measuring the reliability and consistency of the answers, this way we make sure the set of the questions raised measure the same objective. We checked reliability statistics against twelve independent variables and in our case Cronbach's Alpha value is .891.

Usually, this value should be between 0 and 1 so the higher the value of Cronbach's Alpha, the more confidence we have that the questions are internally

consistent and correlate to each other. The value we gained (.891) is close to one so we can conclude that the reliability level is high.

TABLE II. CRONBACH'S ALPHA TABLE

Cronbach's Alpha	N of Items
.891	12

In order to measure the sampling adequacy and to examine how the factors explain each other, a KMO test is conducted. The table below presents two different tests: the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's test of Sphericity which are used to measure whether the variables are correlated and also the sampling adequacy.

The accepted value for the KMO index is between 0.5 and 1. In our case, the KMO index is .772 which indicates that factor analysis is appropriate. A KMO value more than 0.5 and a significance level for the Bartlett's test less than 0.05, as in our case, indicate that there is significant correlation in the data. The degree to which a single variable is associated with other variables is shown by variable collinearity.

TABLE III. KMO AND BARTLETT'S TEST

KMO Measure of Sampling Adequacy		.772
Bartlett's Test of Sphericity	**Approx. Chi-Square**	410.171
	df	66
	Sig.	.000

4) *Rotated Component Matrix:* To find the correlations between items and components or factors, the rotated com- ponent matrix was used. This matrix helps to determine what each of the components means. Small coefficients below .5 were suppressed in order to offer a clearer view of the group factors. The first component has a strong correlation with the lack of proper infrastructure, IT skills and knowledge, internet speed, lack of electricity and institutional support. The second component has a correlation with security, privacy, and reliability, while the third component has a correlation with IT skills and knowledge, management, performance, the price for accessing resources, and internet speed.

5) *Key benefits of adopting Cloud Computing in HEIs:* In the continuation of data collection, the respondents stated there are several benefits of adopting Cloud Computing technology in higher education in the RKS. In this way, we have also answered the second research question: **R2: What do you consider to be the key benefits when adopting Cloud technology in higher education institutions in RKS?**

Some of the key benefits are the availability of resources, better data management, modernization of the way of studying, the widest range of online resources, better cooperation, and the use of green technology.

TABLE IV. CORRELATIONS BETWEEN ITEMS AND COMPONENTS

	Component		
	1	**2**	**3**
Security		.866	
Privacy		.876	
Reliability		.847	
SP support	.729		
Lack of proper infrastructure	.550		
IT skills and knowledge			.581
Management			.556
Performance			.786
Price for accessing resources			.862
Internet speed	.523		.526
Lack of electricity	.855		
Institutional support	.813		

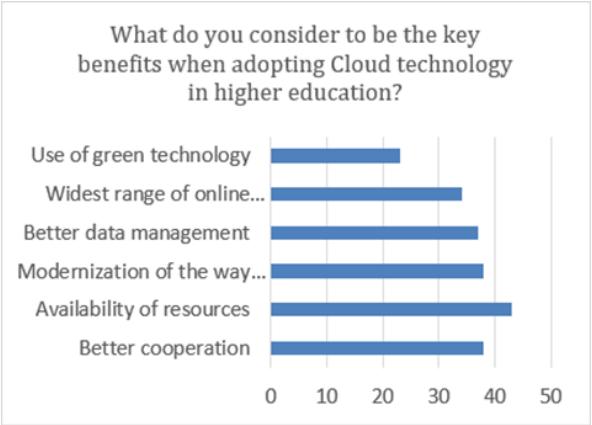


Figure 7. Key benefits of adopting CC in HEIs

Lastly, considering the advantages, benefits, and challenges of the implementation of this technology, 93% of the respondents answered that they are in favor of the integration of Cloud Computing technology in higher education in the Republic of Kosovo, while 7% stated that they have no answer.

C. Conclusions

With all of the investments in the country's technology industry, it should be emphasized that there is still room to be enlightened about Cloud Computing solution, including the benefits it delivers as well as the possible obstacles we may face throughout the implementation process. The goal of this study paper was to identify the important elements or significant challenges that impact Cloud Computing adoption in higher education institutions in the Republic of Kosovo. From the data analysis, it was concluded that all twelve independent variables such as security, privacy, reliability, Service Provider support, lack of proper infrastructure for integration, IT skills and knowledge, management, performance, the price for accessing resources, internet speed, institutional support, and lack of electricity are the key factors or challenges affecting CC adoption in public and private universities in the Republic of Kosovo. This research work aspires to guide and serve as a reference point for educational institutions and other actors in the process of analyzing and examining the circumstances for the adoption of Cloud technology in educational institutions in the Republic of Kosovo as a study case. HEIs in the country as well as the different stakeholders involved in this sector should consider the above-mentioned aspects in order to implement cloud technology in these institutions. In the future, more research work can be done in the context of increasing the targeted sample and expanding the field of study.

REFERENCES

[1] L. Aliu, "Analizë e sistemit arsimor në Kosovë," 2019.

[2] K. H. A. Al-Shqeerat, F. M. A. Al-Shrouf, M. R. Hassan, and H. Fajraoui, "Cloud computing security challenges in higher educational institutions - a survey," *International Journal of Computer Applications*, vol. 161, no. 6, pp. 22–29, Mar 2017. [Online]. Available: http://www.ijcaonline.org/archives/volume161/ number6/27154-2017913217

[3] C. Bulla, B. Hunshal, and S. Mehta, "Adoption of cloud computing in education system: A survey," *international journal of engineer- ing science and computing*, vol. 6, p. 6375, 07 2016.

[4] A. Elgelany and W. G. Alghabban, "Cloud computing: Empirical studies in higher education a literature review," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, 2017. [Online]. Available: http://dx.doi.org/10.14569/ IJACSA.2017.081017

[5] R. Parveen, "Challenges in cloud computing adoption- an empirical study of educational sectors of saudi arabia," *International Journal of Managing Information Technology*, vol. 11, pp. 1–11, 2018.

[6] Y. A. M. Qasem, S. Asadi, R. Abdullah, Y. Yah, R. Atan, M. A. Al-Sharafi, and A. A. Yassin, "A multi-analytical approach to predict the determinants of cloud computing adoption in higher education institutions," *Applied Sciences*, vol. 10, no. 14, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/14/4905

[7] A. Shahzad, A. G. Golamdin, and N. A. Ismail, "Opportunity and challenges using the cloud computing in the case of malaysian higher education institutions," *The International Journal of Management Science and Information Technology (IJMSIT)*, no. 20, pp. 1–18, 2016. [Online]. Available: http://hdl.handle.net/ 10419/178822

[8] M. Ibrahim and M. Elmasry, "Cloud computing for e-learning: A proposed model for higher education institutions in developing countries," *International Journal of Scientific Technology Research*,vol. 10, pp. 408–416, 03 2021.

[9] R. Manoharan, "A systematic review of cloud security challenges in higher education," *The Online Journal of Distance Education and e-Learning*, vol. 5, pp. 1–10, 10 2017.

[10] M. Workineh, N. M. Garcia, and D. Midekso, "Cloud computing as technological solutions for higher education institutions: Adoption readiness assessment model: Reseach in-progress," in *2017 Inter- national Conference on Computing Networking and Informatics (ICCNI)*, 2017, pp. 1–

[11] A. Hussein Alghushami, N. H. Zakaria, and Z. Mat Aji, "Factors influencing cloud computing adoption in higher education institutions of least developed countries: Evidence from republic of yemen," *Applied Sciences*, vol. 10, no. 22, 2020. [Online].
Available: https://www.mdpi.com/2076-3417/10/22/8098

[12] V. Arkorful, "Cloud computing adoption in higher education: A comparative study between public and private universities in sub saharan africa," *Library Philosophy and Practice (e-journal)*, 2019.

[13] C. Alimboyong and M. Bucjan, "Cloud computing adoption among state universities and colleges in the philippines: Issues and chal- lenges," *International Journal of Evaluation and Research in Edu- cation (IJERE)*, vol. 10, p. 1455, 12 2021.

[14] M. Odeh, A. Garcia-Perez, and K. Warwick, "Cloud computing adoption at higher education institutions in developing countries: A qualitative investigation of main enablers and barriers," *Interna- tional Journal of Information and Education Technology*, vol. 7,pp. 921–927, 01 2017.

[15] M. Odeh, "A novel framework for the adoption of cloud computing in the higher education sector in developing countries," 12 2020.

[16] E. Bahçivan and S. Kapucu, "Adaptation of conceptions of learning science questionnaire into Turkish and science teacher candidates' conceptions of learning science," European Journal of Science and Mathematics Education, vol. II, pp. 106-118, 2014.

Augmenting Student Education Using the RealityScan Application for Generating 3D Content

V.V. Kozov and B. Ivanova
University of Ruse "Angel Kanchev", Department of informatics, Ruse, Bulgaria
vkozov@uni-ruse.bg, bivanova@uni-ruse.bg

***Abstract* - Augmented reality in education is surging in popularity in schools worldwide. Through augmented reality (AR), educators are able to improve learning outcomes through increased engagement and interactivity. Augmented Reality (AR) in education features aspects that enhance learning abilities like problem-solving, collaboration, and creation to better prepare students for the future. It is also good for traditional pedagogy focused on technical knowledge and proficiencies.**

An analysis is made on the innovative technology for generating 3D models - the RealityScan application by Epic Games for iOS. Several characteristics and functionalities of the application are described and tested. Methodologies and workshops based on them are created and presented.

A process of including the aforementioned technologies in the curriculum and easing the students into using such tools in their future professional career is created and partially implemented. Several benefits of becoming familiar with such technologies are analyzed. A conclusion is made on several critical points that have to be addressed when implementing the necessary steps for improving the education system using AR.

Keywords - augmented reality; education methodology; 3D models

I. INTRODUCTION

Institutions are currently facing the challenge of motivating and encouraging students to engage in professional careers in AECO-Architecture, Engineering, Construction, and Operations sectors as they do not find jobs appropriate to their degree. One of the reasons is the students' lack of proper intuition and insights as a skilled professional. A solid choice for improving the student's level of engagement in the classroom and simultaneously providing them with the basic necessary experience needed to make them a valuable future cadre is making use of Augmented Reality (AR) applications. One of the available technologies that are proven to fulfill our criteria are Virtual Reality (VR), Augmented reality (AR) and Mixed Reality (MR). Studies on the topic have been analyzed and it has been concluded that our students will benefit greatly if they are able to engage in AR learning.

AR makes learning itself more engaging and fun. It inspires and engages students, thus making the learning process of STEM and coding easier, more enjoyable and better than it has ever been [1].

The next logical step in AR-based learning discovered from reviewing current literature is the development of applications that can enable personalized learning materials for both students and teachers. Several of the future technology challenges are user acceptance, proving the educational effect it has on learners, and further development of the frameworks used to develop these innovative applications. As the cost of hardware and software decreases, AR technology will become more affordable, thus allowing it to be widespread at all educational levels, especially in higher education. [2].

A bibliometric analysis of 1977 articles was performed, obtaining results of scientific productivity. The evidence showed a growing interest in studying the sustainability of AR in higher education. AR has been applied mainly in HE settings and compulsory levels of computer science education, as a resource to motivate students and to promote student learning.[3].

The usage of AR in training and teaching has allowed the development of different possibilities: 1. erasure of content which could make it difficult to obtain relevant information for the student to use; 2. generation of content that makes the information more understandable to the student; 3. giving the student the possibility to observe the object from different points of view and different angles; 4. facilitation of the generation of ubiquitous learning for the student; 5. the students' immersion in an "artificial" environment, such as simulators or laboratories; 6. an enriched printed material through different resources; 7. improvement of the way in which students learn through motivation; 8. learning objects that generate "prosumers" students instead of "consumers" of information. The potential offered by this technology is increased with its incorporation through different disciplines as well as the different educational levels in which they can be used. It should be noted that experiences and studies are most often conducted at the higher education level – in the universities [4].

Studies have shown that AR can have a consistently positive impact on student motivation. There are studies which prove that AR can specifically increase student engagement in science learning. The increase may be largely attributed to the elements of curiosity, fantasy, and control presented using AR technology, as student motivation may be directly influenced by using an attractive or stimulating medium or learning material [5].

A custom solution is considered for a later period in time, as the resources necessary to create one exceed the timeframe of this study. Furthermore, the data gathered using a ready-made solution would become a stepping-stone in furthering our investment in this type of

technology as well as obtaining experience and constructing opinions on how to improve the process.

Our goal is to help students get familiar with 3D object generation and take into account the newest available solutions and software which help with reducing manual work. RealityScan is a perfect example of cutting-edge tool that is still being tested on production. Using it should help students recognize the flow of software development during its final testing phases and innovation. Higher education is the place to indulge in learning sufficiently complicated software for such purposes. Work on similar topics has been progressing at our university for a long time [6], and this research is a part of the next logical steps.

II. LAYOUT

A. *Analysing available technologies*

Our criteria for comparing different tools are several – 1) the cost of the application, 2) the availability of hardware to run the application on, 3) the viability of the application to be used as a learning tool in the curriculum and last, but not least, 4) the learning curve that will be necessary for both teachers and students to pass in order to understand and use the application Table 1. In order to evaluate the learning curve, we have consulted with other professors at our university and based on their extensive experience with working with students, the difficulties of learning the tools have been taken into account.

Application	Cost	Hardware requirements	Learning materials	Learning curve
RealityScan	Free	iOS compatible device (iPhone / iPad)	online video tutorials, no official documentation	Moderate, assuming the students previous experience
Nvidia NeRF	Free	Windows/Linux	A large number of available resources	High
Reality Capture	Paid	Windows	A large number of available resources and official documentation	Moderate-to-easy

Table1. Comparison between different software applications used for generating 3D models

An analysis of some of the available technologies is made and several of their peculiarities are described.

The company creating RealityScan advertises it as a transformative tool, which captures real-world assets for digital experiences. It is put into comparison to the traditionally complicated, technical, and labor-intensive process of doing so. As it released in its' beta version, RealityScan took users preferences about RealityCapture - fast and easy 3D SCANNING – and allowed it to be used by anyone with a smartphone, turning the cameras into an endless source of creativity. Some of the greatest features of this app are: free tool, no equipment needed, ease of use [7].

Neural Radiance Fields (NeRF [8]) are neural networks capable of generating three dimensional images or scenes from a set of two-dimensional images developed by NVIDIA. Using spatial location and volumetric rendering, the model uses the camera angle from the images to render the three-dimensional space of the scene.

All the described applications can export models to SketchFab – modeling platform for publishing, sharing, buying and selling 3D and AR content.

Reality Capture is a photogrammetry software application for Windows, developed by Epic Games. It enables the user to create ultra-realistic 3D models from a set of images and/or laser scans. It is possible to create virtual reality scenes, textured 3D meshes, orthographic projections, geo-referenced maps and much more from images and/or laser scans automatically. One of the limitations of the product is its' high price and the strict requirements to the level of equipment. As the software is heavy on consuming resources, a powerful video card is a must-have [9].

Several limitations of RealityScan are present in the current version. There is a limit to how many models can be exported to SketchFab - the intertwined 3D manipulation tool. Users are limited to 50 models a month per account. For educational purposes, this is not a severe drawback, and a form of licensing can be purchased if there is a need to expand. The application itself has a maximum of 200 photos per object, but our experience indicates that 100 photos are usually sufficient to generate an adequate scene containing it.

The quality is improved gradually the more photos are taken from different angles - the change is visible in the level of detail that is present in the finished model Figure 1.

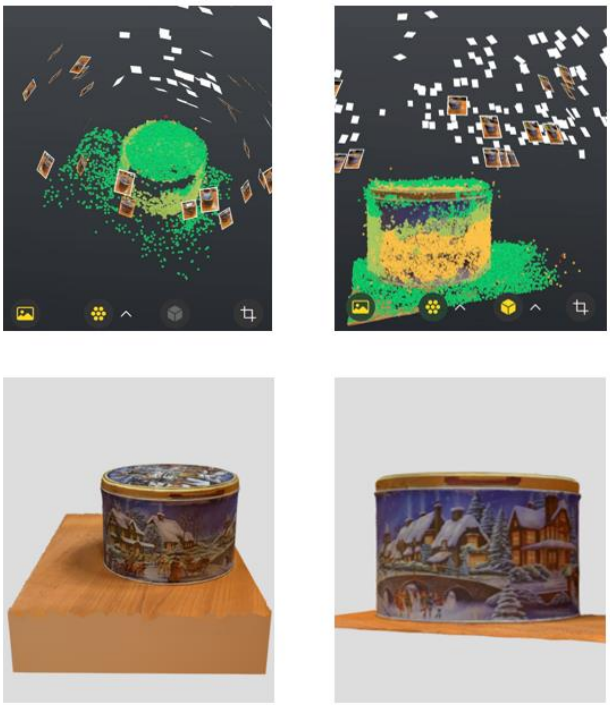


Figure 1. Comparison of three dimensional models generated using 100 photos (left) and 200 photos (right)

The SketchFab application supports a variety of exportable file formats. The default format is .fbx – FBX files typically contain mesh, material, texture and skeletal animation data, all of which are sufficient for fully creating, manipulating and modifying a 3D image. Another supported format is .glTF / .glb – the GL Transmission Format is designed for compact file size, fast loading, run-time independence and complete 3D scene representation. It is hailed by users as "the JPEG of

This research is supported by the Bulgarian Ministry of Education and Science under the National Program "Young scientists and Postdoctoral Students - 2".

3D," and can use both JSON / ASCII (.gltf) and binary file format (.glb) file extensions. SketchFab also allows exporting in .usdz – a 3D file format standard created by Pixar. It has been adopted by Apple as the main format for AR applications on iOS AR Quick Look. All of the formats described above are providing a good variety of choice in applications where the work with the models can advance. The SketchFab application is used as a "container" for the generated 3D models, it is recommended to be used in conjunction with RealityScan by its publishers.

B. Methodology

A methodology has been adapted to suit the needs of improving the students' abilities to work with advanced 3D graphic creation and manipulation software. It can be seen on Figure 2.

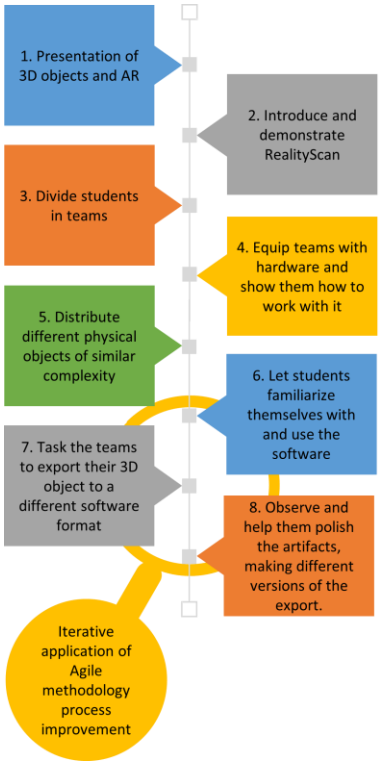


Figure 2. Adapted methodology for teaching and aiding students in their work with 3D creation and manipulation software

The first step of the methodology introduces students to the basic qualities and properties of three-dimensional objects represented in computers. Simple descriptions and types of technologies that make use of these models such as AR, VR, mixed reality, as well as some examples of them are shown. The second part of the theoretical preparation for the workshop introduces them to the software tool they would be using – RealityScan. Several examples on successful use of the software are described. A short tutorial is shown in order to let them gain initial insight into working with the tool.

The students are then divided in groups, distributed evenly among the available hardware. Our experience in separating students in teams for various projects, workshops and tasks has taught us that the best result can be achieved using groups of minimum two or maximum three students. The different teams are provided with a tool, capable of using the application and are educated in practice how to use it and gain advantage of its features. Each group is given a physical object of similar complexity to scan and are left for an appropriate time to experiment, use the tool and complete the task of scanning their assigned object, thus generating a 3D model successfully.

An iterative approach is used for the next part of the workshop – scanning the object and generating a model is the first step among several forming a feedback loop, giving the students the opportunity to practice their knowledge on agile methodology in practice (agile methodologies are processes that favor the iterative approach to finding and optimizing a solution to a problem). The next step requires the teams to export the object in format workable with Blender - a free 3D modeling software. This necessitates the use of their knowledge of different graphical formats and compatibility issues between the different software systems. The process itself is manual work intensive, but it improves their general expertise for working with 3D objects and manipulation tools, as well as stimulating their brain hand-coordination skills.

C. Further steps and expansion in other workshops

Witnessing the entire process of creating images from data is a prerequisite to visualizing the objects using AR. After the teams have completed their assigned tasks, they are required to export the objects to a software that generates QR codes to be used for visualizing the results through their phones. An app is used to show their creations as a part of AR, so that they can see the entirety of the implementation process. Several AR libraries are currently being analyzed for further research. A temporary testing is done using the AR JavaScript library organization in GitHub [10], as well as the "Mind AR JS" subsidiary, forked by HiuKim Yuen [11].

A preliminary version of a methodology has been created to aid the process of including the scanned object to another workshop project. It can be seen on Figure 3.

The process mainly relies on students having studied web development and application deployment in their previous classes, but detailed instructions are being prepared for their benefit. The teams of students have to take their exported object and include it in a scene using the provided open source library. After they finish the task of successfully simulating a scene on their local computer, they are encouraged to publish the application on the web for testing purposes. As Software Engineering students at the University of Ruse, they are already registered in Azure, the Microsoft cloud platform, as their accounts are set up in previous subjects, required in their curriculum. The use of other hosting platforms is also accepted, as many prefer Amazon Web Services or personal hosting environments. After they have succeeded with exposing their web project to the world wide web, the learners have to generate QR codes that will make use of the software library functionality for visualizing the models in AR. The next step is testing whether everything works in the expected manner, and the last step is the live testing using their own mobile device. The students have already been instructed on the use of Agile methodologies in other subjects, so they should be using a task management system to split their obligations in the team, monitor their

work and progress and report on any problems that have arisen during testing or deployment.

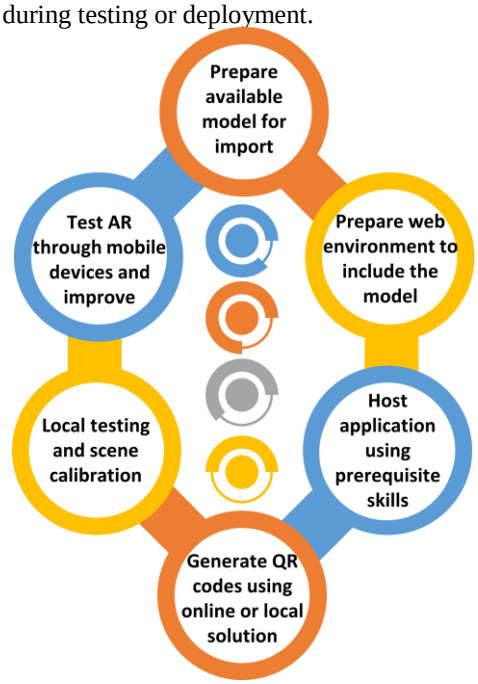


Figure 3. Methodology for training students to apply software engineering skills to the use and deployment of open source libraries aiding the visualization of 3D models in AR

D. Work and study environment preparation

A great implementation of the AR technology within RealityScan is the shown positions of photos that are already taken. This helps the users identify neglected angles that need to be addressed, as well as complements the application with the possibility to split the photogrammetry process in sessions, allowing pause and later continuation of the scanning.

Figure 4. , the entire process of capturing an object using RealityScan can be observed. After a proper object has been selected, it is put in a well-lit area, preferably on surface with color, which does not consume much light. It is also recommended that the object is of different color, compared to its environment in order to achieve the best results. The second part of the process on the figure shows a screenshot of the RealityScan capturing mode. Every photo already taken by the camera is displayed using AR, so that the operator can align the camera with less covered angles. This functionality has immense impact in improving the coverage of the object, especially when the operators are not qualified photographers. This further elevates the applications' qualifications to be used in education, where the students are not professionals and they need to start gaining valuable experience. The supportive use of AR in this case is possibly the most important factor that influenced the decision to include AR in the students' curriculum – seeing its' assistive qualities when doing practical work related to computer graphics and 3D model generation.

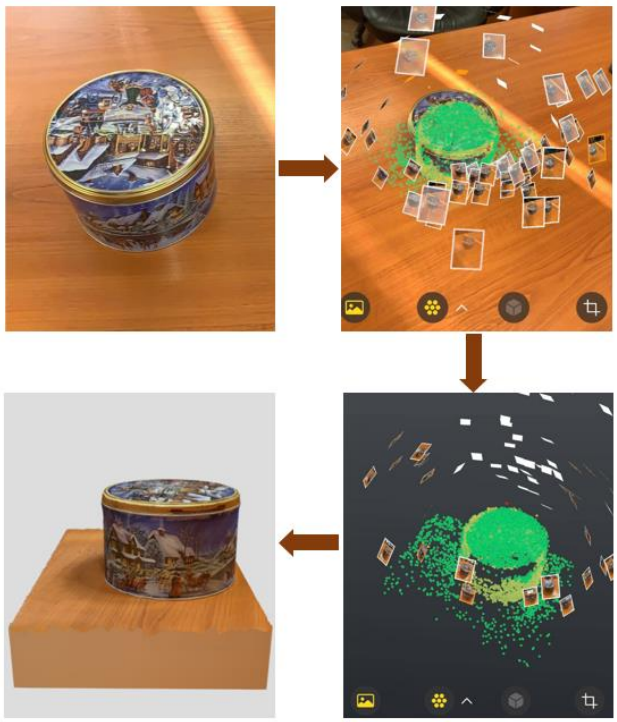


Figure 4. The different stages of converting a real life object to a 3D model in SketchFab using RealityScan

After photos of the object from different angles have been made, the next step is improving the quality of the coverage in the zones with red or yellow indicators. As shown in the third step of the process in Figure 4. , the environment can be turned off, and we can focus on which area of the model has to be worked on. A temporary model can be generated so that we can view the rough sketch and improve upon it.

The last step in the process is the exporting of the model to the connected application – SketchFab. As long as the connection is stable and the quality of photos is of the necessary level, the generated object will be of sufficient quality for educational purposes.

It should be noted that the entire process of capturing photos that are send to the cloud and the feedback returned from the cloud to the mobile device, consumes an amount of data that is currently in the volume of one to two gigabytes. There are variations depending on the quality of the camera of the mobile device, as well as the work area in which the object is being captured, and whether it has been cropped or not. After testing, our best results were obtained using a small work area, using an object that has fewer edges, in a single session. The camera quality of an IPhone 10 is sufficient for our case for education purposes.

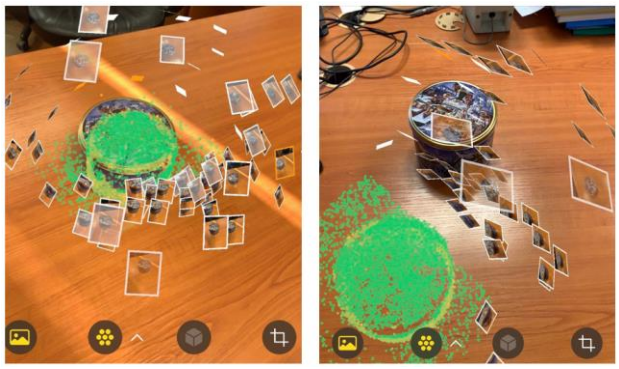


Figure 5. Comparison between the different sessions and the issues with centering the object and its' recognition points

Several difficulties that have been found during the process include correct use of the session functionality of the app. Experiments showed that it is cumbersome to find proper alignment of the object during the subsequent sessions. After the first scan is complete, the images are uploaded to the cloud and an object is generated, the data for the mobile devices' position is lost. A small task about sessions will be included in the students' workshop so that they can also experience the limitations of the software. The problem can be seen on Figure 5. The misalignment between the object in the first session and its virtual representation in the second session is obvious. A prompt centers the work area again, but it is difficult to continue the process without errors. It is meaningful to show the students the current drawbacks in positioning the camera.

E. Advantages and disadvantages in using AR in education

Based on our experience with working with AR in higher education, we have confirmed the benefits and drawbacks of using it for educational purposes.

The wide use of digital assets as well as the dynamics of students' working and living environment reflect the necessities of contemporary learners. We are currently facing a situation of continual search for possibilities for training options beyond traditional classroom approaches. Learners are connected to the virtual world by means of the Internet and their smartphones, having to absorb an enormous amount of data online. Information technologies provide learners with the ability to purposely collect and synthesize data extracted from different sources in a short amount of time [12].

Many educational institutions have realized the need to improve the learning process using AR technology. Some of the most outstanding advantages are the interactive learning process, the heightened student engagement, the transformation of study materials, and the enhancement of distance learning. From an e-learning and student-centered aspect, the key benefits of using AR include support of kinesthetic (tactile) learning, collaborative learning, distance/remote learning, learner-centered learning and creative learning [13].

As AR is a relatively new and developing technology, it introduces difficulties when educators have to operate using it. It must be noted that it has the same disadvantages of any new technology – adoption is difficult, as it is constantly evolving, improving, and changing, but with the development of applications like RealityScan, the process of including it in the classroom should become more streamlined and accessible to everyone.

Another challenge that has to be mentioned is that from the social aspect, the usage of AR can lead to isolation and undermine the element of human interaction that is so vital to successful learning environments. This is one of the main factors which prompted the inclusion of teamwork and team projects, which help in that aspect, as students work together solving challenges, communicating and improving both socially and technologically.

III. CONCLUSION

Students have to face the difficulties of operating such software and develop experience, giving them ideas for solutions and improvements, as well as igniting their interest in working with such applications.

The created methodologies strive to enhance student education and their experience with modern technologies – such as AR, virtual environments, 3D modeling and manipulation.

This study gives us the groundwork in order to further our research on the topic and start experiments with students. After all the preliminary work is completed, it is expected that the workshops should flow smoothly and the students will benefit from realizing the classwork goals.

After analyzing and testing the capability of the tools to be used in the classroom in a small test, it has been concluded that they are a viable solution to the problem of introducing and working with AR. More extensive testing with students is required, so that data and feedback can be collected and analyzed, thus improving the process and the methodologies.

ACKNOWLEDGMENT

This research is supported by the Bulgarian Ministry of Education and Science under the National Program "Young scientists and Postdoctoral Students - 2".

REFERENCES

[1] D. Roopa, R. Prabha, G.A. Senthil, Revolutionizing education system with interactive augmented reality for quality education, Materials Today: Proceedings, Volume 46, Part 9, 2021, Pages 3860-3863, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.02.294.

[2] Iqbal, M.Z.; Mangina, E.; Campbell, A.G. Current Challenges and Future Research Directions in Augmented Reality for Education. Multimodal Technol. Interact. 2022, 6, 75. https://doi.org/10.3390/mti6090075

[3] Abad-Segura, E.; González-Zamar, M.-D.; Luque-de la Rosa, A.; Morales Cevallos, M.B. Sustainability of Educational Technologies: An Approach to Augmented Reality Research. Sustainability 2020, 12, 4091. https://doi.org/10.3390/su12104091

[4] Cabero-Almenara, J.; Barroso-Osuna, J.; Llorente-Cejudo, C.; Fernández Martínez, M.d.M. Educational Uses of Augmented Reality (AR): Experiences in Educational Science. Sustainability 2019, 11, 4990, https://doi.org/10.3390/su11184990

[5] Khan T., Tasneem, Johnston, K., Ophoff, J., The Impact of an Augmented Reality Application on Learning Motivation of Students, 2019, 1687-5893, https://doi.org/10.1155/2019/7208494

[6] Ivanova, G., Ivanov, A., & Kolarov, K. (2013, June). 3D virtual learning and measuring drill tools. In Proceedings of the 14th International Conference on Computer Systems and Technologies (pp. 337-343).

[7] CapturingReality, https://www.capturingreality.com/introducing-realityscan , (accessed February, 3, 2023).

[8] Nvidia Nerf, https://developer.nvidia.com/blog/getting-started-with-nvidia-instant-nerfs/ , (accessed April, 3, 2023).

[9] CapturingReality, https://www.capturingreality.com/realitycapture , (accessed February, 3, 2023).

[10] GitHub, AR js org, https://github.com/AR-js-org/AR.js , (accessed February, 3, 2023).

[11] GitHub, Mind AR, https://github.com/hiukim/mind-ar-js , (accessed February, 3, 2023).

[12] Shoilekova K., Advantages of Data Mining for digital transformation of the educational system Computer Science Online Conference: Artificial Intelligence in Intelligent Systems, 2021, 450–454, doi: 10.1007/978-3-030-77445-5_42

[13] Alzahrani, N.M. Augmented Reality: A Systematic Review of Its Benefits and Challenges in E-learning Contexts. Appl. Sci. 2020, 10, 5660. https://doi.org/10.3390/app10165660

Implementation of a Facilitative Approach to Teaching Mathematical Disciplines to Future Mathematics Teachers by Means of the MAPLE Package

I.V. Shyshenko, Ya.O. Chkana, O.V. Martynenko, O.M. Udovychenko, I.I. Stotskyi and O.V. Semenikhina

Sumy State Pedagogical University named after A.S. Makarenko, Sumy, Ukraine
shiinna@ukr.net, chkana_76@ukr.net, elenamartova21@gmail.com, udovich_olga@fizmatsspu.sumy.ua, igor.stotskiy@gmail.com, e.semenikhina@fizmatsspu.sumy.ua

***Abstract* - The aim of the article is to reveal the experience of using a facilitative approach in teaching future teachers of mathematics in teaching mathematical disciplines using the mathematical package MAPLE. Among the methods of studying the effectiveness of the facilitative approach in teaching mathematical disciplines to future teachers of mathematics, which highlights the advantages and disadvantages of this process, we have chosen SWOT-analysis. We have determined experimentally that one of the appropriate technologies of the facilitative approach when studying mathematical analysis by future mathematics teachers is "World Cafe". The description of its practical possibilities confirms its effectiveness in creating favorable conditions for training future mathematics teachers in mathematics analysis classes using the MAPLE package, to increase the motivational component of educational activities, creative personality and subjective involvement in working on problems and situations.**

Keywords - facilitative approach; teaching future teachers of mathematics; teaching mathematical disciplines; mathematical package MAPLE.

I. Introduction

The current socio-cultural and educational situation in Ukraine is characterized dynamism, uncertainty, variability and low predictability. Under such circumstances the role and importance of professional training of future professionals in various social areas. Educational reform in Ukraine is aimed at humanizing education, democratic and tolerant attitude towards children, ensuring their right to choose the educational trajectory, meeting individual needs and desires, developing the capabilities and abilities of each child [4; 20]. In this context, the development of the New Ukrainian School uses the experience of the functioning of educational systems of the most progressive countries in the world, learning in which is based on the pedagogy of partnership [11; 15]. In this regard, the issue of appropriate training of teachers in the higher education system is relevant.

These fundamentally new conditions put forward an improvement in the quality of education professional training of future teachers of mathematics, able to operate in today's conditions. There is an application of a facilitative approach to the professional training of future mathematics teachers relevant given the need to use interactive learning technologies.

A. How the educational activities of future teachers should be organized

The organization of educational activities of future teachers should combine two important components: on the one hand – subject-oriented approach (learner-centered approach), based on individual characteristics of each participant in this process, on the other – the formation of their ability to interact in a team to achieve a certain result and communicate in a multicultural environment. This leads to the search for and implementation of various interactive methods based on the technology of cooperation and partnership, which help to build joint learning activities, promote socialization of the individual, allow you to successfully master the relevant competencies. One of such methods, in our opinion, is facilitation, which changes the roles of teacher and student, makes them like-minded and equal participants in the pedagogical process.

In the system of professional education of future teachers there is a gradual accumulation of experience in implementing a facilitative approach to the educational process, however, in our opinion, in modern research of domestic scientists this issue is not given enough attention. In particular, it requires the study and understanding of the use of facilitative technologies in the teaching of mathematical disciplines and the development of their methodological content.

B. Formulation of research tasks and methods

The purpose of the article: to reveal the experience of using the facilitative approach in teaching future teachers of mathematics in teaching mathematical disciplines using the mathematical package MAPLE.

A set of interrelated methods of pedagogical research was used to solve the formulated tasks in the research:

- theoretical: analysis, systematization and generalization of pedagogical and psychological research to clarify the content of the concept of "facilitative approach"; modeling to develop a model of mathematical analysis classes using the facilitative approach in teaching future teachers of mathematics in teaching mathematical disciplines using the mathematical package MAPLE;

- empirical: pedagogical observation of the educational process, analysis of the quality and success of training.

The methodological basis of the study are developed in world humanistic philosophy, psychology and pedagogy ideas about the self-worth of the individual and the inherent needs of continuous improvement, the position of unity of consciousness and activity, the idea of the relationship between pedagogical facilitation and universal values, meaning of life, love, freedom, creativity, dignity, the idea of the integrity of the individual. The provisions of pedagogy of self-development as a methodological basis of education and training, the provisions of system-synergetic, competence, personality-activity and eco-facilitative approaches to the organization of the pedagogical process were also system-forming for us.

C. Analysis of scientific sources

Our study is relevant to the article by Jacobs J., Seago N., & Koellner K. [5], which describes a study of video facilitation resources for math teachers (PowerPoint slides, video clips and transcripts, lesson schedules, math assignments and other handouts, a field guide to geometric transformations, congruences and similarities, interactive computer applets, built-in scores and a comprehensive facilitator's guide). Research shows that this presentation of information in the teaching of mathematical similarity on the basis of geometric transformations has a positive effect on the knowledge of teachers of mathematics. Our research concerns the teaching of mathematical analysis, but we have obtained similar results.

Given the work [1; 7; 8], the role and importance of pedagogical facilitation in the training of future teachers of mathematics is that this type of professional and pedagogical activities of teachers provides pedagogical support for personal growth and pedagogical support of his educational achievements, actualizes students' desire for self-identification, self-development self-realization, self-improvement, promotes the implementation of individual characteristics in the study of mathematical disciplines (cognitive abilities, creativity, inner intentions, inclinations, abilities, talents, aspirations).

Also, our study can be used to build pedagogical technology to prepare future teachers for pedagogical facilitation. Our article complements the study of special training of future teachers for pedagogical facilitation, which is defined as a type of pedagogical interaction that provides conscious, intensive and productive development and self-development of its participants [13; 14]. The result of professional training of future teachers for pedagogical facilitation is appropriate training, which is defined as a stable integrative personal education, including professional motives, goals, general and professional knowledge and skills, personal qualities of the future teacher, providing productive interaction with students and providing them with effective pedagogical support.

Our research is also consistent with research on the application of mathematical packages in the training of specialists in mathematics, computer science and physics [2; 3; 6; 12; 17]. It is noted that modern mathematical education of future mathematics teachers is associated with the mastery of theoretical principles, methods and tools of computer science as the basis of teaching mathematics, increases the level of professional activity of mathematics teachers.

II. THE EXPERIENCE OF USING THE MAPLE PACKAGE IN THE IMPLEMENTATION OF A FACILITATIVE APPROACH TO THE TEACHING OF MATHEMATICAL DISCIPLINES FOR FUTURE MATHEMATICS TEACHERS

A. Characteristics of pedagogical facilitation

Based on the concept of K. Rogers [15], pedagogical facilitation we mean a special style of subject-subject interaction, due in particular to the personal qualities of the teacher-facilitator, which promotes personal activity of all subjects of the pedagogical process and ensures learning productivity. The teacher-facilitator plays the role of the unspoken leader of the group, who by his presence and influence helps each of its members to show initiative, makes the decision-making process easier and more efficient, provides positive interpersonal communication.

The defining characteristics of pedagogical facilitation are cooperation; own position; individuality and equality; self-disclosure; organization of educational space [18].

Under the facilitative approach in teaching is understood a set of principles of the facilitator and ways to implement them in educational interaction through certain algorithms and techniques. Among the properties of the facilitative approach, related to the content-target and structural features of facilitative interaction, are their subjective orientation, non-directive interactivity, contextuality, dynamic multitasking [16; 20].

The problem of applying facilitative approaches in the teaching of future teachers determines the search for innovative approaches in higher education, based on the latest learning technologies. The general scheme of implementation of the facilitative approach in teaching is presented in Figure 1.

The analysis of this scheme shows that the implementation of the facilitative approach has an algorithm similar to the algorithm for solving mathematical problems. This determines the feasibility of its use in the teaching of future mathematics teachers.

B. SWOT-analysis of the use of the facilitative approach in teaching

Among the methods of studying the effectiveness of the facilitative approach in teaching mathematical

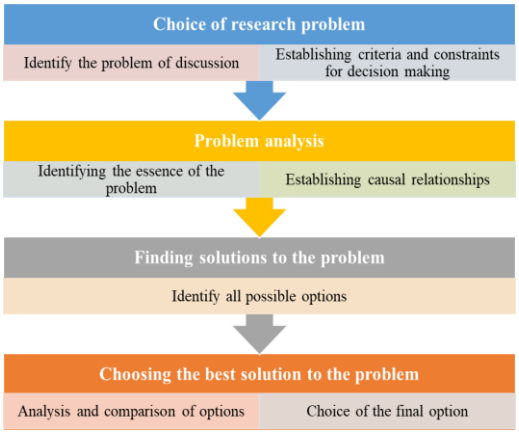


Figure 1. Stages of facilitation in learning

disciplines to future teachers of mathematics, which highlights the advantages and disadvantages of this process, we have chosen SWOT-analysis.

The results of our SWOT analysis are presented in Table 1.

This analysis determines the feasibility of using a facilitative approach in the teaching of future mathematics teachers.

Among the main characteristics of the implementation of the facilitative approach in the training of future teachers of mathematics, we highlight the following:

- directing the interaction of teacher and student to solve educational and creative tasks on the basis of cooperation and mutual understanding;

- taking into account the right of each participant to his own position and tolerant attitude to opinions other than his own;

- involvement of each student and teacher in the organization of joint activities;

- adoption of rules on ways of interaction in the group with simultaneous individual and group responsibility for the results of activities [15];

- directing the teacher's activities of students to make logically sound decisions, taking into account the specifics of the mathematical discipline in which the problem is considered.

TABLE I. SWOT-ANALYSIS OF THE USE OF THE FACILITATIVE APPROACH IN TEACHING

BENEFITS	DISADVANTAGES
following the principle of cooperation development of communicative culture of students actualization of motivational resources for learning independent search and processing of information on a particular mathematical problem creating a favorable psychological climate for solving the problem, showing initiative and independence creating a positive psychological climate of trust and support between teacher and students use of interdisciplinary links selection of information resources appropriate for solving the problem	increase the time to find a solution to the problem the opinions of individual students may not be taken into account the degree of involvement of students in solving the problem can vary significantly the wrong role position of the teacher-facilitator is possible specificity of mathematical knowledge
OPPORTUNITIES	**THREATS**
formation of skills of group activity, cooperation in decision-making formation of facilitative competence manifestation of their own intellectual abilities development of creative thinking development of attractive and assertive personality traits improving the skills of solving competence tasks improving digital skills	the emergence of discomfort in communication between the participants in the process possible loss of the solution option or no solution at all the problem-solving process can turn into a facilitator's own presentation

C. Survey of teachers of fundamental disciplines in the specialty "Mathematics" of pedagogical universities

We conducted a survey on the use of a facilitative approach in the training of mathematics teachers. 28 mathematics teachers took part in the survey. According to the survey results, only 18.7% of teachers use facilitative approaches in all types of classes, 28.6% indicate that they do not take into account the provisions of this approach in their professional activities, and 52.7% of teachers indicate that they partially use it only for time of practical and seminar classes. 54.8% of teachers turn to additional literature (articles and manuals) for the implementation of the facilitative approach, another 37.3% do it sporadically use Internet resources (video) for this purpose. Proportionality in the answers about teachers' preferences for the introduction of a facilitative approach in different types of classes was interesting: 34.6% of respondents prefer the introduction of a facilitative approach to study theoretical issues, and during the training of practical skills 65.4% turn to a facilitative approach. The results confirmed the fact that it is currently difficult to find a teacher who does not use IT in their practice. Yes, to the question: "Do you use IT in your work?" none of the respondents gave a negative answer. 76.9% of participating teachers answered that they use IT quite often, 23.1% - use, but not systematically. These results indicate that teachers in Ukraine today are well aware of existing information technologies and educational resources. All surveyed teachers used in their practical activities presentations and demonstration materials (interactive posters, diagrams, drawings), more than half use videos, interactive practical work, multimedia systems. Almost the same number of respondents also deal with methodological collections and electronic textbooks. One third of the teachers answered that they used the Maple mathematical package in practice.

Thus, the analysis of the results of the survey conducted among teachers revealed that respondents use a facilitative approach in their practical activities and consider its use necessary today. More than two thirds of respondents believe that this process is effective and ready for its implementation.

D. Describe the experience of using a facilitative approach in teaching future teachers of mathematics when teaching a course of mathematical analysis by means of Maple

The facilitative approach to the training of future mathematics teachers can be carried out in a group or microgroups in the process of verbal exchange of each member of the group or individual microgroups whose members perform a common task. At the same time, the application of this approach is expedient both during lectures and in seminars and practical classes in mathematical disciplines, and the direct choice of the appropriate facilitative technology determines the following aspects: the purpose and objectives of joint activities to achieve the educational goal; algorithm of performance of work and forms of cooperation of participants at its various stages; ways to organize feedback; list of involved resources and digital tools.

Another factor in the successful implementation of a facilitative approach in the teaching of mathematical disciplines to future teachers of mathematics, we consider the use of computing power of information technology.

Analysis of the work environment and the realities of modern mathematical training show the feasibility of studying the possibilities of using several packages, but the limited study time and financial costs narrow the range of SCM that should be considered in higher education. The power of symbolic calculations and graphics, a wide range of teams that support the solution of problems in various fields of mathematics led to the choice of mathematical environment MAPLE [9; 19].

We describe the experience of using a facilitative approach in teaching future teachers of mathematics when teaching a course of mathematical analysis.

The course "Mathematical Analysis" is studied by students in the first and second years of pedagogical universities, and in the third and fourth years they use the acquired knowledge to study courses of complex analysis, differential equations, equations of mathematical physics, special sections of mathematical analysis. The material of the course of mathematical analysis has a powerful potential for the systematic formation of their readiness for teaching.

The organization of the educational process at the Pedagogical University for students to study the course "Mathematical Analysis" involves the implementation of two main components: the study of theoretical material and its application to solve problems. When studying the course "Mathematical Analysis" at the Pedagogical University, the student must master a certain system of knowledge and skills on the main issues of the course (master the system of concepts and master the apparatus of differential and integral calculus, learn to use knowledge to solve abstract and applied problems). In addition, students should develop mathematical abilities, form a high level of mathematical thinking. Therefore, there is a need to improve the teaching methods of mathematical analysis, namely to apply the technologies of the facilitative approach.

Note that the correct supervision of the facilitator teacher in the process of finding its solutions is an important point in the implementation of the facilitative approach, as solving mathematical problems involves the use of different methods, but the solution is usually the only possible. In addition, the possession of fundamental mathematical knowledge on the chosen topic, necessary for practical application, is one of the factors in the successful implementation of the facilitative approach.

At the beginning of the lesson the teacher announces the topic and justifies its choice for further research and discussion. Students are divided into groups, each of which receives its own task to find the sum of the degree series. Tasks for groups, depending on the level of students, are determined by the teacher. Students with a lower level of learning opportunities receive intermediate-level tasks, and students with a higher level of learning opportunities have higher-level tasks. They investigate the properties of the proposed series during the time determined by the teacher and determine the method for finding its sum, check the results using the MAPLE package. Then the groups exchange the terms of the tasks without their work. This process continues until each group has completed all the tasks. Then the results are discussed with a possible discussion, if necessary. At each stage, pedagogical support and support of students by the teacher are important. He manages the process itself, but does not put pressure on students, intervenes in the discussion as needed. If students are unable to come up with an idea, the teacher directs their work or solves a similar problem on the board. In the end, students must choose the best solution for each problem.

When solving problems on this topic with the MAPLE package, keep in mind that the task of any sequence is identified with the function of a set on a set of natural numbers, i.e. each natural n corresponds to some $f(n)$. Implementation of typical tasks does not usually require the connection of additional packages. To get a quick answer, it is sometimes necessary to limit the calculations to the actual area.

The following are solutions to typical problems of individual work on the topic.

Problem 1. Decompose into a Taylor series by powers $(x-1)$ a function $f(x)=\ln x$.

Solutions. To schedule functions in a row, the *series* command is provided, which allows you to find the answer to a certain order of accuracy.

> f := ln(x);

$$f := \ln(x)$$

> series(f, x = 1, 5); F := convert(%, polynom);

$$F := x - 1 - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4$$

Let us check the result by constructing a function and its approximation in the form of a polynomial near the point $x = 1$ (Fig. 2).

> plot([f, F], x = .1 .. 3, color = [red, blue]);

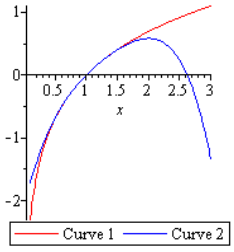


Figure 2. Illustration for the task

Answer:

$$x - 1 - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + O\left((x-1)^5\right)$$

Problem 2. Construct the integration domain $z = x^2 + y^2, z = 0, x = 1, y = 2, x = 0, y = 0$.

Solutions. For construction we will use not only commands of a package of graphics plots, but also commands of assignment of sequences.

Construction on one drawing of all surfaces that limit the body, gives only an idea of its contours. This task can be considered simple due to the simplified version of the sides of the body, it is limited by planes, which can be seen by turning the structure composed of given surfaces (Fig. 3).

>with(plots):

>implicitplot3d([z = x^2+y^2, z = 0, x = 1, y = 2, x = 0, y = 0], x = 0 .. 1, y = 0 .. 2, z = 0 .. 5, color = [red, red, grey, green, grey, green])

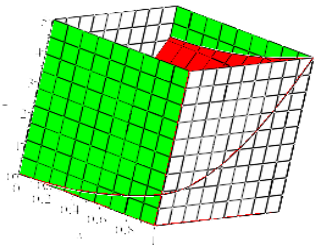


Figure 3. Illustration for the task

To obtain clear intersection curves of each of the pairs of surfaces, we use the intersectplot command and create geometric objects, which are, in particular, separate intersection curves of the paraboloid $z = x^2 + y^2$ with planes.

> p1 := intersectplot(z = x^2+y^2, x = 1, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5, color = black, thickness = 2);

> p2 := intersectplot(z = x^2+y^2, y = 2, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5, color = black, thickness = 2);

> p3 := intersectplot(z = x^2+y^2, y = 0, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5);

> p4 := intersectplot(z = x^2+y^2, x = 0, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5);

Separately we will create geometric objects that delimit the desired body from above and below.

> p5 := implicitplot3d(z = x^2+y^2, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5, color = green);

> p6 := implicitplot3d(z = 0, x = 0 .. 1, y = 0 .. 2, z = 0 .. 5, color = grey);

We will now construct the sides of the body with the *polygonplot3d* commands, for which we will separately calculate the points (sequences) on the intersection curves of the paraboloid with the vertical planes, which will be connected by a closed polyline.

> k1 := seq([1, (1/4)*T, 1+(1/16)*T^2], T = 0 .. 8);

$$k1 := [1, 0, 1], \left[1, \frac{1}{4}, \frac{17}{16}\right], \left[1, \frac{1}{2}, \frac{5}{4}\right], \left[1, \frac{3}{4}, \frac{25}{16}\right], [1, 1, 2], \left[1, \frac{5}{4}, \frac{41}{16}\right], \left[1, \frac{3}{2}, \frac{13}{4}\right], \left[1, \frac{7}{4}, \frac{65}{16}\right], [1, 2, 5]$$

> k2 := seq([(1/4)*T, 2, 4+(1/16)*T^2], T = 4 .. 0, -1);

$$k2 := [1, 2, 5], \left[\frac{3}{4}, 2, \frac{73}{16}\right], \left[\frac{1}{2}, 2, \frac{17}{4}\right], \left[\frac{1}{4}, 2, \frac{65}{16}\right], [0, 2, 4]$$

> k3 := seq([(1/4)*T, 0, (1/16)*T^2], T = 4 .. 0, -1);

$$k3 := [1, 0, 1], \left[\frac{3}{4}, 0, \frac{9}{16}\right], \left[\frac{1}{2}, 0, \frac{1}{4}\right], \left[\frac{1}{4}, 0, \frac{1}{16}\right], [0, 0, 0]$$

> k4 := seq([0, (1/4)*T, (1/16)*T^2], T = 8 .. 0, -1);

$$k4 := [0, 2, 4], \left[0, \frac{7}{4}, \frac{49}{16}\right], \left[0, \frac{3}{2}, \frac{9}{4}\right], \left[0, \frac{5}{4}, \frac{25}{16}\right], [0, 1, 1], \left[0, \frac{3}{4}, \frac{9}{16}\right], \left[0, \frac{1}{2}, \frac{1}{4}\right], \left[0, \frac{1}{4}, \frac{1}{16}\right], [0, 0, 0]$$

> p7 := polygonplot3d([[1, 2, 0], [1, 0, 0], k1, k2, [0, 2, 0]]);

> p8 := polygonplot3d([[0, 2, 0], k4]);

> p9 := polygonplot3d([[1, 0, 0], k3]);

All geometric objects are displayed in one figure with the *display* command (Fig. 4).

> display(p1, p2, p3, p4, p5, p6, p7, p8, p9, orientation = [-30, 60, 0]);

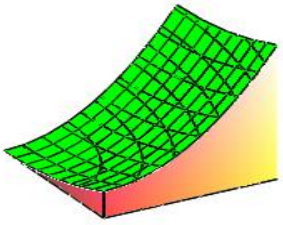


Figure 4. Illustration for the task

If students have not been able to generate any ideas, the teacher directs their work or suggests a solution to a similar problem on the board. As a result, students must choose the optimal method of solving each problem, which leads to the correct result.

The wider use of distance learning caused by the pandemic leads to the involvement of digital tools for synchronous and asynchronous interaction of participants in the educational process. The most popular in these conditions are the platform "MOODLE", electronic textbooks, social networks, cloud technologies (GOOGLE-disk), online boards, sites-designers of educational tasks and more. For lectures and seminars in mathematics, we provide programs "ZOOM" and "Teams", the capabilities of which allow you to transfer facilitative technologies to the online environment. The tools of these services allow you to group students and create closed channels in them (this can only be done by the teacher, students do not have the ability to edit groups and move from one group to another). The facilitator can be in touch with one active communication group at a time and with three others on standby. When completing a task, members of each group can chat, create an instant video conference and collaborate on a document [10].

E. The results of the pedagogical experiment

We were also interested in how the use of a facilitative approach in the teaching of mathematical disciplines of future mathematics teachers with the support of the MAPLE package affect the effectiveness of the educational process. We conducted research on the basis of Makarenko Sumy State Pedagogical University, the study involved groups of students 2-4 courses future teachers of mathematics. The study was conducted for two

years in the first half of the school year, when students studied a course in mathematical analysis with the support of the MAPLE. We compared the performance of these student groups in the first half of the academic year with their performance in mathematics analysis, but in the previous semester, when training was conducted in the usual most common way in the lecture-practical system. The results are shown in table 2.

TABLE II. RESULTS OF EXPERIMENTAL TRAINING

Group of students	Progress of students: success of training	
	ordinary learning	experimental training
2 course (12 students)	48%	54%
3 course (10 students)	53%	68%
4 course (16 students)	56%	72%

As you can see, the use of a facilitative approach in teaching mathematics to future mathematics teachers with the support of MAPLE supports the desire of higher education students to self-development, self-realization, self-improvement, helps to reveal their abilities and cognitive abilities. An important point of successful use of facilitation technologies is the possession of fundamental mathematical knowledge on the chosen topic, necessary for practical application. Therefore, we believe that the experience of using a facilitative approach in the teaching of future mathematics teachers in the teaching of mathematical disciplines using the mathematical package MAPLE is positive and deserves to be disseminated.

III. CONLUSION

Describing the practical possibilities of facilitative technologies confirms their effectiveness in creating favorable conditions for training future mathematics teachers in mathematics analysis classes using the MAPLE package, to increase the motivational component of educational activities, disclosure of creative potential and subjective involvement in working on problems and situations. The facilitative approach in higher education provides the choice and implementation of individual educational trajectories of students, promotes the development of their personal qualities.

REFERENCES

[1] H. Borko, K. Koellner, and J. Jacobs, "Examining novice teacher leaders' facilitation of mathematics professional development", Journal of Mathematical Behavior, vol. 33, 2014, pp. 149–167.

[2] K. Das, "Role of ICT for Better Mathematics Teaching", Shanlax International Journal of Education, vol. 7 (4), 2019, pp. 9-28.

[3] O. Fedorenko, and Yu. Botuzova, "Experience of using ICT tools for teaching mathematical analysis of future teachers of mathematics", Information Technologies and Teaching Aids, vol. 75 (1), 2020, pp. 153–169.

[4] H. Forgasz, "Teachers, equity, and computers for secondary mathematics learning", Journal of Mathematics Teacher Education, vol. 9, 2006. pp. 437–469.

[5] J. Jacobs, N. Seago, and K. Koellner, "Preparing facilitators to use and adapt mathematics professional development materials productively", International Journal of STEM Education, vol. 4 (30), 2017.

[6] B. Kissane, A. McConney, and K. Ho, Review of the use of technology in Mathematics education and the related use of CAS calculators in external examinations and in post school tertiary education settings. Perth, WA: School Curriculum and Standards Authority, 2015.

[7] K. Koellner, and J. Jacobs, "Distinguishing models of professional development: the case of an adaptive model's impact on teachers' knowledge, instruction, and student achievement", Journal of Teacher Education, vol. 66(1), 2015, pp. 51–67.

[8] K. Lesseig, R. Elliott, E. Kazemi, M. Kelley-Petersen, M. Campbell, J. Mumme, and C. Carroll, "Leader noticing of facilitation in videocases of mathematics professional development", Journal of Mathematics Teacher Education, available online: http://link.springer.com/article/10.1007/s10857-016-9346-y/fulltext.html (date of last access: 18.01.2023).

[9] Maple – The Essential Tool for Mathematics – Maplesoft, available on: http://www.maplesoft.com/products/Maple/index.aspx (date of last access: 20.01.2023).

[10] N. Miroshnichenko, "Using Teams to organize collaboration online in math lessons", in Distance Learning: Challenges, Outcomes and Prospects: Guide II. From the experience of educators of the city of Kiev: teaching method. manual, M.F. Wojciechowski, S.V. Ivashneva, O.G. Fidanyan (ed.), Kyiv, 2021, pp. 269-271.

[11] New Ukrainian school concept, available online: https://mon.gov.ua/storage/app/media/zagalna%20serednya/nova-ukrainskashkola-compressed.pdf (date of last access: 15.12.2022).

[12] N. Ponomareva, "The use of mathematical packages for informatics training of future teachers of mathematics", Theory and Methods of Learning Mathematics, Physics, Informatics, vol. 13(3), 2015, pp. 160-169.

[13] O. Popova, O. Kabanska, and V. Popov, "The essence and technology of pedagogical forecasting of development of innovation processes in education", in Educational Studios: Theory and Practice: monograph, Prague-Vienna: Premier Publishing, 2018, pp. 273–279.

[14] O. Popova, and A. Tsapko, "Substantiation of the content of the stages of technology of preparation of future teachers for pedagogical facilitation in professional activity", Pedagogical sciences, vol. 93, 2020, pp. 117-122.

[15] C. Rogers, Way of Being. Boston: Houghton Mifflin, 1980.

[16] A. Sacristán, J. Cortés-Zavala, and P. Ruiz-Arias, "Mathematics Education Across Cultures": in Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Mexico. Cinvestav / AMIUTEM / PME-NA, 2020.

[17] M. Santos-Trigo, and M. Camacho-Machín, "Digital Technologies and Mathematical Problem Solving: Redesigning Resources, Materials, and Extending Learning Envioronments", In K. Newton (Ed.), ProblemSolving: Strategies, Challenges and Outcomes. Hauppauge, New York: Nova Science Publishers, lnc. 2016, pp. 31-50.

[18] K. Shevchenko, "Pedagogical facilitation in the context of teacher's professional competence", Scientific Bulletin of the Nikolaev National University named after V.O. Sukhomlinsky. Series: Psychological Sciences, vol. 13, 2014, pp. 258-263.

[19] A. Syahmarani, and P. Nasution Khairiah, "Training in using Maple and GeoGebra software for mathematics teachers in the environment SMA Negeri 15 Medan", available online: https://www.researchgate.net/publication/345332246_Training_in_using_maple_and_geogebra_software_for_mathematics_teachers_in_the_environment_SMA_Negeri_15_Medan (date of last access: 03.02.2023).

[20] S. Vlasyuk, "Pedagogical facilitation as an effective way to implement the humanistic principles of partnership pedagogy", available online: https://conf.zippo.net.ua/?p=246 (date of last access: 20.12.2022).

Digital Technologies in the Activities of Physical Culture and Sports Specialists: Features of Professional Training

P. Rybalko*, O. Shukatka**, S. Lazorenko*, V. Kyselov*, N. Skachedub* and Yu. Kozeruk Paryshkura***
* Makarenko Sumy State Pedagogical University, Sumy, Ukraine
** Ivan Franko National University of Lviv, Lviv, Ukraine
*** Kyiv National University of Technologies and Design, Kyiv, Ukraine
petrorybalko13@gmail.com, shukatka1973@ukr.net, serglazarenko@gmail.com, vands7kiseliov@gmail.com, nata.skachedub@gmail.com, j.kozruk@gmail.com

Abstract **- The article is devoted to the problem of training specialists in physical culture and sports (PhC&S) for the use of digital technologies in professional activities. Digital technologies characteristics by the professional activity of PhC&S specialists are given. Attention is focused on modern digital health tools. The authors characterize digital technologies in the professional activity of PhC&S specialists and describe the features of training PhC&S specialists for the use of digital technologies in professional activities. The authors argue that the training of future PhC&S specialists for the use of digital technologies in professional activities should be carried out purposefully according to a methodology that integrates specific forms, methods, and means of training. The content of the methodology is the disciplines "Information and communication technologies in the field of PhC&S" (the goal is to form ideas about digital technologies in the field of PhC&S in general) and "Digital health technologies in the field of PhC&S" (the goal is to form: knowledge about health and a healthy lifestyle, digital technologies for maintaining and controlling their health and the health of others; ability to use digital health tools in professional activities; reflection skills about the use of digital health tools). The results of the statistical analysis of the pedagogical experiment are presented, confirming the developed methodology's effectiveness.**

Keywords - specialists of physical culture and sports, digital technologies in the field of PhC&S; digital health facilities; training of PhC&S specialists; professional training.

I. INTRODUCTION

Globalization and digitalization of all spheres of activity, particularly in the field of physical culture and sports (PhC&S), determine the emphasis on the professional training of its specialists. And if the development of physical qualities is carried out, as a rule, at the expense of training, then specialists' self-improvement in the field of PhC&S today is impossible without the involvement of digital technologies and tools. PhC&S specialists today have the opportunity to use specialized software, social networks, and health care services, as well as specialized tools, the indicators of which can be used to monitor the physical state of the body. Therefore, the problem of learning and mastering such tools becomes relevant for PhC&S specialists already at the stage of their professional training.

Analysis of scientific and methodological works in the field of professional training of PhC&S specialists confirms researchers' interest in: problems of using information technologies in the professional training of specialists in the field of physical education and sports, in particular specialized software [1], mobile-oriented fitness technologies [2], innovative technologies [3], information technologies [4; 5], etc.; introduction of courses aimed at developing skills to use IT in professional activities in the educational programs of training of PhC&S specialists in the USA [6], Germany, and Poland [7]; problems of development of PhC&S specialists' information and digital culture [8].

The survey of employers proved the importance of using digital technologies in the professional activity of PhC&S specialists, including specialized software in the field of digital health, in particular, social networks and channels with video content and fitness bracelets, pedometers, etc.

The terminological analysis of the concept of "digital health" proved the lack of unanimity in the opinions of scientists and practitioners regarding its interpretation. According to the analysis of the regulatory documents of the World Health Organization, digital health should be defined as a result of using digital, mobile, and wireless technologies to achieve healthcare goals. The regulatory documents of the Ministry of Health and Ministry of Digital Transformation of Ukraine show no domestic analog of the term in question.

Content analysis of Internet sources allows us to interpret digital health as a result of the integration of digital technologies with the healthcare industry. On the Web, the term "digital health" is used to cover a wide range of technologies used in healthcare, medical informatics, medical education, health promotion, and public health purposes. It includes other concepts such as "eHealth", "mHealth", and "Wireless Health" [9; 10; 11].

Note that these concepts include, among other things, the use of social services as platforms for cooperation and

interaction between specialists and clients/patients: blogs (simple websites with short messages that reflect different points of view and contain information that is presented in a narrative form; thematic blogs are used for health care, for example, nikko.ua/blog); wikis (web pages for joint editing of information; special wiki pages highlight health issues, for example, sport-wiki); social networks (web pages that connect users and organize communication between them; for example, FB or other social networks pages for communication between fitness trainers); podcasts (audio recordings in the form of lectures, for example, on health topics, for example, podcasts.apple.com/us/podcast/ted-health/id470623173); chat rooms (environments for rapid exchange of text messages in real-time, for example, mixsport.pro/chat-rooms); video sharing (a web page that allows private or professional users to upload video files).

Based on the results of the content analysis, the directions of using social networks in the context of digital health support by PhC&S specialists were characterized as 1) the creation of special health-related pages (e.g., on Facebook) to receive answers to questions of interest; 2) sharing knowledge and receiving feedback (e.g., creating pages on Twitter, where daily information about healthy eating habits is provided); 3) using the information in a new practical context (e.g., sharing healthy food recipes and eating healthy food).

It should also be noted that PhC&S specialists actively use the YouTube service:

- cardio training (https://www.youtube.com/channel/UCQkaczRlyBjl3UKBH59W3XQ, USA);
- cardio training, yoga, fitness (https://www.youtube.com/channel/UCHJBoCDxaCTRrwCHXEBA-BA, Germany);
- tabata (https://www.youtube.com/channel/UCrom-lPe2Ekby1kWLOAnL1A, USA);
- tabata (https://www.youtube.com/channel/UCq7bR6RxqqOx8cptc1-0AVQ, Korea);
- fitness (https://www.youtube.com/channel/UCfTC21JEugC0O9tNSPQrbqQ, USA);
- running (https://www.youtube.com/channel/UCNFi_fSAzGT7ZK_xsfLboig, Estonia);
- swimming (https://www.youtube.com/channel/UCU1eyHLlMZXu0TjmI887FyA, Ukraine);
- Thai boxing (https://www.youtube.com/channel/UCF6qYExQuDsBBJrKyTl5Yig, Ukraine)
- mountainboard, skiing (https://www.youtube.com/user/skiandschool/, Ukraine).

Digital health technologies are directly related to the use of mobile communication technologies and allow monitoring of clients and providing them with qualified recommendations from PhC&S specialists [12].

Digital health tools include hardware (wireless devices, e.g., pulse oximeter, fitness bracelets, Smart Band tracker, microcontrollers, etc.) and software (e.g., MyFitnessPal, Waterbalance), social networks, and health information technologies.

Also, an important trend today is the emergence of research on the use of augmented and virtual reality [13]

Today, operating the listed technologies and tools is important in the professional activities of PhC&S specialists. Stakeholders noted the importance of forming appropriate training results within educational programs: the ability of PhC&S specialists to use digital technologies in their professional activities, not at a household level, but at a professional level.

The purpose of the study is to describe the methodology for training PhC&S specialists for the use of digital technologies in professional activities.

II. PECULIARITIES OF PHC&S SPECIALISTS TRAINING FOR THE USE OF DIGITAL TECHNOLOGIES IN PROFESSIONAL ACTIVITIES

The process of training future PhC&S specialists for the use of digital technologies in professional activities takes place under the influence of digital technologies and tools and should be carried out not spontaneously, but purposefully according to the developed method of forming skills to use digital technologies in professional activities. The methodology integrates five components (purpose, content, forms, techniques, and means of teaching) and is implemented considering several methodological approaches (technological, BYOD approach, visual-digital).

The technological approach is involved due to the need to integrate various technologies (educational, informational, digital) into the professional training of PhC&S specialists. This will make it possible to develop skills in using various digital services to create one's own professional content and distribute it through electronic communication channels. The BYOD approach is used to ensure the intensification of training and access to electronic resources for educational and professional purposes, direct mastering of digital health tools on one's own portable devices, which are often more powerful than the material and technical support of higher education institutions. We use a visual and digital approach to visualize dynamically the external, internal, and associative connections of the main concepts and processes (including the processes that occur inside the human body) that combine PhC&S and digital health fields through the use of digital technologies of the specialized direction.

The content of the model is the disciplines "Information and communication technologies in the field of PhC&S" (the goal is to form ideas about digital technologies in the field of PhC&S in general) and "Digital health technologies in the field of PhC&S" (the goal is to develop: knowledge about health and healthy lifestyle, digital technologies for maintaining and controlling one's health and the health of others; ability to use digital health tools in professional activities; reflection skills about the use of digital means health).

The discipline "Information and communication technologies in the field of PhC&S" involves the formation of knowledge about digital resources for educational and sports purposes, knowledge about specialized software and cloud services for the creation of various types and formats of sports materials (photos, audio, and video); skills to search for the special digital resources in all the kinds of sports, skills to use DT and cloud services to create/process author's content (photos, videos) in their own sports specialization. The discipline is studied in the first year, has a minimum volume (3 ECTS credits), and does not allow a fully demonstrate of the potential of digital technologies and means for specialists in the field of PhC&S. The course program involves mastering a number of topics.

The discipline "Digital health technologies in the field of PhC&S" is studied in the 3rd year of training (3 credits). Her program involves mastering such topics (Figure 1).

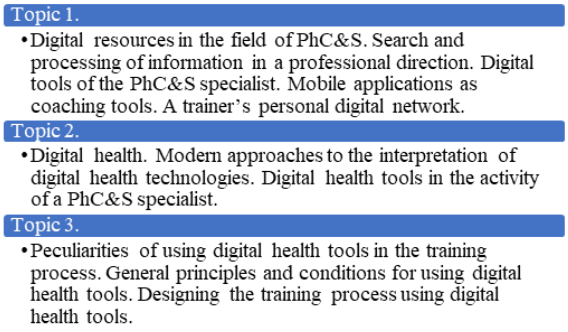


Figure 1. The program topics of discipline "Digital health technologies in the field of PhC&S"

Note that the introduction of the topic "Digital resources in the field of PhC&S. Search and processing of information in a professional direction" is due to a cursory consideration of the methods of searching for information on the Internet in the general course "Information and communication technologies in the field of PhC&S", since it is believed that the school course of Computer Science provided relevant knowledge and skills. However, as our experience shows, the level of general preparation of students majoring in "017 Physical Education and Sports" in the field of digital technologies is mostly low and average, and therefore we consider it important to develop knowledge and ability to use information search and processing methods, including those of professional direction. For this, search systems and queries, peculiarities of searching for various types of information, etc. are considered and the following types of tasks are offered: search for information about interesting (non-standard) facts from the history of Olympic sports and summarize them in the form of essays, presentations (e.g., on doping scandals), a compilation of tables (e.g., on characteristics of specialized software).

Among the methods of learning in the conditions of traditional forms of the educational process (practical classes) organization, we single out a problem-based method as one of the leading methods.

Among the tools, we consider practice-oriented exercises and exercises for the development of critical thinking of the youth to be effective. Exercises for filling in tables with blanks are classified as exercises of the first type. For the second type of exercise for the formation of skills using Internet resources, specialized social networks, or YouTube channels, "Find the mistake" exercises were offered: e.g., a video of performing a forward roll was provided, after which it was necessary to characterize possible consequences of incorrect performance of the exercise, to give approximate physical loads indicators, describe digital health indicators if a certain type of hardware was used. With reference to Internet resources, students had to justify the error and the reason for its appearance, propose possible ways to eliminate it, provide recommendations for its prevention, and describe the digital characteristics of digital health tools. The exercise ended with a multimedia presentation of the results and their discussion.

Future PhC&S specialists conducted a content analysis of the specialized resources, analyzed the possibilities of placing the author's materials and published their own video on social networks. Further discussion of video materials in social networks and interaction in web communities with classmates and teachers performed the function of reflection and contributed to the development of critical thinking. The practice has shown that such tasks contribute to the formation of students' abilities to argue their own views, search, select and analyze information, and develop communication skills.

An important element of training was mastery of various digital health tools aimed at recording individual physical characteristics of health and various software in the field of health. Among them: are pedometers; electronic bracelets, where it is possible to measure the amount of physical exertion (number of kilometers traveled, number of floors climbed, etc.), physical parameters of the body (pulse, pressure, weight, sleep characteristics, etc.); calorie counters (MyFitnessPal, FatSecret, Diet&Diary); fitness apps (Endomondo, Nike Training Club, InShapeFree, RunKeeper, SamsungGearFit). The purpose of mastering such tools was the formation of reflection skills regarding the restoration and improvement of one's own health by monitoring the dependence between a person's health and physical exertion.

When selecting the exercises, we took into account that they should be multifunctional, and meet such requirements as comprehensibility; professional orientation; availability of instructions for a certain type of action with specialized material; novelty.

III. CHECKING THE EFFECTIVENESS OF THE AUTHOR'S METHODOLOGY

To test the effectiveness of the described methodology, a pedagogical experiment was conducted, which was based on a comparison of the average in two samples (control (102 people) and experimental (81 people) groups). Students of the specialty " PhC&S " were acquainted with the features of the experiment, its

goals, and methods of data collection. The data was collected during 2019-2022.

We proceeded from the fact that groups can have different levels of their ability to use DT in professional activities. Therefore, at the beginning of the pedagogical experiment, the similarity of samples regarding the ability to use digital technologies in professional activities was confirmed. We used the rating points of students in the discipline "Information and communication technologies in the field of PhC&S". For the statistical hypotheses "H0: the averages of both groups are statistically the same" and "Ha: the averages of both groups are different statistically" the criterion of chi-square Pearson was applied. At the significance level of 0.05, the null hypothesis was confirmed by the statistical similarity of the averages [5; 14].

Then the students of the experimental group studied the special course “Technology of digital health”, after which the students of both groups were asked to complete individual tasks: to develop a program of health activities for a given age group with a certain sports specialization, in which to foresee the monitoring of possible physical parameters using digital health tools and information support of the program in social networks.

The task was aimed at testing the skills to use DH in the professional activities of future PhC&S specialists. The evaluation criteria for the final task were: taking into account age (1 point); consideration of sports specialization (1 point); use of digital health technologies (1 point), number of physical health indicators monitored (up to 2 points), information support of the program (up to 2 points). The maximum possible result is 7 points.

For statistical analysis, the Pearson chi-square criterion with the same hypotheses was again used ("H0: the averages of both groups are statistically the same" and "Ha: the averages of both groups are statistically different"). The critical value of the criterion for a significance level of 0.05 and three degrees of freedom is 7.815. The calculated value of the criterion is 8.94. Therefore, the alternative hypothesis Ha is accepted: the averages of the two groups are different, and in the experimental group, the average score is higher (72 in control group vs. 76 in experimental group). Statistical analysis proves that in the experimental group, the differences in the results are not random, they are higher in relation to the results of the control group. The analysis confirmed the effectiveness of the developed methodology, focused on the formation of the skills of future specialists of PhC&S to use digital technologies in professional activities.

IV. Conclusion

Active development of digital technologies and tools determines their use in the professional activities of PhC&S specialists. Among the variety of digital technologies and tools in the field of PhC&S, digital health should be noted as the integration of digital technologies with the field of health care.

Digital technologies in the professional activities of PhC&S specialists are connected with the use of mobile communication technologies and access to Internet resources. Digital health tools include hardware (wireless devices, e.g., pulse oximeter, fitness bracelets, Smart Band tracker, microcontrollers, etc.) and software (e.g., MyFitnessPal, Waterbalance), social networks, and health information technologies.

Mastering the listed technologies and tools is an important result of the training of future specialists of PhC&S, which integrates traditional forms of training (practical classes) with problem-based methods and tools (practically oriented exercises and exercises for the development of critical thinking) and provides for the modernization of the content of professional training.

The conducted study actualizes the problem of training future physical education specialists to use digital technologies in the conditions of non-formal/informal education. The study of the problems that stand in the way of successful training of future PhC&S specialists to use digital health tools and the development of recommendations for their avoidance also remain relevant.

References

[1] Y. I. Petrenko, V. M. Zigunov, D.I. Balashov, D.V. Bermudes, and I.V. Shyshenko, “Use of the software biomechanics in physical culture and sports specialists' preparation,” In: 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO, 2019, pp. 795-799. https://doi.org/10.23919/MIPRO.2019.8757068.

[2] N. V. Chukhlantseva, L. V. Shuba, and V. V. Shuba, “Mobile oriented fitness technologies as a means of influence on students’ physical activity,” Information Technologies and Learning Tools, vol. 75, No 1, 2020, pp. 253–268. https://doi.org/10.33407/itlt.v75i1.2581.

[3] S. Atamanyuk, O. Semenikhina, and I. Shyshenko, “Theoretical fundamentals of innovation of higher education in Ukraine,” Pedagogy and Education Management Review (PEMR), Tallinn, Estonia, Issue 2(4), 2021, pp. 30-36. https://doi.org/10.36690/2733-2039-2021-2-30.

[4] O. Semenikhina, A. Yurchenko, P. Rybalko, O. Shukatka, D. Kozlov, and M. Drushlyak, “Preparation of future specialists in physical culture and sports for the use of digital health means in professional activity,” Information Technologies and Learning Tools, vol. 89, No 3, 2022, pp. 33–47. https://doi.org/10.33407/itlt.v89i3.4543.

[5] S. Atamanyuk, Semenikhina O., and I. Shyshenko, “Possibilities of information technologies in training of specialists of physical culture and sport for the use of innovative types of motor activity,” New York. TK Meganom LLC. Innovative Solutions in Modern Science, vol. 2(46), 2021, pp. 196-210. https://doi.org/10.26886/2414-634X.2(46)2021.15.

[6] F. Hayward, “Internationalization of U.S. Higher Education,” Preliminary Status Report 2000. Washington, DC: American Council on Education, 2000.

[7] Program Rozwoju Sportu do roku 2020. Retrieved from: http://www.msport.gov.pl/sport.

[8] S. A. Lazorenko, and O. V. Semenikhina, “Development of Information and Digital Culture of Future Specialists in Physical Culture and Sports as a Modern Problem of Education,” Science and Education a New Dimension. Pedagogy and Psychology, vol. VIII (95), issue 239, 2020, pp. 29-32.

[9] L. Laranjo, D. Ding, B. Heleno et al., “Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression,” British Journal of Sports Medicine, vol. 55, issue 8, 2021, pp. 422-432. http://dx.doi.org/10.1136/bjsports-2020-102892.

[10] J. M. Raja, C. Elsakr, Sh. Roman, B. Cave, I. Pour-Ghaz, A. Nanda, M. Maturana, and R. N. Khouzam, “Apple Watch, Wearables, and Heart Rhythm: where do we stand?” Annals of

Translational Medicine, vol. 7, issue 17, 2019. https://doi.org/10.21037/atm.2019.06.79.

[11] R. Sedhom, M.J. McShea, A.B. Cohen et al., "Mobile app validation: a digital health scorecard approach," Npj Digital Medicine, vol. 4, issue 111, 2021. https://doi.org/10.1038/s41746-021-00476-7.

[12] D. Lupton, "Digital Health Technologies and Digital Data: New Ways of Monitoring", Measuring and Commodifying Human Embodiment, Health and Illness. Retrieved from: https://www.researchgate.net/publication/305434534.

[13] V. Thakare, G. Khire, and M. Kumbhar, "Artificial Intelligence (AI) and Internet of Things (IoT) in Healthcare: Opportunities and Challenges," ECS Transactions, vol. 107, issue 1, https://doi.org/10.1149/10701.7941ecst.

[14] V. Petruk, Yu. Rudenko, A. Yurchenko, I. Kharchenko, S. Kharchenko, and O. Semenikhina, "Analysis of the Results of the Pedagogical Experiment on the Integrated Analysis of the Average and Dispersions," International Journal of Modern Education and Computer Science (IJMECS), vol. 14, No 6, 2022, pp. 25-34. https://doi.org/10.5815/ijmecs.2022.06.03.

Exploring the Pedagogical Use of AI-Powered Chatbots Educational Perceptions and Practices

E. Rakovac Bekeš*, V. Galzina **
* I. gymnasium Osijek, Osijek, Croatia
** Faculty of Education Osijek and Faculty of Kinesiology Osijek
ella.rakovac@skole.hr, vgalzina@foozos.hr

Abstract - **By providing personalized and adaptable learning experiences, artificial intelligence (AI) has the potential to revolutionize education. One promising development in this field is the use of generative artificial intelligence technology, such as the ChatGPT conversational agent. Chatbot technology has the potential to revolutionize the way educational content is accessed, created, understood, and implemented. In this research paper, we examine the current state of AI-powered chatbots in education and discuss the advantages and disadvantages of using chatbots in this context. A case study of an AI unit planner (Copilot) built on top of the GPT-3 application in K–12 settings will also be presented, examining how chatbots are used to assist teachers in lesson design and the extent to which teachers are familiar with them. Finally, the implications of chatbot technology for the future of education are discussed, along with recommendations for future research directions.**

Keywords - artificial intelligence; chatbots; education; ChatGPT; educational content

I. INTRODUCTION

Artificial Intelligence (AI) has been hailed as a transformative technology that has the potential to revolutionize a variety of fields of human endeavor. AI algorithms and techniques have been developed and implemented in a variety of human endeavors, and the impact of AI on society is likely to increase as the field continues to advance. [1]

As stated in [1], there are numerous examples of using AI for a good purpose. AI has been utilized in the medical field to diagnose diseases, provide personalized treatment plans, and monitor patient health. Algorithms based on artificial intelligence have been trained to recognize patterns in medical imaging, allowing for earlier diagnosis and more precise test results. AI has the potential to provide physicians with real-time patient data, allowing them to make informed treatment decisions.

The financial industry is another area where AI has made substantial progress. Several financial procedures, including credit scoring and fraud detection, have been automated using AI algorithms. Financial institutions and investors can make more informed decisions with the assistance of AI systems that analyze market trends and offer investment advice.

The transportation industry also enjoys the advantages of AI. The development of autonomous vehicles that can transport people and goods safely and efficiently is currently underway. Additionally, AI systems can analyze traffic patterns and optimize routes, thereby reducing congestion and enhancing transportation efficiency overall.

AI has had an impact on the field of education as well. AI systems have been used to personalize student learning experiences and provide real-time performance feedback. AI systems can also analyze student data to identify problem areas, allowing teachers to concentrate their efforts on these areas.

AI has the potential to revolutionize a variety of other fields, such as retail, manufacturing, and energy. The promise of artificial intelligence is that it can help us make better decisions, increase efficiency, and more effectively solve complex problems. Nevertheless, it is essential to be mindful of the limitations of AI and to ensure that its application is ethical, responsible, and sustainable.

It has the potential to be a transformative technology that can enhance a variety of facets of life. Even though it's important to be aware of AI's limits and possible risks, its benefits are clear, and as the technology gets better, its effects on society are likely to grow. As AI technology continues to advance, it's essential that we develop guidelines and protocols for its use.

Given the new trend of people using digital technologies in their vacant time and the fact that a big part of daily life takes place on social networks and the Internet in general, large companies saw this as an investment opportunity and began investing substantial financial resources to contribute to the development of these technologies while increasing their revenues. Because technology has made natural interaction possible, it is inevitable that it has also become a trend in education.

II. ARTIFICIAL INTELLIGENCE AND CHATBOTS IN EDUCATION

Artificial intelligence is a science and research aimed at creating intelligent machines, especially intelligent computer programs, according to [2]. When asked if artificial intelligence is a copy of human intelligence, McCarthy says that it can be, but not always. There are many definitions of artificial intelligence, but this is the most widely accepted one. That there is no consensus on the definition can be concluded from [3]. The issue is that intelligence itself is a problem. The same issue arises with the definition of artificial intelligence due to the complexity and varying understandings of intelligence. According to [2], artificial intelligence refers to the research and development of intelligent devices, particularly intelligent

computer systems. McCarthy responds that artificial intelligence can imitate human intelligence, but not always. There are numerous acknowledged definitions of artificial intelligence, but this one is the most prevalent. It can be concluded from [3] that there is no consistency over the definition. Problematically, intellect itself is the issue. The similar problem emerges with the definition of artificial intelligence due to the complexity and diversity of intelligence definitions.

AI is, at its most basic level, the use of computers and machines to do tasks by emulating human experience, decision-making, and other processes. In other words, AI happens when systems perform complex pattern matching and acquire knowledge in the process. There are numerous ways to understand the nature of artificial intelligence. There are two types of artificial intelligence: rule-based and machine learning-based. The second provides a recommendation or solution proposal based on decision-making guidelines. In this sense, it is the most fundamental shape. A system of intelligent tutoring is an example of this technology. It may provide students with detailed and individualized feedback [4]. AI fueled by machine learning is more effective since the computers can actually learn and improve over time, especially when presented with large, complicated datasets.

A. Chatbots

The term "chatbot" represents two defining factors: "chat" in lieu of conversational capabilities and "bot" shortly for robot. [5] Chat-bots are automated programs meant to perform instructions based on certain inputs and deliver feedback in a conversational manner. [6]

A chatbot is a computer program based on artificial intelligence that replicates human communication via textual and/or auditory means. However, different chatbot-related terms are used today that don't share the same interpretation: chatter bot, smart bot, educabot, quizbot, digital assistant, personal assistant, virtual tutor, conversational agent, etc. Typically, they are excellent applications of artificial intelligence (AI).

The evolution of chatbot technology has accelerated over the past few decades, due in part to recent developments in Natural Language Processing (NLP) and machine learning [7,8]. The origins of chatbots can be traced back to the 1950s, when Alan Touring wrote a paper in Computing Machinery and Intelligence posing the challenging issue "Can machines think?" (1950). Since then, numerous chatbots have been developed, including Eliza (1966), Parry (1972), Racter (1983), Jabberwacky (1988), and A.L.I.C.E. (1995), some of which are still in use. Originally, these chatbots were text-based and preprogrammed, based on Question - and - answer scripts, so their reactions were deemed predictable, and the interaction was not viewed as natural by human standards. Modern chatbots include new functions such as speech recognition and synthesis technology, customizable interaction, integration with third-party applications, omnichannel deployment, context-awareness, and multi-turn capacity [7,9,10]. As a result, a variety of chatbots are now integrated into many electronic devices, programs, and applications, such as messaging platforms (WhatsApp, Telegram, Kik, Slack), interactive media and gaming platforms (Xbox, Roblox), and social media (Instagram, Facebook Messenger, Twitter). There are essentially no areas where chatbots do not currently exist or will not exist in the near future. As Fryer et al. [10] emphasized: 'During the first three decades of their existence, chatbots evolved from exploratory software to a wide variety of potential internet-based friends, advisors, and merchants. In the two decades that followed this initial rise, advancements in text-to-speech-to-text and the increasing usage of mobile devices and home assistants have made chatbots an integral part of the daily lives of many people. (p. 17) The word chatbot can be deceiving, as it can refer to a large number of programs with varying styles and functions.

Conversational AI, further known as conversational agents, can be defined as intelligent software programs that learn through communication in a manner similar to humans. Originally, conversational agents provided standard responses to common questions, but they have since evolved into more complex software with the help of Natural Language Understanding (NLU), neural networks, and deep-learning technologies. They represent a sophisticated dialogue system that replicates written and/or spoken discussion with human users, generally through the Internet. They may be based on text or be read aloud, and they may respond with speech, visuals, virtual motions, or haptic-assisted physical gestures. In addition, some of them, such as IBM Watson, Apple Siri, Samsung Bixby, and Amazon Alexa, are intelligent personal assistants (IPAs) trained to perform specific duties [11]. Fryer and Carpenter [12] studied the interaction between two chatbots and 211 students and concluded that they might be utilized well for self-practice through regular conversation, although the authors considered that these programs are more beneficial for advanced language learners. Correspondingly, Hill et al. [13] compared 100 instant messaging interactions to 100 exchanges with the chatbot named Cleverbot across seven dimensions: words per message, words per conversation, messages per conversation, word uniqueness and use of profanity, abbreviations, and emoticons. The authors discovered notable differences, such as the fact that people communicated with the chatbot for longer periods of time (but with fairly short texts) than they did with another person, and that human–chatbot communication appears to lack much of the vocabulary found in conversations between people, which is consistent with later research on the use of chatbots by children [14]. Ayedoun et al. [15] employed a semantic approach to illustrate the positive influence of deploying a conversational agent on the willingness to communicate (WTC) in the setting of English as a Foreign Language (EFL) by exposing users to a variety of everyday conversation contexts. In their study, the conversational agent created an immersive environment in which students could imitate a variety of everyday English interactions in order to lessen their fear and boost their self-confidence. Until now, numerous articles on conversational agents in language acquisition have been published. Io and Lee [16] found various study gaps related to "the paucity of research from the human perspective" (p. 216) and stated that modern types of chatbot such as mobile chat apps, embedded website services, and wearable devices should receive special attention in future research.

Shawar [17] reviewed the usage of several chatbots in language learning and emphasized some of its benefits, including student delight, less language anxiety, endless repetition opportunities, and multimodal capabilities. In a separate assessment, Radziwill and Benton [18] analyzed a selection of qualities and quality concerns associated with the creation and deployment of chatbots. The authors considered criteria such as efficiency, efficacy, and customer happiness before concluding that chatbots can positively enhance learning but can potentially be used to engineer societal harm (misinformation, rumors). In a similar vein, Haristani [19] evaluated various types of chatbots and highlighted six benefits: lessened language anxiety, widespread availability, multimodal practice, novelty effect, rich range of contextual vocabulary, and useful feedback. Bibauw et al. [20,21] presented a comprehensive review of 343 dialogue-based systems-related papers. The authors described the interactional, instructional, and technological components as well as the favorable effects on student self-confidence and motivation. Recently, Huang et al. [22] examined 25 empirical studies and found five pedagogical uses of chatbots: as interactants, simulations, helplines, transmission, and recommendations. The authors highlighted various advantages, such as timeliness, usability, and personalization, as well as disadvantages, such as the perceived unnaturalness of the computer-generated voice and communication failures.

Dokukina and Gumanova [23] underlined the originality of employing chatbots in a new scenario based on microlearning, learning process automation, and an adaptive learning environment. However, some disadvantages of using chatbots in education have also been highlighted. Criticism is primarily focused on the planned character of their responses, which makes them somewhat predictable, as well as their restricted understanding (vocabulary range, purposeful meaning) and inefficient communication (off-topic, meaningless words) [24,25]. Students' (lack of) interest in such predictable technologies and the (in)effectiveness of chatbot–human interactions in language acquisition are the two most frequently stated obstacles. In the first scenario, [26] examined chatbot–human versus human–human talks across a variety of tasks and assessed the effect on student engagement. The researchers stated that human associate tasks can predict future course interest of students, whereas chatbot partner circumstances decrease this interest. In other words, after a certain period of time, students lose interest in chatbot conversations, a phenomenon known as the novelty effect.

B. GPT-3

More than four years ago, OpenAI academics and engineers published their initial work on generative models—language models—artificial intelligence systems that could be pre-trained with a massive and diverse corpus of text via datasets, a method they termed generative pre-training (GP). [27] Language understanding performance in natural language processing (NLP) was enhanced in GPT-n through "generative pre-training of a language model on a diverse corpus of unlabeled text, followed by discriminative fine-tuning on each individual task," as reported by the authors. This reduced the requirement for human oversight and time-consuming hand-labeling. [27]

Microsoft introduced their Turing Natural Language Generation (T-NLG) system in February 2020 [28], claiming that it was the "biggest language model ever published with 17 billion parameters." It improved performance over any other language model on a range of tasks, including text summarization and question answering.

The 2020 version of Generative Pre-trained Transformer 3 (GPT-3) is an autoregressive language model that employs deep learning to generate text that resembles human language. Given an initial text as a prompt, it will generate a continuation of the prompt. GPT-3, which was in beta testing as of July 2020, is part of a movement in natural language processing (NLP) systems toward pre-trained language representations. [29]

The structure is a decoder-only transformation network with 2048-token-long contexts and an extraordinary 175 billion parameters, needing 800GB of storage space. The model was trained through generative pre-training; it is trained to anticipate the future token based on the previous tokens. On numerous tasks, the model showed impressive zero-shot and few-shot learning. [27]

It is the third-generation language prediction model in the GPT series, produced by OpenAI, a San Francisco artificial intelligence research center, as the successor to GPT-2. [29]

The exceptional quality of the writing produced by GPT-3 makes it impossible to identify whether it has been written by a person, which has both advantages and disadvantages.

A team of 31 researchers and engineers at OpenAI described the construction of GPT-3, a third generation "state-of-the-art" language model, in an arXiv preprint dated May 28, 2020. [29] The team enhanced GPT-3's capacity by more than two orders of magnitude relative to its predecessor, GPT-2, making it the largest non-sparse language model to date. (In a sparse model, many parameters are fixed, thus even though there are extra overall parameters, there is fewer meaningful data.) Because GPT-3 is similar in structure to its predecessors, its improved capacity and number of parameters account for its greater accuracy. GPT-3 is 10 times more capable than Microsoft's Turing NLG, the next largest NLP model at the time. [30]

In this article they cautioned of the possible risks of GPT-3 and advocated for further research to decrease risk. David Chalmers, an Australian philosopher, referred to GPT-3 as "one of the most fascinating and significant AI systems ever developed." A review published in The New York Times in April 2022 stated that GPT-3 is capable of writing unique language with human-like fluidity.

Since the training data for GPT-3 was exhaustive, it doesn't require additional training for specific language tasks. [30] The trained data contains occasional harmful language, and GPT-3 generates occasional harmful language while imitating its training data. According to research from the University of Washington, GPT-3 produced similarly harmful language as GPT-2 and CTRL, which are also natural language processing models. OpenAI has adopted a variety of ways to decrease the

amount of harmful language produced by GPT-3. As a result, GPT-3 generated less harmful language than its predecessor, GPT-1.

C. ChatGPT

OpenAI introduced ChatGPT (Chat Generative Pre-Trained Transformer) [31] in November 2022. It is built upon OpenAI's GPT-3 family of big language models and is fine-tuned (a transfer learning technique) utilizing both controlled and reinforcement learning approaches.

It was released as a concept on November 30, 2022, and soon attracted notice for its thorough responses and clear answers spanning a wide range of academic fields. Its inconsistent factual accuracy was noted as a major flaw. After the launch of ChatGPT, OpenAI was valued at $29 billion.

Although a chatbot's primary goal is to simulate human conversationalists, ChatGPT is multifunctional. For instance, it can write and debug computer programs, compose music, teleplays, fairy tales, and student essays, answer test questions (sometimes at a level above the average human test-taker), write poetry and song lyrics, emulate a Linux system, simulate an entire chat room, play games like tic-tac-toe, and simulate an ATM. [32] The training data for ChatGPT consists of person pages and data about Internet phenomenon and programming languages, such as message board systems and Python.

ChatGPT has received extensive and strong criticisms from educators, reporters, artists, researchers, and public advocates in the months following its introduction. James Vincent of the website The Verge viewed ChatGPT's viral success as sign that artificial intelligence has entered the mainstream. But users have noticed on ChatGPT's "hallucinate" habit, since it fabricates facts, even references. [33]

III. An impact on teaching and learning

When considering the broader issue and what it means to teach and learn in the age of AI, embracing teaching and learning with the help of AI is a difficult and inadequately elaborated topic that teachers need to address. In fact, researchers such as Seldon and Abidoye [34] refer to education as the "Cinderella of the AI tale," pointing to the underdeveloped and generally overlooked issue of employing AI in learning and teaching settings. In contrast, Holmes et al. [35] believed that it would be presumptuous to assume that AI will have no effect on teaching and learning, not just from a technological standpoint but also from a pedagogical and ethical one, as well as from the perspective of the competencies of educators. The primary distinction between AIED and other educational technology solutions is that AIED aims to facilitate the creation of adaptable and personalized learning environments so each pupil has their own educational opportunities. AIED systems would ideally be capable of making computational conclusions that aid teachers in gaining a comprehensive picture of how students learn best and how optimal learning is influenced by past knowledge, teaching methods, and learning and physical settings.

ChatGPT writing scientific article introductions and abstracts raises ethical concerns. It has started to be stated as having co-authored several publications. [36] In The Atlantic, Stephen Marche wrote that its implications on academic circles, especially application essays, is unknown.

Since January 4, 2023, the NYC Department of Education has blocked ChatGPT from public school internet and devices.

One of the lead data scientists and creators of a Croatian recruitment app made for successful hiring tested ChatGPT on the Croatian language graduation exam. The output test was corrected by a professor of the Croatian language who corrects real graduation exams. [37] ChatGPT performed better on literature tasks containing text (80%), while it performed slightly worse on tasks without text (73%), confirming once again that this AI tool requires context for more accurate conclusions.

The author of this test is convinced that the language results would be improved if ChatGPT received more information about them, and he demonstrated this on a question which required selecting the answer with a vowel from the four provided options. ChatGPT generated only 150-200 words, some of which were in English and Serbian, so after multiple attempts it switched to "classic GPT" i.e., GPT-3, which can generate longer texts due to its fewer character limitations.

GPT-3 tore the essay apart from the beginning, and the aforementioned portal also published a thorough correction of it, in which the professor who made the correction reached the following conclusion: The essay focuses primarily on the work's content. The interpretation of the characters' relationships and the description of the events are extremely precise. Literary-historical and literary-theoretical concepts, textual quotations proving claims, and comparisons to other literary texts (reader experience) are absent. Spelling and grammar rules are strictly adhered to, while syntactic rules are slightly disregarded.

While technology cannot impact the educational system on its own, synergy between technology and educators is required for optimal results. Technology is not a hazard to the educational process or the methods of instruction; it serves to provide possible solutions to specific issues and concerns, and the best results will arise from technology and teacher cooperation.

IV. Educational Copilot – AI unit planner

Educational Copilot is an AI tool that is helpful for educators. It allows teachers to create unit plans, presentations, quizzes, handouts, projects, writing prompts, student reports and even generate a parent email within a few minutes. The main objective of its creators [38] is to provide educators with high–quality, AI–generated lesson planning and preparation templates and resources. They target to save educators time and energy so that they can concentrate on what truly matters: their students.

This tool saves teachers a lot of their time mining Google for valuable information, and it is significantly

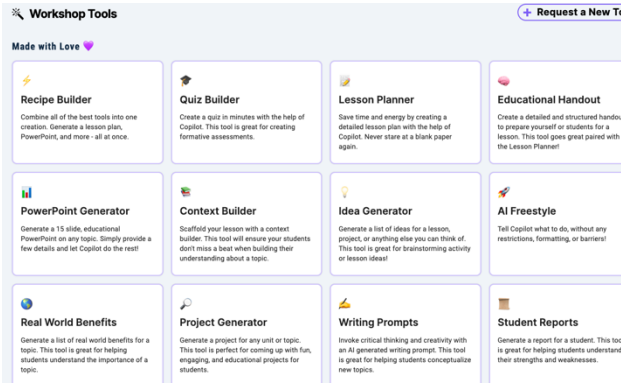


Figure 1. Screenshot of educational CoPilots' available tools at the moment

better used. Copilot, like most other tools, is developed on GPT-3 large language models. This means it is created and trained on massive amounts of data. It interprets the prompts you give it (instructions) to generate outputs. Therefore, the more detailed the prompts, the better the outputs.

The tools in Copilot cannot search the internet, they generate the outputs from training data. This means there might sometimes be incorrect information or misconceptions. The tools in Copilot basically act as automated prompting and formatting interfaces, to save teachers work.

Periodically, Copilot will 'fine-tune' the model based on the input and output data in the system. Over time, this should make it more tailored to educational use.

A. Evaluation study

We conducted a study to evaluate the quality of educational Copilot and better comprehend teachers' willingness to use it. The primary objective was to assess the instructional quality of created lesson plans and effects of AI on teachers' preparation time and overall satisfaction. This study investigated the hypothesis that a teacher's use of Copilot would increase the teacher's willingness to use AI-supported tool and to evaluate effects on planning and preparation. A modified lesson plan analysis protocol (LPAP) was used as a tool for evaluation [39].

Participants were 34 middle school teachers from four schools and three subjects (math, history, and biology). All participating teachers had at least five years of experience teaching at the middle school level and no prior experience using AI tools in or for the classroom.

B. Usual lesson plans vs Copilot lesson plans

Using Copilot, we created three lesson plans for each of the listed subjects (Mathematics, Biology, History). The initial data entry consisted of the titles of the currently used textbooks in the classrooms and the curriculum outcomes stated for a specific topic. Presentations, key-questions, handouts, and additional worksheets for homework were included in the lesson plans created for two hours of class time. We provided teachers with a choice between lesson plans developed with Copilot and those developed by other co-experts which were offered in the methodological manual (non-AI). Teachers were not informed that AI created some of these offered lesson plans. Offered plans were identical in general metadata (title, summary, keywords, etc.) in terms of definition of educational objectives (subject domain, topic domain, general learning outcomes, curriculum standards) and needed follow-up content. Teachers had to evaluate both lesson plans for their subject and decide which one they would prefer to use in their teaching. They filled out a modified LPAP Evaluation Questionnaire which includes 5-point Likert items for evaluating the instructional quality of each lesson in the following categories: educational objectives, educational resources, lesson stages, lesson approaches, lesson evaluation. At the end of the questionnaire there was a selection criterion for proposed lesson plans.

All three lesson plans from the teacher's manual (non-AI) were rated "good" or "very good" by 76% of teachers. Only 4% rated these lesson plans as excellent in regards to their pedagogical value. Lesson plans created using Copilot achieved a slightly better average rating, with 81% of teachers rating them "good" or "very good" and 7% rating these lesson plans as "excellent". Furthermore, when they had to choose which of the two proposed lesson plans they would use in their teaching, 78% stated they would choose the one created using Copilot, 10% those from the manual, and the rest were not sure about their selection.

C. Effects

The teachers were informed, following the initial selections, that some of the selected lesson plans were created using an AI tool. The Copilot tool was identified and a link to it was provided. Teachers were required to create working materials for classes as required, evaluate the quality of the material and the tool using the same modified LPAD protocol, and list the advantages and disadvantages of the aforementioned. Nearly 90% of teachers stated that the tool is extremely useful and simple to use, 78% that they will use it frequently, and 95% that it is a significant time saver that will give them more time to spend with students and on personal growth. As for disadvantages, all History teachers agreed that the facts were slightly inaccurate and required corrections. Most teachers (63%) find the language barrier to be the greatest disadvantage; inputs can be in any language, but outputs are only in English. It takes time to translate them if they are going to be implemented.

Concerning the perception of using AI in education, alarming percentages have been observed. 74% believe that artificial intelligence is a threat to the educational sector and their professions, and 86% have never used AI in and for teaching.

The values indicate that teachers valued the Copilot tool for generating lesson plans as preferable and that content made and given utilizing AI capabilities is of good to very good instructional and pedagogical quality but that perception of it is not so positive amongst the teacher population.

V. CONCLUSION

The integration of AI and chatbots, such as ChatGPT and GPT-3 based tools, in education has the potential to

revolutionize the way teachers teach and students learn. However, despite this potential, many teachers are still ignorant of the possibilities of AI and the role it can play in education. This ignorance is a barrier to the effective use of AI in education and is something that must be addressed. By providing training and support to teachers, they can gain the knowledge and skills they need to use AI ethically and effectively in their classrooms. This will not only benefit students by providing them with more engaging and interactive learning experiences but will also help teachers to better understand and make use of the latest technology in education. The use of AI in education is still in its early stages, but as it becomes more widespread, it is important to ensure that teachers have the knowledge and skills they need to use it responsibly and effectively. With the right support and training, teachers can harness the power of AI to transform education for the better.

REFERENCES

[1] B. Marr, "10 Wonderful Examples Of Using Artificial Intelligence (AI) For Good", Forbes, 2020. Accessed on: December 13, 2022. [Online].

[2] J. McCarthy, "What is artificial intelligence", 2007. Accessed on: January 6, 2023. [Online].

[3] P. Wang, "On Defining Artificial Intelligence", Journal of Artificial General Intelligence, vol. 10, 2019, pp.1-37.

[4] J. A. Kulik, J.D. Fletcher, "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review," Review of Educational Research, vol. 86 (1), pp. 42-78, 2015.

[5] Chocarro, R.; CortinÃÉas, M.; Marcos-MataÃÅs, G. "Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics". Educ. Stud. 1-19. 2021.

[6] Ischen C., Araujo T., Voorveld H., van Noort G., Smit E. "Privacy concerns in chatbot interactions. Chatbot research and design". CONVERSATIONS 2019. Lecture Notes in Computer Science, vol 11970 (pp. 34–48). Springer, 2020.

[7] Luo, B.; Lau, R.Y.; Li, C.; Si, Y.-W. "A critical review of state-of-the-art chatbot designs and applications". Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2022.

[8] Adamopoulou, E.; Moussiades, L. "Chatbots: History, technology, and applications". Mach. Learn. Appl. 2020.

[9] Okonkwo, C.W.; Ade-Ibijola, A. "Chatbots applications in education: A systematic review". Comput. Educ. Artif. Intell. 2021.

[10] Fryer, L.; Coniam, D.; Carpenter, R.; Lăpus, neanu, D. "Bots for language learning now: Current and future directions". *Lang. Learn. Technol. 24*, 8–22.. 2020.

[11] de Barcelos Silva, A.; Gomes, M.M.; da Costa, C.A.; da Rosa Righi, R.; Barbosa, J.L.V.; Pessin, G.; De Doncker, G.; Federizzi, G. "Intelligent personal assistants: A systematic literature review". *Expert Syst. Appl. 147.* 2020.

[12] Fryer, L.; Carpenter, R. "Bots as language learning tools". Lang. Learn. Technol. Vol.10, 8-14. 2006.

[13] Hill, J.; Ford, W.R.; Farreras, I.G. "Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations". Comput. Hum. Behav. Vol. 49, 245-250. 2015.

[14] Xu, Y.; Wang, D.; Collins, P.; Lee, H.; Warschauer, M. "Same benefits, different communication patterns: Comparing Children's reading with a conversational agent vs. a human partner". Comput. Educ. 161. 2021.

[15] Ayedoun, E.; Hayashi, Y.; Seta, K. "A conversational agent to encourage willingness to communicate in the context of English as a foreign language". Procedia Comput. Sci. 2015.

[16] Io, H.N.; Lee, C.B. "Chatbots and conversational agents: A bibliometric analysis". Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, 10-13 December 2017; IEEE: Piscataway, NJ, USA, pp. 215-219. 2017.

[17] Shawar, B.A. "Integrating CALL systems with chatbots as conversational partners". Comput. Sist. 21, 615-626. 2017.

[18] Radziwill, N.M.; Benton, M.C. "Evaluating quality of chatbots and intelligent conversational agents". arXiv 2017, arXiv:1704.04579.

[19] Haristiani, N. "Artificial Intelligence (AI) chatbot as language learning medium: An inquiry". J. Phys. Conf. Ser. 2019.

[20] Bibauw, S.; FrancÃßois, T.; Desmet, P. "Computer Assisted Language Learning". Routledge: London, UK, pp. 827-877. 2019.

[21] Bibauw, S.; FrancÃßois, T.; Desmet, P. "Dialogue Systems for Language Learning: Chatbots and Beyond". The Routledge Handbook of Second Language Acquisition and Technology; Routledge: Abingdon, UK, pp. 121-135, ISBN 1-351-11758-0.2022.

[22] Hwang, G.-J.; Chang, C.-Y. "A review of opportunities and challenges of chatbots in education". Interact. Learn. Environ. 2021.

[23] Dokukina, Irina & Gumanova, Julia. "The rise of chatbots – new personal assistants in foreign language learning". Procedia Computer Science. 169. 542-546. 2020.

[24] Allouch, M.; Azaria, A.; Azoulay, R. "Conversational Agents: Goals, Technologies, Vision and Challenges". *Sensors* 2021, *21*. 2021.

[25] PeÃÅrez-MariÃÅn, D. "A Review of the Practical Applications of Pedagogic Conversational Agents to Be Used in School and University Classrooms". Digital. 18-33. 2021.

[26] Huang, W.; Hew, K.F.; Fryer, L.K. "Chatbots for language learning-Are they really useful? A systematic review of chatbot-supported language learning". J. Comput. Assist. Learn. 38, 237-257. 2022.

[27] Radford, Alec; Narasimhan, Karthik; Salimans, Tim; Sutskever, Ilya. "Improving Language Understanding by Generative Pre-Training". p. 12. 2018.

[28] Sterling, B. "Web Semantics: Microsoft Project Turing introduces Turing Natural Language Generation (T-NLG)". Wired. ISSN 1059-1028. 2020.

[29] Brown, T. B.; Mann, B. et al. "Language Models are Few-Shot Learners". Advances in Neural Information Processing Systems. Curran Associates, Inc. 33: 1877–1901.2022.

[30] Sagar, R. "OpenAI Releases GPT-3, The Largest Model So Far". Analytics India Magazine. 2020.

[31] Roose, K. "The Brilliance and Weirdness of ChatGPT". The New York Times. 2022.

[32] Edwards, B. "No Linux? No problem. Just get AI to hallucinate it for you". Ars Technica. 2022.

[33] Mintz, S. "ChatGPT: Threat or Menace? Are fears about generative AI warranted?". Inside Higher Ed. 2023.

[34] Seldon, A.; Abidoye, O. "The Fourth Education Revolution: Will Artificial Intelligence Liberate or Infantilise Humanity". The University of Buckingham Press: Buckingham, UK, 2018.

[35] Holmes, W.; Bialik, M.; Fadel, C. "Artificial Intelligence in Education: Promises and Implications for Teaching and Learning". Center for Curriculum Redesign: Boston, MA, USA, 2019.

[36] Bushard, B. "Fake Scientific Abstracts Written By ChatGPT Fooled Scientists", Study Finds. Forbes. 2023.

[37] Židov I. N1info. " Znanstvenik testirao ChatGPT na državnoj maturi iz hrvatskog" Accessed on: February 05, 2023. [Online].

[38] Educational Copilot, Decatur, Texas, https://educationcopilot.com/

[39] Ndihokubwayo, K. et al. „Lesson plan analysis protocol (LPAP): A useful tool for researchers and educational evaluators". Heliyon 8. 2022. https://doi.org/10.1016/j.heliyon.2022.e08730

Automatic Evaluation of Student Software Solutions in a Virtualized Environment

M. Fabijanić, G. Đambić, B. Skračić, and M. Kolarić
Algebra University College/Software Engineering, Zagreb, Croatia
mario.fabijanic@algebra.hr

***Abstract* – Universities are using Automated Programming Assessment Systems (APAS) to minimize problems that emerge by manually managing student software solutions: subjectivity, inefficiency in case of many solutions, and lack of fast and rich feedback for the student. Lately, handling code executions in a secure way has become a requirement. Researchers addressed this topic and proposed generic security models: user-level restrictions, process-level restrictions, and virtualization. In the case of virtualization, commonly suggested solutions are virtual machines and containers. With all the benefits of virtual machines, but less resource-demanding, containers are becoming widely used last years. This paper proposes a newly developed assessment model and an information system and proposes a solution for the automatic evaluation of student software solutions in a virtualized environment. It analyzes the key parts of the student solution source codes and commands against a combination of the test case input data, using the pattern-finding method. This method was used in the context of finding the required lexical structures. The proposed system is used during the exam evaluation of the Programming course at the College of Algebra, and the relational database courses, for evaluating SQL solutions in combination with the previously developed system at the College of Algebra.**

Keywords* – *APAS; computer science education; educational technology; automatic evaluation; assessment; virtualized environment; partial marks; configurable assessment process

I. INTRODUCTION

With the increase in the popularity of computer science, as well as university and professional studies in the field of computing, the enrollment trend at the corresponding higher education institutions has also increased. In correlation with this positive trend, the number of students who take exams and other forms of testing skills and knowledge of programming each semester has also increased. Teachers, teaching assistants, and other staff members spend a lot of time manually and repetitively correcting student solutions due to the volume of tests and the number of students that take them.

However, the problem is not new, the first solution to this problem was created about 60 years ago [1]. In the meantime, the requirements changed, and the solutions for automatic correction of the program code met the needs.

This paper describes a software solution that enables students to submit solutions to exam tasks, reduces the time consumption of exam correctors, and minimizes the possibility of human error during evaluation and scoring. For security reasons, reading and executing submitted user files should be considered untrusted and executed within a virtualized environment. The goal is to isolate the processes that occur when compiling and executing the code from the environment of the operating system that serves the service itself. Functional and non-functional requirements, architecture, and system domain entities are defined, as well as an analysis of the functionality of each system component. In the end, an example of the application of the proposed solution is shown.

The implemented service satisfies two groups of requirements: code execution requirements and solution evaluation requirements.

A. *Code Execution Requirements*

Code execution is a prerequisite for a successful solution evaluation process. The expected system capabilities are:

- translation of the original program code
- execution of compiled program code while passing input data
- stopping the process after a defined time limit.

The main purpose is to translate and execute the source code submitted as a solution to a specific task. The source code is translated using a compiler, and only if the code can be translated into an executable, the program moves to the execution phase. After executing, it is necessary to read the stream of standard output (standard output) and the stream of standard error (standard error). The compilation or execution of the program must be stopped after a certain time limit to save processor and memory resources and to ensure the smooth operation of the system.

Non-functional requirements are not related to examples of use but to system characteristics, and in the context of code execution they are the following [2], [3]:

- error tolerance
- compilation and execution within a protected environment
- agnosticism towards programming languages
- ease of setup and configuration.

When compiling the source code and executing the program, numerous errors are possible, and some of the

causes of these errors [2], [3] are the impossibility of compiling the source code, infinite compilation time, exceptions, and infinite loops. Source code with syntax error cannot be compiled, therefore the program ends with the transfer of the cause of the error. If it is an infinite translation time, the translator does not report an error, but the resources of the environment in which the program is translated are increasingly occupied.

During execution, the program may generate an exception (for example, accessing a non-existent array index) that is passed to the standard error stream, which needs to be read and the error passed. Like the case of infinite compile time, continuous execution of a program caused by faulty logic within the code does not raise an error but consumes resources of the environment in which the program is executed.

Due to the variety of software transfer formats, the submitted source code received by the system for security reasons should be considered untrusted, malicious, and a potential cause of the errors. Moreover, the code execution unit could be used within different exam correction systems, so it is important to separate the process of compiling and executing the program from the specifics of the programming language and related tools. This condition will be satisfied by the correct implementation of a protected virtualized environment.

II. METHODOLOGY

Evaluation of student program solutions consists of two separate tasks: the evaluation of execution results based on defined test cases, and static analysis of the source code, i.e., checking the existence of certain lexical structures.

Different methods can be applied during the automatic evaluation [4]: unit tests, sketching synthesis and error statistical modeling, peer-to-peer feedback, test cases with random input data, and pattern matching.

A. Test cases methods

To develop a software solution evaluation service, a combination of the test case method with random input data and the pattern-finding method was chosen [5]. Each test case contains predefined input and expected output data. The pattern-finding method was used in the context of finding the required lexical structures. The examiner defines the test-ordered pair of input-output data in the format of the standard input that will be used during the execution of the program and the standard output that will be compared with the actual output of the executed program. Based on the coincidence of these two data, the number of points will be awarded.

B. Static analysis

Static analysis of the source code is a check of syntactic rules. Each rule is defined as the presence of specific lexical constructions and the number of occurrences within the source code. For each task, the examiner defines a set of rules related to a specific request. Any non-compliance with the rules is punished with a defined point penalty, which is deducted from the final sum of points for that task.

C. Service requirements for automated evaluation of software solutions

There are several basic requirements of the service for the automated evaluation of software solutions:

- The examiner defines the test tasks: question text, name, test cases, grading rules, and the test taker receives feedback on the number of points earned and the reason for possibly deducted points.
- The examiner defines rules and test cases for each task.
- The examinee transfers his solution to the evaluation service in the form of a compressed file. The service creates a directory structure that meets the project pattern. During the exam, the examinee can upload his current solution several times to check the evaluation results of the currently solved tasks. At the end of the exam, the user confirms that his solution application is final and receives a complete detailed analysis of the scoring of each task.

D. Proposed solution architecture

The software solution includes a server web service that communicates with a database and a protected virtualized environment. The client communicates with the server by sending HTTP (Hypertext Transfer Protocol) requests according to the defined routes of the server web service.

REST (Representational State Transfer) is a program architecture model that contains a certain set of restrictions [6]. A system that implements REST is called RESTful and is based on the client-server concept. The RESTful system does not store states (stateless) and is characterized by uniform accessibility to every resource through HTTP requests. The backend service passes data from the request component to the evaluation component for code execution. The code execution component, communicating with the virtualized environment, translates the source code and if the translation is successful, executes the program and monitors the success of the implementation of these two phases. The server service stores all assessment requests in a relational database, which also stores data on defined exams, tasks, test cases, and rules.

E. Virtualized environment

There are more tools and techniques available to create a protected isolated environment. The most popular techniques are creating virtual machines and creating containers. The main difference between them is the hypervisor used by virtual computers. The hypervisor manages virtualized resources and guest operating systems, thus creating virtual machines with running complete operating systems and associated processes. Unlike virtual machines, starting a container does not require the existence of a hypervisor but a container engine installed as part of the host operating system.

Using containers with the necessary libraries within the container [7] makes it possible to start application processes independently of the host operating system. Starting containers is less time-consuming than starting virtual

machines [6], [7] thanks to the lack of a hypervisor and the absence of starting the entire virtual operating system.

To ensure the processing of submitted files without consequences for the host computer, during the operation of the service, it is necessary to repeatedly destroy and recreate the virtualized environment. For ease of management, a protected virtualized environment is implemented using containers. Large community support, ease of installation and configuration, and many libraries for integration with programming languages and working environments [8], [9], [10], [11], [12], [13] to implement a virtualized environment for the code execution component, Docker was chosen. Each container is started with the help of a Docker image, a file that contains instructions for creating a Docker container and at the same time represents the basis for the file system that will be located inside the container. Each image has multiple layers that speed up the construction process. The Docker image used to start the virtualized environment in which the program code will be compiled and executed is g++:4.9, which contains the GNU C++ compiler g++.

After successfully starting a container using the docker start command, the container is ready to receive commands. If the execution time of the program inside the container exceeds the given time limit, the docker stop command will stop the container and its processes. After stopping the container, the container is deleted with the docker rm command so that it does not occupy resources unnecessarily (Figure 2).

F. Implementation of code evaluation services

In the data layer, a database is used to store relevant data about exams and assessment results. PostgreSQL was chosen as an open-source tool with strong community support and many libraries available for integration with a variety of programming languages and frameworks.

One exam can have one or more tasks. Each task has one or more test cases and one or more rules. For test cases, a text is defined that represents the standard input and the text of the standard output that will be compared with the actual output of the program. One or more rules can be defined for each task. Each rule contains a textual identifier, a description, and a regular expression of the check, the number of points that are deducted if the specific rule is not satisfied, and the number of occurrences of the lexical construct described by the rule. During the evaluation, a record is created that contains information about the year, semester, and exam, as well as information that indicates the status of the submission of the solution. An individual exam task contains the evaluation status waiting, which is true if the task has been fully evaluated. Each evaluation creates records of the result of checking test cases and rules.

To implement the business layer, the Go programming language and the Echo framework were used, which allows the creation of a REST web server service with all the necessary functionalities, such as controllers for accepting requests and defining processing methods, authentication filters, and logging management [8], [9], [10], [11]. Web service consists of two main packages: Executor and Marker. Inside the Executor package is the executor.Executor structure and associated methods for executing code. Each of these methods serves as a wrapper for calls made by the Docker agent. Some of the important methods are: ContainerList, ContainerStart, CopyToContainer, ContainerExecAttach, ContainerKill, and ContainerRemove [8], [9], [10], [11]. The Marker package defines structures and methods for evaluating individual tasks. The verification of test cases is carried out by calling methods for compiling and executing code from the Executor package, while the verification of rules is carried out by checking the existence of, certain lexical constructions in the source code using regular expressions. After the evaluation is complete, the results are stored using the Store package method.

Two values must be sent within the body of the initial POST request at the /submission path: exam_id, which represents the integer value of the universal identifier of the exam stored in the database, and project_zip, which is a compressed file of the MS Visual Studio Solution folder and related projects. If some value is missing, an HTTP response with the status code 400 Bad request is returned to the user. If the exam with the submitted identifier is not found, the status code 404 Not found is returned. If the request received the correct parameters, the compressed file is extracted, and the extracted files are stored on the server's file system. A directory is then created whose keys are the names of the tasks, and whose values are a list of files associated with that task, which includes a file with the extension cpp and optional files with the extension txt and csv used in the tasks.

The first step is to create a record that is the result of the assessment, and the user is returned an HTTP response with the status code 201 Created and the universal identifier of the result of the submission of the solution, submission_result_id. The method is not terminated after creating the HTTP response but continues with the evaluation. First, by calling the CompileCheck method, the possibility of translating the source code of the solution is checked. If the source code cannot be translated, the evaluation stops, and 0 points are awarded.

This is followed by verifying the test cases by executing them with the passed input data and verifying the expected and actual output data. When starting execution, as well as when compiling source code, a timer is started. Using the lexical construct select of the Go programming language, the process is stopped if the counter ends before the program is executed or compiled.

Rule checking is performed by finding patterns in the text of the source code using regular expressions [14], [15], e.g., to find the definition of a for loop, the regular expression is used: [^a-zA-Z0-9_]for [\t\n]*\(. The search is for matching text that contains for loop definition.

The evaluation process can take a long time, so the system evaluates solutions simultaneously using the goroutines of the Go programming language. Each goroutine is active while the assessment is in progress and ends with the recording of the obtained points. The client application continuously send a GET request to the /submission/:id path to display the grading status and current scoring results. This process can be stopped when the waiting attribute of the JSON response body is set to false (Figure 1).

The initial POST request to the path /submissions returns the HTTP response that contains submission identifier. The submission_result_id attribute is a universal identifier of the entity that represents the result of the evaluation of the solution, and it must be sent as a parameter when calling the GET request according to the path /submission/:id where id is the value of the submission_result_id attribute (Figure 3).

Next, the GET request to the specified path returns a JSON response in which the waiting attribute is specified for each task, which indicates the status of the evaluation of the solution:


```
{"id": 101,
  "created_at": "2022-05-
18T22:58:12.814289+02:00",
  "exam": {
    "id": 1,
    "name": "Midterm Exam 1"},
  "final": false,
  "task_results": [{
      "id": 62,
      "waiting": true,
      "compiles": true,
      "scored_points": 0,
      "message": "",
      "task_id": 1,
      "test_case_results": [{
            "id": 41,
            "passed": true}],
  "test_cases_passed": 1,
  "rule_results": [{
    "id": 22,
    "satisfied": false,
    "points_affected": 0.5}],
  "rules_satisfied": 0}]}
```


III. RESULTS

The basic application of the presented solution is in the automated correction of exams from the course Programming. The course contains six learning outcomes; therefore, it is essential that the system can support the definition of rules and conditions that the student needs to meet the learning outcomes.

The conditions are:

- defining different types of variables
- defining for, while and do while loops
- defining methods with different signatures
- defining and using structures
- defining and using pointers and references

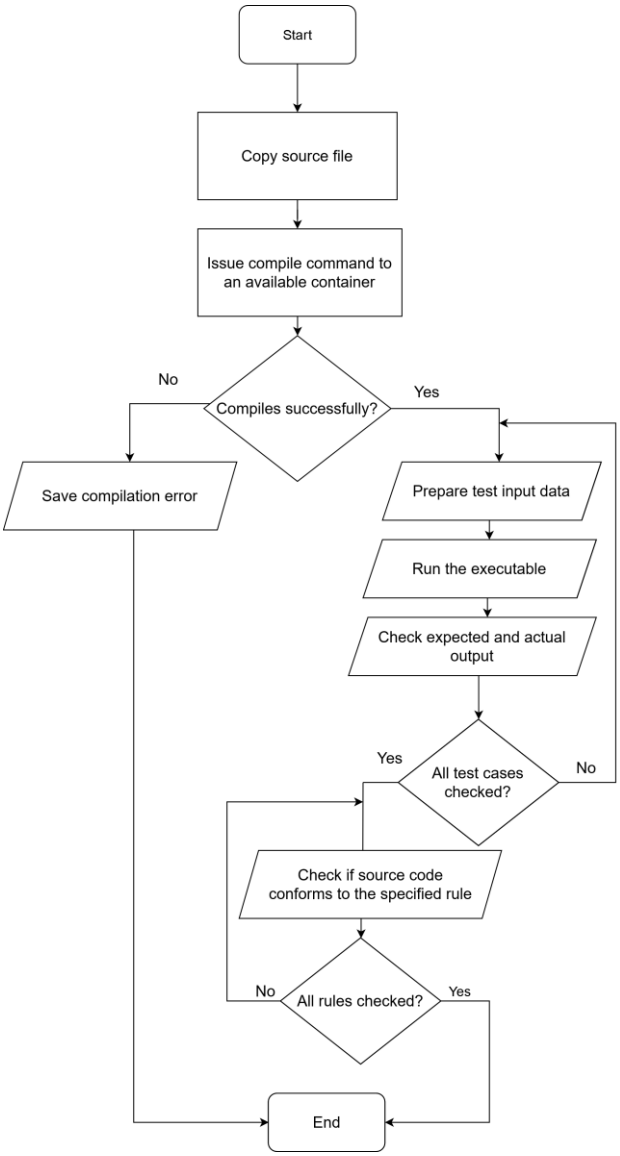


Figure 1. Solution evaluation process

- defining and using ifstream and ofstream file management types.

A. *Example of use for assessing program solutions*

During the exam, each student creates his own program solutions based on the assigned tasks. The following tasks were selected as an example of using the proposed solution:

1. It is necessary to ask the user to enter two whole numbers and print their sum. (2 points)
2. It is necessary for the user to load the word and print "yes" if it is a word, and print "no" otherwise. (3 points)
3. Ask the user to enter the height of an isosceles triangle and to print the corresponding triangle using the star symbol "*". An example of a printout (4 points):

```
*
***
*****
*******
```

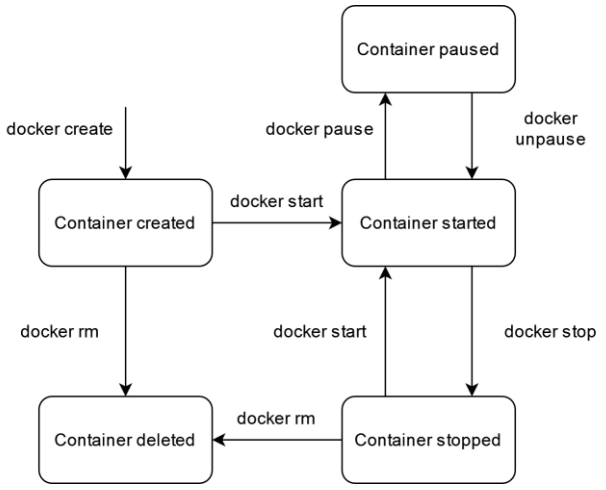


Figure 2. Docker container life cycle

Students use the MS Visual Studio 2022 Integrated Development Environment and select the Empty project form, thus defining the project that will contain the source code of the solution to the first task. In addition to the project, the student also defines a solution that represents a group of projects connected to one exam solution. The next step is to create projects for each remaining task.

Students can check the current solution while solving the exam tasks. The folder created by the project should be compressed and transferred to the evaluation system as such. After finishing the evaluation, an overview of the scoring details is available (Figure 4 for the first task). The source code of the solution is translated, and all test cases and rules are satisfied, and the solution of the task achieves the maximum number of points. Figure 5 shows the scoring details of the second task. Although the solution is translatable and satisfies all the rules for the existence of lexical constructions, the source code contains a logic error that results in only one test case being satisfied. The source code of the solution of the third task cannot be translated and the task achieves 0 points. The program execution procedure and the checking of rules and test cases were not carried out (Figure 6).

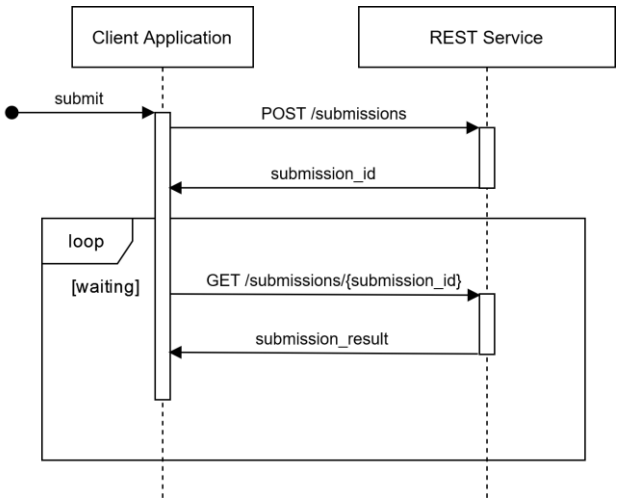


Figure 3. Communication between the client application and the server service



Figure 4. First task evaluation results



Figure 5. Second task evaluation results

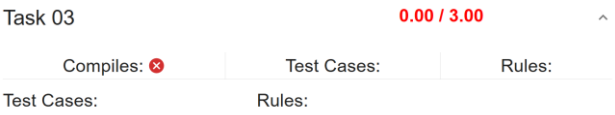


Figure 6. Third task evaluation results

B. *Example of use with a system for grading sql solutions*

Analogous to the process of evaluating C++ program solutions, it is possible to evaluate SQL solutions using the system for evaluating SQL solutions previously developed at the College of Algebra [16]. Parameters marked with the value of the Content-Type header set to multipart/form are submitted to the system via an HTTP request on the path /submissions: *config* with the task settings, and *sql* with the SQL file of the student's solution.

Task settings define rules that indicate the presence of certain lexical constructions of SQL queries, such as keywords SELECT, FROM and ORDER BY, as well as the course of evaluation if a certain rule is not respected. The SQL file contains the solutions of several tasks that are separated within the source code by the comment mark -- LOXTY, where X is the sequence number of the learning outcome, and Y is the sequence number of the task.

After successfully receiving the request, the SQL solution evaluation system performs a static analysis of the code as well as checking whether the SQL query returns the necessary rows defined by the task. A database container is used when executing SQL queries. If the specified container is not started, a command is given to start the container, within which an SQL script is then executed to create a database filled with the initial data needed to test the student's solution.

The JSON response body for the HTTP POST request on the path /submissons contains a universal identifier of

the solution that needs to be sent during each subsequent HTTP GET request on the path /submissions/:id. The response to the specified HTTP request contains the number of points achieved for each task, as well as feedback on the reasons for possibly deducted points.

C. Disadvantages of the proposed system

Naive manual testing of the system revealed the shortcomings of the developed solution. The name of each task within the student's solution must explicitly match the defined task name within the exam. The output data in the standard output stream must explicitly match the expected output data.

An improvement to such an approach would be to implement metrics when checking task names and output data. An example is the Levenshtein distance [8], [17]. When checking the test cases, it would be necessary to define the permissible value of the Levenshtein distance between two strings of characters to avoid deducting points due to random errors. For example, it is possible to define a Levenshtein distance of 1, if one wants to ignore a misspelled letter in the solution.

Static analysis of the presence of lexical constructions is performed using regular expressions, which are not an optimal tool when it comes to performance. Due to their limitations, regular expressions are not the most suitable tool for finding specific lexical constructions [14], [15]. The development of a parser that creates an abstract syntax tree from the given source code, and whose search is more precise than using regular expressions would improve proposed system.

The presented solution has a monolithic architecture, and two microservices should be developed instead: for code execution and for evaluating the solution. With such an approach, the system would enable a larger set of functional requirements for different test tasks. Such an agnostic architecture would more simply implement a system for evaluating software solutions of other programming languages within a common virtualized environment.

IV. Conclusion

The developed grading system allows examiners to automate grading. The key parts of the code and commands are analyzed, and the use of the developed system is described along with the corresponding textual description and graphical representation of the user interface. The proposed architecture enables the development of different types of clients that can call the developed system. After analyzing the examples of the developed system use, shortcomings were observed, and possible system upgrades were suggested.

References

[1] J. Hollingsworth, “Automatic graders for programming classes,” Commun. ACM, vol. 3, no. 10, pp. 528–529, Oct. 1960, doi: 10.1145/367415.367422.

[2] H. Z. Dosilovic and I. Mekterovic, “Robust and Scalable Online Code Execution System,” in 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Sep. 2020, pp. 1627–1632, doi: 10.23919/MIPRO48935.2020.9245310.

[3] I. Mekterovic, L. Brkic, B. Milasinovic, and M. Baranovic, “Building a comprehensive automated programming assessment system,” IEEE Access, vol. 8, pp. 81154–81172, 2020, doi: 10.1109/ACCESS.2020.2990980.

[4] H. Aldriye, A. Alkhalaf, and M. Alkhalaf, “Automated grading systems for programming assignments: A literature review,” Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 3, pp. 215–221, 2019, doi: 10.14569/IJACSA.2019.0100328.

[5] A. Z. Javed, P. A. Strooper, and G. N. Watson, “Automated generation of test cases using model-driven architecture,” Proc. - Int. Conf. Softw. Eng., 2007, doi: 10.1109/AST.2007.2.

[6] R. T. Fielding, “Architectural Styles and the Design of Network-based Software Architectures,” 2013.

[7] D. Merkel, “Docker: lightweight Linux containers for consistent development and deployment,” Linux Journal, vol. 2014, no. 239. p. 2, 2014, Accessed: Jan. 26, 2023. [Online]. Available: https://www.seltzer.com/margo/teaching/CS508.19/papers/merkel14.pdf.

[8] T. Combe, A. Martin, and R. Di Pietro, “To Docker or Not to Docker: A Security Perspective,” IEEE Cloud Comput., vol. 3, no. 5, pp. 54–62, 2016, doi: 10.1109/MCC.2016.100.

[9] R. Zhang, A. M. Zhong, B. Dong, F. Tian, and R. Li, Container-VM-PM Architecture: A Novel Architecture for Docker Container Placement, vol. 10967 LNCS. Springer International Publishing, 2018.

[10] T. Bui, “Analysis of Docker Security,” 2015, [Online]. Available: http://arxiv.org/abs/1501.02967.

[11] A. M. Potdar, D. G. Narayan, S. Kengond, and M. M. Mulla, “Performance Evaluation of Docker Container and Virtual Machine,” Procedia Comput. Sci., vol. 171, no. 2019, pp. 1419–1428, 2020, doi: 10.1016/j.procs.2020.04.152.

[12] P. Gkikopoulos, V. Schiavoni, and J. Spillner, “Analysis and Improvement of Heterogeneous Hardware Support in Docker Images,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2021, vol. 12718 LNCS, pp. 125–142, doi: 10.1007/978-3-030-78198-9_9.

[13] N. Vermeir, “Application Architecture,” Introd. .NET 6, pp. 259–273, 2022, doi: 10.1007/978-1-4842-7319-7_9.

[14] P. Bille and M. Farach-Colton, “Fast and compact regular expression matching,” Theor. Comput. Sci., vol. 409, no. 3, pp. 486–496, Dec. 2008, doi: 10.1016/J.TCS.2008.08.042.

[15] A. Backurs and P. Indyk, “Which Regular Expression Patterns Are Hard to Match?,” Proc. - Annu. IEEE Symp. Found. Comput. Sci. FOCS, vol. 2016-December, pp. 457–466, Dec. 2016, doi: 10.1109/FOCS.2016.56.

[16] M. Fabijanic, G. Dambic, and B. Fulanovic, “A novel system for automatic, configurable and partial assessment of student SQL queries,” 2020 43rd Int. Conv. Information, Commun. Electron. Technol. MIPRO 2020 - Proc., pp. 832–837, 2020, doi: 10.23919/MIPRO48935.2020.9245264.

[17] L. Yujian and L. Bo, “A normalized Levenshtein distance metric,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 6, pp. 1091–1095, 2007, doi: 10.1109/TPAMI.2007.1078.

Digital Literacy of Students at the Faculty of Humanities and Social Sciences Zagreb: An Empirical Study

R. Vrana
Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb, Croatia
rvrana@ffzg.unizg.hr

***Abstract* – The paper presents the results from the empirical research study of students at the Faculty of Humanities and Social Sciences, Zagreb, Croatia (FHSS) about their perception of digital literacy, one of key literacies in the 21st century essential for employment and participation in digital society. The results indicate that the students were acquainted with the term digital literacy for the first time early in their education. Now when they are part of the higher education system, the students recognized the most valuable characteristics, activities, skills, and competencies related to digital literacy. They also indicated the preferred ways of acquisition of new knowledge about digital literacy. Finally, the respondents highly rated the value of digital literacy for their success in the study, employment, and life in digital society.**

Keywords – digital literacy, students, higher education, empoloyment, digital society

I. INTRODUCTION

As future employees in different industries, students like being prepared and like being equipped with knowledge and skills appropriate for the employment within the framework of the Industry 4.0 and beyond. "The exponentially increasing proliferation of new digital technologies has steered the individuals of the digital society to consume more digital content, reproduce it, create new content, share it and provide a critique of it when needed" [1]. Today students in Croatia aim to acquire knowledge and skills that will secure them employment amid rapid emergence and deployment of new technologies necessary both in professional and personal life.

Demonstrating the knowledge and skills of mastering digital tools and services has become a priority in modern developed societies for every individual. The digitalization of services in every area of life has rendered the ability to use digital tools an important skill and keeping oneself updated in the world of digital tools and services has become a priority [2], especially for young people like students entering local and global labor market.

The pandemic of the virus SARS-CoV-2 and the time after the pandemic, demonstrated a need for digitally literate individuals ready to start teaching, learning, and working online immediately, almost without any additional education or training. The current general situation in the world seeks education from individuals different from that of the previous century [3]. Therefore, digital literacy has been given a high priority in education.

In the last two years, the emphasis of scientists and professionals working in the field of education was on many different types of literacies, including digital literacy. During that period, digital literacy has evolved in parallel with rapid technological advancements, so it has become necessary in academic life to prepare students for long-term learning [4]. But, possessing the digital literacy knowledge and skills or merely "knowing" is not enough. Today's education demands focusing on the development of additional skills like soft skills, communication, creative thinking, and adaptability [5]. Still, the focus of this paper will remain only on digital literacy.

II. DIGITAL LITERACY

Digital literacy is one of many literacies and is considered being a part of the umbrella term - information literacy. There is no general consensus on what exactly constitutes the term digital literacy, however, many definitions of that term are available. One of the first definitions of digital literacy originates to Gilster [6]: "the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers". Audrin and Audrin [7] did a systematic literature review on digital literacy definitions and aspects of definitions and found out plentitude of approaches suggesting the development of notions about digital literacy since 1997. Tinmaz, Fanea-Ivanovici and Baber [8] also did a systematic review on digital literacy and one of their conclusions is that digital literacy is an umbrella term for many different technologies seeking case studies which would explore different implementations of this term.

Digital literacy is seen as "a technical skill to operate digital devices, such as computers, smartphones, smartboards, projectors, social media and website or blog design" [9]. Other authors like Kesici [2] suggests that digital literacy is not only an ability to use digital tools and services, but it also includes high-order thinking skills such as analysis, synthesis evaluation, and critical and creative thinking. For other authors, digital literacy is not related exclusively to use of ICT per se. Martin [10] interpreted digital literacy as "the ability to succeed in encounters with the electronic infrastructures and tools that make possible the world of the twenty-first century". Digital literacy "provides an individual with core capabilities to achieve

valued outputs in life. It is a critical enabler of economic transformation as it promotes employment opportunities through the ability to access digital content and online services" [11]. "Young people are encouraged to use the internet, because they are told that digital skills are crucial in order to be successful in the information society" [12]. According to the American Library Association [13], digital literacy is"the ability to use ICT to find, evaluate, create, and communicate information, requiring both cognitive and technical skills". Generally speaking, digital literacy is "about being able to operate in the digital world in its widest sense and this embraces a range of skills that are needed for life" [14]. Another view of digital literacy is content oriented. Heitin [15] suggested that digital literacy "also refers to content creation. That includes writing in digital formats such as email, blogs, and Tweets, as well as creating other forms of media, such as videos and podcasts".

Jones introduced and explained another related term - critical digital literacy - literacy is not solely about how you think but and what you do after reading and writing online (and 'critiquing' what we read) and that your actions are part of wider societal actions and rooted in communities and that it depends on people and contexts [16].

Some of the skills covered by the term digital literacy include: internet search, hypertextual navigation, content evaluation, assembling knowledge, creating content, organising and sharing content, reusing content, filtering and selecting material and self-broadcasting in social networks [8]. Furthermore, digital literacy includes using ICT, processing information, and engaging with media and they all are related to general literacy and numeracy, social and emotional skills, critical thinking, complex problem solving, and the ability to collaborate [17]. International Telecommunication Union [18] in its document "Digital insights 2021" differentiate between three levels of digital skills - basic, intermediate and advanced; "Basic skills refer to foundational skills for performing basic tasks and cover hardware, software and basic online operations. Intermediate skills refer to skills that enable the use of technologies in more meaningful and beneficial ways. They include skills required to perform work-related functions and they expand to account for changes in technology. Advanced digital skills are needed by specialists in ICT professions such as computer programming and network management. They include: skills related to emerging technologies, such as artificial intelligence, big data, cybersecurity, the Internet of Things (and app development), and are acquired through advanced formal education". According to Pavey [14] digital literacy incorporates media literacy; communications and collaboration; career and identity management; ICT literacy; learning skills; digital scholarship; information literacy.

Digital literacy competencies include using, understanding, accessing, managing, collecting, coordinating, estimating, providing information, and evaluating information from ICT-based sources [8].

Digital literacy is also related to digital inquiry in digital environment: connecting ideas to personal interests and a desire to know, asking questions that probe beyond simple fact gathering, investigating answers from multiple perspectives, constructing new understandings, expressing the new ideas through a variety of formats, and reflecting on both the process and product of learning [19].

TABLE 1. CHARACTERISTICS OF DIGITAL ENVIRONMENT [19]

Opportunities	**Challenges**
Amount of Information	
Information available on practically anything that students want to know Multiple formats and languages Levels from novice to expert	Too much information pressures students toward passive acceptance of whatever they find to avoid an unending search Information accepted if found quickly and easily; little in-depth probing
Organization of Information	
Participatory organization (tagging and organizing text for own personal and academic use) Natural language searching Platform of tagging provided by others	Lack of sequence and hierarchy in information (order has nothing to do with time, place, or even synchrony with researcher's main idea) All information presented with equal importance
Type of Information	
Diverse formats (visual, oral, graphic, video, audio) Diverse types, from personal opinion to research	Specific information with little or no contextualization Lack of general overview information
Access to Information	
Multi-tasking Consolidation of devices—one-stop information access	Pressure for speed Widening digital divide Environment favors access over reflection
Interactive Environment	
Collaboration and shared learning Deepened learning through social interchange of ideas High engagement and participation by all, even the shy or reluctant	Individual voice is lost in group dynamic Ability of individuals to post and publish leading to assumption of authority ("If I said it and it's on the Web, then I must be an expert.")

III. RESEARCH

During their education at the Faculty of Humanities and Social Sciences, University of Zagreb (FHSS), the students are educated about ICT as they enroll specific undergraduate or graduate courses. The purpose of the research study in this paper was to get an insight into perceptions about digital literacy by the students at the FHSS. The goals of the research study were: 1.) establishing the most important characteristics and aspects

of digital literacy recognized by the students at the FHSS and 2.) establishing importance of digital literacy for success in their studying, employment, and life. The basic presumptions (hypotheses) were: 1.) students of the recent generations at the FHSS are almost fully aware of the term of digital literacy and its characteristics and aspects; 2.) surrounded by ICT, they are aware that mastering digital literacy has become essential for the success in their studying, employment, and life. A self-administered online questionnaire consisting of 12 closed type questions was implemented using as research tool. The online questionnaire was available from January 13th till January 27th 2023. An invitation to participate in the research was sent on two occasions to the students' mailing list. The total of 105 students participated in this research study. The limitations of the study are that it includes only students at the FHSS who actively participated in the research study and, based on that fact, the results of the research study cannot be generalized on the whole population of the students at the FHSS (more than 7000 students). However, the findings provide an important understanding of digital literacy by the participants in the research study. It should be noted that the questions in the questionnaire did not include gender or year of study because the sample of the research was not stratified and did not include the specific number of students (participants) according to each year of study. Instead, the call for participation aimed at inviting as many students from the FHSS as possible. During an academic year, the students at the FHSS are invited to participate in many research studies and therefore are possibly becoming reluctant to participate in another new research on any topic. For that reason, this research study used non probabilistic convenience sample.

IV. FINDINGS

The following part of the paper presents the results of the empirical research study of the students at the FHSS.

Q1. In what phase of your education did you encounter the term digital literacy for the first time? (N=105)

TABLE 2. ENCOUNTERING THE TERM DIGITAL LITERACY FOR THE FIRST TIME

	N	%
High School	30	28,6
Elementary school	31	29,5
University	28	26,7
Not sure	16	15,2

The answers to the first question indicate that almost two thirds of the students participating in this research encountered the term digital literacy as early as elementary or high school, before they enrolled the FHSS, and approximately one third encountered the term as late as university. The earlier they had encountered the term digital literacy the better insight into it they got. Simultaneously they acquired necessary knowledge and skills as society has become more intensively digital over the last two decades influencing lives of the respondents on many levels.

Q2. Which of the following characteristics do you relate to the term digital literacy? (multiple answers possible) (N=105)

TABLE 3. DIGITAL LITERACY CHARACTERISTICS

	N	%
Management of content in digital environment	87	82,9
Analysis of content in digital environment	84	80
Awareness about content in digital environment	82	78,1
Access to content in digital environment	80	76,2
Identification of content in digital environment	76	72,4
Evaluation of content in digital environment	65	61,9
Synthesis of content in digital environment	50	47,6
Creation of new knowledge	50	47,6
Communication of knowledge to other people	46	43,8
Integration of content in digital environment	35	33,3

The question 2 focused on the list of relevant characteristics of digital literacy found in the selection of literature in the references of this paper. Management, analysis, and awareness (all content related) are the top three choices of the respondents. The students' choice of answers gives us an insight into their view on relevance of the characteristics of digital literacy in the table 3. Their choices have been possibly influenced by the study programs at the FHSS focusing increasingly on the content in digital environment for daily student activities. One of the limitations of the online self-administered questionnaire was inability of researcher to ask additional questions about reasons for making choices like these.

Q3. Which of the following activities do you relate to the term digital literacy? (multiple answers possible) (N=105)

TABLE 4. DIGITAL LITERACY RELATED ACTIVITIES

	N	%
Use of electronic devices	91	86,7
Browsing and searching information on the internet	85	81
Collaboration in online environment	83	79
Presentation creation	82	78,1
Selection of data and information	80	76,2
Text editing	79	75,2
Understanding of personal security online	78	74,3
Fake news detection / recognition	74	70,5
Interpretation of information and digital content	74	70,5
Critical thinking when processing information	72	68,6
Interaction by use digital technology	71	67,6
Uploading of information and digital content on the internet	69	65,7
Sharing of content by used of digital technology	65	61,9
Information storage on digital medium	65	61,9
Selection of software for a task	63	60

Understanding of a device security	62	59
Information processing and management	62	59
Understanding of role of ICT in society	61	58,1
Data management	60	57,1
Creation and use of databases	56	53,3
Ethics in use of ICT	56	53,3
Data evaluation	52	49,5
Technical problem solving	53	50,5
Creation and use of spreadsheets	50	47,6
Backup creation and storing it on different locations	51	48,6
Communication with other people by use of Web 2.0 tools	49	46,7
Creation and editing of sound	48	45,7
Creation and editing of video	48	45,7
Visualizing content in software	43	41
Web page creation	42	40
Collaboration with other people by use of Web 2.0 tools	41	39
Programming	29	27,6

This question offered a long list of activities related to digital literacy and use of ICT in general. Use of electronic devices, browsing and searching information and collaboration online are at the top of the list based on the number of answers given by the respondents while once popular Web 2.0 tools, Web page creation and even programming are at the very bottom of the list. Judging from the results, Web 2.0 tools are no longer attractive i.e. the respondents are willing to communicate and collaborate with other people but are not bound to Web 2.0 tools for doing so. Web 2.0 is now an old paradigm and seems started becoming unattractive to the current generations of students. Nevertheless, the respondents highly valued collaboration in online environment which, in many cases, still presume use of Web 2.0 related software. The answers in the table 4 suggest that the respondents chose more frequently more content oriented answers and less frequently technology oriented answers.

Q4. Which of the following skills do you relate to the term digital literacy? (multiple answers possible) (N=105)

TABLE 5 DIGITAL LITERACY RELATED SKILLS

	N	%
Independent research	83	79
Being acquainted with terms related to ICT and popular digital platforms	82	78,1
Adoption of new technologies	73	69,5
Collaboration in digital environment	71	67,6
Creation of multimedia for presentation on the internet	63	60
Explaining technology to other people	51	48,6

In this question the students were asked about skills related to digital literacy they recognized. They chose independent research as the most highly valued skill related to digital literacy. The students at the FHSS spend a significant amount of time during the day doing research supported by the ever-growing number digital information resources. The FHSS belongs to the scientific area of social sciences and humanities, so it is rather surprising to see the students' strong orientation towards knowledge related to ICT and popular digital platforms and adoption of new technologies instead of collaboration in digital environment which is immanent to younger generations of students at non-technical faculties. Collaboration at the workplace in the digital environment between employers and employees and between employees themselves is one of the prerequisites of the modern age. Being acquainted with the terms related to ICT and popular digital platforms and adoption of new technologies are also important skills because of the rapid changes of ICT suggesting that the students wish to be up to date with the latest developments.

Q5. Which of the following competencies do you relate to the term digital literacy? (multiple answers possible) (N=105)

TABLE 6. DIGITAL LITERACY RELATED COMPETENCIES

	N	%
Acquisition of new knowledge about existing ICT	88	83,8
Reading and understanding of multimedia and hypermedia	80	76,2
Connecting information	78	74,3
Appropriate communication in online environment	78	74,3
Finding and critical thinking about information	75	71,4
Reading of visual symbols	60	57,1
Personal ideas management in online environment	57	54,3
Collaborative work to communicate information	47	44,8
Synthetic thinking – creating new combinations out of existing information	44	41,9
Management of teams by use of ICT	37	35,2

This question aimed at finding out about competencies related to the term digital literacy. While there is no surprise in seeing acquisition of new knowledge at the top of the list of answers, it is surprising to see collaborative work, reading of visual symbols and management of teams positioned lower on the list of the chosen answers, compared to answers in questions 3 and 4. All these lower ranked activities have become dominant during the pandemic of the virus SARS-CoV-2, for practical reasons.

Q6. In what ways do you supplement your knowledge about digital literacy? (multiple answers possible) (N=104)

TABLE 7. SUPPLEMENTING KNOWLEDGE ABOUT DIGITAL LITERACY

	N	%
Independent learning by use of different information resources	94	90,4
At the FHSS	56	53,8
By reading literature on a specific topic	34	32,7

Webinars on the internet	19	18,3
Lectures (non-study related) at the FHSS	8	7,7
Lectures (non-study related) outside the FHSS	7	6,7

Independent research is one of the most important students' activities at universities in general. The internet has only expanded the capacities for this activity by offering different digital information resources and tools available to the students at the FHSS. Independent research as a form of supplementing knowledge about digital literacy is followed by reading literature and by participating in Webinars. The latter have become popular during the time of the SARS-CoV-2 virus pandemic and are still popular. Other type of lectures at the FHSS and outside of it are less represented in the results being less attractive to the students.

Q7. Which of the following answers describe important aspects of digital literacy? (1=least important; 5=most important) (N=105)

TABLE 8. RATING IMPORTANT ASPECTS OF DIGITAL LITERACY

	1	**2**	**3**	**4**	**5**
Selecting a technology	4	9	45	41	6
Searching information on the internet	2	4	10	34	55
Evaluation and selection of information resources on the internet	0	1	10	29	65
Use of a digital technology	3	4	23	47	28
Behavior while using a technology	0	2	17	44	37

In this question the students were asked to rate aspects of digital literacy from the least important to the most important. Evaluation and selection of information accompanied by searching information on the internet were chosen as having the most significant impact on the respondents. These results are similar to those in questions 3 and 6 in which browsing and searching information on the internet and independent learning by use of different information resources were highly rated by the respondents. Three other aspects related to technology in general were also important to the students but were also rated a bit lower in comparison to use of information resources. Still, the technology (understanding of it and use of it) is important to the students, which was confirmed in questions 3, 4 and 5.

Q8. According to your estimation, which of the two ways of acquisition of knowledge do you think is better for acquisition of knowledge and skills about digital literacy? (N=105)

TABLE 9. QUALITY OF WAYS OF ACQUISITION OF KNOWLEDGE AND SKILLS RELATED TO DIGITAL LITERACY

	N	**%**
Independent way of learning (apart from lectures at the FHSS)	53	50,5
Formal way of learning (as a part of lectures at the FHSS)	52	49,5

Quality of independent learning and formal learning were represented almost identically in this question, while in question 6 independent learning as form of supplementing knowledge about digital literacy gained significantly greater popularity among the students in this research. One could ask, if the quality of both independent learning and formal learning was almost equal, why is independent learning so much more popular for the students and what are the shortcomings of formal learning?

Q9. Rate the importance of digital literacy for success at the FHSS (1=not important at all; 5= most important). (N=105)

TABLE 10. RATING IMPORTANCE OF DIGITAL LITERACY FOR SUCCESS AT THE FACULTY

	1	**2**	**3**	**4**	**5**
N	0	3	15	35	52
%	0	3	14	33	50

The students (half of them) participating in this research rated digital literacy most highly for achieving success at the FHSS and four fifths of the students ranked digital literacy highly (4 and 5 on the scale), which is very significant.

Q10. Rate your own level of knowledge and skills of digital literacy (1=no knowledge about digital literacy at all;5=extensive knowledge about digital literacy at all). (N=105)

TABLE 11. RATING LEVEL OF KNOWLEDGE AND SKILLS OF DIGITAL LITERACY

	1	**2**	**3**	**4**	**5**
N	0	2	31	56	16
%	0	2	30	53	15

Almost two thirds of the students in this research rated their knowledge 4 one the scale of 5 which is high and overall ratings of digital literacy in this question are high. This is an estimation made by the students themselves and is not measured precisely by some type of a test. However, it shows an optimistic picture about the students' preparedness for activities related to digital literacy.

Q11. Rate the importance of digital literacy for successful employment (1=completely insignificant;5=completely significant). (N=105)

TABLE 12. RATING THE IMPORTANCE OF DIGITAL LITERACY FOR SUCCESSFUL EMPLOYMENT

	1	**2**	**3**	**4**	**5**
N	1	5	9	43	47
%	1	5	9	41	45

This question offers results similar to the results in question 10, but this time about importance of digital literacy for successful employment of students. The students in this research rated digital literacy highly

important for success in their employment (over four fifths of them by choosing 4 and 5 on the scale).

Q12. Rate the importance of digital literacy for life in digital society (1=not important at all; 5= most important). (N=105)

TABLE 13. RATING THE IMPORTANCE OF DIGITAL LITERACY FOR LIFE IN DIGITAL SOCIETY

	1	**2**	**3**	**4**	**5**
N	1	3	9	24	68
%	1	3	9	23	65

The last question aimed at rating the importance of digital literacy for life in digital society. Based on the results (approximately two thirds of the students chose the highest rank), digital literacy is very important for life in digital society which reflects the students' view on their role in digital society now and in future.

V. CONCLUSION

The importance of digital literacy has increased because of the invention of myriad of electronic devices and internet technologies and because of the growing digitalization of society in general. The research about digital literacy of students at the FHSS achieved both research goals and confirmed both hypotheses. It showed that almost two thirds of the students who participated in this research have encountered the term digital literacy early, in elementary or high school which prepared them for the studying at the university where use of digital technologies intensifies even more. In three consecutive questions, the students chose the most important activities, skills and competencies related to digital literacy. The results also created a picture of what is important to the students at the beginning of 2023 regarding digital literacy. A shift from Web 2.0 and technology itself towards content related activities, skills and competencies can be observed. The students in the research like to supplement their knowledge by themselves and the during the formal education at the FHSS. Finally, the students in the research see digital literacy as highly important for their studying at the FHSS, employment and life in digital society. They rated their knowledge of digital literacy rather highly. Digital literacy is important to all students because it is one of the basic literacies today. It has entered the education system in Croatia at an early age, preparing school students for their education at the university level, employment and finally, for life.

The future research about digital literacy and university students should:

1.) be carried out regularly (annually) to provide more data about trends in digital literacy acquisition at the FHSS

2.) include additional aspects of ICT especially emerging technologies that will be used at the FHSS or carry a potential for used in education and research at the FHSS.

REFERENCES

[1] P. Reddy, K. Chaudhary, B. Sharma and S. Hussein, "Essaying the design, development and validation processes of a new digital literacy scale", Online Information Review, vol. ahead-of-print no. ahead-of-print, 2022. Retrieved from: https://doi.org/10.1108/OIR-10-2021-0532

[2] A. Kesici, "The effect of digital literacy on creative thinking disposition: The mediating role of lifelong learning disposition", Journal of Learning and Teaching in Digital Age, vol. 7 no, 2, pp. 260-273, 2022.

[3] J. Varghese and M.N.M. Musthafa, "Integrating digital literacy skills and technological intelligence in the higher education curriculum of india: A new paradigm", Issues and Ideas in Education, vol. 10 no, 1, pp. 31-38, 2022.

[4] S. Wahjusaputri and T.I. Nastiti, "Digital literacy competency indicator for indonesian high vocational education needs", Journal of Education and Learning (EduLearn), vol. 16 no. 1, pp. 85-91, 2022.

[5] N. Khan, A. Sarwar, T.B. Chen and S. Khan, "Connecting digital literacy in higher education to the 21st century workforce", Knowledge Management & E-Learning, vol. 14 no 1, pp. 46-61, 2022.

[6] Gilster, P. (1997). Digital Literacy. New York, NY: John Wiley & Sons, Inc.

[7] C. Audrin and B. Audrin, "Key factors in digital literacy in learning and education: A systematic literature review using text mining", Education and Information Technologies, vol. 27 no. 6, pp. 7395-7419, 2022. Retrieved from: https://doi.org/10.1007/s10639-021-10832-5

[8] H. Tinmaz, Y. Lee, M. Fanea-Ivanovici and H. Baber, H., "A systematic review on digital literacy", Smart Learning Environments, vol. 9, 2022. Retrieved from: https://doi.org/10.1186/s40561-022-00204-y

[9] F. Fauzan, F. Arifin, M.A. Lubis and F.M. Firdaus, "Lecturer's digital literacy ability in the pandemic", Cypriot Journal of Educational Sciences, vol. 17 no. 4, 2022., pp. 1130-1142.

[10] A. Martin, "A european framework for digital literacy", Nordic Journal of Digital Literacy, vol. 1, pp. 151- 161, 2006.

[11] K. Chetty, L. Qigui, N. Gcora, J. Josie, L. Wenwei and C. Fang, "Bridging the digital divide: measuring digital literacy", Economics, vol. 12, no. 1, 2018, pp. 2018-23, 2018.

[12] S. Vandoninck, L. D'Haenens and V. Donoso, "Digital Literacy of Flemish Youth: How do they handle online content risks?" , vol. 35 no. 4, 2010, pp. 397-416.

[13] What is Digital Literacy? American Library Association Retrieved from: https://alair.ala.org/handle/11213/16260

[14] S. Pavey, "Information Literacy and Digital Literacy. In: C. Roche, B. Band, N. Cavender, L. Chambers, A. Everall, E. Krajewski, et al. (Authors), Creating a School Library with Impact: A Beginner's Guide, 2022, pp. 67-86. Retrieved from: doi:10.29085/9781783305551.008

[15] L. Heitin, "Digital literacy: Forging agreement on a definition: Digital literacy: An evolving definition", Education Week, vol. 36 no. 12, 2016.

[16] R.H. Jones, "Commentary: Critical digital literacies as action, affinity, and affect", TESOL Quarterly, vol. 56 no. 3, pp. 1074-1080, 2022. Retrieved from: https://doi.org/10.1002/tesq.3153

[17] R. Bandura, E.I. Mendez Leal, "The digital literacy imperative", 2022. Retrieved from: https://www.csis.org/analysis/digital-literacy-imperative

[18] Digital Skills Insights 2021, International Telecommunication Union, 2021. Retrieved from: https://academy.itu.int/sites/default/files/media2/file/21-00668_Digital-Skill-Insight-210831_CSD%20Edits%206_Accessible-HD.pdf

[19] B. Stripling, "Teaching students to think in the digital environment: Digital literacy and digital inquiry", School Library Monthly, vol. 26 no. 8, pp. 16-19, 2010.

Effective Educational Ukrainian Practices of the Formation of Media Literacy

Yu. Rudenko*, S. Ahadzhanova*, K. Ahadzhanov-Honsales*,
O. Bieliaieva**, A. Korovai** and O. Semenikhina**
* Sumy National Agrarian University, Sumy, Ukraine
** Sumy State Pedagogical University named after A. S. Makarenko, Sumy, Ukraine
yango641@ukr.net, svitlana.ahadzhanova@snau.edu.ua, karen.ahadzhanov-honsales@snau.edu.ua,
Lenyse4ka-85@ukr.net, a.korovai@fizmatsspu.sumy.ua, e.semenikhina@fizmatsspu.sumy.ua

***Abstract* - The article touches on the problem of forming the younger generation's media literacy. It is shown that in order to solve it, it is necessary to: develop the critical thinking of young people regarding the consumption, verification, and evaluation of media products; to form the ability to resist manipulation, destructive influences, and hostile propaganda; to develop a culture of interaction in the Internet space and the ability to protect one's own information space.**

A workshop developed by a team of authors and the methodology of the practicum are described: the logical sequence of classes (motivation, encouragement to study; teaching practical techniques; checking the results; reflection); forms and methods of learning (quizzes, competitions, training, master classes). It is shown that the following are important for the successful implementation of the workshop: interactivity, game methods, learning techniques for emotional relief, learning by example, and real media products.

The effectiveness of the workshop has been confirmed. The dynamics of achievements have shown significant positive changes in the ability to distinguish facts from judgments and to identify inaccurate information. The number of students who realized the responsibility for spreading unverified information and the importance of creating an adequate own information space also increased.

Keywords - media literacy; Info media literacy; development of critical thinking; media products; professional training; information propaganda; destructive influence.

I. INTRODUCTION

Recently, there is an increased interest in media literacy in the world. As soon as society realized the scale of the impact of information technologies on the psyche, behavior, preferences, and values of citizens, the task of forming media literacy of the population to prevent destructive political, social, and other influences became actual. In the world community, the problem of the formation of media literacy is solved in different ways. Some countries interpret media literacy as the ability to read, interpret, and create informative content and therefore consider the development of critical thinking important. Others believe that media literacy is not only the development of critical thinking but also the formation of behavioral and psychological skills [1].

The global experience of teaching media literacy is of considerable interest. Theoretical aspects of media literacy were considered by N. Prykhodkina [2]. She describes the basic theories of media education: semiotic approach, ideological approach, media as popular arts approach, practical approach, uses and gratifications approach, protectionist approach, ethic approach, critical thinking approach, social and cultural approach, and media-didactic approach. Jin Yang [3] guides by the constructivist or social-cultural approach to learning and knowledge development that knowledge is learned best from participating in discussions and from contributing to the knowledge-generating. His study proposes and discusses the thesis that social media have the potential to facilitate dialogic learning and thus assist the accomplishment of educational goals and objectives. The guide [4] stresses teaching inductively and beginning whenever possible with students' direct experiences with the media. The basic method of media studies is that of a "spiral curriculum", a concept developed by educator Jerome Bruner [5] raise the issue of youth media literacy from a political point of view in retrospect and notes that its relevance decreases over time, which is unacceptable.

The practical aspect of the problem and the development of media literacy among young people were studied in [6]

In the paper [7] authors touch upon the problem of developing a tool that can evaluate the level of media literacy in terms of a person's ability to understand, analyze and evaluate media messages. M.Hameleers [8] analyses media literacy at point-view misinformation and accents that evidence-based misinformation is seen as more accurate than fact-free misinformation, and the combination of news media literacy interventions and fact-checkers is most effective in lowering issue agreement and perceived accuracy of misinformation across countries. In the course of the German language practical classes and in the optional course the following media education methods P.Pokrovska [9] uses the media-biographic method, the method of narrative analysis of media text, and L. Kolberg's method of moral dilemma and proves its effectiveness. I. Waterman [10] presents psychological

and pedagogical features of media literacy. The methodology of teaching media literacy in Ukraine is highlighted in scientific works [11; 12].

In Ukraine, due to Russian military aggression, the problem of the formation of media literacy of the population has become significantly aggravated, since the war is accompanied by information attacks of enemy propaganda, a huge number of fakes, and aggressive, offensive media reports. Especially vulnerable are young people who are the most active users of information content, and at the same time are subject to the most destructive psychological and emotional influences due to the lack of life experience and age-related psychological characteristics [6].

In the conditions of the multiplicity of existing information threats, we consider it expedient to develop the following components of media literacy among young people:

1) critical thinking (be critical of media products, have practical skills to verify them, the ability to distinguish facts from judgments, detect manipulative content, fakes, propaganda);

2) a responsible attitude to the creation, distribution, and consumption of information in the Internet space;

3) the ability to counteract destructive influences, psychological manipulation, propaganda, and disinformation;

4) the ability to ignore "informational" noise.

In the conditions of an educational institution, such development is possible if students pass a special workshop on media literacy, which was developed by individual authors of this paper [1; 13-15]. *The aim* of the paper is to develop and implement effective educational Ukrainian practices of the formation of media literacy. Research objectives: to develop the content and methodology of the workshop on media literacy; to check the effectiveness of the practicum in baccalaureate.

II. Methodology for Workshop on Media Literacy

The media literacy workshop is a cycle of practical classes that are organized logically and consistently and consists of the following stages: creating interest and motivation for learning; practice training; verification of results; reflection.

In the first lesson, the teacher needs to substantiate the role of media literacy, and its significance in war conditions and identify the real level of students' media literacy.

The introductory lesson "Is media literacy important?" was organized to justify the importance of media literacy, and motivate students to study. It is held in the form of a discussion in which media reports are discussed. Students are shown several sample messages and asked to answer the following questions:

– What is the truth value of this message?

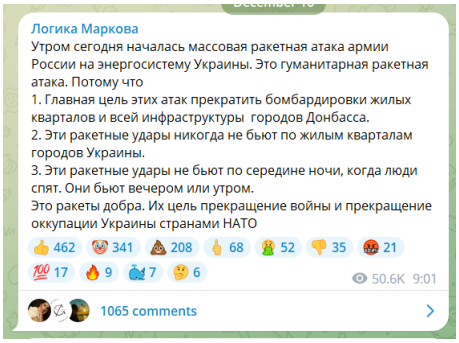
– What techniques are used to attract attention?

– Can people interpret this message in different ways?

– What points of view and values are presented or omitted from this message?

– What answers to these questions will be given by a person who has media literacy skills, and a person who considers them unnecessary?

An example of a media message is shown in Fig. 1. The message refers to “humanitarian rocket fire that does not hit civilian targets and contributes to a ceasefire at the front” (Figure 1a). At the same time, real events and the results of rocket attacks testify to the death of 13 people and the destroyed houses in Kirovograd (Figure 1b).



a)



b)

Figure 1. Distorted and real information about the missile attack

The result of the discussion should be the conclusion: the opinions of a media-literate person and a person who does not yet possess media literacy skills will be different. The habit of checking information and being critical of the content will make it possible to assert that the message was created to justify massive missile strikes on Ukraine, which is beneficial to the Russian occupiers [14]. Basic human values are hushed up and replaced. And although this message justifies unacceptable cruelty, it finds the support of a large number of people, as evidenced by the likes under the post.

Similar examples with false information should show the role of media literacy for a modern person. Students must understand the threats of information wars and the need for critical thinking. Students should have the desire to develop critical thinking and practical skills for checking information for accuracy

The purpose of the lesson "How media literate are you?" is to identify students' media literacy levels. The lesson includes interactive exercises (tests, quizzes, games) that allow students to identify their personal media literacy level. The lesson may include:

Test "What do you know about media literacy" to identify general knowledge about media literacy;

The quiz "Fact or Opinion?" reveals the ability to distinguish between facts and judgments, the level of critical thinking;

Game-task "True or False?" determines the ability of students to check media reports for accuracy;

The exercise "News from my feed" determines the level of responsibility in the consumption and distribution of media in social networks;

Exercise "How attentive are you?" (students are offered instructions in which important information is present among the information noise). The attentiveness when reading the message, and the student's ability to identify manipulations and fakes are checked.

At the end of the lesson, the results are announced. As a rule, they encourage students to increase their level of media literacy in the following lessons.

The "*Fact Check*" lesson is designed to develop the skills to identify false information presented in the form of text, photos, and videos. It is a master class, where methods for identifying false information are demonstrated using examples of various content [15].

An example of the task is "Refutation of a photo fake." The teacher focuses on manipulations in the title, teaches to check the date of publication and photo for authenticity, to find primary sources. Demonstrates how to use Google Lens to find similar photos and check when they were posted. Figure 2 shows a montage of photos from 2014 of Ukrainians honoring those who died on Maidan and a manipulative caption that Ukrainians “kneel before Biden.” Checking the publication date with Google Lens disproves the title.

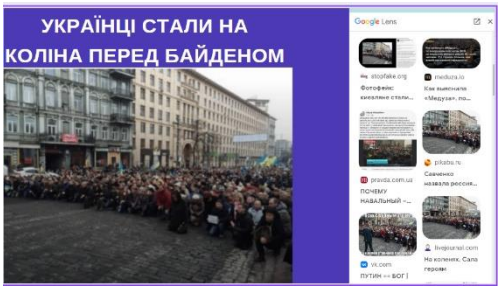


Figure 2. Photomontage as an example of photo-fake

Conclusion to the lesson: it is important to acquire the skills of a media literate person:

– to distinguish manipulation elements in headers;

– to read full information, check links to primary sources and their reputation;

– to analyze the authority and opinions of experts (general phrases or specific facts, judgments, or emotional statements)

– to be able to use programs that allow you to identify a photo and the date of publication (Google Lens service, Tineye browser extension), identify the place of an event (Google Maps, Google Street View services), verify videos (YouTube DataViewer, InVID).

The "*Information Evening*" class follows the goal of forming the skill of a conscious and responsible process of information consumption.

During the lesson, a situation observed in every student's everyday life is recreated. For 2-3 minutes, it is suggested to view your feed in the social network, set favorites, and share information. After a certain time, students' actions are discussed through answers to the questions: What events are covered in the tape? Do you get the impression that you are surrounded only by like-minded people?

In most cases, students answer these questions in the affirmative, which leads to the conclusion that the student is in an "information bubble." The appearance of the information bubble is explained (the operation of algorithms that shape the information space of the user depending on his requests, comments, preferences, and actions with the distribution and removal of the content). The result of such algorithms is a formed news feed that corresponds to the user's worldview and preferences, which may be biased. The negative consequences of falling into an information bubble are discussed: lack of balance of opinions, confidence in one's rightness, omission of really important messages, and falling under the influence of manipulation.

Particular attention is paid to the dissemination of information. It is recommended to review the content and check it in other sources. The importance of disseminating information consciously, not under the pressure of emotions, as well as in compliance with copyright, and ethical and legal norms, is emphasized.

Recommendations are given on how to avoid the information bubble:

– use the browser in incognito mode;

– periodically delete search history and cookies;

– consciously find other points of view;

– read news directly from news sites, and not through links in social networks;

– distinguish and avoid "junk pages" (a distinctive feature of "junk pages" is that they do not have contact details and information about the editorial staff, they write about absolutely everything, and the texts are full of emotions);

– do not participate in tests, games, or flash mobs (most often they are used not only for advertising but also

for fraudulent schemes and collecting personalized information);

– do not interfere in discussions and disputes on unknown sites, with strangers, or on unknown topics (their goal is to provoke people, and raise the rating of a particular post).

Conclusion to the lesson: the consumption of information by a media-literate person is a conscious, balanced, and responsible process. The presence of critical thinking and the ability to control one's behavior in the information space allows one to avoid psychological manipulation and the destructive effects of propaganda.

Lesson "Emotional pause" on how to learn to get rid of excessive emotionality when consuming negative media messages. It is explained to students that any message pursues certain goals. To achieve these goals, the authors are trying to influence the emotions of the consumer, to encourage reckless impulsive actions.

At the lesson, students are offered a message of an overly emotional nature. After reading the message, students should mark this emotion on the wheel of emotions (an adaptive model of identification and strength of emotions developed by American psychologist) [16]

After that, the teacher distracts students to a calm psychological topic (music, humor, nature) for 20-30 seconds and again asks them to indicate the strength of the emotion on the wheel of emotions. As a rule, the results have the following form (the asterisk of emotions has shifted to the left on the scale) (Fig. 3).

As the images show, the first emotion "Anger" is marked with an asterisk almost in the center of the wheel, and after an emotional pause, the mark moved to the edge of the wheel.

Both labeled drawings are shown on the screen so that the training participants can see for themselves the effectiveness of the "emotional pause" technique, in reducing the strength of the emotion or changing the focus of attention. The mechanism of information destructive attacks from the standpoint of the human psyche and effective ways to master oneself are explained.

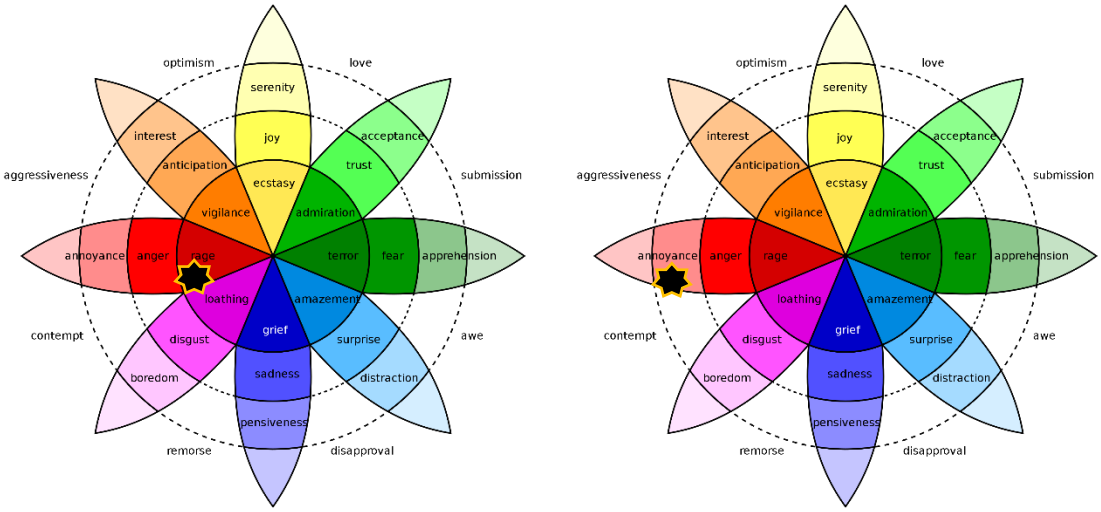


Figure 3. Identification of emotion and its strength before (on the left) and after (on the right) reading media

Conclusion: it is important to develop the ability to consume information in a rational and balanced manner, to be able to take emotional pauses, and to objectively and critically evaluate content.

The "Reflection" class is designed to test knowledge and consolidate the skills acquired during the practicum.

Students are invited to take the test again and write a short review about their impressions of the completed practicum. Feedback should contain detailed answers to the following questions:

Do you consider "media literacy" an important element of the education of a modern person?

Do you know how to verify media information?

Is it important to be conscious and responsible about the dissemination of information in social networks?

Was the internship useful for you?

The described workshop is dynamic and, if necessary, some classes, for example, "Fact-checking", can be increased in volume.

III. Description of the Pedagogical Experiment and Its Results

To check the effectiveness of the workshop, a pedagogical experiment was conducted, which lasted throughout October-November 2022. 64 students (1st year) of Sumy National Agrarian University and Sumy State Pedagogical University participated in the experiment. Four tools for checking educational achievements were used:

1) test "What do you know about media literacy"

2) quiz "Fact or opinion?"

3) task "Reliable-unreliable information"

4) exercise "How attentive are you?"

The tools are the author's development of Yu. Rudenko.

The results at the beginning of the workshop and after its completion are presented in Table 1.

TABLE I. RESULTS OF THE PEDAGOGICAL EXPERIMENT

Task Type	Verification	**Number of questions**	**Number of correct answers (average)**		**Dynamics, number of questions**
			Before the workshop	After the workshop	
test	understanding the essence, components, and requirements of media literacy	7	4	6	2
quiz	The ability to distinguish judgments and facts	20	10	18	8
task	Skills to analyze information (identify primary sources, check content and photos)	5	2	4	2
exercise	Ability to separate information noise		20	54	34

Interpretation of results. The “What do you know about media literacy” test and the “True-false information” exercise revealed a positive performance trend: on average, the number of correct answers increased by 2 for both tools. During the Fact or Judgment quiz, the dynamic was +8 correct answers. 20 students before the lesson and 54 after it correctly completed the exercise to identify attentiveness when reading instructions.

Thus, the test results indicate a low level of understanding of the essence of media literacy, an inability to distinguish between facts and judgments, as well as ignorance of methods for checking information for reliability at the beginning of the experiment and a significant improvement in knowledge and skills at the end of the experiment. Thanks to the last tool, we state that attention, the ability to concentrate, and logical thinking have increased.

We were interested in student reflection. Therefore, at the beginning and the end of the workshop, it was proposed to write detailed answers to questions about the quality of the workshop.

Analysis of responses revealed a positive result. For example, to the question “Do you consider “media literacy” an important element in the education of a modern person?” The number of "Yes" answers doubled. 48 students believe that thanks to the workshop they learned to check media information for authenticity. Before the practicum, 18 students were able to check it (the question “Can you check media information for authenticity?”). 56 students (62.5% more than at the beginning of the experiment) realized their responsibility for the dissemination of information (the question "Is it important to consciously and responsibly treat the dissemination of information in social networks?").

All 64 students found the practice useful. Among the comments were: “Thanks to the workshop, I realized the responsibility for creating and disseminating media information”, “I really liked the workshop, now I feel like a media literate person”, “media literacy is very important, especially in times of war”, “I realized how important it is to think critically, not to succumb to emotions,” etc.

IV. CONCLUSION

Thus, the media of all formats are a powerful factor in both positive and destructive impacts on the psyche and behavior of young people. Therefore, it is important to develop the media literacy of students. In modern conditions, media literacy should include the ability to think critically, the availability of practical skills for checking media messages; a responsible attitude to the creation, distribution, and consumption of information in the Internet space; the ability to counteract destructive influences, psychological manipulation, propaganda, disinformation; cybersecurity skills.

In the conditions of an educational institution, such development is possible if students undergo a special workshop on media literacy. For the workshop, the logical sequence of lessons is important (interest, motivation, practical skills; checking student progress; reflection); use of interactive and game methods (quizzes, competitions, training, master classes); learning from real examples, and real media messages of all kinds.

According to the results of the pedagogical experiment, the improvement of critical media consumption skills was recorded, and the skills to distinguish facts from judgments, and to identify fakes and propaganda were formed. The workshop made it possible to study the techniques of emotional unloading and methods of counteracting psychological manipulation. There are grounds to assert that students have formed a responsible attitude to creating and disseminating information in the Internet space.

Conducting such a workshop requires a teacher to have a high level of media literacy, developed communication skills, and skills in using various digital interactive teaching methods. Therefore, in the perspective of further research, profile training of teachers is planned.

REFERENCES

[1] Yu. Rudenko, and V. Ovcharenko, “Analysis of abroad experience of developing youth media literacy,” Osvita. Innovatyka. Praktyka – Education. Innovation. Practice, Vol. 10, No 6, 2022, pp. 45-51. https://doi.org/10.31110/2616-650X-vol10i6-006.

[2] N. O. Prykhodkina, “Osnovni teoretychni kontseptsii mediaosvity [Basic theoretical concepts of media education],” Pedahohika vyshchoi ta serednoi shkoly – Pedagogy of the high and secondary school, Issue 3, 2016. pp. 264-277. http://nbuv.gov.ua/UJRN/PVSSh_2016_49_28.

[3] Jin Yang, “Social Media's Potential to Facilitate Dialogic Learning”, 2014. https://doi.org/10.4018/978-1-4666-6046-5.ch068.

[4] B. Duncan, “Canada Offers Ten Classroom Approaches to Media Literacy,” https://www.medialit.org/reading-room/canada-offers-ten-classroom-approaches-media-literacy.

[5] R. Wallis, and D. Buckingham, “Media literacy: the UK's undead cultural policy,” International Journal of Cultural Policy, 2016. https://www.tandfonline.com/doi/full/10.1080/10286632.2016.1229314.

[6] Yu. Rudenko, N. Dehtiarova, S. Petrenko, and V. Horokhova, “Development of media literacy of the young generation: a practical aspect of the problems,” Physical and mathematical education, vol. 37, No 5, 2022, pp. 56–63. https://doi.org/10.31110/2413-1571-2022-037-5-008.

[7] A. Arooj, G. Saima, S. Noshina, H. Mian, and A. Muhammad, “Revisiting media literacy measurement: Development and validation of 3-factor media literacy scale,” Journal of Computer Assisted Learning, vol. 38, 2022. https://doi.org/10.1111/jcal.12682.

[8] M. Hameleers, “Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands,” Information, Communication & Society, vol. 25, 2020, pp. 1-17. https://doi.org/10.1080/1369118X.2020.1764603.

[9] P. L. Pokrovska, “Formation of media competence of students of humanitarian faculties during the study of specialized disciplines,” Information Technologies and Learning Tools, vol. 71(3), 2019, pp. 215–226. https://doi.org/10.33407/itlt.v71i3.2558.

[10] I.M. Watermann, “Pflege 4.0. Auswirkung auf die praktische Berufsausbildung von Pflegekräften,” Hausarbeit (Hauptseminar) – Pädagogik – Berufserziehung, Berufsbildung, Weiterbildung – Arbeiten publizieren: Bachelorarbeit, Masterarbeit, Hausarbeit oder Dissertation, München: GRIN Verlag, 2017. https://www.grin.com/document/383437.

[11] D. Verbytskyi, Yu. Sikora, and O. Usata, “Mediahramotnist yak odyn zi skladnykiv protsesu pidhotovky suchasnoho pedahoha profesiinoho navchannia [Media literacy as one of the components of the training process of a modern professional teacher],” Naukovyi visnyk Uzhhorodskoho universytetu. Seriia:«Pedahohika. Sotsialna robota» – Scientific Bulletin of Uzhhorod University. Series: "Pedagogy. Social work", №1 (48), 2021, pp. 69-72.

[12] O.O. Rekun, “Mediahramotnist u systemi vyshchoi osvity [Media literacy in the system of higher education],” Visnyk Natsionalnoho universytetu «Chernihivskyi kolehium» imeni T. H. Shevchenka – Bulletin of the Chernihiv Collegium National University named after T. G. Shevchenko, Vol. 8 (164), 2020, pp. 41-44. https://doi.org/10.5281/zenodo.3905534.

[13] O. Semenog, O. Semenikhina, P. Oleshko, R. Prima, O. Varava, and R. Pykaliuk, “Formation of Media Educational Skills of a Future Teacher in the Professional Training,” Revista Românească pentru Educaţie Multidimensională, Volume 12, Issue 3, 2020, pp. 219-245. https://doi.org/10.18662/rrem/12.3/319.

[14] Yu. Rudenko, N. Dehtiarova, and N. Tutova, “Practical aspects of mind-shaping the development of critical health,” Current Science Nutrition, No 3(3), 2022, https://doi.org/10.52058/2786-6300-2022-3(3)-316-327.

[15] M. Drushlyak, O. Semenog, N. Hrona, N. Ponomarenko, and O. Semenikhina, “Typology of internet resources for the development of youth’s Infomedia literacy,” Information Technologies and Learning Tools, vol. 88, No 2, 2022, pp. 1-22. https://doi.org/10.33407/itlt.v88i2.4786.

[16] Plutchik’s Wheel of Emotions: Exploring the Emotion Wheel. Retrieved from: https://www.6seconds.org/2022/03.

The Formation of Infomedia Literacy of Students in a Media Tournament

M. Yachmenyk*, I. Kharchenko**, O. Semenog*, N. Kyrylenko*, M. Ostroha*, S. Bohoslavskyi* and O. Semenikhina*
* Sumy State Pedagogical University named after A. S. Makarenko, Sumy, Ukraine
** Sumy National Agrarian University, Sumy, Ukraine
jachmenyk.m@gmail.com, kharchenko-inna@ukr.net, olenasemenog@gmail.com, mariia.ostroha@gmail.com, serhii.bohoslavskyi@gmail.com, e.semenikhina@fizmatsspu.sumy.ua

***Abstract* - The article touches upon the problem of forming Infomedia literacy of high school students in a media tournament. The concept of info-media literacy is described. It is substantiated that the modern form of extracurricular work is a media tournament, participation in which can affect the level of info-media literacy of young people. The experience of organizing and conducting a media tournament within the framework of the grant project of the Council for International Research and Exchanges (IREX) is described, which included a number of events (Bar-camp "Online Tools", Workshop "Media literacy tools in the study of literary works", Workshop "Formation of skills of evaluation and interpretation of media information", Training "Language media master" and four checkpoints). The effectiveness of the media tournament has been proven statistically. According to the results of the selection of teams for the Media Tournament, it was revealed that Ukrainian schools do not have a policy of conscious formation of information literacy of students.**

Keywords* - *Infomedia literacy; media literacy; media education; digital skills; media tournament; high school students.

I. INTRODUCTION

The Concept for the introduction of media and information education in Ukraine emphasizes the development of media education as a fundamental component of the state's information security [1]. After a full-scale Russian invasion, Ukrainian society needs citizens with a strong immunity to manipulation and propaganda. It is extremely important to form critical thinking in young people to combat disinformation. Therefore, Infomedia literacy is becoming a relevant and important personal characteristic of each member of Ukrainian society.

The problem of the formation of information and then media-information literacy was actualized at the beginning of the 21st century as a response to the increased information influence of media on the consciousness of ordinary citizens. The analysis and systematization of the documents show the attention of UNESCO [2; 3] and the European Commission [4] to such issues as the need to include media education in the educational process of secondary schools, the influence of the media text on the students' thinking, the influence of media education and media literacy on the formation of the outlook of members of society.

Scientists also analyze the problems of the formation of information literacy in the context of the development of the information society. In particular, the paper presents a review of scientists' research on approaches to the interpretation of information literacy [5].

The growth of literacy in the world is generally noted in the report of The UNESCO Institute for Statistics [6]. In the research work [7] media education in the USA is analyzed.

A significant amount of research on media literacy and information literacy shows the importance of this problem for the world in general and Ukraine in particular. Practice shows that in the conditions of distance learning, young people began to spend more time online, and there were risks of information impact (both positive and negative). Therefore, it became important for us to: (1) examine the available results for Infomedia literacy and its formation of it; (2) find ways to develop Infomedia literacy of young people outside educational institutions; (3) check their effectiveness.

II. INFO-MEDIA LITERACY IN THE RESEARCH RESULTS

The modern strategy UNESCO describes "Media information literacy" as the combined set of knowledge and skills necessary today for the life and work of the individual in modern society [8]:

– media literacy. It contributes to the understanding of the role and functions of media in a democratic society and the conditions under which the media function and influence democratic processes. It involves a critical understanding and analysis of content by the individual and the presence of digital skills, necessary to create a high-quality media product for the user;

– information literacy. It consists of the ability to find the necessary information, determine its reliability, systematize/summarize it, critically evaluate it, realize its significance, provide its own and informational needs. It assumes the availability of digital skills in information processing and its dissemination.

Given the information challenges faced by Ukrainian society in cyberspace, we see it is important to explore the Infomedia literacy of Ukrainian youth. In the project "Learn and Distinguish: Infomedia Literacy" [9], Infomedia literacy is considered as a combination of eight components (Table. 1).

TABLE I. COMPONENTS OF INFOMEDIA LITERACY

Cipher	Name	Characteristic
K1	*Media literacy*	Understanding how media, owners, and editorial policy work
K2	*Critical thinking*	Ability to ask questions
K3	*Social tolerance*	Ability to identify and counteract hate speech
K4	*Fact-checking*	Ability to identify manifestations of propaganda
K5	*Information literacy*	Ability to effectively search for information
K6	*Digital security*	Understanding the digital footprint, cyber-bullying
K7	*Visual literacy*	Ability to analyze photos, video content
K8	*Creativity*	Ability to implement ideas through online tools

Source: Own work

The formation of some of these components is described in [10] The method of integrating Infomedia literacy with various school subjects is presented in [11]. Features of language education and Infomedia education are described in the article [12]. In [13], the authors consider the typology of Internet resources for the development of Infomedia literacy of young people. Article [14] and [15] presents the features of the formation of media educational skills among teachers and the preparation of teachers for the formation of such skills in students. The publication [16] touches on the formation of skills to resist informational influences.

At the same time, the analysis of Ukrainian scientific sources showed that insufficient attention is paid to the formation of information and media literacy of young people in extracurricular activities, one of the modern forms of which is a media tournament. Therefore, the purpose of the study is to substantiate the positive influence of the media tournament on the formation of students of Infomedia literacy.

III. MEDIA TOURNAMENT AND FEATURES OF ITS ORGANIZATION

One of the innovative technologies for the formation of Infomedia literacy in the conditions of extracurricular activities is the media tournament. The tournament, as a form of work with gifted youth, is a team competition in which the priority is not a person, but a collective method of solving problems [17]. A media tournament is a team competition that requires, in addition, solving the tasks (case studies), as well as a creative approach to creating, presenting, and sharing media products using digital tools.

We organized and held the All-Ukrainian tournament on media literacy of school teams of border regions (hereinafter - the Media Tournament; border regions are regions on the border of Ukraine with the Russian Federation and Belarus) [13]. We describe the features of its implementation.

Teams of 10-12 students in grades 6-10 can participate in the Media Tournament. The composition of the team cannot change during the tournament. The team is headed by the captain, who is its official representative. The team is accompanied by a school coordinator (teacher, administration representative, and media tournament mentor).

The media tournament takes place in several stages.

Stage I – selection of teams. Teams submit applications and motivation letters. Incoming monitoring takes place with a selection block of questionnaires, the results of which determine the participants of the second stage. Five teams go to the second round.

Stage II – training-competition of teams. This stage combines four pieces of training, and four competitions (discussion meetings) of participants to solve tournament problems. Teams have to prepare a creative team presentation.

Stage III – final. It includes the presentation of team creative media developments.

In general, the tournament involves the participation of teams, each of which alternately plays one of three roles – "speaker", "opponent", or “reviewer". For the Media Tournament, we added the roles of "case author" and "observer", since the ability to create media text and the ability to work with the information received is important for the development of Infomedia literacy.

Each of the roles involves group training using the brainstorming method:

1) the team "author" creates a case on the subject of previously studied material;

2) the team "speaker" provides an argumentative statement of the team's position on the case task (case from the team "author");

3) the team "opponent" provides a reasoned statement of the counter-position, and analyzes the shortcomings of the team's "speaker" performance;

4) the team "reviewer" provides an express analysis of the positions of the speaker and the opponent, determines the advantages and disadvantages of the theses expressed by them, summarizes the results of defending certain positions;

5) the team "observer" summarizes the information received and provides a preliminary assessment to all teams.

The tournament table is formed (Table 2).

TABLE II. FEATURES OF EVALUATION

Role	Author	Speaker	Opponent	Reviewer	Observer	Total
Maximum points	5	4	3	2	1	15

During the competition, the teams work independently, accompanied by a teacher-coordinator and mentor of the tournament. The competition is subject to regulations (Table 3).

TABLE III. REGULATIONS FOR SOLVING THE PROBLEM

Types of actions	Time
The author offers the Speaker a task, the Speaker accepts or rejects the task	2 min.
Preparing a response	2 min.
Report	7 min
Clarifying the Opponent's questions to the Speaker and Answers	2 min
Preparing for the process of an opponent	2 min
Opponent's speech	5 min
Speaker's response to the Opponent's comments and the discussion between them	3 min
Clarifying questions of the Reviewer to the Opponent and the Speaker and the answers	2 min
Preparation for peer review	2 min
Reviewer's speech	3 min
Answers of the Speaker and Opponent to the reviewer's comments and discussion between them	5 min
General team discussion	5 min
Clarifying questions of the presenter to the teams	3 min.
Concluding Words by the Reviewer, Opponent, and Speaker	2 min
Grading by the Observer team. Performances	5 min

If the Team did not find the answer, incorrectly solved the problem, or refuses to perform, then the team loses the point.

The winner of the Tournament is the team that scored the most points.

IV. ALL-UKRAINIAN TOURNAMENT ON MEDIA LITERACY OF SCHOOL TEAMS OF BORDER REGIONS

The media tournament aims to develop student's skills of critical perception of information necessary to detect disinformation, propaganda, and other manipulations. The initiator of the All-Ukrainian media tournament for high school students was the Laboratory of Infomedia Literacy of the Sumy State Pedagogical University named after A.S. Makarenko [18]. Students of 9-11 grades of educational institutions of Sumy, Kharkiv, Zaporizhia, Khmelnytsky, and Cherkasy regions took part in the Media Tournament.

Total 63 participants (students, teachers, tutors).

Student teams wrote a motivation letter in the first round and passed an introductory survey.

The second round is implemented in the form of four training.

Bar-camp "Online Tools" (protection of personal data, mastering the principles of media security on social networks, awareness of the digital footprint and its consequences). Bar camp's coach proposed to determine: what relates to "personal" and "non-personal" data and why. There was an acquaintance with computer programs for verification of images and videos, and interactive developing online games were offered, which are posted on the website of the School of Responsible Parenthood "Media Puzzles" [19] ("Top 5 books of your family", "Top 5 useful Internet resources", WebQuest).

Workshop "Media literacy tools in the study of literary works" (fact-checking, analysis of media texts and their critical assessment, value analysis of texts, etc.). The event was based on the integration of Infomedia literacy into the school subject – Ukrainian literature. The teams visited a virtual excursion at the museum estate of Ukrainian writer Ivan Kotliarevsky in Poltava and identified fakes and truthful reports on examples of facts from Ukrainian folk art. The teams were asked to conduct a mini-study of the pseudonyms of Ukrainian artists (Lesia Ukrainka, Panas Myrnyi, Ostap Vyshnia, Marko Vovchok, and Oleksandr Oles).

Workshop "Formation of skills of evaluation and interpretation of media information" (recognition of manipulations on examples of historical texts). The students worked with photo and video materials, archival records, and Google maps. They analyzed images and historical documents and determined their veracity. Also, participants traveled through the virtual Middle Ages, got acquainted with 3D-reconstructions of a medieval city, and had the opportunity to virtually visit the modern cities of Europe, which were made during the Middle Ages.

Training "Language media master" (work with dictionaries as the ability to work with information and the ability to ethically tolerant communication). The participants of the tournament discussed why it is necessary to use dictionaries, critically perceive information, and pay attention to headlines and sources in the process of reading messages. The trainer especially accented the teams' attention the "emotionality" of news content.

Also, the second round included four checkpoints - competitions for solving cases on the topics of "Safe media surfing", "Interregional literary media trip", "Historical anti-fake quest", and "Consumption of quality information through the skills of its creation". The participants of the tournament presented the author's tasks related to the definition of fakes, the detection of bots, and the refutation of false information.

In the final of the media tournament, the teams presented their media developments – video presentations and electronic newspapers (Fig. 1) for the development of critical thinking and info-media skills (skills of tolerant network communication, detection of fake materials, safe stay on the Internet, creation of high-quality media texts, etc.).



Figure 1. The first page of the Team's newspaper

The meeting of the commandant at the Media Tournament is presented in figure 2 (Fig. 2).



Figure 2. Meeting of the commandant at the Media Tournament

Another team's work – The interactive guide "Safe media surfing of teenagers" – is shown in figure (Fig. 3)

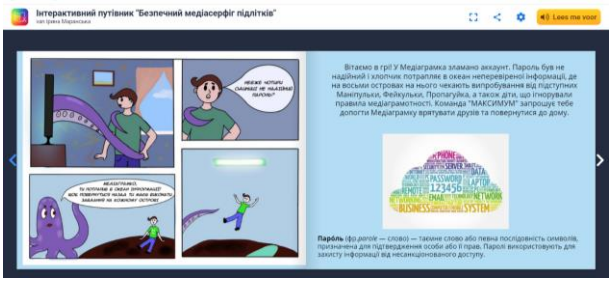


Figure 3. Interactive guide "Safe media surfing of teenagers"

V. Analysis of the Media Tournament Results

To confirm the effectiveness of the Media Tournament, we initiated two surveys (at the beginning of the tournament and after it). The participants of the tournament answered the following questions.

1. Would you check the information you learned from the internet site about the postponement of the celebration from September 1 to August 31?

2. Do you check the information obtained from social networks and media?

3. Do you think that a bot is a person who behaves destructively on the Internet (insults, humiliates, etc.)?

4. Is the position "I rarely change the passwords of access to my social media accounts" correct?

5. Do you take comic tests on the Internet (for example, what kind of cartoon character am I?)

6. Is it correct to assume that the algorithm of the news feed adapts to the consumer, and therefore it is a manipulation?

7. Is it correct to believe: if a person has many followers, then he knows the questions that he talks about, and therefore he can be trusted?

Each positive answer to questions 1, 2, and 6 and a negative answer to questions 3, 4, 5, and 7 was estimated at 1 point.

The total number of students is 50 (5 teams, 10 students in each team). From the set of answers, 15 results are arbitrarily selected, which are statistically processed according to the criterion of signs (Table 4).

TABLE IV. Survey Results

Respondent's cipher	First poll	Second poll	Difference
1	3	4	1
2	4	3	-1
3	2	5	3
4	2	5	3
5	4	3	-1
6	4	5	1
7	4	6	2
8	4	6	2
9	3	3	0
10	3	3	0
11	2	3	1
12	3	5	2
13	3	5	2
14	5	4	-1
15	3	3	0

After the calculations, we recorded general changes in the results of 12 respondents. Positive changes – in 10 respondents.

Hypotheses H0 (the media tournament does not affect the development of information media literacy of students) and Ha (media tournament affects the development of information literacy of students) were built.

The empirical value of statistics is T = 10. The segment of acceptance of H0 is [20; 21] for the significance level of 0.05

According to the rule of acceptance of the hypothesis, the data indicates the favor of an alternative hypothesis. Since the empirical value of statistics has gone beyond the critical interval on the right, there is reason to consider such an impact positive.

VI. Comments on the Organization of the Media Tournament

Dieter Baakke's approaches to media education of young people became important for our project [22]. He considers four aspects of its formation.

1. Criticism of the media. The acquisition of competence to be critical of the content of information while applying an analytical approach to social processes. Reflexive correlation of the ideas of the analyzed text with previously acquired experience, and personal attitude to the content. Ethical dissemination of the information, social responsibility for the widespread content.

2. Media training. Acquisition of knowledge about the structure of the media system, and features of the media. Knowledge of informational (knowledge of processes and structures, for example, how journalists work) and technology (knowledge of work process or technical processing) training.

3. Use of media. To disseminate information from different types of media, it is necessary to learn the basics of their activities in an interactive mode.

4. Development of a media product. It is important to be able to create a quality media product, so you use an innovative and creative approach, changing and developing the media system.

Special preparation for the tournament was not provided. Schools did not introduce Infomedia literacy courses into the educational process. However, the media tournament was attended by teams from those schools that provided a certain type of activity in working with the media. We agree with the words of L. Naidonova: "For a child to correctly consume media, it is necessary to put it in the position of the creator" [23].

Therefore, in our opinion, the creation of the school newspaper "D. A. R-media" in the communal school "Giftedness" contributed to the education of such a media consumer who effectively satisfies his interests using the media, critically analyzes and evaluates media messages, has the skills of collective activity and dialogue communication, actively defends life position through media creation. A school newspaper is a set of news and articles that have undergone editorial processing and are intended for reading by schoolchildren. Its purpose is to provide relevant, interesting, useful information for the student, at a convenient time, in a convenient place. Together with the children, teachers determined the target audience (the school publication focuses specifically on students, one group of whom participates in the active creation, and the second – in the "consumption" of information, so it is important not to make a publication with focuses on parents, teachers, the director). The organization of the production of the school newspaper required the distribution of responsibilities among students of school self-government according to creative abilities and preferences. The articles' authors were schoolchildren who belong to the target group which will read the newspaper. The immersion in conditions that are close to real, and using journalistic methods and techniques for creating print media, students quickly determined the concept and purpose, the composition of the editorial board, the content model, features of filling sections, distribution, and promotion. The created school newspaper, as the practice has shown, solved several tasks: (1) creating a school community (students, teachers, parents); (2) moral and ethical education of the student audience, the formation of a worldview and the education of media taste; (3) formation of students' communication skills, ability to work with information; (4) the formation and development of students' media literacy; informing the community about events in the life of the school.

We also see the prospect of our study in the context of the program "Education 4.0: Ukrainian sunrise" [24], which supports the development of Infomedia literacy.

As a collective form of work, the media tournament develops in high school students such soft skills as critical and analytical thinking, leadership, teamwork, time management, and the ability to form and defend their opinions, while maintaining an atmosphere of mutual support and cooperation, the need to think quickly during the game, defending the results of their developments and commenting on the opponent's answers based on analysis, applying arguments and assertive communication.

VII. CONCLUSION

The media tournament showed that it has a positive impact on the development of information literacy of young people, which is characterized by media literacy, critical thinking, social tolerance, ability to fact-check, information literacy, digital security, visual literacy, and creativity.

The ability to analyze media text develops with the help of discussions (the teaching method) and of digital tools (working with photo and video materials, archival records, Google maps, and historical documents contribute to the development of critical thinking and critical assessment skills.

At the same time, according to the results of the Media Tournament, it was revealed that there is no purposeful policy for forming information media literacy of students in Ukrainian schools. However, most students are aware of the importance of acquiring critical thinking skills, verifying the reliability of the information, counteracting cyberbullying, and creating and protecting their own information space in social networks.

According to the results of the selection of teams for the Media Tournament, it was revealed that schools do not have a policy of conscious formation of information literacy of students. At the same time, the students themselves are aware of the importance of acquiring critical thinking skills, and therefore they would like to continue to participate in such activities. This encourages us to do further research.

The prospect of further scientific research we see in the search and classification of other forms of extracurricular work on the formation of Infomedia literacy in students.

REFERENCES

[1] The concept of introducing media education in Ukraine (new edition). Retrieved from: https://ms.detector.media/mediaosvita/post/16501/2016-04-27-kontseptsiya-vprovadzhennya-mediaosviti-v-ukraini-nova-redaktsiya/.

[2] C. Robinson, “Aspects of literacy assessment: topics and issues from the UNESCO Expert Meeting,” UNESCO Expert Meeting on Aspects of Literacy Assessment, Paris, 2003. Retrieved from: http://unesdoc.unesco.org/images/0014/001401/140125eo.pdf.

[3] The participants in the High Level Colloquium on Information, "The Alexandria Proclamation on Information Literacy and Lifelong Learning," UNESCO, IFLA, & NFIL, 2005. Retrieved from: https://milunesco.unaoc.org/mil-resources/the-alexandria-proclamation-on-information-literacy-and-lifelong-learning-2005/.

[4] European Commission, "European Report on Quality Indicators of Lifelong Learning: Fifteen Quality Indicators Report Based on the Work of the Working group on Quality Indicators," Brussels: EC Directorate-General for Education and Training, 2002. Retrieved from: http://www.aic.lv/bolona/Bologna/contrib/EU/report_qual%20LLL.pdf.

[5] S. Garner, "High-Level Colloquium on Information Literacy and Lifelong Learning", Report of a meeting sponsored by the United Nations Education, Scientific, and Cultural Organization (UNESCO), National Forum on Information Literacy (NFIL) and the International Federation of Library Associations and Institutions (IFLA), Bibliotheca Alexandrina, Alexandria, Egypt November 6-9, 2005. Retrieved from: http://www.ifla.org/publications/high-level-colloquium-on-information-literacy-and-lifelong-learning.

[6] F. Horton, "UNESCO Information for All Programme Understanding Information Literacy: A Primer," Edited by The Information Society Division, Communication and Information Sector, Paris, Unesco. 2007. Retrieved from: https://unesdoc.unesco.org/ark:/48223/pf0000157020.

[7] D. Considine, "Media Education in United States of America," Educating for Media and the Digital Age. Country Reports. Austrian Federal Ministry of Education and Cultural Affairs & UNESCO, 1999.

[8] Media and Information Literacy, Retrieved from: https://www.unesco.org/en/media-information-literacy.

[9] Learn and Distinguish: Infomedia Literacy. Retrieved from: https://www.aup.com.ua/onlayn-prezentaciya-mediagramotnis/.

[10] Infomedia Literacy Advisor for Students and Teachers of Modern Educational Institutions: Workshop, Compilers: S.M. Gergul et al. Chernihiv: Desna Polygraph, 2021.

[11] Methods of integration in the project "Learn and distinguish: info-media literacy": a collection of materials / Ed.: V.F. Yvesanov (chairman ed.). Kyiv: Academy of Ukrainian Press, IREX, Free Press Center, 2022.

[12] O. Kucheruk, "Integration of infomedia literacy into the space of language education," Ukrainian language and literature at school, No3, 2020, pp. 9-12.

[13] M.G. Drushlyak, O.M. Semenog, N.V. Grona, N.P. Ponomarenko, and O.V. Semenikhina, "Typology of internet resources for the development of infomedia literacy of youth," Information Technology and Teaching Aids, Vol. 88, No. 2, 2022.

[14] M.M. Yachmennyk, "Preparation of the future teacher of the Ukrainian language and literature: integration of the IREX project "Learn and distinguish: information and media literacy"," Physical and mathematical education, Issue 3(25), Part 2, 2020, pp. 113-118.

[15] O. Semenog, O. Semenikhina, P. Oleshko, R. Prima, O. Varava, and R. Pykaliuk, "Formation of Media Educational Skills of a Future Teacher in the Professional Training," Revista Românească pentru Educație Multidimensională. Volume 12, Issue 3, 2020, pp. 219-245. https://doi.org/10.18662/rrem/12.3/319.

[16] Yu. Rudenko, M. Drushlyak, O. Naboka, V. Proshkin, and O. Semenikhina, "Special Course On Information Hygiene As A Tool For Developing Youth's Ability To Resist Informational Influences," E-learning in the Transformation of Education in Digital Society Scientific Editor Eugenia Smyrnova-Trybulska "E-learning", 14, Katowice–Cieszyn, 2022, pp. 268–287. https://doi.org/10.34916/el.2022.14.20.

[17] Yu.V. Grebenyuk, and O.M. Zaritsky, "Tournament as a game," H.: View. group "Basis", 2010. 176 p. (B-ka zhurn. "Physics in schools of Ukraine"; Vol. 3 (75)).

[18] Laboratory of information and media literacy of Sumy State Pedagogical University named after A.S. Makarenko. Retrieved from: https://mediaspaceteacher.blogspot.com/p/blog-page_23.html.

[19] School of Responsible Parenthood "Media Puzzles". Retrieved from: https://sites.google.com/view/sspu-mediapazzl/.

[20] Infomedia Literacy Advisor for Students and Teachers of Modern Educational Institutions: Workshop, Compilers: S.M. Gergul et al. Chernihiv: Desna Polygraph, 2021.

[21] V. Petruk, Yu. Rudenko, A. Yurchenko, I. Kharchenko, S. Kharchenko, and O. Semenikhina, "Analysis of the Results of the Pedagogical Experiment on the Integrated Analysis of the Average and Dispersions," International Journal of Modern Education and Computer Science (IJMECS), vol. 14, No 6, pp. 25-34, 2022. https://doi.org/10.5815/ijmecs.2022.06.03.

[22] D. Baacke, "Medienpädagogik (Nachdr.)," Max Niemeyer Verlag, 2007, pp. 97–98.

[23] L. A. Naidyonova, "Mediapsykholohiia: osnovy refleksyvnoho pidkhodu [Media psychology: the basics of a reflexive approach]," National Academy of Pedagogical Sciences of Ukraine, Institute of Social and Political Psychology. Kirovohrad: Imex-LTD, 2015.

[24] Prohrama velykoi transformatsii. Osvita 4.0: ukrainskyi svitanok [Program of great transformation. Education 4.0: Ukrainian dawn]. Retrieved from: https://mon.gov.ua/storage/app/media/news/2022/12/10/Osvita-4.0.ukrayinskyy.svitanok.pdf.

Digital Technologies in Teaching Physics: An Analysis of Existing Practices

A. Yurchenko*, Yu. Khvorostina*, V. Shamonia*, M. Soroka* and O. Semenikhina*
* Sumy State Pedagogical University named after A. S. Makarenko, Sumy, Ukraine
a.yurchenko@fizmatsspu.sumy.ua, y-y-y@fizmatsspu.sumy.ua, shamonawg@gmail.com, ms.mikro.1@gmail.com, e.semenikhina@fizmatsspu.sumy.ua

***Abstract* - The article raises the problem of using digital technologies (DT) in teaching physics. To identify scientific and methodological results on the use of DT in teaching physics, an analysis of publications contained in the Scopus / Web of Science databases and scientific and methodological publications of Ukraine is presented. The following directions have been recorded: the use of various digital tools as tools for the study of physical processes; the use of applications to solve physical problems, the use of virtual and digital physical laboratories, the use of virtual and augmented reality to teach physics, and the use of specialized environments to model physical processes. Separately, virtual physical laboratories are noted – computer environments that allow you to simulate the behavior of real-world objects in conditions that are close to real ones. The PhET resource, which contains online laboratories to support the study of various topics of the physics course, is briefly described. Based on the results of the survey of teachers, Ukraine's practices of using digital technologies in teaching physics are summarized. In particular, there revealed a certain teacher bias regarding using virtual physical laboratories to teach physics.**

Keywords* - *teaching physics; digital technologies in teaching physics; virtual physics laboratories; computer simulations; physics teacher.

I. INTRODUCTION

The development of IT has influenced the educational sector. Today there are a large number of subject-oriented software (mathematics, physics, biology, etc.) designed to facilitate the perception of various concepts, phenomena, and processes. Among such tools, a separate group includes tools to support the teaching of physics, which simplify calculations, record data and build dependencies, simulate various physical processes, etc. A special group is virtual physical laboratories – computer environments that allow you to simulate the behavior of real-world objects in conditions that are close to reality. In such laboratories, the processes of measuring and processing data are undertaken by the computer, and the user's task is only to correctly configure the research parameters and see the existing dependencies.

Teachers in Ukraine, due to the lack of funding for schools with equipment for conventional physical experiments, are looking for alternative ways to teach physic, including the use of digital technologies and means. Methods of their use are not well-established. The degree of their impact on learning outcomes is not unambiguous, so it is advisable to see an existing practice for their use.

The purpose of the study is to characterize the state of implementation of digital technologies and tools in teaching physics, which are presented in scientific publications, and to identify the opinions of physics teachers and students, future physics teachers, on the possibility of using digital technologies in school.

The research methodology provided for analyzing and systematizing scientific publications to identify educational trends in the use of DT in teaching physics, as well as conducting surveys and studying teachers' opinions through conversation. Physics teachers and university students took part in the survey (Sumy region, Ukraine). The total number of respondents is 106 people.

II. ANALYSIS OF WORLD SCIENTIFIC AND METHODOLOGICAL RESULTS OF THE USE OF DIGITAL TECHNOLOGIES IN TEACHING PHYSICS

To identify scientific and methodological results on the use of digital technologies in teaching physics, we analyzed scientific articles contained in the Scopus/Web of Science databases for the last 5 years and devoted them to the use of information technology in teaching physics.

We found trends as follows.

A. The use of digital tools as tools for the study of physical processes

Thus, the work [1] presents the experience of using a high-speed video camera during demonstration experiments in physics. The authors provide examples of experiments used in introductory physics courses that may be relevant in the analysis of the slow reproduction of the filmed material. Article [2] dedicated to using a smartphone as a portable laboratory to perform a variety of physical experiments. In particular, several experiments are described to determine the acceleration of gravity *g*.

B. Using specialized applications to solve physical problems

The resource [3] analyzes the "Physics Master" (https://play.google.com/store/apps/details?id=co.alexis.gbbxs) application, developed by Education Alexis Media, an online coaching institute in India. The application includes a variety of physical quantities dependences and

can serve as an assistant in calculating speed, acceleration, and other quantities. Physics Master has the function of a physical calculator, which helps calculate tasks of varying complexity and shows the steps to complete tasks. This mobile application has many topics for studying and generalizing the theory: fundamental quantities, derived quantities, vector, and scalar quantities, measurements, and errors, basics of kinematics, average and instantaneous speed, medium and instantaneous acceleration, uniformly accelerated motion, laws of dynamics, etc. The application also contains online physics courses that will help online solve physical tasks. The application contains recorded video lectures by practicing physicists, typical thematic tasks, online tests, discussion groups, etc. The application "PhysicsMaster – Physics Calc" (https://play.google.com/store/apps/details?id=com.terracciano.physics_master), which is presented in the article [4], allows you to calculate already known values, and organize quizzes that make it possible to study physics in a game form.

C. Use of virtual and digital physical laboratories

The article [5] describes the development of various types of virtual laboratories in the field of physics over the past 30 years. The study aims to evaluate the virtual laboratories used in the teaching of physics, in terms of the purpose of use and teaching methods. The authors compare the features of the use of various physical laboratories, the convenience of their interfaces, the tools laid, and the ability to adapt such programs to the target audience. Work [6; 7] is a review and presents the results of content analysis on the choice of virtual physical laboratories in teaching physics by teachers, depending on the topic being studied. Another publication [8] describes the use of online platforms on which, without any equipment, only with the appropriate software and hardware, it is possible to conduct virtual and physical experiments. The authors propose the use of virtual platforms to teach an applied physics course at the Faculty of Natural Sciences and Engineering. Article [9] describes the results of the study of the effectiveness of teaching physics using a virtual laboratory. In the example of the study of geometric optics, the effectiveness of virtual simulations for successful learning outcomes was confirmed. At the same time, there was no significant difference in the attitude of students to the study of physics.

D. Using virtual and augmented reality to teach physics

Based on immersive learning technologies, the authors of the article [10] explore the problem of teaching abstract physics topics (in particular, the concept of modern physics). The results of teaching the Gravity Assist concept in a virtual educational environment are presented (Gravity Assist, also known as Gravitational Slingshot or Swing-By, is a method of accelerating or slowing down space probes by exchanging impulses with the planet). This topic requires students to have the skills of abstract multidimensional thinking that has been formed in a virtual environment. Separately, we note that augmented reality can be perceived data generated by application sensors for measuring temperature, distance, acceleration, illumination, noise level, magnetic properties, etc. (Fig. 1).

E. Use of specialized environments for modeling physical processes

In the article [11] the method of computer modeling as a teaching method is investigated. The authors prove that this is a full-fledged method of teaching physics. In the article [12], the author demonstrates how you can effectively use the GeoGebra environment (https://www.geogebra.org/) to create animation and modeling in the study of physics. The author presents several examples of digital learning materials: animations, simulations, and computer games that can help to understand terms, phenomena, and processes, as well as improve the assimilation of knowledge and skills in several branches of physics, including motion, dynamics, electrical circuits, and swings. The study highlights the connection between these digital learning materials and the theory of physics.

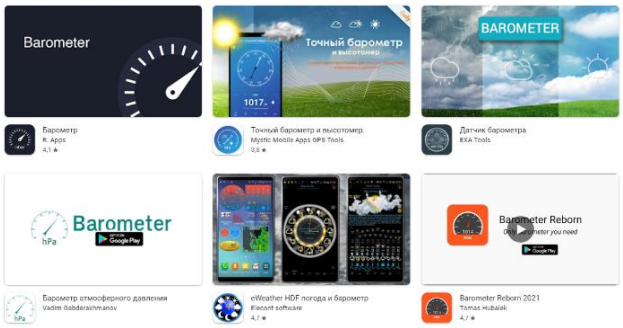


Figure 1. Variety of barometer applications

So, according to the analysis of scientific publications, the following areas of use of digital technologies in teaching physics have been identified:

- the use of digital tools as tools for the study of physical processes;
- the use of applications to solve physical problems;
- the use of virtual and digital physical laboratories;
- the use of virtual and augmented reality in teaching physics;
- the use of specialized environments for modeling physical processes.

III. ANALYSIS OF UKRAINIAN SCIENTIFIC WORKS ON THE USE OF DIGITAL TECHNOLOGIES IN TEACHING PHYSICS

The analysis of the Ukrainian scientific space on the methodology of teaching physics at school confirmed the using the above directions of digital technologies in teaching physics. In particular, to prepare for external non-assessment in physics the work [13] offers a mobile application of physical orientation "Physical formulas. Physics" ((https://play.google.com/store/apps/details?id=com.ua.physics) developed by a group of Ukrainian specialists BbytesLab). The application includes basic formulas and tables from the Ukrainian school course of physics (grades 7-11) and can work without access to the Internet.

The authors of the article [14] propose to use Lab4Physics (https://play.google.com/store/apps/details?id=com.lab4u.lab4physics) to conduct an educational physical experiment. Lab4Physics is a

program that, along with the measuring sensors embedded in the program, allows you to use the student's phone as a laboratory tool. The authors claim that thanks to this it is possible to conduct a significant number of experiments without special physical equipment. To experiment, you need to select and run the desired virtual tool and point the smartphone at the object whose parameters are measured. The authors note that the advantages of using mobile devices and applications in school lessons include: mobility, accessibility, compactness, and speed.

The paper [15] describes the composition and scope of use of the worldwide digital laboratories Einstein (http://bit.ly/3Xovod9) and Archimedes (http://bit.ly/3wegVEx), which involve the use of various digital sensors, with the help of which a wide range of research, demonstration, and laboratory work in physics can be carried out.

The article [16] describes the creation of a computer laboratory for the use of complex (for example, an oscilloscope) or hazardous devices (for example, a nuclear reactor), which are necessary when studying a physics course. Laboratory robots are programmed (Object Pascal). The author demonstrates an example of modeling a laboratory installation "Study of the phenomenon of electromagnetic induction" and notes that it is important for students to check the models for adequacy – to compare the result and work of the program with experimental data.

The generalization of scientific and methodological results indicates the active use of specialized environments for the visualization of physical experiments and their simulation. Thus, the use of the interactive simulator of physical processes STEP (https://apps.kde.org/uk/step/) is indicated in [17]. The authors describe the method of its use in the study of the elastic pendulum, mathematical pendulum, resonance, mechanical wave, and Brownian motion and confirm its effectiveness by the results of a pedagogical experiment. The article [17] discusses the Proteus environment (https://www.labcenter.com/), which is used to simulate physical processes at the micro level. The authors note that Proteus is a powerful computer-aided design system, which provides the ability to simulate analog and digital devices and simulate their work to detect design and tracing errors. The publication [18] touches on the problem of using Proteus as a means of virtual and augmented reality to visualize the work of individual elements of the information system. The work confirms the expediency of its use in the training of IT specialists.

Separately, we note the popularity of the resource PhET (https://phet.colorado.edu/), which presents online laboratories for teaching physics [19], which visualize many physical experiments, laws, and dependencies. Here are some examples of online laboratories of the PhET resource:

- laboratory work on the topic "Determination of focal length and optical force of a fine lens" (Fig. 2) (https://phet.colorado.edu/sims/html/geometric-optics/latest/geometric-optics_uk.html). Students can study the lenses and the features of their action, analyze the formula of a thin lens, independently establish the relationship between the object and its image under the action of the lens, determine the optical power of the lens;

- laboratory work on the topic "Laws of equilibrium" (https://phet.colorado.edu/sims/html/balancing-act/latest/balancing-act_uk.html). Students can experimentally check what the balance of strength and their shoulders should be so that the lever is in balance.

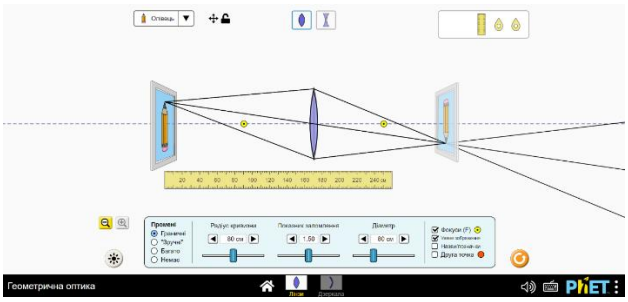


Figure 2. Interactive simulation of the PhET laboratory "Geometric Optics"

The simulation helps students independently compare the relationship of the forces acting on the lever, and the ratio of his shoulders, formulate the equilibrium condition of the lever and analyze what factors affect the accuracy of measurements (Fig. 3);

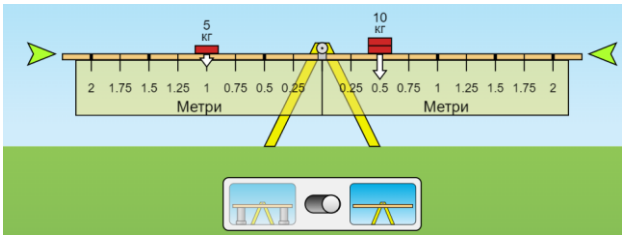


Figure 3. Interactive simulation of the PhET laboratory "Balancing"

- laboratory work on the topic "Boyle–Mariotte, Gay-Lussac and Charles laws for different thermodynamic states of gas" (https://phet.colorado.edu/sims/html/gases-intro/latest/gases-intro_uk.html). Students experimentally investigate isothermal, isobaric, and isochoric processes. Configuring different simulation parameters for each thermodynamic state of the gas in the tank, students observe the change in the concentration of gas molecules and calculate the required physical quantities (Fig. 4).

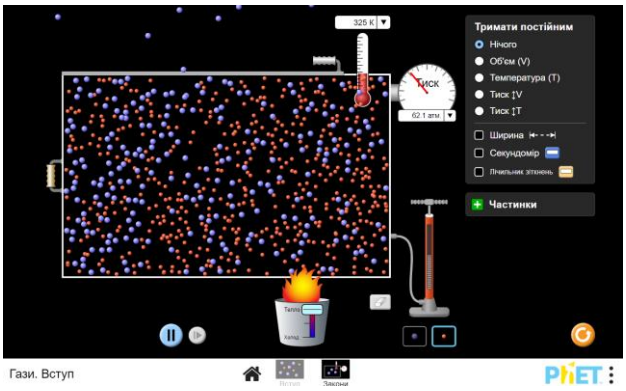


Figure 4. Interactive simulation of the PhET laboratory "Gases"

Thus, the analysis of scientific and methodological publications shows that Ukrainian teachers use digital technology in teaching physics, and according to the content analysis, we note the popularity of using virtual laboratories for the visualization/ simulation of physical processes.

IV. UKRAINIAN PRACTICES OF USING DIGITAL TECHNOLOGIES IN TEACHING PHYSICS - SURVEY

We investigated the issue of using digital technologies in teaching physics in practice. Physics teachers of the Sumy region and students of the Sumy State Pedagogical University named after A.S. Makarenko (specialty "Secondary education. Physics", level of higher education bachelor's and master's) participated in the survey (Table 1).

The total number of respondents is 106 people: 62 physics teachers and 44 students.

Since we recorded five areas of use of digital technologies in teaching physics, and we were interested in the opinion of teachers and students in each of the areas, we offered respondents the following questions:

1) Do you know about the certain direction of using digital technologies in teaching physics?

2) Do you have your own experience in using certain directions of digital technologies in teaching physics?

3) Based on the results of the experience gained, do you want to continue to use digital technology in teaching physics?

4) Do you want to deepen your knowledge and ability to use a certain direction of digital technology in teaching physics?

5) Which of these directions do you consider the most effective in teaching physics (2 positions)?

Questionnaires had the form of a table (according to table 1), where rows are the directions of use of DT, and columns are question numbers. Respondents had to mark the positive or negative answer for every position. The last column required only two positive answers.

The summary results are shown in Table 1.

TABLE I. SURVEY RESULTS

Question	**Question No1**		**Question No2**		**Question No3**		**Question No4**		**Question No5**	
The direction of the use of DT	*Teachers*	*Students*	*Teachers*	*Students*	*Teachers*	*Students*	*Teachers*	*Students*	*Teachers*	*Students*
The use of digital tools as tools for the study of physical processes	37%	43%	13%	25%	13%	25%	82%	100%	85%	95%
The use of applications to solve physical problems	13%	16%	2%	11%	1%	11%	45%	62%	11%	25%
The use of virtual and digital physical laboratories	8%	66%	2%	50%	2%	46%	73%	100%	76%	86%
The use of virtual and augmented reality in teaching physics	5%	14%	0%	2%	0%	2%	31%	68%	29%	25%
The use of specialized environments for modeling physical processes	13%	45%	5%	43%	5%	39%	42%	61%	40%	45%

We will comment on the results.

To question 1 "Do you know about the certain direction of using digital technologies in teaching physics?" we have a situation that, in general, students are more aware of the use of digital technologies in teaching physics than teachers. We explain this by the fact that today every educational program for the training of physics teachers necessarily involves familiarizing themselves with specialized software in the field of physics. Physics teachers in most cases learn about such technologies through advanced training courses or through self-education, which does not guarantee knowledge of all available ways of using digital technologies in teaching physics. During live communication with teachers, it was revealed that it was during the transition to distance learning that teachers' ideas about the range of digital technologies and means of teaching physics expanded.

To question 2 "Do you have your own experience in using a certain direction of digital technologies in teaching physics?" we again have a situation where students give more positive answers than the teachers. We explain this by the fact that students already have experience in learning digital technology in physic, as well as experience in industrial practice, where they introduced such technologies in teaching physics. During personal communication with teachers, it was found that the use of DT based in schools is quite limited: the Internet is not everywhere, access to computer classes is not always available, and not all students and teachers have smartphones where specialized software is installed. Therefore, teachers note a lack of experience. Teachers do not accept applications for solving physical problems as a means of learning (they believe students should solve physical problems independently). Such applications are used mainly for self-testing. Teachers know little about virtual and digital physical laboratories, unlike students who meet them at the university. The use of virtual and augmented reality in teaching physics is new for everyone. Teachers do not use such tools, and students have experience using only complementary realness. Teachers believe that it is not rational to use specialized environments and to model physical processes since it requires a lot of additional time to master the means of learning (it is better to spend this time-solving problems).

Answers to question 3 "Based on the results of the experience gained, do you want to continue to use digital technology in teaching physics?" testify: those who have

had experience in using digital technology in teaching physics, will use them in the future. During the conversation, the teachers noted that this was an "interesting experience", and "the students liked it". At the same time, we found that there is some bias among teachers about using virtual physical labs to teach physics: there is no access to computers or the Internet at the right time; incomprehensible interface language of apps or computer programs, the inability to install the application on a specific operating system; the inability to install programs on phones, etc.

To question 4 "Do you want to deepen your knowledge and ability to use a certain direction of digital technology in teaching physics?" we recorded significant percentages for students (62% to 100% depending on the direction) and smaller for teachers (from 31% to 82% depending on the direction). It turned out that when teachers are aware of the availability of a certain direction of digital technology and imagine their use in the realities of their professional activities, they have the desire to learn how to use this direction. But after getting acquainted with the peculiarities of conducting a virtual physical experiment in school conditions, they noted that "it cannot completely replace the real one, where it is possible to manually assemble the electrical circuits or arrange in certain way instruments to demonstrate certain physical phenomena."

To question 5 "Which of these directions do you consider the most effective in teaching physics (2 positions)?" the unanimity of the opinions of teachers and students in favor of the first (the use of digital tools as tools for the study of physical processes) and the third (the use of virtual and digital physical laboratories) directions was recorded.

The first direction was chosen because each physics teacher will surely use available digital tools to confirm or disprove any hypotheses, provided they know how to apply them. The third direction was chosen through high-quality visualization of processes and the ability to organize an independent study of certain topics of physics Teachers note that a virtual experiment allows you to: carry out certain studies, which cannot be reproduced in real life; repeat the study many times, while changing the parameters; save study time on the organization of a real physical experiment; test the model in various, including boundary cases, which in reality can lead to extreme situations.

V. CONCLUSION

The use of digital technologies in teaching physics is in demand not only because of the need for distance education. It is in demand because of the development of technologies and means, the emergence of immersive learning technologies, and the availability of mobile tools that can accumulate and process empirical data to understand the dependencies of the physical world.

The analysis of theoretical developments proves the spread of experience in introducing digital technologies into physics training, but the practical state of their use in Ukraine testifies to isolated practices. The reasons are not only a weak material and technical base, but also ignorance of the ways to use digital technologies in teaching physics and the unwillingness of working teachers to use such technologies. Therefore, the urgent task of universities is the advanced training of physics teachers who will want to be familiar with various digital technologies and means of teaching physics and are ready to introduce effective methods of their use in their professional activities.

In the future, it is planned to develop a special course designed to form the ideas of physics teachers about the available digital technologies and tools that are already used to teach physics, and develop the skills of their implementation in educational practice.

REFERENCES

[1] M. Milner-Bolotin, O. Aminov, W. Wasserman and V. Milner, "Pushing the boundaries of science demonstrations using modern technology," Canadian Journal of Physics, vol. 98, No 6, 2019, pp. 571-578. https://doi.org/10.1139/cjp-2019-0423.

[2] M. Anni, "Quantitative Comparison between the Smartphone Based Experiments for the Gravity Acceleration Measurement at Home," Education Sciences, vol. 11(9), 2021, 493. https://doi.org/10.3390/educsci11090493.

[3] Education Alexis Media, "Dodatok v Google Play – Physics Masters [Application on Google Play – Physics Masters]," Retrieved from: https://play.google.com/store/apps/details?id=co.alexis.gbbxs.

[4] Carlo Terracciano – MasterApps, "Dodatok v Google Play – PhysicsMaster – Physics Calc [Application in Google Play – PhysicsMaster – Physics Calc]," Retrieved from: https://play.google.com/store/apps/details?id=com.terracciano.physics_master.

[5] O. K. Mircik, and Z. S. Ahmet, "Evaluation of research related to virtual physics laboratory applications," Canadian Journal of Physics, vol. 96, issue 7, 2018, pp. 740-744. https://doi.org/10.1139/cjp-2017-0747.

[6] V. Shamonia, O. Semenikhina, M. Drushlyak, and S. Lynnyk, "Computer visualization of logic elements of the information system based on," 15th International Conference on ICT in Education, Research, and Industrial Applications (ICTERI 2019), Kherson, pp. 459-463, June 12-15, 2019.

[7] O. K. Mircik, and Z. S. Ahmet, "Virtual laboratory applications in physics teaching," Canadian Journal of Physics, vol. 96, issue 7, 2018, pp. 745-750. https://doi.org/10.1139/cjp-2017-0748.

[8] L. F. NegliaV, and D. Llulluy-Nuñez, "Use of Virtual Platforms for the Experiences of the Course of Applied Physics to Engineering II at the Universidad de Ciencias y Humanidades Lima – Peru," 2021 IEEE World Conference on Engineering Education (EDUNINE), Guatemala City, Guatemala, 2021, pp. 1-4, https://doi.org/10.1109/EDUNINE51952.2021.9429136.

[9] S. Erdogan, and E. Bozkurt, "The effect of virtual laboratory applications prepared for Geometrical Optics Lesson on students' achievement levels and attitudes towards Physics," Pegem Journal of Education and Instruction, Vol. 12, No. 2, 2022, pp. 226-234. https://doi.org/10.47750/pegegog.12.02.22.

[10] G. Monahan, M. Cossoul, S. Harris, I. Humer, C. Guetl, and C. Eckhardt, "Gravity Assist: An Immersive and Interactive Visualization," 2021 International Conference on Advanced Learning Technologies (ICALT), Tartu, Estonia, 2021, pp. 401-405, https://doi.org/10.1109/ICALT52272.2021.00128.

[11] Z. Pucholt, "Effectiveness of simulations versus traditional approach in teaching physics," European Journal of Physics, vol. 42, No. 1, 2021. https://doi.org/10.1088/1361-6404/abb4ba.

[12] M. Jancheski, "GeoGebra animations, simulations and computer games in teaching and learning physics," INTED2019 Proceedings, 2019, pp. 7613-7623. https://doi.org/10.21125/inted.2019.1865.

[13] BbytesLab, "Dodatok v Google Play – Fizychni formuly. Fizyka [Application in Google Play - Physical formulas. Physics],"

Retrieved from: https://play.google.com/store/apps/details?id=com.ua.physics.

[14] O. Kolesnikova, N. Myslitska, and D. Semeniuk, “Vykorystannia tekhnolohii BYOD dlia formuvannia eksperymentalnykh znan ta umin uchniv z fizyky [Use of BYOD technologies for formation experimental knowledge and life of physics masters],” Fizyko-matematychna osvita – Physical and Mathematical Education, Issue 2(20), 2019, pp. 48-53. https://doi.org/10.31110/2413-1571-2019-020-2-008.

[15] A. Yurchenko, “Tsyfrovi fizychni laboratorii yak aktualnyi zasib navchannia maibutnoho vchytelia fizyky [Digital physical laboratories as a important mean of training of future teacher of physics],” Fizyko-matematychna osvita – Physical and Mathematical Education, Issue 1(4), 2015, pp. 55-63.

[16] O. Sytnikov, “Zasoby stvorennia ta vykorystannia virtualnoi fizychnoi laboratorii [Means Of Creating And Using A Virtual Physical Laboratory],” Fizyko-matematychna osvita – Physical and Mathematical Education, Issue 1(5), 2018, pp. 298-301. https://doi.org/10.31110/2413-1571-2018-015-1-057.

[17] L. Kovalov, M. Medvedieva, and I. Poberezhets, “Vykorystannia interaktyvnoho imitatora fizychnykh protsesiv Step v osvitnomu protsesi u zakladakh vyshchoi osvity [Use of the interactive simulator of Step physical processes in the educational process in higher education institutions],” Fizyko-matematychna osvita – Physical and Mathematical Education, Issue 3(29), 2021. pp. 68-73. https://doi.org/10.31110/2413-1571-2021-029-3-011.

[18] V. H. Shamonia, O. V. Semenikhina, V. V. Proshkin, O. V. Lebid, S. Y. Kharchenko, and O. S. Lytvyn, “Using the virtual environment to train future IT professionals,” CEUR Workshop Proceedings 2547, 2020, pp. 24-36. Retrieved from: http://ceur-ws.org/Vol-2547/paper02.pdf

[19] PhET Interactive Simulations. Retrieved from: https://phet.colorado.edu/en/simulations/filter?subjects=physics&type=html,prototype.

Fostering Critical and Computational Thinking in the Field of Primary and Secondary Education in non-STEM Subjects by Using Data Sets and Applications

Ines Vlahović, Ivana Ogrizek Biškupić
ALGEBRA University College, Zagreb, Croatia
ines.vlahovic@algebra.hr ; ivana.ogrizekbiskupic@algebra.hr

***Abstract* – Critical and computational thinking in primary and secondary education in recent years shows growing importance in methodical approaches used in the classroom. Although many examples exist for using critical and computational thinking in STEM educational area, the social sciences i.e., non-STEM areas were somehow left out due to the relatively more difficult design of such content, especially in the part of the development of computational thinking. In that context, in this paper we present one way of using critical and computational thinking in non-STEM education, more specifically on the example of historical data sources, through the use of programming in Python. As an example, we use historical data for Trans-Atlantic Slave Trade routes, that were the largest long-distance coerced movement of people in history up to the mid-nineteenth century, for connecting concepts of databases, Data Science and programming for development of critical and computational thinking in context of history science. This way of using modern approaches in classroom should give teachers and pupils a broader picture of importance of interdisciplinary education for critical and computational thinking development through STEM and non-STEM classes that give pupils novel skills needed for future labor market.**

Keywords – Critical Thinking; Computational Thinking; STEM, non-STEM; Data Science; programming; Python; history, Trans-Atlantic Slave Trade routes

I. INTRODUCTION

In this paper we analyze prerequisites for substantiation of critical thinking and present one possible approach to foster critical and computational thinking in non-STEM courses by using historical data sources by programming in Python. The idea is to develop new approach and present one possible model to promote STEM in non-STEM courses as an upgrade of gamification method in the early stage of education (primary) with an aim to introduce STEM basics to wider number of students thus creating an opportunity for new learning and skills. Starting from the very beginning it's extremely important to persist in finding a way to encourage critical thinking at an early age of life, especially in primary education. This mode of thinking represents the skill that improves intellectual values by processing information and experiences due to observing and analyzing all factors throughout the process of synthesizing, reasoning, and conceptualizing. We could name many arguments why critical thinking is important in the life of an individual. One of the most important among others is boosting-up the research skills, enhancing the problem-solving ability, improving decision making, stimulating creativity and curiosity. Siburian et. al [1] published the findings confirming significant correlation between critical and creative thinking skills on cognitive learning with 72.80% contribution of critical thinking skills and creative thinking skills simultaneously to cognitive learning results. On the other hand, Amanda Hiner [2] argues that critical thinking is often not effectively or intentionally taught in college classrooms, resulting in disappointing learning outcomes for students and that educators should be trained on how to teach critical thinking and should integrate focused instruction in critical thinking into their classrooms in order for students to make significant, positive gains in their critical or analytical abilities. Hence, the teacher's role is invaluable in this process.

Student critical thinking skills can be improved through adjusting curriculum to encourage these skills [3]. Bailin and Battersby [4] suggest that critical thinking requires a careful critical examination of the subject issue. They provide an example of the questions for the inquiry process e.g.:

- What is the issue?
- What kinds of claims or judgments are at issue?
- What are the relevant reasons and arguments on various sides of the issue?
- What is the context of the issue?
- How strong is each of the arguments?
- Weighing and balancing the evaluated arguments, what judgment should we come to?

In order to even begin to inquire, students need to be clear about the issue which is to be impetus for the inquiry [5]. This requires that they understand the criteria for a well-formed and appropriate issue for inquiry. To foster critical thinking in elementary school some authors [6] suggest applying a 3 Step Method. The 1st element of this method is setting a cognitive challenge, something that is not easily answered and, in the case of philosophical inquiry, probably has no definite answer. Having a problem set, the 2nd thing is how to approach it, what method to use to move towards a solution (e.g.,

collaborative process inside the group) and the 3rd one is social construction – using the kind of dialogue described earlier to investigate the problem.

II. TECHNOLOGY AND APPLICATIONS IN NON-STEM COURSES

One of the possible approaches to foster creativity in education we see in acquiring STEM knowledge and skills. The data suggests that more students prefer non-STEM courses, hence, the idea was to develop a method of using STEM knowledge in a problem-solving exercise based on a critical thinking in the history course (primary and secondary school).

Regarding the idea, Nurzengky et. al [7] have already developed one possible model of student critical thinking skills in history education. Their survey study was empirically tested on N:124 junior high school students in South Jakarta and even though they imply that model should be slightly modified, results confirmed that student interpretations, inference, evaluation, explanation, and monitoring have been positively associated with student critical thinking skills in history education. Machotka & Papapetrou [8] have presented project output, the conceptual frame and exploratory workshop to probe the field and to build a constructive interdisciplinary dialogue between two research areas: Data Science and Art History with conclusion that Data Science can enrich art studies while analysis of visual data can have a positive impact on Data Science. In their paper the authors present examples of future research directions from the Data Science perspective of Digital Art History that include (a) the development of new dimensionality reduction methods and feature-extraction mechanisms from complex Digital Art History data (e.g. digitized manuscripts and images), (b) the construction of novel neural network architectures based on CNNs and LSTMs with attention mechanisms to address the time and space variables, (c) the design of user-interactive tools with advanced machine learning and data management capabilities that can be deployed in real-time scenarios, such as museum navigation, and (d) explicitly maintaining the users and human experts to provide interactive feedback to the learning process via techniques such as reinforcement learning or active learning [8].

Looking towards the data sets and by analyzing its content we have followed the discussion of Damerow & Wintergrün [9] regarding challenges of the initial data collection process and how it fits into the overall research data life cycle. They suggest creating a data corpus much more explicit to contextualize data considering that decision on what particular data set, method, or technology we use has a great impact on the results of an analysis.

In recent years a huge amount of data has been collected for scientific and educational purposes. A huge amount of data has become an integral element of contemporary society [10]. For educational purposes a vast number of data set emerged that can be used as additional tools in classroom to improve learning such as data sets from given links:

- https://data.world/datasets/education
- https://libguides.css.edu/EDU6190/DataSets
- https://dataschools.education/resource/useful-datasets-for-data-education-in-schools/

although, most of them are applicable to higher education. Available data sets, if prepared for non-STEM teachers, could be used for critical and computational thinking in schools as an example of Slave Voyages show.

In this paper in the Introduction, we provide the main challenges of critical thinking to be introduced in the classroom. Next, the Technology and applications in non-STEM courses section give overview of relevant examples of using STEM knowledge of problem-solving principles in non-STEM courses. The Conceptualizing data sets for history courses section shows an example of Slave Voyages data set based on collected historical data where we describe the process of preparing and analyzing the example data set for research questions for which we developed program in Python programming language. In Results and discussion, we show ways in which the program can be used for development of critical thinking. In Conclusion we emphasize the potential future applications and further development of this approach for applications in the curriculum of primary and secondary schools, on the basis of which students and professors will be able to develop computational and critical thinking.

III. CONCEPTUALIZING DATA SETS FOR HISTORY COURSES

A. History and Methodology of Slave Voyages data set collection and preparation

After collapse of the Amerindian population, the New World has experienced a shortage of labor at plantations (sugar, tobacco, other agricultural goods) and at gold exploitation in 16th century. As a result, Europeans and Euro-Americans made a slave trade routes with Africans in order to maintain the supply chain with new tropical goods and gold to the old world through slavery [11]. It was this slavery that has driven the growth of the economy of America and other countries (empires) involved in the slave trade and has triggered changes on the political global stage with an impact on society and culture and other aspects of the daily lives of the stakeholders involved.

Data collected in Voyages Database contains data about voyages that occurred from 16th to mid-19th century from which almost all are published and archival information. As noted by the authors of the database, there are records of more than 36000 voyages documented through one or more history sources from which one sixth are based on at least one record, while ~60% of the voyages have three or more separate sources each. List of the sources are given at the link Trans-Atlantic Slave Trade - Understanding the Database (slavevoyages.org). Collected slave trade data gives a broader picture of history causes and consequences through examining cultural, demographic, economic and

technological changes, although through examining newly discovered history sources, new evidence and new records are occurring in database. Currently data set shows that 12520000 captives were sent to America from Africa in period of 16th to mid-19ht century in around 41190 voyages from which 10.7 million are disembarked mostly in America. The validity of this data set was also confirmed with genetic research shown in paper [12].

Data set used for Trans-Atlantic Slave Trade is published on the web page of Slave Voyages project (https://www.slavevoyages.org/) whose purpose is to collect archival data on slave-trading voyages from unpublished and published history sources and create dataset for further scientific research. The whole data set is available at link https://www.slavevoyages.org/voyage/database. Variables used in this database are divided into eight groups: 1) Ship, nation, owners – 14 variables, 2) Voyage Outcome – 5 variables, 3) Voyage Itinerary – 11 variables, 4) Voyage Dates – 9 variables, 5) Captain and Crew – 4 variables, 6) Slave (numbers) – 12 variables, 7) Slave (characteristics) – 9 variables, 8) Source – 1 variable, making total of 55 variables. The complete description of variables (fields) can be found on the link https://www.slavevoyages.org/voyage/about#variable-list/2/en/. Due to the large number of used variables in the database, when creating customized data tables (data sets) based on specific research question suitable for use in classes, records should be cleaned and prepared according to the methods and concepts used in Data Science. We must emphasize the importance of knowing the description of the variables when interpreting the data used, i.e., we must be aware of the way they are collected and the methods used especially when working with historical sources where estimated and calculated values are also used.

B. Application of ready-made program solutions in Python for history class - Slave trade routes example

For developing critical thinking and computational thinking outside of traditional STEM classes we build small program based on Python programming language ((version 3.7.9) in Spyder scientific environment) for connecting basic concepts in Data Science (preparing and cleaning data, libraries used for Big Data analysis in Python) with history data of slave trade routes through selected research questions suitable for K-12 students.

1) What research questions can be conducted based on historical sources about slave trade routes?

As we already mentioned, this data set has 55 variables and from those variables a vast number of research questions can be asked in context of history of Trans-Atlantic Slave Trade Routes and critical thinking. For this paper we choose to present a few:

1. For specific selected range of years of arrival at port of disembarkation (in overall period of 16th to 19th century) and country, discuss ratio of total embarked and total disembarked slaves in Middle passage (voyage of enslaved Africans across the Atlantic Ocean to the New World) for every individual voyage?
2. For specific selected range of years of arrival at port of disembarkation (in period of 16th to 19th century) and selected country discuss ratio of total embarked and total disembarked slaves in Middle passage grouped by specific year?
3. What does comparison of those numbers for different countries tell you about the role of individual countries in the slave trade voyages for that selected period range?

To give an answer to these questions, we prepared custom data set that include fields with the variables Year of arrival at port of disembarkation, Total embarked, Total disembarked, Flag (imputed). From initially 36073 records and after cleaning the data (in this case of cleaning the data we think about discarding those records that don't have data for one of the variables in individual record) we got for our customized data set 34444 records, of which 2209 records have unknown assigned country. Also, we sorted data by year and country (Flag (imputed) variable) because of practical usage for educational purposes, although this step is not mandatory (part of the data table is shown on Fig. 1).

Year of arrival at port of disembarkation	Total embarked	Total disembarked	Flag (imputed)
1606	90	77	Denmark / Baltic
1641	150	130	Denmark / Baltic
1646	347	269	Denmark / Baltic
1646	260	212	Denmark / Baltic
1647	392	320	Denmark / Baltic
1647	251	174	Denmark / Baltic
1657	325	259	Denmark / Baltic
1660	328	285	Denmark / Baltic
1679	290	245	Denmark / Baltic
1683	334	254	Denmark / Baltic
1686	250	200	Denmark / Baltic

Figure 1. Example of prepared and cleaned customize data set according to research questions and Data Science principles.

For this customized data set and research questions proposed, we write a Python program shown below in Fig. 2.

```
#import of libraries for Data Science
import pandas as pd
import matplotlib.pyplot as plt

#import of dataset
df = pd.read_csv("Questions_dataset.csv",sep="\t")
#printing dataset (all columns are not shown)
pd.set_option('display.max_columns', None)
print(df)
#printing dataset (all columns are shown)
pd.set_option('display.max_columns', None)
print(df)
#List of counties
Countries={"1":"Denmark / Baltic","2":"France","3":"Great Britain",\
           "4":"Netherlands","5":"Portugal / Brazil","6":"Spain / Uruguay",\
           "7":"U.S.A.","8":"Unknown"}
#Insert range of years (1500-1870)
Start_period=int(input("Please enter a start year:\n"))
End_period=int(input("Please enter a end year:\n"))

print(Countries)

country= input("Select country:")

#Criteria for filtering data by Years and selected country for bar plot
years = df.loc[(df["Year of arrival at port of disembarkation"]>=Start_period)\
               &(df["Year of arrival at port of disembarkation"]<=End_period)\
               &(df['Flag (imputed)']==Countries.get(country))]
#print(years)
#print(years.sum(axis=0))


#creating bar plot with filtered data for every record in data
ax =years.plot.bar(x="Year of arrival at port of disembarkation",y="Total embarked",\
                   color="red", label="Total embarked",figsize=(12,5))
years.plot.bar(x="Year of arrival at port of disembarkation",y="Total disembarked",\
               color="blue", label="Total disembarked",ax=ax)
ax.set_xlabel("Year of arrival at port of disembarkation")
ax.set_ylabel("Total embarked/Total disembarked")
plt.title('Slave trade routes')
plt.show()
```

```
#creating bar plot for filtered data with group by function
gy = years.groupby('Year of arrival at port of disembarkation')
GroupedYears=gy.sum()
print(GroupedYears)
ax =GroupedYears.plot.bar( figsize=(14,5))
ax.set_xlabel("Year of arrival at port of disembarkation")
ax.set_ylabel("Total embarked/Total disembarked")
#labels of years on bar plot
for container in ax.containers:
    ax.bar_label(container)

plt.title('Slave trade routes')
plt.show()

years2 = df.loc[(df["Year of arrival at port of disembarkation"]\
                >=Start_period)
                &(df["Year of arrival at port of disembarkation"]\
                <=End_period)]


#creating graph for countries in selected year range
gf=years2.groupby("Flag (imputed)")
GroupedFlag=gf.sum()
print(GroupedFlag[["Total embarked","Total disembarked"]])

ax =GroupedFlag[["Total embarked","Total disembarked"]].plot.barh(figsize=(12,5))
ax.set_xlabel("Year of arrival at port of disembarkation")
ax.set_ylabel("Total embarked/Total disembarked")
for container in ax.containers:
    ax.bar_label(container)
plt.title('Slave trade routes')
plt.show()
```

Figure 2. Custom made code for investigating research questions.

In summary, this program as input data takes our custom data set in csv format. For user selected options for available records of countries that were included in slave market, we select a range of years (for example for option 6-Spain /Uruguay, we select range of years from the year 1800 to the year 1805) and with this input parameters we plot three bar plots. First of the plots describe individual voyages, second bar plot shows grouped data for those criteria for range of years and third bar plot shows for selected years number of embarked and disembarked captives for involved countries in Trans-Atlantic slave trade routes. For working with data set, we use pandas and matplotlib libraries, one of the most common libraries used in Data Science field. On the Fig. 2 in code in front of each bar plot code is comment that separates parts of the program according to proposed questions so users could change arguments as needed.

IV. Results and Discussion

In order to give an answer to questions proposed, after running the program, we select period from the year 1800 to the year 1805, and for country option, option 6-Spain /Uruguay. As a result of running the code we got bar plots shown on Fig. 3-5.

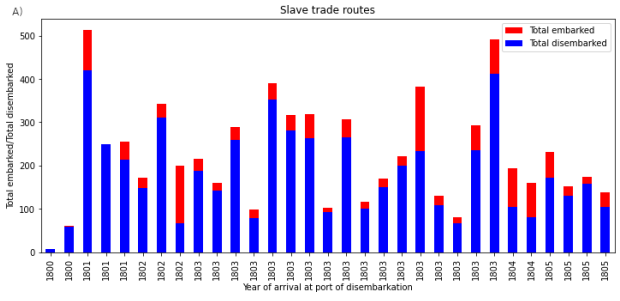


Figure 3. Bar plots obtained with python programming language for year range 1800-1805 and county Spain/Uruguay. A) Individual voyages in selected year range and county.

According to Fig. 3, for purpose of interpreting obtained bar plots, students could argue and conclude, that individual voyages have small amount of embarked and disembarked slaves, that for some voyages the difference between embarked and disembarked slaves is large so interpretation of this data should use critical

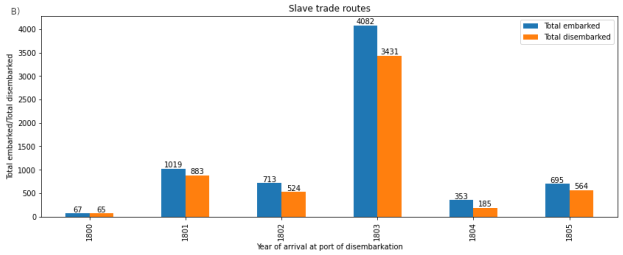


Figure 4. Bar plots obtained with python programming language for year range 1800-1805 and county Spain/Uruguay. B) Grouped data by year for selected year range and country.

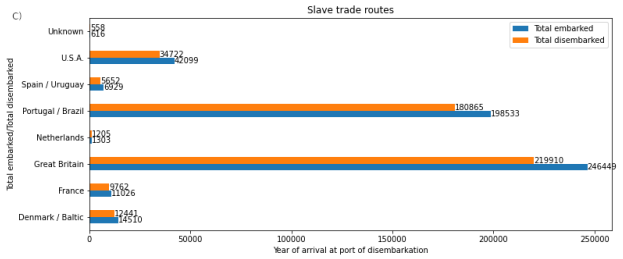


Figure 5. Bar plots obtained with python programming language for year range 1800-1805 and county Spain/Uruguay. C) Horizontal bar plot for countries in data set for selected range of years.

thinking in order to connect those results with historical circumstances (storms, rebellion, diseases, pirates as an example) for conducting the realistic interpretation based on historical data. Also, students could discuss which year had the highest frequency of voyages under a certain flag, which in this case had been the year 1803, as is seen from the selected parameters.

In Fig.4 for a selected country, students can track data grouped by specific year in a selected range of years and discuss the historical circumstances that led to an increase in the global slave trade for a given year, but also and circumstances that led to decrease in it. For our case of selected years that was also the year 1803 when traffic had been almost four times greater than for other years for selected country (Spain/Uruguay).

We investigate for answering the third question from list of countries that were included in Trans-Atlantic slave trade routes (Denmark / Baltic, France, Great Britain, Netherlands, Portugal / Brazil, Spain / Uruguay U.S.A., Unknown), which one of them had the highest traffic in that specific year range. From the year 1800 to the year 1805, the highest number of embarked and disembarked slaves in Middle passage had Great Britain with 246449 embarked and 219910 disembarked slaves. If we compare this short-selected range of years with the total number of captives embarked and landed from the year 1500 to the year 1870 for Great Britain, the traffic in that selected short range of years accounts for 8% of total slaves embarked and 8.47% of slaves landed. If we compare that short time span for Great Britain with the slave trade of each listed country from the 16th to the mid-19th century, for which data exists, it accounts for 2.32% of all embarkations and 2.39% of all slave landings. Overall numbers for each country can be obtained by using this program for input values from the year 1500 to the year 1870. This program, except for bar plots, also

prints and summary table for selected range of years as Table 1 shows. It is interesting that the highest traffic had Portugal/Brazil so analyzing historical context using critical thinking could be discussed through themes of economy, society, politics and technology for involved countries.

TABLE 1. SUMMARY NUMBERS OF EMBARKED AND DISEMBARKED SLAVES THROUGH 16TH TO MID-19TH CENTURY.

Flag (imputed)	Total embarked	Total disembarked
Denmark / Baltic	104273	86648
France	1307821	1113711
Great Britain	3082495	2596168
Netherlands	594375	510009
Other	3686	3176
Portugal / Brazil	3883856	3471023
Spain / Uruguay	569126	497734
Sweden	191	176
U.S.A.	377764	313401
Unknown	718580	594350
Sum	**10642167**	**9186396**

As an example of usage of this program for critical thinking and computational thinking development we recommend to compare for different countries how slave trade took place over the span of twenty years and discuss which of the country were dominant from 16th to mid-19th century in slave trade and put it in history circumstances relevant for domains of technology, economy, society and politics. An example of using program in this way for grouped data are shown in Fig. 6 to Fig. 12 for the period from the year 1650 to the year 1670 and countries Denmark / Baltic, France, Great Britain, Netherlands, Portugal / Brazil, Spain / Uruguay and U.S.A.

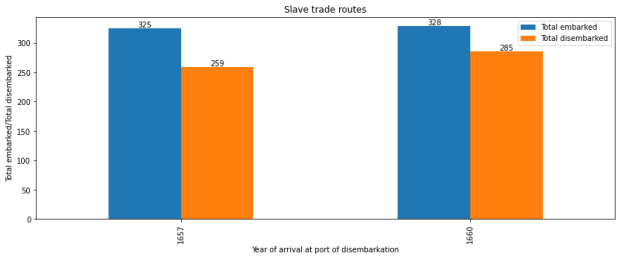


Figure 6. Denmark / Baltic (1650-1670)

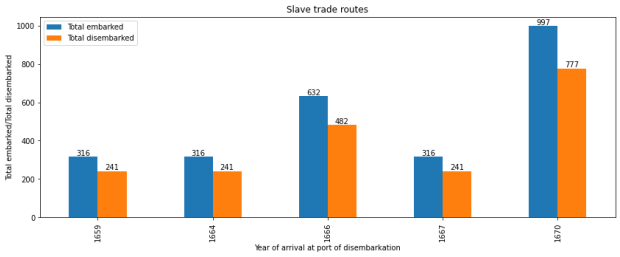


Figure 7. France (1650-1670)

By comparing bar plots for this period, we can conclude that for given data set the smallest traffic with captives were under flags of U.S.A. and Denmark/ Baltic and the highest occurred in that period for Netherlands (66999 embarked slaves, and 53993 disembarked slaves) as Table 2. shows, that is realistic conclusion specifically because US gained their independence a century later.

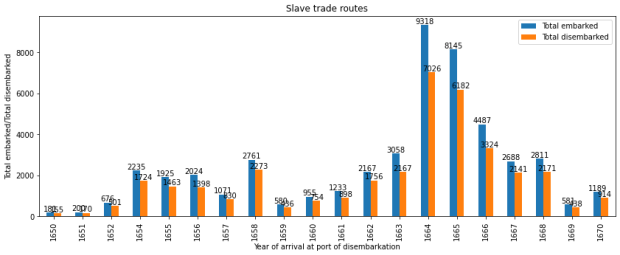


Figure 8. Great Britain (1650-1670)

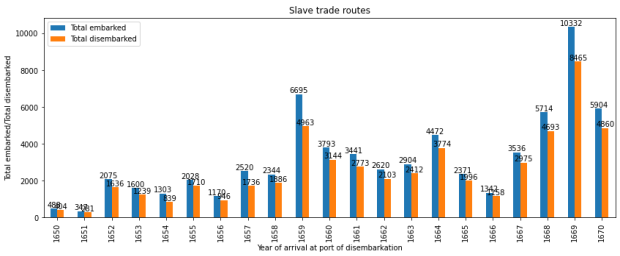


Figure 9. Netherlands (1650-1670)

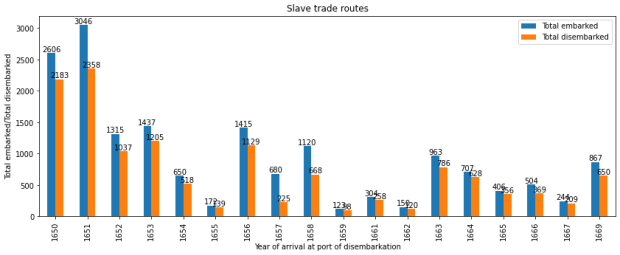


Figure 10. Portugal / Brazil (1650-1670)



Figure 11. Spain / Uruguay (1650-1670)



Figure 12. U.S.A(1650-1670)

TABLE 2. SUMMARY NUMBERS OF EMBARKED AND DISEMBARKED SLAVES THROUGH THE YEAR 1650 TO THE YEAR 1870.

Flag (imputed)	Total embarked	Total disembarked
Denmark / Baltic	653	544
France	2577	1982
Great Britain	48285	36721
Netherlands	66999	53993
Portugal / Brazil	16709	12936
Spain / Uruguay	6493	5305
U.S.A.	181	155
Unknown	2466	1804
Sum	**144363**	**113440**

For this data - in selected range of years, students can comment the trends in traffic for individual country (for example Portugal/Brazil had the highest traffic of slaves at the beginning of that period (1650) but at the end of that period was much lower and in decline (1670), while in contrast, Netherlands had at the beginning of that period the lowest traffic but till the end of the period the traffic grew). This comparison of traffic with slaves could also be a theme for discussion in classroom with including the history context of that period - what history circumstances in economy, society, politics, technology happened in each country for that specific period that had led to increase or decrease in slave trade traffic.

V. CONCLUSION

Connecting interdisciplinary research areas and their concepts in primary and secondary education and beyond give students and teachers a broader opportunity for knowledge-based learning using modern approaches. Teachers can use research methods to give students tasks from real world problems to develop a critical thinking that will prepare them for their professional future. As the world moves towards the digital age, students and teachers should adapt to the changes by using critical and computational thinking for their future jobs. One way of adapting can be done by introducing in educational curriculum such interdisciplinary content, as we show in this paper on the history example of Trans-Atlantic trade slave routes, a real history data set that is used for exploring history by introducing research questions that can lead to critical and computational thinking development. To increase the level of critical and computational thinking, more variables from this data set can be used in a particular research question in order to see increasingly complex historical facts through critical thinking, while respecting the limitations that historical sources (data sets and variables) and Data Science methods may have. For example, we could use in data set a variable for type of the ship that was carrying slaves to see if there is some correlation between died slaves or we can use the same logic for making conclusions about correlation between length of the voyage and died slaves.

With introducing more variables in data set related to research questions, because of null values in records in some fields (due to lack of history record details), the number of overall number of records could be reduced in which case we have to discuss validity of our data interpretation using critical and computational thinking. In this paper we propose a way of fostering critical and computational thinking in the field of primary and secondary education in non-STEM subjects by using data sets and their possible applications as a starting point. Although for this specific example of Slave trade routes, one could comment that maybe the better options would be usage of different kind of visualization tools in non-STEM classes, but the aim of this example is that teacher and student should gain knowledge of Data Science concepts such as cleaning and preparation of the data sets and building customized data sets that could be used in ready-made Python scripts on all educational levels even if they at the beginning don't have appropriate digital skills. Also, according to Digital Education Action Plan (2021-2027, https://education.ec.europa.eu/focus-topics/digital-education/action-plan), 39% of teachers don't use digitalized contents in classroom because of lack of digital skills so with prepared and detailed instructions, they should advance their digital skills in classroom by adjusting examples and expanding the given ones. Our future work will include expanding examples and data sets and Python scripts for history and for other non-STEM subjects, so modern and advanced primary and secondary curricula could implement them in future education.

REFERENCES

[1] Siburian, Jodion; Aloysius, Duran Corebima; Ibrohim; Murni, Saptasari. The Correlation Between Critical and Creative Thinking Skills on Cognitive Learning Results. Eurasian Journal of Educational Research 81, 2019, pp. 99-114

[2] Hinter, Amanda. Equipping Students for Success in College and Beyond: Placing Critical Thinking Instruction at the Heart of a General Education Program (Chapter 11) // Critical Thinking and Reasoning – Theory, Development, Instruction and Assessment (Fasko, D. & Fair F. Eds.), Brill Sense, Boston, 2020, pp. 188-208

[3] Pieterse, T.; Lawrence, H.; Nel, H. F. Critical thinking ability of 3rd year radiography students. Health Sa Gesondheid, Vol 21, 2016, pp. 381-390

[4] Bailin, S.; Battersby, M. Reason in the balance: An inquiry approach to critical thinking (2nd ed.). Cambridge, MA: Hackett, 2016

[5] Bailin, S.; Battersby, M. Inquiry: Teaching for Reasoned Judgment (Chapter 2) // Critical Thinking and Reasoning – Theory, Development, Instruction and Assessment (Fasko, D. & Fair F. Eds.), Brill Sense, Boston, 2020, pp. 46

[6] Cleghorn, P. Critical Thinking in the Elementary School: Practical Guidance for Building a Culture of Thinking (Chapter 9) // Critical Thinking and Reasoning – Theory, Development, Instruction and Assessment (Fasko, D. & Fair F. Eds.), Brill Sense, Boston, 2020, pp. 153

[7] Nurzengky, Ibrahim; Desy, Safitri; Sri, Nuraini; Taufik Rihatno, Edwita; Arita, Marini; Apri, Wahyudi. Introducing a Model of Student Critical Thinking Skills in History Education, International Journal of Advanced Science and Technology Vol. 29, No. 8s, 2020, pp. 88-96

[8] Machotka, Ewa; Papapetrou, P. Method in Interdisciplinary Research: Data Science for Digital Art History. International Journal for Digital Art History vol. 4, 2019, pp. 5.03-5.08

[9] Damerow, J.; Wintergrün, D. The Hitchhiker's Guide to Data in the History of Science. Isis, volume 110, number 3, 2019, pp. 513-521

[10] Neil Selwyn. Data entry: towards the critical study of digital data and education, Learning, Media and Technology, 40:1, 2015 64-82, DOI: 10.1080/17439884.2014.921628

[11] Eltis, D. A brief overview of the Trans-Atlantic Slave Trade. Voyages: The Trans-Atlantic Slave Trade Database, 2007, pp. 1700–1810.

[12] Steven J. et al. Genetic Consequences of the Transatlantic Slave Trade in the Americas, The American Journal of Human Genetics, Volume 107, Issue 2, 2020, pp. 265-277

Gamification in Education: Building an Escape Room using VR Technologies

A. Staneva*, T. Ivanova, K. Rasheva-Yordanova*, D. Borissova**
* University of Library Studies and Information Technologies, Sofia, Bulgaria
** Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences, Department of Information Processes and Decision Support Systems, 1113 Sofia, Bulgaria
a.staneva@unibit.bg, t.ivanova@unibit.bg, k.rasheva@unibit.bg, daniela.borissova@iict.bas.bg

***Abstract* - The rapid expansion of media diversity and the growing popularity of "virtual worlds", created the need and opportunity to apply various innovative teaching methods. There is a change in the needs of learners who are no longer in the role of passive listeners. Interactive games can significantly increase learning effectiveness, motivation, and student achievement and turn students from passive to active participants in the learning process. This paper examines the place of gamification in formal education and its role in supporting traditional learning. The purpose of the report is to present a gamified lesson based on virtual reality technologies. The article is focused on two points: (1) introducing innovation in the education of students at the high school stage, and (2) motivating students to think critically, synthesize available information, and solve problems. The paper presents a model and implementation of a VR Escape room as a support tool for acquiring new knowledge in a digital environment. The innovative lesson model is applicable in both traditional and electronic learning environments. The model is independent of the study subject and educational level but is adapted to the attitudes, expectations, and needs of students in high school education.**

Keywords - virtual reality; escape room; online learning; e-learning; distance education; serious games

I. INTRODUCTION

Transformations in the education system are most often provoked by rapidly changing circumstances on a global scale. In recent years, we have witnessed a permanent change in the e-learning model from asynchronous (via e-mail) to synchronous interactive learning (through the introduction of popular educational electronic platforms such as Moodle, Teams, etc.) [1]. According to Zunic et al, a catalyst for this process is the pandemic caused by COVID-19, which has significantly intensified the IT sphere of education, imposing e-learning as part of the measures to cope with the pandemic.

For a period of almost 2 years, for many of the training institutions, distance learning and working has turned out to be the only option. In a period of uncertainty, there was a need to search for appropriate solutions to ensure the adaptability of the educational process with a view to reducing inequalities, guaranteeing equal opportunities for quality education, ensuring the sustainability of acquired knowledge and skills and reducing the risk of students dropping out of school.

It is known that the ways of interaction between a teacher and a student directly affects the effectiveness and efficiency of the educational activity [2]. This case is especially relevant when it comes to training carried out through a digital channel. This interaction simultaneously involves technical indicators (such as Internet connection, access to a training platform, availability of technical means, including a computer, tablet, phone, etc.), non-technical indicators (educational resources and adapted to the form of training learning materials), as well as psychological indicators (motivation of learners and trainers).

Today, more and more teachers and teacher communities are beginning to use media technologies to facilitate the work process in a network and in the presence of asynchronous work [3]. On the other hand, the use of digital learning environments related to the creation of fascinating and engaging learning experiences lead to an increase in the use of opportunities for gamification of education and learning [4]. They in turn lead to a higher commitment of the students to the learning process, better acquisition of the material, while at the same time the training is implemented in a much more interesting and motivating way for them.

The aim of this paper is to show the possibilities of virtual reality (VR) technologies to support the learning process. The paper is organized into three relatively self-contained sections. The first section examines virtual education as a new form of interaction imposed by virtual reality (VR). The second section examines gamification as a possible and necessary approach for teaching high school students. The third section presents the construction of a VR escape room for the purpose of training students in the last year of high school education.

II. VIRTUAL EDUCATION – A NEW FORM OF INTERACTION IN THE EDUCATIONAL PROCESS

We most often associate e-learning with courses taught entirely online [5, 6], where the physical distance between teacher and student is irrelevant [7]. Usually, this form of learning takes place through an online platform [6, 8] for teacher-student interaction [6].

There is also the so-called hybrid learning approach that focuses on creating a cohesive learning experience combining face-to-face sessions with online learning materials and activities [8]. Along with the already

established practices imposed by electronic distance learning and the existing hybrid methods of interaction, we consider the presence of a third form combining the peculiarities of the previous two groups. We consider e-learning as a specific form of hybrid learning, where the simultaneous presence of a teacher and a student, or a teacher and a group of students in a common virtual space in real time is possible. For simplicity of expression, we will call this form of learning "VR education" or simply "virtual education".

A. Place of virtual education in modern conditions

Virtual education as a form of learning has the potential to partially or completely replace existing traditional learning. This is because in virtual reality the main advantages of traditional learning are preserved (we have real-time interaction and feedback). At the same time, features are added here that guarantee higher training efficiency. In the form of a real experience, virtual education will increase the interest and motivation of even the most disinterested and at-risk group for dropping out of the education system. On the other hand, virtual education guarantees real-time learning, stimulates team work, motivates the performance of individual tasks, based entirely on project-based learning.

It is important to note that in order to ensure an effective educational process, it is necessary to create a useful digital environment that meets the needs of learners and copes with the transforming challenges in education [9]. It has been proven that the use of innovative media increases students' interest in the learning process [10] and creates a beneficial learning atmosphere. Along with the existing technical limitations, we observe limitations imposed by purely psychological factors. The development of students' motivation in the conditions of distance learning is possible when creating an effective electronic educational environment, which in turn should provide quality content, interactivity and the possibility of individualizing the learning process.

Analyses show that distance learning is a challenge not only for students, but also for their teachers [8]. The reasons for this are complex, but are most often associated with: (1) the need to introduce new educational approaches to engage learners, (2) the need to develop new electronic educational resources adapted to the form of learning and (3) the need to provide a suitable environment for interaction between learner and trainer. The introduction of virtual reality in education will work simultaneously on all three indicators as: (1) it creates an innovative form of learning and interaction; (2) meets the conditions imposed by the learning environment and (3) provides an environment for interaction between a teacher and a student, between a teacher and a group of students, or only between a group of students in the process of their self-training.

The use of innovative media [11, 12] will increase students' interest in the learning process [12], create a fruitful learning atmosphere, and stimulate students' acquisition of more knowledge in a shorter period of time. This in turn leads to an increase in the effectiveness and quality of training.

Virtual reality, as an innovative media technology, allows the integration of gaming approaches. That is why VR finds its place in education, significantly increasing the efficiency of the educational process and giving the opportunity to upgrade, enrich and visualize the educational content. Moreover, virtual education has the potential to transform students from passive consumers of information into discoverers and creators of digital content.

III. GAMIFICATION AS PART OF VIRTUAL EDUCATION

In recent years, researchers as well as various industries have increasingly expanded their focus and interest from the pragmatic issues of human-computer interaction to include aspects such as user experience, emotion, satisfaction, and motivation [13]. One method that includes this concept is called "Gamification": the use of game design elements in a non-game context [14]. The goal of the gamification methodology is to stimulate user engagement.

Gamification in education is a modern method of organizing the learning process, using game elements for learning, which can attract students to the active learning process [15]. The integration of game approaches in learning affects the behavior of students, their commitment and motivation, which in turn leads to the improvement of the knowledge and skills acquired by them.

In the last decade, gamification has started to become popular and gain more and more attention in educational contexts. The development potential of educational games arises from their ability to teach and reinforce not only knowledge, but also important skills (including problem solving, cooperation and communication). Their motivational power is conveyed in a number of mechanisms to encourage people to engage with them – without reward, just for the joy of playing and the possibility of winning [16].

According to Kapp, gamification is "using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems." [17]. The specific factor here is that game metaphors, game elements and ideas are used in a context different from that of standard games.

The use of game mechanics improves students' abilities to acquire new skills [18], and gamification as a process allows students to become protagonists of their own learning process [19]. This advantage can be used in disciplines where there is little interest or lack of motivation on the part of learners.

A. Applying escape rooms to training

One emerging popular serious game format is the educational escape room [20, 21]. Escape rooms are increasingly used in all levels of formal, vocational and corporate adult education for multiple educational purposes [22, 23].

In practice, escape rooms are popular, interactive, group, live adventure experiences in physical spaces [24]. In the field of education, they can be presented as multisensory games that facilitate the perception of the

material, increase the ability to remember and accelerate the learning process, while allowing the achievement of educational goals. They create a sense of fun and relaxation in the learning process, and thanks to their mysterious format, escape rooms subtly develop memory, perception, ability to concentrate and contribute to the overall development of students.

Most often, the task of the participants is to perform a series of tasks of a different nature [25]. These can be puzzles, finding hidden objects, breaking codes, ciphers or opening padlocks. In them, every symbol, word or part of the environment can be significant and lead the participants to the end of the game. Exiting the room before the end of the specified time is perceived as equivalent to victory [24]. During a game, users are required to apply multiple skills, including analytical and critical thinking, problem solving, communication, teamwork and cooperation, ability to communicate in a group [26, 27]. The game also builds time management and stress skills – leading skills for certain fields of science.

It is noteworthy that digital educational escape rooms are mainly implemented in 2D and are considered a suitable tool to promote a deeper understanding of the material in online distance learning [22]. We believe that society is ready to embrace change through the introduction of both virtual education and virtual reality in education.

IV. GAMIFICATION OF TRAINING THROUGH A VR ESCAPE ROOM

There is currently a rapid development of a series of emerging exponential technologies including augmented reality [28]. and the metaverse. In recent years, there has been a particular interest in VR technologies, even in the field of education [29, 30]. Multi-user VR environments and virtual worlds enable the application of different forms of game-based learning [31].

The application of VR gamefication in the learning process will provide advantages among which: (1) a favorable environment for knowledge acquisition in the presence of low stress and the possibility of teamwork; (2) good opportunity for theoretical, practical knowledge and acquisition of new technological skills and last but not least (3) active self-learning that is critical and self-regulated with improved concentration and high motivation.

Our proposal in this direction is connected with the construction of an escape room realized through virtual reality technologies. We believe that its application in the educational process will lead to the following benefits:

- A higher level of commitment and motivation of players (students) to educational activities and processes.
- Gradual upgrading of knowledge in the process of passing the different levels in the game.
- Implementation of knowledge control by imperceptibly for the players (students) assessment of their incoming and outgoing skills.
- Encouraging student engagement and teamwork through collaborative and competitive work.

The lesson we use to build the escape room is interdisciplinary and combines knowledge from the applied fields of history, arts and informatics. The target group is students from the last high school stage, and the purpose of the resource is to exercise material related to Java Script programming. Based on the set goal, the game type is deep immersion, with a low-threshold approach to knowledge consolidation.

What is valuable for us in the design phase is to highlight the existing challenges for the learning resource design team. As a result of analysis, three main groups of challenges stood out:

- ***Pedagogical aspects*** - game design has the main mission of forming new knowledge and serving as an aid in the learning process. For this reason, the teaching methodology must be available, the purpose of the lesson must be clear. It is necessary for the designer to have sample scenarios of the traversal of the various scenes.
- ***Design aspects*** - as multimedia, an interactive product is created, it must meet certain design concepts and correspond to best practices in user experience. A well-designed graphic design would help build motivation, create engagement and lead to the achievement of the expected educational outcomes.
- ***Aspects related to the content*** – the educational resource must be provided with the necessary theoretical base and opportunities for its sharing and assimilation. A set of interactive elements is needed to provide learning information from the platform and at the same time provide mechanisms to request it in reverse order (from student to platform) in order to evaluate the acquired knowledge.
- ***Storytelling*** – allows interdisciplinarity and formation of new knowledge. Through storytelling, we introduce the players (students) into a real or fictional plot. The role of this element is fundamental – according to Lee Scheldon, it is here that the player‘s desire to discover what happens next in the game is unlocked. [32].

The design of the escape room follows a linear structure. Solving one clue will provide the object needed to solve the next clue, and so on until students can solve the quiz as a team and escape the room. Problems in a linear room must be solved in a certain order. The average game duration is 60 minutes. For the correct implementation of the tasks in the specified time, the emphasis is on the organization, communication and cooperation between the teammates. The route can be traveled in a different way, but its purpose is to test the knowledge of the participants. For this purpose, test riddles have been prepared for each step, based on the study material.

A. Storytelling and technical realization

The thematic Escape room with Java Script quizes has been named FunClassTomb. Students immerse themselves in the virtual cultural-historical world from the beginning of the 3rd century BC in the Thracian tomb located in the city of Kazanlak, Bulgaria. There the participants are "buried" together with the beloved wife of the ruler Roigos – the heir of Sevtus III or also Sevtes III (Thracian ruler from 330 BC to 300 BC and the last great ruler of the Odris Kingdom). The players' mission is to find the codes hidden in the mural and jewelry, decipher them and open the two doors of the underground tomb.

The game welcomes players with a screen featuring a map of the tomb and the puzzles to be solved. The musical background is Thracian music, and the start of the game is counted down by a timer that is activated by the start button. This initial screen of the game is shown in Fig. 1.



Figure 1. Home screen of the escape room

The game can be for both teams and for achieving individual results. At the entrance, the entry-level participants are tested using a closed questionnaire of 4 questions. Only those who pass the test are given information on how to enter the tomb. Successfully solving the main quiz unlocks the next level. The answer to each of the quizes is based on several keys in the corresponding level, built using hotspots. The general scheme that follows the educational escape room is presented in Fig. 2.

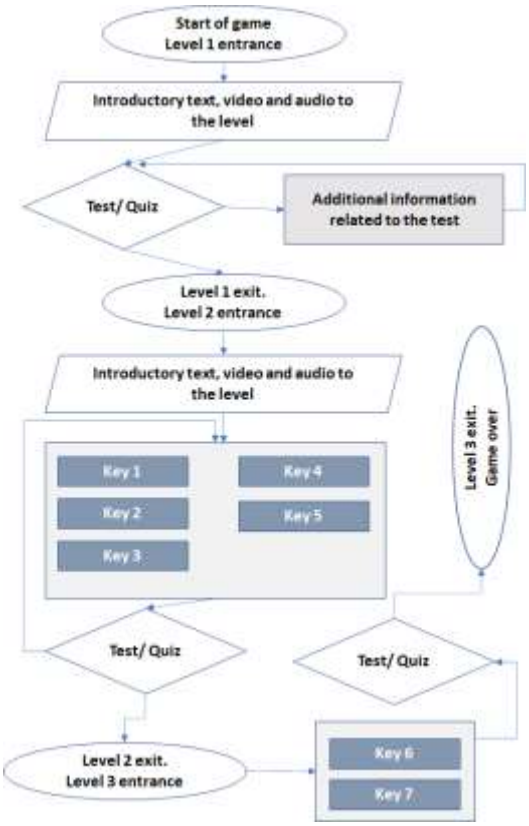


Figure 2. General scheme of the escape room game

It is worth mentioning that the current game is realized on several levels. These levels are related to the level of knowledge required. Appropriate study materials including text, video, audio, and relevant questions have been implemented for each level. Each successfully completed level allows entry to a higher level of the game.

The time limit to solve all the quizes and get out of the tomb is 45 minutes (equivalent to one lesson). In case the time runs out before solving the puzzle, the game ends with the message "Timeout". Since the room was created for educational purposes, the possibility of replaying the game has been provided. The screen with such a message is shown in Fig. 3.

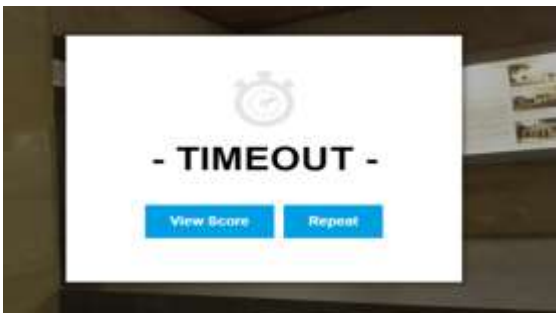


Figure 3. Timeout screen

In the course of the game, the student alone or in a team must correctly complete all the tests and collect all the keys to the levels.

The main keys (archaeological objects) serve to assist the player during solving the quiz. Discovery of the keys

does not necessarily occur in a fixed sequence. They include information that forms new knowledge needed to complete the game, but also indicate a letter that is involved in forming the answer to the final quiz. Each level ends with a specific puzzle, the solution to which leads to the entrance to the next level. The total number of levels is 3 (Fig. 2).

To achieve a realistic feeling and full immersion, VR is used, built on the basis of 360-degree images. The screen of the final product is adaptive: the game automatically adapts according to the size of the device used by the student. Also, it doesn't matter if the player is viewing the picture horizontally or vertically.

Quizzes are hidden by hotspots, and upon discovering their location, a Quiz appears with questions or a key to the final quiz. In order to acquire new knowledge, but also for better immersion in the storytelling of the game, hidden clues have been created (through HotSpot points), each of which opens multimedia (text fields, audio and/or video information). Such hotspot points are visualized in Fig. 4.



Figure 4. Distribution of hotspots at Level 2

It should be mentioned, that the use of virtual reality allows the created elements of the puzzle to be changed regularly, without the need for a complete change. In this way, the reuse of already defined elements is obtained.

Successful completion of all tasks leads to the end of the game and displays a message about the achieved result (see Fig. 5).

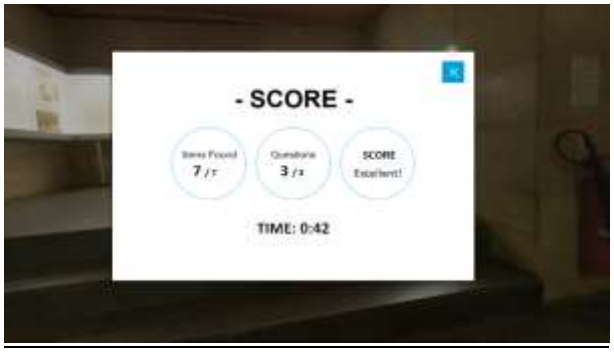


Figure 5. Successful completion of the game screen

The game presented in this section has a demonstration character and combines knowledge from several academic disciplines. In the course of the pilot test, satisfactory results were reported, meeting our preliminary expectations. We observed higher motivation on the part of the learners and more recent results of the learning process, reported already during the game. The created example demonstrates that gamification through escape rooms is applicable and certainly could be applied to other subject areas of the educational system regardless of the age groups.

V. Conclusion

Virtualization in education is a process that has already begun. By introducing VR, the learning process will become more dynamic, more efficient, more interesting, and much more economical. The implementation of VR technologies benefits students, teachers, educational institutions, and the learning process itself. Teachers need to quickly transfer traditional classrooms to virtual ones, replace traditional teaching methods with interactive ones and reorganize the learning process in the form of research games.

The described approach in the current article can be seen as working in the right direction toward the coming age of active learning. On the other hand, virtual reality adds additional advantages related to the learned objects. The proposed model and implementation of VR Escape room as a tool for acquiring new knowledge in a digital environment is implemented on the example of a Thracian tomb located in the city of Kazanlak, Bulgaria. The used tools for the realization of such educational game based on thematic escape room using VR Technologies rely on JavaScript and HTML. Therefore, this approach can easily be used in any school as it only requires the presence of a browser.

Future directions for the development of such themed escape rooms involving VR technologies are related to the development of both appropriate questions and determining their difficulty to engage with an appropriate level of educational play.

Acknowledgment

These research findings are supported by the National Scientific Research Fund, Project N КП-06-Н67/1.

References

[1] L. Zunic, F. Zunic, A. Džaferagić. Beneficije e-nastave. Educa . Dec2022, Vol. 15, p79-85. 7p. 2022

[2] N. Sudarwati, R. Rukminigsih. Evaluating E-Learning as a Learning Media: A Case of Entrepreneurship E-Learning Using Schoology as Media. International Journal of Emerging Technologies in Learning (iJET), 13(09), pp. 269–279. https://doi.org/10.3991/ijet.v13i09.7783, 2018.

[3] S. A. Smith. Virtual reality in episodic memory research: A review. Psychon. Bull. Rev. 2019, 26, 1213–1237.

[4] D. Borissova, T. Ivanova, N. Buhtiarov, N. Naidenov, K. Rasheva-Yordanova, R. Yoshinov, M. Garvanova and I. Garvanov. Application of Information Technology in the Teaching of Mathematics when Study of 2D Geometric Shapes". The 45th International ICT Hybrid Convention MIPRO-Croatian Society for Information, Communication and Electronic Technology, May 23 - 27, pp. 638 - 643, 2022.

[5] C. J. Brahler. Technology Paved the Road for Students in a High-School Dropout Recovery Program to an Online College Class. Handbook of Research on Learning Outcomes and Opportunities in the Digital Age. 2016, DOI: 10.4018/978-1-4666-9577-1.ch026

[6] R. M. P. Salazar, G. R. T. Arcos, K. Á. Á. Cadena. Academic E-Branding Focused on Higher Education Institutions: Analyzing Digital Visibility in Latin American and European Universities.

Improving University Reputation Through Academic Digital Branding. 2021, DOI: 10.4018/978-1-7998-4930-8.ch006

[7] J. K. Hope. Could Educational Technology Replace Traditional Schools in the Future? Encyclopedia of Information Science and Technology, Fourth Edition. 2018, DOI: 10.4018/978-1-5225-2255-3.ch211

[8] L. Addimando. Distance Learning in Pandemic Age: Lessons from a (No Longer) Emergency. International Journal of Environmental Research and Public Health, 2019. 19.16302.10.3390/ijerph192316302.

[9] A. M. Toda, R. M. C. do Carmo, A. P da Silva, I. I. Bittencourt, S. Isotani. An approach for planning and deploying gamification concepts with social networks within educational contexts. International Journal of Information Management. 2018 https://doi.org/10.1016/J.IJINFOMGT.2018.10.001

[10] S. Islam, H. Baharun, C. Muali, M. I. Ghufron, M.M.E.I. Bali, M. Wijaya, I. Marzuki. To Boost Students' Motivation and Achievement through Blended Learning. Journal of Physics: Conference Series, 1114(1), 1–11. 2018. https://doi.org/10.1088/1742-6596/1114/1/012046

[11] M. Bali, Musrifah. The Problems of Application of Online Learning in the Affective and Psychomotor Domains During the Covid-19 Pandemic. Jurnal Pendidikan Agama Islam, 17(2), 137–154. https://doi.org/https://doi.org/10.14421/jpai.2020.172-03, 2020

[12] M. Muhammad Mushfi El Iq Bali, H. Baharun, A. Madanibillah, C. Muali et al. Innovative Learning Media Based on e-Learning in the New Normal Era. Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore, March 7-11, IEOM Society International, 2021.

[13] S. Deterding. Situated motivational affordances of game elements: A conceptual model. CHI Gamification Workshop, 2011.

[14] S. Deterding, D. Dixon, R. Khaled, L. Nacke. From Game Design Elements to Gamefulness: Defining "Gamification". CHI Gamification Workshop 2011. Available from: doi: 10.1145/2181037.2181040, 2011.

[15] K. Seaborn, D. I. Fels. Gamification in theory and action: A survey. International Journal of Human Computer Studies, 74, 14-31. https://doi.org/10.1016/j.ijhcs.2014.09.006, 2015.

[16] K. Kapp. Games, Gamification, and the Quest for Learner Engagement. T+D, 66(6), 64-68, 2012.

[17] K. Kapp. The gamification of learning and instruction: Game-based methods and strategies for training and education, 2012.

[18] V. Giang. Gamification" Techniques Increase Your Employees' Ability To Learn By 40%. Retrieved from Business Insider, 2013.

[19] D. N. Ardoy. Metodologías activas y gamificación en educación secundaria [Active methodologies and gamification in secondary education]. In A. García de Alcaraz (Ed.), El futuro de la educación física: Nuevos enfoques y aplicaciones metodológicas [The future of physical education: New approaches and methodological applications] (pp. 95-122). University of Murcia, 2017.

[20] M. Grande-de-Prado, S. García-Martín, R. Baelo, V. Abella-García. Edu-Escape Rooms. Encyclopedia 2020, 1, 12–19, 2020

[21] M. Stohlmann. Mathematical digital escape rooms. School Science and Mathematics. 10.1111/ssm.12564, 2023.

[22] A. Makri, D. Vlachopoulos, R. A. Martina. Digital Escape Rooms as Innovative Pedagogical Tools in Education: A Systematic Literature Review. Sustainability, 13, 4587, 2021A. Staneva, K. Rasheva-Yordanova, D. Borissova. Integration Multimedia and Virtual Reality in the Online Teaching of Fine Arts.In:DIPP'2022, pp. 2535-0366, ISSN: 1314-4006 (Print), eISSN: 2535-0366 (Online), 2022.

[23] V. Adams, S. Burger, K. Crawford, and R. Setter. Can you escape? creating an escape room to facilitate active learning. J. Nurses Prof. Develop., vol. 34, no. 2, pp. E1–E5, Apr. 2018.

[24] S. Nicholson. Creating Engaging Escape Rooms for the Classroom. Child. Educ, 94, 44–49, 2018.

[25] A. Manzano-Leon, J.M. Rodriguez-Ferrer, J.M.Aguilar-Parra, A.M. Martinez, A. Luque de la Rosa, D. Salguero Garcia et al. Escape rooms as a learning strategy for special education Master's degree students. Int J Environ Res Public Health. 2021;18(14):7304–321

[26] H. E. Vidergor. Effects of digital escape room on gameful experience, collaboration, and motivation of elementary school students. Computers & Education, nr 166, s. 1–14, 2021.

[27] S. Stollhans. Designing escape game activities for language classes, (w:) Plutino A., Borthwick K., Corradini E. (red.), Innovative language teaching and learning at university: treasuring languages, s. 27–32. Online: https://www.pedocs.de/volltexte/2021/21845/pdf/Stollhans_2020_Designing_escape_game.pdf [DW 09.08.2022].

[28] S. Misra, C. Roy, A. Mukherjee. Introduction to Industrial Internet of Things and Industry 4.0; CRC Press: Boca Raton, FL, USA, 2021; ISBN 9781003020905.

[29] N. Pellas, S. Mystakidis, I. Kazanidis. Immersive Virtual Reality in K-12 and Higher Education: A systematic review of the last decade scientific literature. Virtual Real. 25, 835–861, 2021.

[30] N. Yee, J. N. Bailenson, N. Ducheneaut. The Proteus Effect: Implications of Transformed Digital Self-Representation on Online and Offline Behavior. Commun. Res. 36, 285–312, 2009.

[31] N. Pellas, S. Mystakidis, A. Christopoulos. A Systematic Literature Review on the User Experience Design for Game-Based Interventions via 3D Virtual Worlds in K-12 Education. Multimodal Technol. Interact, 5, 28, 2021.

[32] L. Sheldon. The Multiplayer classroom designing coursework as a game, Second edition, CRC Press, 2020

The Ubiquity of Mathematics in the Practical Basics of Computing and Programming

M. Bednjanec, D. Bednjanec and B. Kuhar
Elektrotehnička škola, Zagreb, Croatia
magdalena.bednjanec@skole.hr, doroteja.bednjanec@skole.hr, biljana.kuhar@skole.hr

***Abstract* - Most students of secondary technical schools find it difficult to learn the material from subjects that require logical and abstract thinking. Such students often resist and struggle to master material that requires this type of thinking, such as, for example, geometry. When acquiring content from programming subjects, for such students, abstract and logical thinking does not represent such an obstacle in solving tasks as is the case with mathematics. In practical subjects, such as the practical basics of computing, the correlation with mathematics did not give them problems. The research was conducted using a questionnaire containing 15 open and closed questions. The respondents are students of a high school that educates students in electrical engineering and computer science.**

The goal of the research was to investigate the cause of such students' approach to mathematical tasks and to find out which teaching methods and teaching aids would help students in presenting abstract tasks. The conclusion is that students are more motivated when solving tasks by writing computer programs because running the written code makes it easier for them to make logical connections and abstract thinking. Proposals for teaching methods and teaching aids for teaching mathematics were also made to facilitate students' acquisition of teaching content.

Keywords* - *mathematics, programming, teaching topics, students, algorithm

I. INTRODUCTION

The article describes the research on the causes of the difficult acquisition of mathematical material where logical and abstract thinking is required, as well as the application of different teaching methods and aids for presenting tasks. The research was carried out in a high school where students are educated in various technical professions.

We paid the greatest attention to Computer Technicians.

With the new professional curriculum in 2017 for obtaining the qualification of Computer Technician in the educational sector of Electrical Engineering and Computing, Workshop exercises were abolished and programming courses were added. It was noticed that after that year there are minor problems with the acquisition of mathematical topics. Students often complained in math classes about abstract concepts like vectors. In the lessons on algorithms and programming, students could see the application of vectors such as movement coordinates in an example of coding a computer game. After several such examples, the students noticed that their mathematical knowledge is the basis for other subjects at school.

After mutual communication between experts in mathematics and programming, it was decided to implement new teaching methods and aids when presenting and solving mathematical problems in mathematics. At the same time, students in the second grade solved mathematical topics on algorithms and mathematics during the same period.

The results were achieved with the positive motivation of the students and the desire to understand mathematical topics that are useful to them in programming.

We followed the students for five months while teaching the topics of quadratic equations, quadratic inequality, and quadratic function in the second grade of Computer Technician at the School of Electrical Engineering in Zagreb.

When acquiring new knowledge from the topic of quadratic equations, quadratic inequalities, and then quadratic functions, students have behind them nine years of school mathematics, acquired knowledge and skills that need to be upgraded in such a way that the students themselves are motivated for knowledge and have the will to work.

For students to learn the material more easily, according to Mayer, the following principles should be followed. The dual-channel hypothesis holds that learners can divide information into two distinct cognitive structures, the visual channel, and the aural channel. The second supposition is that one channel's capacity for processing information is constrained. To avoid a cognitive overload, it is advantageous for learning settings to promote the engagement of both the visual and auditory channels. This can be accomplished, for instance, by combining spoken texts or spoken images with written texts or visual visuals. The third supposition holds that to absorb new information, students must actively interact with the learning material [1],

Before starting the research, the students learned how to solve quadratic equations in mathematics class. General equation solving is the basis of solving all problems, so the goal is for students to apply equation solving to real-life problems. The problem tasks are reduced to the text of solving life problems, however, the biggest problem was reduced not to set but to solving equations. They independently created algorithms for solving quadratic equations.

The research included second-grade students of the Electrical Engineering School in Zagreb.

Participation in the research was anonymous and voluntary, conducted through a Google Forms survey. Students were asked 15 survey questions, 5 open-ended and 10 closed-ended.

On three closed-ended questions, students could justify their answers. The answers to the questions indicate their own choices that the students had at the time of filling out the survey. 45 students participated in the survey.

In response to the first question, which read "Knowledge of programming helps me to learn the teaching contents of mathematics: " the respondents had to indicate their level of agreement with this statement. The answer can be seen in Figure 1.

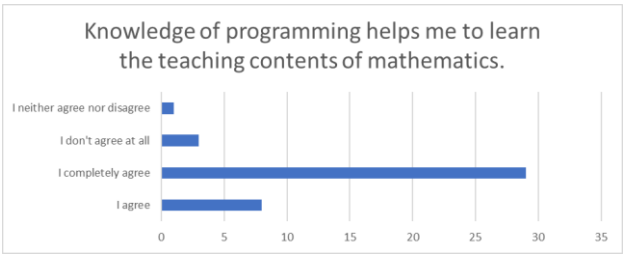


Figure 1. Knowledge of programming helps me to learn the teaching contents of mathematics

The next question, also of closed type, related to the visualization of mathematical problems, read: "Visualizing mathematical problems helps me to solve tasks." The answers offered were "Yes" and "No".

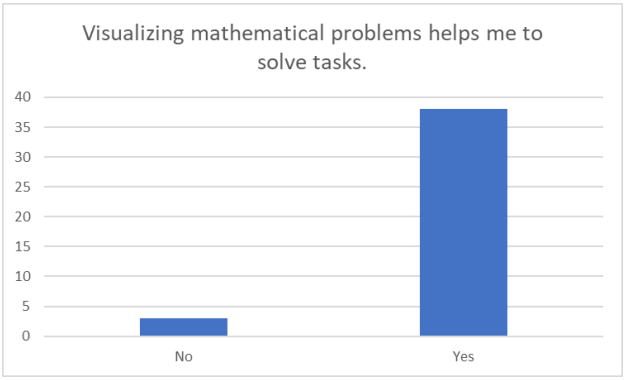


Figure 2. Visualizing mathematical problems helps me to solve tasks

Figure 2. shows that 38 respondents answered Yes while 3 answered No.

Then followed another closed question, "I often use online tools and applications to solve mathematical problems more easily." The ratio of "Yes" and "No" answers was almost equal, as can be seen in Figure 3.

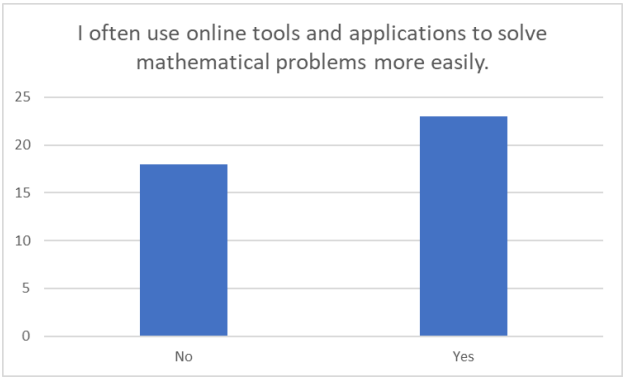


Figure 3. I often use online tools and applications to solve mathematical problems more easily

Figure 4. shows the ratio of respondents who often use online tools and applications when solving mathematical problems and those who do not.

The fourth question related to the method of solving the quadratic equation. The answers were offered: formulas, factorization, calculators, PhotoMath, Desmos, programs in C, I don't solve.

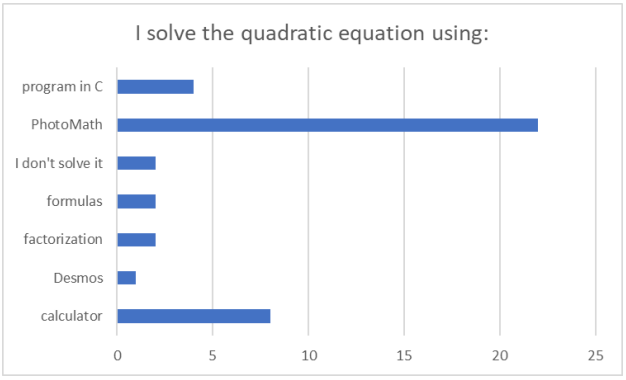


Figure 4. I solve the quadratic equation using...

Figure 5. The graph shows the percentage of respondents regarding the survey question: "I solve a quadratic equation using..."

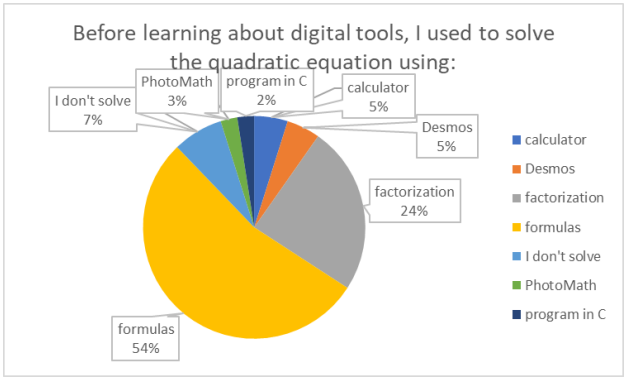


Figure 5. Before learning about digital tools, I used to solve the quadratic equation using…

II. DIGITAL ENVIRONMENT

The digital environment consists of all ways of using computers and the Internet. These include, among others,

smartphones, computer games, online tools for solving tasks, and social networks.

According to [2], the study of student problems takes place within natural science and computer, informatics subjects according to the approved curriculum of the Ministry of Science and Education. During the research, the students were introduced to a smart calculator with which they can solve quadratic equations of any type[3]. The students were introduced to the PhotoMath application, which automatically displayed the procedure for solving and solving the given quadratic equation. Continuing with the mathematical topics of quadratic equations and quadratic inequalities, the quadratic function and its graph follow. Understanding a quadratic function, drawing it, and researching it takes time, so students are allowed to use a calculator, tablet, mobile phone, or computer to solve the quadratic equation as quickly as possible.

After the students were taught to use numerous online tools and applications to solve math problems, they were again asked the fourth Question. Unlike last time, the results turned towards digitalization, as we can see from Figure 2. Most of the students use a calculator, some use a mobile phone with PhotoMath, and a very small number stayed on the formula for the quadratic equation.

Figure 6. The graph shows the percentage of respondents regarding the survey question: "I solve a quadratic equation using..."

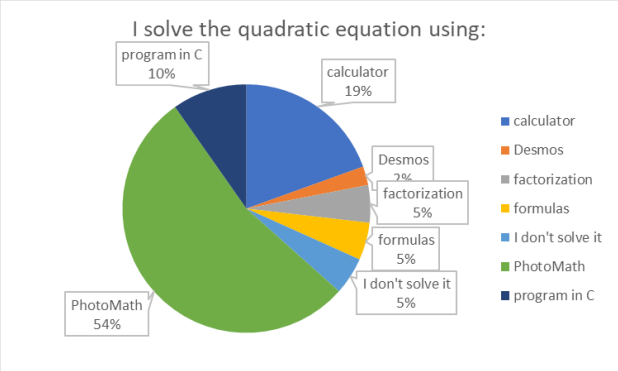


Figure 6. I solve the quadratic equation using...

We divided the students into teams when drawing graphs of quadratic functions and their research.

Students are divided into six teams:

team A with tasks and solutions on a piece of paper, team B with tasks using Geogebra, team C with tasks on a piece of paper and a calculator, team D with tasks on a piece of paper and mobile phones, team E without tasks - creative creation of tasks with a calculator, team F without tasks - creative creating tasks with a mobile phone. The eighth question is open-ended and related to working in teams, respondents had to fill in how many tasks they were able to solve together (as a team) during one hour, considering which team they belonged to.

It turned out that most tasks were solved in teams that independently created tasks and solved them. Students were surveyed after working in teams.

Teams that solved tasks using PhotoMath, calculators, or Geogebra solved more tasks than students who did not use technology, while students without digital tools had the lowest number of solved tasks.

Figure 7 shows the number of solved tasks by teams.

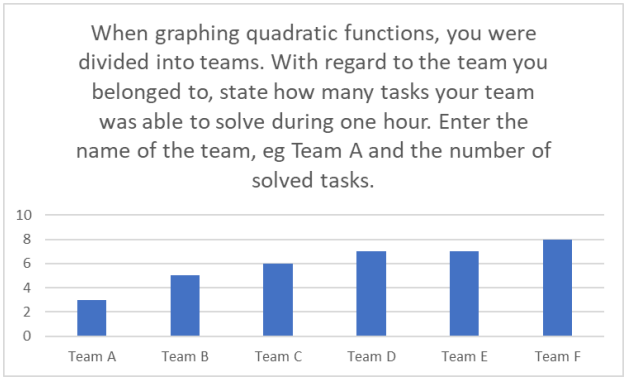


Figure 7. Tasks solved by teams

A Kahoot quiz was used to check the understanding of mathematical topics. As the survey showed, the quiz within the class also showed that students who use a calculator or mobile phone find solutions more easily and understand the teaching topics better. According to [4], the motivation to use digital technologies is increased because we all live in the digital age and behave that way.

The ninth and tenth questions refer to the application of programming in C as a motivation for solving mathematical problems.

Thanks to the excellent cooperation and communication of science teachers and vocational teachers, at the same time, students in the Algorithms and Programming course used the discriminant of a quadratic equation, which they had difficulty connecting with the quadratic function in mathematics. [5] In the course Algorithms and Programming, students learned the concept of function in programming. Their task was to write a function called Discriminant, which receives as arguments the coefficients of the quadratic equation (a, b, and c) and calculates the discriminant according to the entered parameters. After that, the program should have printed a message about how many solutions the quadratic equation with such parameters has. By running the program multiple times and entering different parameters a, b, and c, the response of the program was different depending on the calculated discriminant. Using the example of that task, the students understood the application of the quadratic function.

In the article, research was conducted on the use of digital technologies, teaching methods, and means for improving the understanding of teaching topics in mathematics.

The eleventh and twelfth questions refer to the use of teaching methods in teaching mathematics. The eleventh question read: "In mathematics classes, my favorite teaching method is:", and the answers offered were individual work, work in pairs, teamwork, or work in groups.

The respondents prefer teamwork, which is evident from the graph in Figure 8.

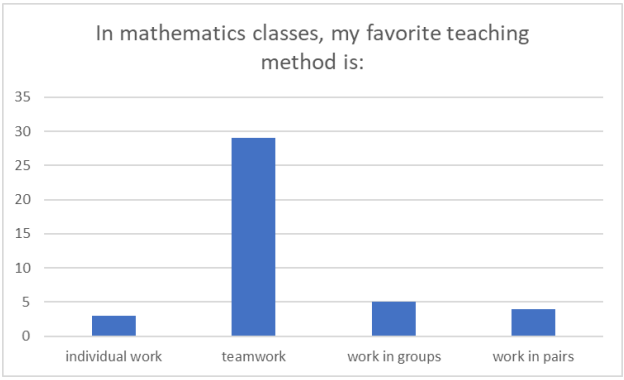


Figure 8. Favorite learning methods

The thirteenth question was open-ended and read: "Which digital tool for solving mathematics problems do you use most often?"

Respondents most often use PhotoMath, while they use Desmos a little less, which is visible from the graph in Figure 9.

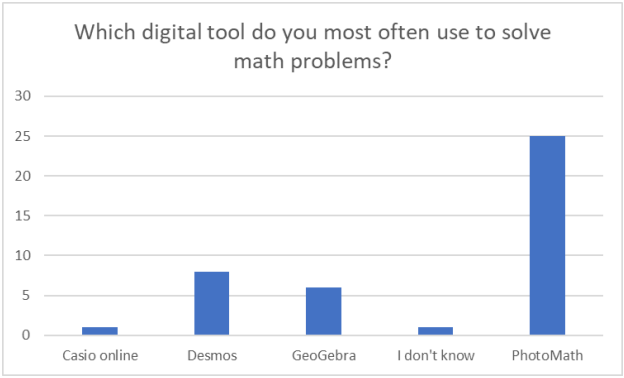


Figure 9. Digital tools

The fourteenth and fifteenth questions were related to the assignment of project tasks in mathematics classes. It was concluded that students prefer project tasks in mathematics because they stimulate the spirit of inquiry, are interesting, and always learn something new from them.

III. Research Evaluation

The research included second-grade students of the Electrical Engineering School in Zagreb this article aims to show an example of good practice and to encourage other teachers to work on cross-curricular topics because they make it easier for students to learn teaching content.

.Participation in the research was anonymous and voluntary, conducted through a Google Forms survey. Students were asked 15 survey questions, 5 open-ended and 10 closed-ended.

On three closed-ended questions, students could justify their answers. The answers to the questions indicate their own choices that the students had at the time of filling out the survey. 45 students participated in the survey.

In general, although it is not required, prior arithmetic knowledge might be helpful when studying programming. Logic, abstract thought, problem-solving, and an understanding of mathematical concepts like algebra and geometry—which are also utilized in programming—are just a few of the abilities and knowledge that learning mathematics can provide you.

To learn to program, you do not need to have any prior arithmetic expertise, though. Since programming is a distinct ability, it may be learned and understood without prior mathematical understanding. Students can learn programming using a variety of techniques, such as visual-based tools, and block programming, or by starting with simple languages with fewer mathematical prerequisites, like Python or Scratch.

On the other hand, programming, and mathematics frequently support one another. Students can use programming to apply mathematical ideas to real-world issues like computation, data analysis, statistics, or simulations. Additionally, mathematics can be helpful in more complex programming disciplines like algorithms, cryptography, machine learning, and video game development.

The ease or difficulty of learning programming ultimately depends on a variety of variables, including the student's particular aptitude, learning preferences, hobbies, and motivation. Although it can be helpful, prior mathematical knowledge is not required to learn to program.

References

[1] R.E. Mayer (Ed.), The Cambridge Handbook of multimedia learning (2nd ed.), Cambridge University Press, New York, NY (2014), pp. 31-48

[2] Gjud, M. and Popčević, I. (2020). DIGITALIZATION OF TEACHING IN SCHOOL EDUCATION. Polytechnic and design, 8 (3), 154-162. https://doi.org/10.19279/TVZ.PD.2020-8-3-04

[3] https://repozitorij.pmf.unizg.hr/islandora/object/pmf%3A10312/datastream/PDF/view

[4] https://op.europa.eu/hr/publication-detail/-/publication/f3bd0532-0255-11ed-acce-01aa75ed71a1

[5] Radovan, Aleksander & Mihaljevic, Branko. (2016). A proposal for the optimal use of mathematics materials for improving programming skills.

From Classroom to Online Environment – The Comparison Analysis of the E-Learning Standards Before and during the COVID-19 Pandemic

K. Aleksić-Maslać, P. Vranešić and B. Sinković

Zagreb School of Economics and Management/Department of Management, Entrepreneurship and Digital Transformation, Zagreb, Croatia

kaleksic@zsem.hr
pvranesi@zsem.hr
bsinkovi@zsem.hr

***Abstract* - The Zagreb School of Economics and Management (ZSEM) has systematically used e-learning in teachings since its establishment in 2002, which has greatly helped in maintaining the quality of education even during the crisis that was caused by COVID-19. In order to systematically follow the quality of e-learning, ten standards were developed that are divided into four groups: static, dynamic, administrational and other. In this paper, the evaluation results of 134 courses will be analyzed of the undergraduate level of three academic years: 18/19, 19/20 and 20/21 in 4 groups of students - the Economics and Management groups in Croatian and English, the Business Mathematics and Economics group and Elective courses group. Regardless of the difference in the quality of the developed courses, all groups behave similarly and have an increase of 13-17% in regards to quality standards during the COVID-19 pandemic. The static and administrative standards have unsignificant changes, while the dynamic and other standards have a significant rise. The analysis of individual standards shows that the usage of certain elements has highly increased. This example of good practice shows how, with good preparation and continuous improvement, it is possible to move from classical to online classes without major problems.**

Keywords – E-learning standards, COVID-19, LMS, student, quality

I. INTRODUCTION

Based on the UNESCO data regarding COVID-19, the pandemic has hit over 1.5 billion students [1] which needed to transfer from the classical or hybrid way of learning towards online education. Although some institutions already had high standard criteria and used some sort of e-learning, this was still a concern for the majority of institutions over the world [2, 3, 4]. Thankfully, the transfer towards e-learning education was easier for institutions that were already open to the technologies that enabled distance learning. Student readiness to transfer to online lectures depends on so called mental condition and the variety of elements such as motivation, technical skills, equipment, etc. [5, 6], [2], and there are a number of different researches that shows how students are aware of the positive sides that e-learning brings and in the majority of cases, the consensus is that it shows general satisfaction on transferring to e-learning and online classes during lockdown [7, 8, 9, 10].

Figure 1 shows the connection between distance learning, e-learning and online learning [11]. Although they are considered as different aspects, in majority of cases they are considered the same – just as it was perceived during the lockdown.

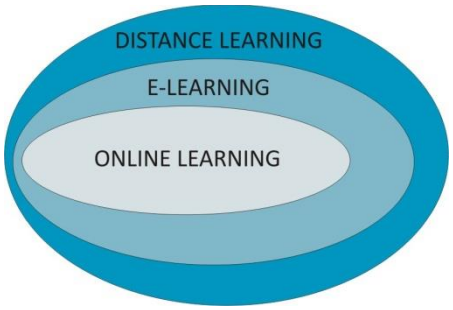


Figure 1. Connection between online learning, e-learning and distance learning

Learning Management System (LMS) plays a very important role as it is used for sharing different lecture materials, diversified assignments and ongoing communication between the student and the professor, but also between the student peers as well [12, 13, 14]. All LMS's have the same SCORM (Sharable Content Object Reference Model) standard, which is a model that accepts different interconnected technical specifications and instructions in designing educational content which is used for learning through the web. Every SCORM surrounding needs to indulge six basic requests: interoperability, accessibility, reusability, durability, maintainability and adaptability [15].

II. QUALITY MANAGEMENT OF E-LEARNING COURSES ON ZSEM

Zagreb School of Economics and Management has systematically used e-learning in education since its foundation of 2002 [16]. It is lecturers' obligation, which is in harmony with developed standards, to maintain and constantly develop e-learning courses on the LMS. Also, it is students' obligation to track all announcements and to do all of the needed assignments related to the enrolled course.

ZSEM has used WebCT as its first LMS, after which it transferred to Blackboard from 2011 to 2015. Since that time, the main LMS has been, and still is, Loomen, a Moodle version created and offered by CARNet (Croatian Academic and Research Network), which also offered a significant technical support throughout the years [17]. As lecturers and students on ZSEM are very familiar with new technologies, there are a lot of other elements that are used to enhance the online experience, such as intensive usage of gamification [18, 19], flip-classroom [20], different simulations and other technological possibilities. It's because of this readiness and well-established framework that made lockdown and transfer to distance learning easily shifted on ZSEM [21, 22].

Table 1 shows ten standards that were developed in 2006 on ZSEM and since that time, the team for evaluation of e-learning courses has been conducting yearly evaluations every academic year. Throughout the years, the standards have been changing and adapting, depending on the LMS usage and other factors [16]. The standards can be divided into four main groups:

- Static standards which are connected to main elements such as lectures that need to be updated regularly, syllabus and the overall design of the course on the LMS.
- Dynamical standards which are connected to the communication of lecturers and students through discussions, professor-student, student-student [23], different types of tests and announcements on the course.
- Administrative standards which are connected to maintaining the data base of students and lecturers.
- Other standards which are connected to additional elements such as synchronous and asynchronous distant learning, gamification, assignments, etc.

Faculty was trained to use the system with every transition, and all learning materials were centrally transferred to a new LMS. Joint workshops are held among the lecturers after the yearly evaluations overview in which experiences are shared and new ideas brainstormed. Online quizzes were already used live in computer lab during pre-COVID period, therefore question banks developed for years contained significant number of entries.

Students are trained to use the LMS from the very beginning of their program, making it mandatory tool for fulfilling their assignments, following lectures, and communication with professors and colleagues.

Announcements regarding the lectures, assignments and exams are published on LMS, with email notifications automatically forwarded. Majority of the assignments are submitted using LMS, and some courses have also implemented the grading system.

TABLE I. STANDARDS FOR ONLINE COURSE QUALITY EVALUATION

STANDARD	DESCRIPTION
S1 – Syllabus (max – 10)	Syllabus needs to be regularly updated. – 5 Syllabus needs to be on the title page. – 5
S2-Lectures (max – 10)	There needs to be a folder Lecture with learning materials. – 5 Lectures need to be regularly updated. – 5
S3 – Design (max 10)	Pages are well organized, intuitive and easy to use. – 5 Pages are well designed and visually attractive. – 5
D1 – Calendar (max – 5)	Used for important events such as mid-terms, demonstrations, variety of assignments, etc. – 5
D2 – Notice (max – 5)	Regularly used announcements for news about the course. Not sending attachments but hyperlinks. - 5
D3 Discussion (max – 10)	Discussion is used for lecture materials and for communication on all levels – professor-student, student-student and student-professor. - 5 Developed discussions (at least 5 topics with min. 10 posts per topic) – 5
D4 – Online Tests (max 10)	Online quizzes used for homework, knowledge check, mid-term and exam periods. – 5 Developed question banks and online exams. – 5
A1 – Number of Students – 5	Regular update of student database. Tracking designer roles on the course. – 5
A2 Self-registration- 5	Enrollment key setup and enabled self-registration for students. – 5
O – Other (up to 10)	Synchronous elements used for distance learning (Zoom, Meet, Teams, etc.) - 5 Recorded lectures and adapted lectures for asynchronous distance learning (ADL). - 5 Gamification Assignments, etc.

III. RESEARCH RESULTS

The evaluation of e-learning courses has continuously been conducted since 2006 and at the end of each academic year, they are presented during the lecturers' council, and the feedback discussed is archived on a Moodle Loomen course "Notice board for lecturers". The importance of the evaluation of e-learning courses is shown through an initiative which is based on a strategic plan and states that no more than 10% of courses should be below the 50% average.

A. *Analyzed sample*

Table 2 shows the sample of N=398 analyzed courses before and during pandemic, while Figure 2 shows the distribution based on analyzed standards on N=134 courses in the academic year of 18/19 (before the pandemic), N=131 in the academic year of 19/20 (with second semester being during pandemic) and N=133 in the academic year of 21/22 (both semesters during pandemic).

TABLE II. ANALYZED SAMPLE OF N=398 COURSES

	18/19	19/20	20/21	Total
EM cro	43	43	43	129
EM eng	43	43	43	129
BME	24	23	24	71
Elective	24	22	23	69
	134	131	133	**398**

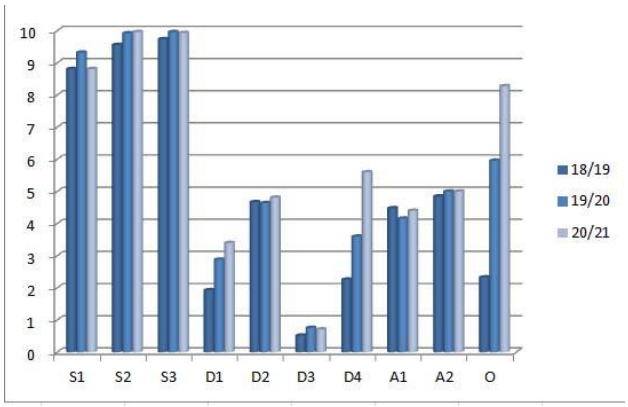


Figure 2. Distribution based on standards before and during pandemic

Static standards S1, S2 and S3 slightly oscillate, similar to administrative standards A1 and A2. Significant increase can be seen in D4 standard which is connected to online tests and O which is connected to a variety of elements regarding synchronous and asynchronous distance learning. It is interesting to notify that D3 standard, which is connected to discussions, did not have significant increase during the transition towards distance learning during the pandemic. Similar behavior of courses can be seen in the further analysis.

B. *Analysis by program groups*

Figure 3 shows an average of 4 different undergraduate study groups of academic year of 18/19 (before the pandemic), academic year of 19/20 (second semester was influenced by the pandemic) and academic year of 20/21 (both semesters influenced by the pandemic). The analysis represents these four groups:

- Economics and Management – Croatian group (EM cro)
- Economics and Management – English group (EM eng)
- Business Mathematics and Economics (BME)
- Elective courses (Elective)

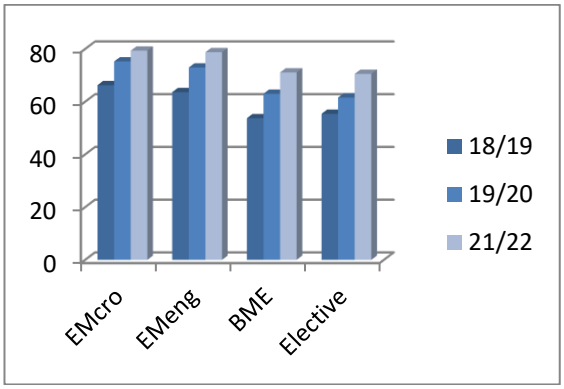


Figure 3. Quality average of e-learning courses by groups in different academic years

In the academic year of 18/19, before the COVID-19 pandemic, the best result is represented by the undergraduate study of Economics and Management Croatian group with a solid average of 66.63%. An evaluation has been made of all 43 regular courses, with only two courses that had an average less than 50%.

Similar situation can be recognized on the undergraduate study of Economics and Management in English, as there has also been 43 evaluated regular courses with two courses being below 50%, however, an overall average was 63.64%, which makes it lower than the Croatian group.

On the program of Business Mathematics and Economics group there has been 24 evaluated courses where most of the courses are of quantitative nature with mostly mathematics and programming, which are only held on the BME program. However, there are several shared courses that the BME students enroll with EM English group, but those results are a part of the evaluation of that program group. An assumption is that it also might be the reason why an average is lower in the last year evaluation before pandemic as it resulted with only 53.75%.

In the group of elective courses, there has been an evaluation of 24 courses, out of which there were even 5 courses that had an average less than 50% and an overall average is also significantly lower with the result of 55.4%.

Unlike regular courses, which are conducted regularly and already had been in use for years, which made it easy for further development and reaching a certain point of high quality, some of the elective courses have been only conducted once as it was their first year as part of the academic schedule and therefore need more time to reach better results.

Furthermore, some of the lecturers on elective courses are experts and specialists from business, which means that they have less academic experience in setting the syllabus, LMS course and other elements. This means that they need more time and support for higher quality development of the course.

During the pandemic, the second semester that was influenced in the academic year of 19/20, there has been a significant development and improvement of all groups,

especially in the ones that had a lower score before pandemic.

EM Croatian group, in accordance with the previous average, had a significant increase of 13.06%, making the final average of 79.36% and there was no course that was scored lower than 50%.

The EM English group had an overall average of 78.78%, which makes another significant increase of 15.17% in accordance with the previous average, and brought the score closer to the EM Croatian groups final score.

The most significant increase was BME group by 17.34%, making it a new overall average of 71.09%.

Elective courses also have a high increase of 15.17% and the result of 70.57%.

The continuous increase in evaluation results even after the lectures returned to the classrooms, points to the conclusion that once adopted and developed, e-learning became fully and equally embraced by lecturers and students.

Since e-learning was more in focus during the pandemic, it has become more important to improve its quality and appearance and therefore usual topic during the faculty meetings, which also reflected in improved evaluation results.

C. Analysis by standard groups

Table 3 shows all unique standards in details over all three academic years, the year of 18/19, before pandemic, and two being during pandemic, 19/20 and 20/21 in all four analyzed program groups.

All groups show an increase, before and during the pandemic. In the category of dynamic and the rest of the standards by 6% on elective courses, by 18.3% on EM cro group, 20% BME and 20.46% on EM en group.

The largest share is referring to the D4 standard which is connected to online tests, which were intensively used in classes during the pandemic for mid-terms and exams.

The increase in O (other) category can be attributed to mandatory scheduling of online sessions during the pandemic, as well as publishing recorded lectures and adapted lectures for asynchronous distance learning.

TABLE III. STANDARDS FOR ONLINE COURSE QUALITY EVALUATION

	S1	**S2**	**S3**	**S**	**D1**	**D2**	**D3**	**D4**	**D**	**A1**	**A2**	**A**	**O**
E&M C 18/19	9.55	9.89	10	**29.44**	2.5	5	0.68	3.64	**11.82**	3.98	4.89	**8.87**	**3.18**
E&M C 19/20	9.65	9.88	10	**29.53**	3.26	5	1.05	5	**14.31**	3.84	5	**8.84**	**7.56**
E&M C 20/21	8.72	10.00	9.77	**28.49**	3.95	5	0.93	7.44	**17.32**	3.84	5	**8.84**	**8.84**
E&M E 18/19	8.64	9.89	10	**28.53**	2.27	4.77	0.23	2.84	**10.11**	4.55	4.89	**9.44**	**2.84**
E&M E 19/20	9.30	10.00	10	**29.3**	3.37	5	0.58	4.19	**13.14**	3.84	5	**8.84**	**7.09**
E&M E 20/21	8.49	10.00	10	**28.49**	4.07	5	0.58	6.63	**16.25**	4.53	5	**9.53**	**8.72**
BME 18/19	9.2	9.2	9.4	**27.8**	0.6	4.4	0.2	0.00	**5.2**	5	5	**10**	**0**
BME 19/20	10	10	10	**30**	1.88	4.58	0.42	1.88	**8,76**	4.79	5	**9.79**	**1.88**
BME 20/21	8.96	10	10	**28.96**	2.5	4.38	0.42	3.96	**11.26**	4.58	5	**9.58**	**7.08**
ELEC 18/19	7.27	8.64	9.09	**25**	1.59	4.09	1.14	0.91	**7.73**	4.77	4.55	**9.32**	**2.27**
ELEC 19/20	8.10	9.76	9.76	**27.62**	2.38	3.33	0.95	1.19	**7.85**	5	5	**10**	**3.81**
ELEC 20/21	9.38	9.79	10	**29.17**	2.08	4.58	0.83	2.08	**9.57**	5	5	**10**	**7.71**

Table 4 and Figure 4 show the differences between specific groups of standards prior and during the pandemic. In the static standards the changes are minimum, regardless if it and increase or decrease, and only significant increase that is worth mentioning is group of Elective courses which used the pandemic timeframe for further development of basic lecturing materials and content.

TABLE IV. STANDARD GROUPS BEFORE AND DURING PANDEMIC

	Δ S %	Δ D %	Δ A %	Δ O
EM cro	1.19	6.88	0.038	7.07
EM eng	0.05	7.68	0.11	7.3
BME	1.45	7.58	0.56	7.08
Elective	5.21	2.3	0.85	6.8

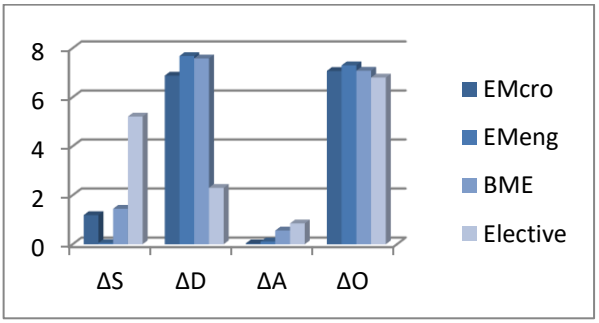


Figure 4. Differences between standard groups before and during pandemic

Dynamic standards in the first 3 groups have a significant increase, while the Elective group again shows deviation and lower rise than the static standards. A more significant contribution in all groups is D4 standard which relates to different types of online tests which are used for

home works, knowledge checks, mid-terms and exam periods. The expectations are that D3 standard is connected to discussions, but should also have a significant increase during the distance learning time, however, that didn't happen due to the new generation and their habits in quick chatting, which make their posts shorter than previous generations. The new generation uses different channels like Instagram, Messenger and Whatsapp. The changes in administrative standards are not significant at all, and in the category of other groups, they all act similar and have a certain increase. Usually this reflects on the synchronous distance learning and usage of the communication tools like Zoom, the one that is used on ZSEM and became a standard, but also the elements of asynchronous distance learning like recorded lectures.

IV. CONCLUSION

The goal of this paper is to quantify the changes in LMS usage and development in the transition from classroom to online environment during the pandemic. Since LMS has been in use at ZSEM from 2002, the accumulated experience of creating content and regular use of the system by students enabled smooth shift towards distance learning.

Ten standards in four groups already developed in 2006 to evaluate e-learning courses, have also been used to quantify the change in the volume and quality of the course content between academic years preceding and during the pandemic. All four groups of programs show an increase, before and during the pandemic on all of the four undergraduate study groups followed by research, with the exception of elective courses whose lecturers have less academic experience and needed more time to adopt.

Most significant increase was recorded in D4 (Online tests) standard, while the expected increase was not present in D3 (Discussion) standard due to change in communication habits among previous and current generations of students.

REFERENCES

[1] UNESCO, Education: from school closure to recovery, https://www.unesco.org/en/covid-19/education-response (accessed January, 23, 2023)

[2] S. Widodo, Y. Wibowo and W. Wagiran, "Online learning readiness during the Covid-19 pandemic", Journal of Physics: Conference Series ICOVEMAT 2020, 1700 (2020) 012033.

[3] E. E. Kabanova, E. A. Vetrova, T.A. Evstratova and L.E. Radchenko, "Remote education development trends during pandemic in Russia". Rev. Univ. Zulia 2020.

[4] S. R. Hawley, J. K. Thrivikraman, N. Noveck, T. S. Romain, M. Ludy, L. Barnhart, W. S.S. Chee, M. J. Cho, M. H. Z.Chong, C.Du, J. I. Fenton, P. Y. Hsiao, R. Hsiao, L.Keaver, H. S. Lee, W. Shen, C. Lai, K. Tseng, W. Tseng and R. M. Tucker, "Concerns of college students during the COVID-19 pandemic: Thematic perspectives from the United States, Asia, and Europe", Journal of Applied Learning & Teaching, Vol.4 No.1 (2021)

[5] S. Borotis and A. Poulymenakou, "E-learning Readiness Components: Key issues to consider before adopting eLearning interventions", in J. Nall, and R. Robson (eds.), Proceedings of world conference on e-Learning in corporate, government, healthcare, and higher education, Association for the Advancement of Computing in Education (AACE), 2004.

[6] N. Kalkan, "Investigation of e-learning readiness levels of university students studying in different departments", African Educational Research Journal Vol. 8(3), pp. 533-539, August 2020.

[7] G. Ilieva, T. Yankova, S. Klisarova-Belcheva and S. Ivanova, „EffectsofCOVID-19PandemiconUniversityStudents'Learning", Information 2021,12,163.

[8] A. Hania, Y. Hijazeina, H. Hadadina, A. K. Jarkasa, Z. Al-Tamimia, M. Amarina, A. Shataratb, M. Abu Abeeleha, R. Al-Tahera, „E-Learning during COVID-19 pandemic; Turning a crisis into opportunity: A cross-sectional study at The University of Jordan", Annals of Medicine and Surgery, 70(2021)102882.

[9] I. Alyoussef, „E-LearningSystemUseDuringEmergency:AnEmpiricalStudyDuringtheCOVID-19Pandemic", Journal Frontiers in Education, June2021, Volume6, Article 677753.

[10] L. Ochnio, T. Rokicki, K. Czech, G. Koszela, M. Hamulczuk and A. Perkowska, „Were the Higher Education Institutions Prepared for the Challenge of Online Learning? Students' Satisfaction Survey in the Aftermath of the COVID-19 Pandemic Outbreak", Sustainability 2022, 14(19), 11813

[11] J. Anderson, „IT, e-learning and teacher development", International Education Journal, ERC2004 Special Issue, 2005, 5(5), 1-14

[12] M. Munoz-Organero, P. J. Munoz-Merino and C. Delgado Kloos, „Student Behavior and Interaction Patterns With an LMS as Motivation Predictors in E-Learning Settings", IEEE Transactions on Education, Volume: 53 Issue: 3

[13] R. Conijn, C. Snijders, A. Kleingeld and U. Matzat, „Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS", IEEE Transactions on Learning Technologies, Volume 10, Issue 1, 2017.

[14] A. Chaubey and B. Bhattacharya, „Learning Management System in Higher Education", IJSTE - International Journal of Science Technology & Engineering | Volume 2 | Issue 3 | September 2015.

[15] A. Parmar, „Paper Review on Sharable Content Object Reference Model (SCORM): Framework for E-learning Standard", 2012 Second International Conference on Advanced Computing & Communication Technologies.

[16] K. Aleksic-Maslac, M. Korican, and D. Njavro, "E-Learning Course Development – Quality Standards", International Conference on Education and Information Systems, Technologies and Applications (EISTA 2008), Orlando, USA, June, 2008.

[17] CARNET – https://www.carnet.hr/

[18] K. Aleksić-Maslać, B. Sinković, and P. Vranešić, "The Role of Competition and Reward Regarding student motivation in the gamification process of different age groups", EDULEARN18 Proceedings , IATED Academy , 2018. Pp 1991-1999.

[19] K. Aleksić-Maslać, M. Rašić, and P. Vranešić, "Influence of gamification on student motivation in the educational process in courses of different fields", 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018.

[20] K. Trojnar , D. Vasić and A. Rauker, „New era of teaching finance: using an inverted classroom teaching approach in the course personal finance", 10th Higher Education Institutions Conference, September 22-23, 2022, Dubrovnik, proceedings.

[21] M. Koričan Lajtman, G. Oblaković and M. Nakić, „Impact of professors' personal factors and digital literacy on students' satisfaction in COVID-19 Pandemic", Virtual 8th Higher Education Institutions Conference, November 26-27, 2020, proceedings.

[22] B. Debič, K. Aleksić-Maslać and P. Vranešić, „Lessons from developing and delivering an artificial intelligence course during the Covid-19 pandemic", 44th International Convention on Information, Communication and Electronic Technology (MIPRO), 2021.

[23] K. Aleksic-Maslac, M. Korican, and D. Njavro, "Important Role of Asynchronous Discussion in E-Learning System", International Conference on Eng. Education and Research 2007 (ICEER 2007), Melbourne, Dec. 02-07, 2

Validating a Model of Smart Service System, Supporting Teachers to Create Educational Maze Video Games

Albena Antonova,
Sofia University/Faculty of Mathematics and Informatics, Sofia, Bulgaria
a_antonova@fmi.uni-sofia.bg

***Abstract* - Intelligent technologies can encourage the design and implementation of personalized learning solutions. In traditional school settings, teachers still hesitate about how to use educational video games in class. In this context smart services can support teachers to design personalized and adaptable learning games by providing more data-enabled, user-centered and context-oriented recommendations. The paper presents the smart service concepts and explores their role in improving teachers' capacity to design educational video games. It steps on the platform APOGEE to construct educational enriched maze video games and proposes a smart service validation model, which is verified and tested. The paper starts with a short introduction of the pilot smart service system, based on the APOGEE platform. After that, it analyzes the existing platforms that teachers can use for learning game creation. Next, a validation model of the smart service system is presented. The validation process is explained in detail, covering two testing rounds with 9 lecturers and 15 students. The discussion part verifies that the proposed smart service pilot implementation can efficiently support teachers to conceptualize educational maze video games.**

Keywords - smart service, educational video games, learning gamification, validation

I. INTRODUCTION

Intelligent systems in educational context can support teachers to improve the general design of personalized learning process, focusing both on the learning outcomes and on the learning delivery [1]. Especially after the pandemic, teachers are increasingly expected to deliver challenging and enjoyable "phygital" learning solutions, combining activities both in the physical and in the digital space [2]. By providing more data-enabled, user-centered and context-oriented assistance, intelligent systems such as smart service platforms can facilitate design of the effective and efficient gamified learning experiences [3]. In the same time, implementing games in the classroom settings is widely discussed as student' oriented approach, ensuring student' engagement and satisfaction, improving learners' creative, critical and problem-solving skills [4]. However, the main reason for the low adoption of educational video games in class still remains the low involvement of teachers in video game design, the lack of adaptation strategies to fit in the classroom learning scenario and the dominant misconceptions about game-based learning among teachers [4, 5].

The present research aims to outline the testing and validation of a pilot implementation of a smart service design model for supporting teachers to build educational video games. By comparing and discussing existing approaches for design and development of educational video games, the model of the smart service platform is implemented in the framework of the APOGEE platform (http://www.apogee.online/). More specifically, the current research proposes a two-step testing and validation process, exploring how low-code smart service model can facilitate and engage teachers, lecturers and non-IT professionals to build 3D educational maze video games on the APOGEE platform.

The paper is structured as follows. The first part makes a short overview of the main problems for applying game-based learning in class, making a brief summary of the popular learning games platforms and the key features to support teachers as game-designers. Next, the APOGEE platform and the smart service model are presented, exploring their approach and characteristics to support teachers to design complex and personalized learning games. The fourth part describes the smart service platform validation model, analyzing the quantitative data and qualitative outcomes obtained from the participants during the testing sessions. At the end are outlined the future work and next stages of smart service platform for assisting teachers in their efforts to create educational video games.

II. BACKGROUND RESEARCH

A. Study Background

Although the advantages of Game-based learning (GBL) and educational video games in literature reviews [6], teachers in general still hesitate to implement educational computer games in their teaching practice. The benefits for using educational video games are identified as presenting an immersive process, where students acquire knowledge during the play. Furthermore, GBL can be applied in various subject areas to improve student motivation, facilitating knowledge acquisition,

behavior change and developing high-order skills such as creativity, critical thinking and problem-solving skills [7].

To support teachers for adopting GBL in their teaching practice, some authors as [8] recommend a practical matrix, addressing six main aspects of the GBL implementation, covering: 1) teacher' own motivation to use GBL, 2) learners' needs and specifics, 3) game content and features, 4) practical classroom settings, 5) design of learning scenarios, 6) GBL results evaluation. All these aspects demonstrate that applying GBL in educational design is a complex issue and learning game design covers only one minor part of all tasks that teachers have to consider before implementing GBL.

As a short summary of the main difficulties and barriers for teachers to implement games and GBL proposed in [9] focus on: 1) the time for learning game development; 2) the limited resources for game creation; 3) the costs and efforts to produce learning games; 4) the lack of direct impact to the learning process; 5) the lack of support from schools and/or parents. As barriers for teachers can be cited as well the inadequate technology support, the lack of supporting materials and best practices for using games for learning, and the lack of time [3]. In addition, many in-service teachers still are not using English and prefer local or translated software game design platforms and GBL supporting materials.

B. Smart Services and Intelligent Systems

The smart services are intelligent systems, usually explored in the framework of industry automation and Industry 4.0 and can refer to terms: smart product-service systems and cyber-physical systems [10, 11]. Built on data-analysis, interactive interfaces and personalized configurations, smart services can be used interchangeably with "smart web services", "intelligent services", "intelligent product-service systems", and others. Therefore, smart services can be defined as "real-time, user-oriented, adaptable and context-oriented services, delivered within networked systems and platforms "[12]. Smart services can be applied to facilitate different educational processes and can cover different learning scenarios [3]. By delivering data-enabled, user-centered and context-aware recommendations and automatic configurations, smart services can assist teachers to better understand students' needs and to improve their learning experiences. Supporting teachers to conceptualize and design educational video games for learning can engage them in the design of more complex and interactive designs.

In the context of education, smart services can be applied in different perspectives: on educational level, on school level, on class level, on teacher level, and on student level. Furthermore, smart services can facilitate learning personalization and learning adaptation, assisting teachers to apply various approaches that can be realized through the use of technological means [13]. Focusing on the practical aspects, smart services have to address the 4C [14]: 1) connectivity "improving access to people, devices, and educational environments"; 2) collection "of (study-related) data"; 3) computation of "key factors, influencing learning and student satisfaction"; 4) communication "within or through technology-equipped education device and environment for increasing value co-creation between students, teachers, and other stakeholders" [14].

C. Exploring Game Design Features of Popular Educational Video Games

Before exploring more sophisticated models to assist teachers to easily create and implement educational video games and GBL in their teaching practices, a short overview is provided for some popular GBL solutions. The outlined game platforms in table 1 are among the most used and recognized game platforms by local teachers, identified in teachers' surveys, interviews and discussions. The list of GBL platforms in table 1 is not exhaustive and its goals are mainly to illustrate and analyze the most common approaches to assist teachers in the design of educational games.

Table 1. Summary of popular learning games platform features.

Learning platform	Game elements	Teachers' support
Kahoot!	Online quiz	2D quiz; predefined templates examples;
Genially	Interactive presentations, quizzes	2D games, combine off-line/online elements, pre-defined games & example;
Learning Apps	Puzzles, quizzes	2D games, examples, pre-filled templates, multiple languages;
Scratch	Programming games	2D game, online editor, examples, gallery;
H5P.org	Interactive exercises	Puzzles, pre-filled examples; integrated with Moodle;

Furthermore, some screenshots of the game platforms are provided on fig. 1, focusing to outline the game designing process. It should be noticed, that most of the game platforms support 2D puzzles and interactive learning activities. The most popular type of games covers quizzes and presentation-type display (Kahoot!, Genially). Most of the game platforms provide examples of the games that can be used as pre-filled templates, that users can modify and adapt with own learning content (learning apps, h5p). Additionally, some of the platforms support online drag and drop visual editors, along with examples of exercise scripts, that teachers can directly explore in their class activities (scratch). Most of the games can be played with a school computer equipment (computer and projector), some of the games allow students to use their own hand-held devices such as mobile phones and tablets (Kahoot!).

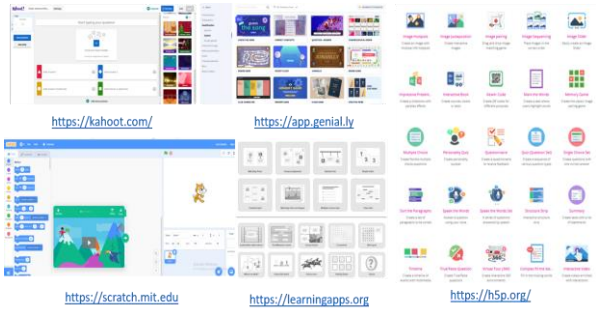


Figure 1. Screenshots of the popular learning game platforms supporting GBL, as described in table 1.

Among the key problems with these models can be summarized as follows:

- Neither of the game platform provides information and guidance how teachers can apply games or GBL in the design of their learning scenarios.
- The process is not interactive, it usually imitates some common approach (building presentations).
- Teachers have to learn how to design the games in every specific game platform (as lock-in factor).
- Dominant 2D puzzles and video games can be easily performed with traditional means (and not need digital platforms);
- Language support (most of the GBL platforms are on English)
- Lack of support for learning personalizing and adaptation of the games to the students' needs.

III. The APOGEE Video Maze Game Platform

The APOGEE platform for construction and generation of rich educational video maze games is based on the Unity editor and supports the design of educational 3D video mazes enriched with mini-games [15]. The enriched maze video-games are defined as a maze 3D video game with multimedia learning content presented within the maze halls on learning boards and within personalized and adaptive puzzle mini-games of various types, embedded into each hall [15]. The APOGEE platform supports design of personalized and adaptable learning experiences as the mini-games could be tailored to the end user profile. The learning content can also be personalized upon the player/learner model attributes such as demographic characteristics, demonstrated outcomes, learning/playing style, and emotional status and further adapted to the learning scenario.

To facilitate teachers to create educational video games in the APOGEE maze design platform, a pilot model of smart service platform is designed. The main objectives of the smart service platform are to support teachers, non-IT professionals and experts, parents and students to design educational video games. Based on recommendations, dynamic data statistics and best practices, teachers can not only learn how to design adaptable and personalized maze video games but improve their GBL skills and gain experience from others.

The process for building educational games in APOGEE starts with submission of an XML document, formally describing the maze, the mini games and the learning content. There game designers should describe all game elements, combining text, images, and audio files, together with the mini-games placed in the maze halls. In these settings, teachers have to describe and upload all the features of the educational video game – such as the number of halls, the structure of the maze, the learning panels, and the mini-games.

Thus, in order to facilitate and support teachers for game design process, a pilot low-code solution is provided. This pilot Smart Service platform, described in [16] aims to propose an alternative approach for teachers to design and structure educational games, providing additional information about GBL, learning scenarios, learners' preferences and some additional statistics [16]. On this platform, the APOGEE game design process is structured on three main parts. The first part introduces the general parameters of the video game, the second is dedicated on game halls design and the third part allows teachers to make a request for game generation. Every hall in the maze is separately generated, including all elements such as educational panels and mini-games. A special page explains types of the mini-games supported in the model, highlighting how they can be adapted and used to personalize student experience. There, teachers just have to fill out several interactive Google forms, making an easy step-by-step approach, selecting answers, and uploading game elements. When the files are submitted, users receive by mail a link to their data, allowing them to make multiple revisions and modifications. The process is highly iterative, requiring many versions and modifications, before completing the final version for a game generation [16].

A. *Smart Service Platform Validation Model*

A detailed validation process is organized to ensure that smart service model platform can efficiently support teachers and end-users, observing their experiences and interest to create and implement educational video games in class.

To validate the model of the pilot smart service platform, a two-step validation process is organized. The testing of the design and delivery of the smart service prototype is made via two workshops: with teachers/ tutors and with students/teachers-to-be. The quantitative data are collected via online questionnaire, and the qualitative data are obtained from semi-structured interviews and discussion, focusing on the concepts of interactivity, personalization and further modification of the platform to support educational video games.

B. *Methodology for Validating the Pilot Smart Service Platform*

In order to test and validate the pilot smart service platform, there are investigated the following indicators of success. Focusing on the needs of the educators to support the game creation processes, the factors of success are based on the matrix of Ney et al. [8]. Based on the self-reported data and external observation, a smart service system for supporting teachers to design educational video games should improve:

1) the teacher's own motivation to use and create educational video games: to what extent the developed system has supported and improved the teachers' understanding and interest for creating educational enriched maze video-games.

2) the needs and specifics of the learners: to what extent the smart system has improved the teachers' understanding of the needs of the learners and the possibilities for personalization.

3) the game content and features: the what extent the educators can easily create and understand approaches to design game and learning content, including as well content for the minigames.

5) design of learning scenarios: to what extent the system improves the understanding of the role and place of games in the learning process.

6) evaluation of the results of game-based learning: to what extent the teachers understand how they can evaluate the learning outcomes of the players.

The only indicator listed in the model of [8], which is not explicitly considered in the presented model, is (4) the practical classroom settings, since the pilot implementation of the system does not address issues related to technological characteristics, practical and logistic challenges for the actual implementation of the developed games in the classroom settings.

In addition to these indicators, complementary features are the practical parameters of creating educational video games. For example, these consist of: time for the first draft and the final version of the game development, dedicated resources and efforts for game design, the satisfaction of the process, the models to link educational games with the learning process and the opportunities for group work and interaction with other teachers.

C. Practical Settings of the Validation Process

The smart service system pilot implementation and the models for creating personalized and adaptive games were validated and tested in practice on two separate settings. All validation activities took part in December 2022 and followed the same procedure.

1) Validation on the e-Creha Tutors' Workhop

First, the testing and validation round was organized with the participants of the e-Creha project workshop (https://www.ecreha.org), taking place on December 9, 10 and 11, 2022 in Sofia. The Tutors' workshop was dedicated on the design and use of educational games and games-based learning approaches, and was closely related with the tasks of the Bulgarian team - development of four educational maze video games on the climate change and the preservation of the architectural heritage. A total of 15 university professors and lecturers from Bulgaria, Italy, Turkey, France and the Netherlands took part in the Tutors' workshop.

Within the Tutors' workshop settings, participants created a pilot educational game by using the tools and instruments, provided in the smart service system. They have to select and define the learning objectives, the learning content and the mini-games for an educational video game. Provided with initial concepts notes and texts for the maze video games, in a participative manner, tutors defined the first draft of the game design within less than 2 hours. Based on the delivered learning content, the first draft of the game was ready for testing, generated only for several hours after this session. This way, tutors tested the game on the third day of the seminar, and then they were asked to provide structured feedback by filling a questionnaire. A more detailed insights and qualitative data were collected within the group discussion and final feedback session.

The feedback questionnaire included a total of 17 questions, out of which 14 questions concerned the game creation process and 3 questions addressed the satisfaction with the workshop. In total 8 participants out of 15 attendees completed the structured questionnaire.

Considering the collected quantitative data, after the workshop the attitude of educators for creating content for computer games is improved fig.2. (before seminar: mean 3,5; st.dev:0,92; after the seminar: mean 4,75; st.dev: 0,46), Likert scale (1 min.-5 max., 3 neutral). Tutors appreciated the working process with the smart service platform as rather easy, including submitting the content in structured forms. As the easiest functionality for tutors is to fill the ready templates and the ability to upload additional files. The selection of learning objectives and the design of the learning content were identified as the most challenging tasks for the game development process.

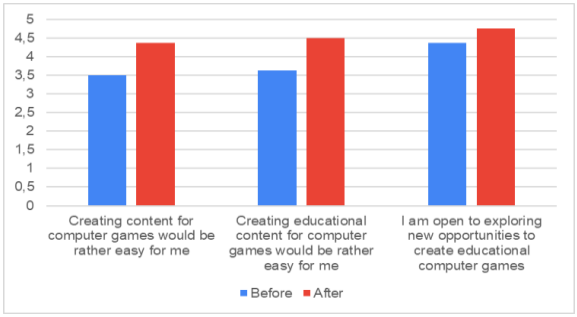


Figure 2. The tutors' attitudes for game design before and after the workshop.

In terms of game personalization approaches, the tutors highlighted that they would apply learning personalization approaches mainly considering the learners' age/ grade, their preferred learning and playing styles, and the level of interest in their specific discipline. All of the respondents stated that additional statistics for learners' preferences, recommendations for good practice, teacher and learner experiences, and individual preferences are to a large degree useful when designing and using educational video-games.

The open-ended questions revealed that the process of developing educational games, using smart-service platform is easy for teachers, they see the potential to create educational games and better understand appropriate approaches for game use and development. One of the participants noted that the workshop motivated her to explore more game-based learning approaches in her own field - architectural design. More research is needed to identify how different type of games and

gamification approaches can be applied in different subjects and discipline areas.

2) Validation with University Students

The second experiment for testing and validating the pilot smart service system was organized with students, enrolled in the BSc teachers' program at Sofia University. The main task for the students was to make an educational video game, by selecting and systematizing learning content according to the chosen topic, including texts, images, diagrams, and audio content, and to define the content of the selected puzzle minigames. They had to complete the task for two to three learning hours, providing learning and game content (including texts and multimedia files such as audio and images, processed to fit the system requirements) for the game development.

Within this experiment, three testing games were successfully designed and developed by the students - a game dedicated to learning styles (the VARK game), a game on Geography, and a game for Triangles. All of the games are operational, can be easily up-graded and adapted to a specific situation and learners.

After the experiments, 15 students filled the same questionnaire for their experience evaluation. The conclusions are that students easily navigated the smart service system, they understood well how to generate an APOGEE video-game, and they received enough additional information about the benefits of using computer games in the educational process.

When comparing data between tutors and students, it can be identified (fig.3), that students are somehow more critical and demanding for using more visual video-game creation technologies comparing to the educators. This can be explained by the fact that university students are better prepared to learn from visual tutorials and they are better adapted to use learning-by-doing systems and technologies.

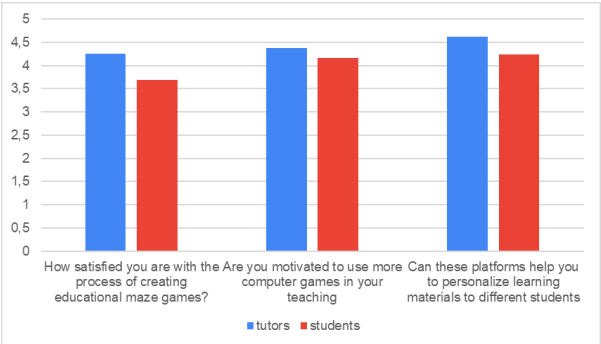


Figure 3. Comparing the attitudes of tutors and students on game design process.

The evaluation of the students' quantitative and qualitative feedback shows that the creation of educational three-dimensional maze video-games is both interesting and in the same time challenging for them and requires a totally new mindset and way of thinking. Some of the students who worked independently encountered difficulties in understanding how to proceed with the APOGEE gaming platform and could not finish their games.

IV. Discussion

The objectives of the selected smart services are to support teachers, providing additional functionalities to improve the effectiveness and efficiency of the learning process. In this perspective, smart services can facilitate teachers to create more interactive, personalized and differentiated digital game solutions, better assessing and adapting them to the individual characteristics of the learners.

From user experience perspective, it is important to underline that both tutors and students had positive attitude for the smart service platform seminar. Most of them enjoyed the participative process of working with the system, collaborating with their peers, thinking about best gamification approaches and its visual representation. Reducing stress on learning a new computer system for game design allowed both teachers and students to dedicate more time and efforts to specify appropriate learning and game content and learning objectives and to focus on mini-games experiences. After playing the games, students became even more motivated to investigate how to use GBL approaches in class.

The general outcomes of the validation demonstrated that the proposed smart-service system model supports the fast and smooth design and development of an APOGEE 3D educational maze game. This experiment for validation session demonstrated that a group of non-trained university lecturers and students succeeded to create several fully functional three-dimensional enriched maze video-games on the platform APOGEE for just few hours of work. It allowed participants to: (1) easily understand the concepts of game-based learning and APOGEE maze game design elements; (2) identify learning goals and objectives of the educational games; (3) provide in participative manner didactic content for the game design, including texts, images and sound, as well as to structure mini-games content (4) identify models and strategies for learning personalization, supported by the APOGEE platform.

V. Conclusion

The most important factor for introducing GBL on practice is to make the process of game design enjoyable, efficient and satisfactory for teachers. By reducing the stress from using a new platform, both students and tutors gained confidence and interest to design and use more educational video games. Further, it is recommended that game design platforms should support more participative approaches for users, allowing them to work in groups, to share experiences, data and learning analytics on best practices for learning scenarios and learners' preferences and to allow interactive and multistage development process.

Acknowledgment

The research leading to these results is partially based on the outcomes of the APOGEE project, funded by the Bulgarian National Science Fund, Grant Agreement No. DN12/7/2017 and supported by the Erasmus+ e-CREHA project, Number: 2020-1-NL01-KA203-064610.

REFERENCES

[1] A.Antonova, Y. Dankov, “Smart Services in Education: Facilitating Teachers to Deliver Personalized Learning Experiences”, In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds) Data Science and Algorithms in Systems, CoMeSySo 2022, Lecture Notes in Networks and Systems, vol 597. Springer Cham., pages: 108-117, 2022.

[2] P.Vate-U-Lan, D. Quigley, P., Masouras, “Phygital Learning Concept: From Big to Smart Data”, in 13th IC on eLearning for Knowledge-Based Society, 15 December 2016, Thailand, pp. 9.1—9.5, 2016.

[3] A.,Antonova, B. Bontchev, “Designing scenarios for personalized learning: enabling teachers to apply educational video games in class”. International Journal of Education and Learning Systems, 4., pp.20-26, 2019.

[4] R.W. Mee Mee, T.S.T Shahdan, M.R. Ismail, K.A. Ghani, L.Pek, W.Y. Von, A. Woo, and Y.S. Rao, “Role of Gamification in Classroom Teaching: Pre-Service Teachers' View”, International Journal of Evaluation and Research in Education, 9, no. 3, pp. 684-690, 2020.

[5] E. Paunova-Hubenova, V. Terzieva, “Information Technologies in Bulgarian School Education”., Proc. of the 13th Int. Technology, Education and Development Conference (INTED 2019), 11-13 March, Valencia, Spain, IATED, ISBN: 978-84-09-08619-1, 5226-52. 2019.

[6] S. Vandercruysse, M. Vandewaetere, G. Clarebout. “Game-based learning: A review on the effectiveness of educational games”. Handbook of research on serious games as educational, business and research tools, pp.628-647. 2012.

[7] S. Vandercruysse, J. Elen, “Towards a Game-Based Learning Instructional Design Model Focusing on Integration”. In Instructional Techniques to Facilitate Learning and Motivation of Serious Games; Advances in Game-Based Learning; Wouters, P., van Oostendorp, H., Eds.; Springer: Cham, Switzerland, 2017.

[8] M. Ney, V. Emin, J. Earp, “Paving the way to Game Based Learning: a question matrix for Teacher Reflection”. Procedia Computer Science, 15, pp.17-24. 2012.

[9] E. Koh, Y.G. Kin, B. Wadhwa, J. Lim, “Teacher perceptions of games in Singapore schools. Simulation & gaming”, 43(1), pp.51-66. 2012.

[10] M. Hermann, T. Pentek, B. Otto. Design principles for Industrie 4.0 scenarios, Proc. of 49th Hawaii International Conference on System Sciences (HICSS), IEEE, 3928-3937. 2016.

[11] Y. Yang, Ying, H., Jin, Y., Cheng, H. K., Liang, T. P.. Special Issue Editorial: Information Systems Research. In book The Age of Smart Services. Journal of the Association for Information Systems, 22(3), 10, 2021.

[12] D. Beverungen, O. Müller, M. Matzner, J. Mendling, J. Vom Brocke. “Conceptualizing smart service systems”. Electronic Markets, 29(1), pp.7-18. 2017.

[13] M. Marienko, Y. Nosenko, A. Sukhikh, V. Tataurov, M. Shyshkina. "Personalization of learning through adaptive technologies in the context of sustainable development of teachers education." arXiv preprint arXiv:2006.05810, 2020.

[14] C. Lim, P. Maglio. Clarifying the Concept of Smart Service System. Handbook of Service Science, Volume II, (2019)349-376.

[15] B. Bontchev, D. Vassileva, Y. Dankov, “The APOGEE Software Platform for Construction of Rich Maze Video Games for Education”, Proc. of 14th International Conference on Software Technologies (ICSOFT'2019), Ed. Marten van Sinderen, Leszek Maciaszek, INSTICC, Prague, Czech Republic, 26-28 July, ISSN: 2184-2833, ISBN: 978-989-758-379-7, pp.491-498. 2019.

[16] A. Antonova, B. Bontchev, “Designing Smart Services to Support Teachers to Create Personalized and Adaptable Video Games for Learning”, ERIS, 2022 (in print), 2022.

.

STREAM Education - Potential for Engaging Students in Generating Innovative Green Ideas and Development of Transversal Skills

Kr. Dyulgerova*, D. Atanasova* and M. Milanova*
* University of Ruse„Angel Kanchev"/Department of IIT, Ruse, Bulgaria
** English Language School "Geo Milev", Ruse, Bulgaria
kdyulgerova@uni-ruse.bg, datanasova@uni-ruse.bg, m.milanova@elsruse.eu

Abstract **- The article presents pedagogical strategies in the design of a STREAM approach-based initiative. The details in the training scenario are examined, the main stages in the design of the individual lessons and the tools for studying the progress of the students are described.**

An interdisciplinary Project Based Learning (PBL), in which students are researchers of real problems, creators of authentic content and creative products, is implemented in an innovative high school with foreign language training and additional profiling in the fields of Mathematics, IT and Entrepreneurship. A team of two teachers, participants in a Professional Learning Community (PLC), develop the design of this type of training, change the working conditions and develop a learning environment different from traditional formal learning spaces. Teachers act as mentors and facilitators, engaging 14- and 15-year-old students in the adventure of identifying problems, proposing solutions, and advocating for a position. The focus is on the development and evaluation of transversal skills acquired through the methods of foreign language learning and computer science.

Optional IT learning groups have the potential to free space for extracurricular learning. Organized that way, students acquire and master competencies in the field of computer science. The development shares the details regarding the idea of integrating learning approaches from the two subject areas - computer science and foreign language learning. The assessment and self-assessment tools provided in the development are easy to use and can provide quick feedback to students and teachers and measure the effectiveness of learning and teaching. Some conclusions are drawn and the proposed methodological approaches are commented on, the challenges in the work are described and options for overcoming them are indicated.

Keywords – STEM education, innovative green ideas, educational system, transversal skills

I. INTRODUCTION

Discussions about the development of educational systems are on the agenda in every country and society, and the most important topics discussed are: Do children acquire the necessary knowledge in schools? Which school subjects are related to the real problems outside the classrooms? Can students develop their potential, reveal their talents, discover their strengths, acquire adequate skills necessary for success in the modern world and meeting the challenges of the labour market?

The correct answers to these questions will effectively direct the necessary resources to build an appropriate legal framework of the formal education system in which young people can acquire qualifications according to the needs of the economy.

The "European Skills Agenda" was developed by the European Commission in 2020 and identifies measures to strengthen sustainable competitiveness, ensure social justice and sustainable growth. According to it, training institutions should focus their efforts on acquiring the right skills in a new work environment with the possibility of flexible transition from one sector of the economy to another. Nicolas Schmidt, Commissioner for Employment and Social Rights, says that upskilling the workforce is the key response to the demands of a rapidly evolving labour market. Twelve "actions" are specified in the Programme, including "EU support for national strategies for upskilling" and "Increasing STEM graduates and promoting entrepreneurial and transversal skills". [1].

STEM training develops abilities in students to apply concepts from Science, Technology, Engineering and Mathematics to solve problems that cannot be solved using only one discipline. Transversal skills aim at solving complex life situations. This type of skills has a supra-subjective character, which is acquired outside the subject limitations. [2]

The current educational model lags behind the demands of society, the economy and the labour market. Progress in the field of education includes the development of state policies, imposes common school strategies, goes through the development of the professional skills of teachers and requires them to take individual actions - the implementation of new approaches in teaching.

The current research focuses on exploring innovative methods, strategies and tools for teaching, learning and assessing progress and builds working models for developing key skills at secondary level.

II. PLACE OF IMPLEMENTATION

Geo Milev English Language School is a secondary municipal school with a language profile and additional profiling in the fields of Mathematics, IT, Entrepreneurship, Biology, Chemistry, History, etc. The school has been part of a state network of innovative schools for five years. During the last three academic years, a partnership has been developed with Angel Kanchev Ruse University. Part of the classes are held in the university laboratories by university lecturers.

A "STEM centre for eco-innovations" has been operating at the school since 2021. It is an integrated

educational environment for the implementation of PBL in the areas of Health and Nature, Technology and Engineering and Sustainable Development. Teachers have permanent attitudes that STEM methods can be transferred to all subject areas and through them non-material skills are taught, since the latter are necessary in order to deal with the challenges of the 21st century - problem solving, logical thinking and critical thinking.

Several Professional Learning Communities (PLCs) are developed in the team with a different focus, which is determined at the start of each academic year, and teachers, according to their own interests and regardless of their own narrow professional expertise, choose the teams in which to be included.

III. THEORETICAL BASIS OF THE PRESENTED SCENARIO

A. *Interdisciplinary Training*

Interdisciplinary learning is a modern scientific approach that opens up the possibility of studying a given problem from the point of view of multiple scientific perspectives. Interdisciplinary courses are organised around a theme, problem, region, cultural or historical period, institution, figure, or idea. [3]

Multidisciplinary methods are approaches from different scientific fields, for which it is possible to be combined, blended and applied to extract the most useful information on a given scientific problem. Interdisciplinarity is seen as a different, combinative, actionable approach to applying multiple methods.

Karri Holley [4] describes three types of interaction between the different scientific fields.

"**Cross-disciplinarity or interdisciplinarity**" is the most general term referring to the crossing of boundaries between individual disciplines. Teachers work within their own disciplinary discourse and norms.

In the case of "**multidisciplinary**" knowledge full integration is also missing. This approach is "encyclopaedic". In multidisciplinary research, students accumulate knowledge in the course of a given inquiry, but no effort is made to synthesise these contributions. Students acquire related knowledge about learning content from different disciplines. They can successfully compare topics from different disciplines and find connections between them.

'**Interdisciplinarity**' is a term referring to the integration of knowledge from multiple disciplines in pursuit of an outcome that is not possible from a single discipline. In this type of training, key skills are built by integrating two or more academic disciplines into a common curriculum.

Interdisciplinary acquired knowledge can be categorised into separate types of knowledge [5]

- Orientation knowledge – helps students to find their own way through different topics: I know what.
- Activity knowledge – helps students to adopt subject-related methods, techniques or strategies: I know how.
- Explanatory knowledge – provides students with arguments that explain why something is the way it is: I know why.
- Reference knowledge – teaches students where to find additional information on a specific topic: I know where.

Most often, in school practice, two or more teachers with different professional expertise formulate a basic idea, set common goals, and educational activities on individual subjects are planned with a main focus on them. In this approach, project-based learning strategies are applied in full force. Students' participation in academic thematic units develops strategic planning and evaluation skills; introduces them to real situations and engages them in a natural non-violent way to the issues.

B. *The acronyms STEM and STREAM*

In the world of education, the acronym STEM lists the disciplines of Science, Technology, Engineering and Mathematics. STEM disciplines are important as a central part of human knowledge and a foundation for other subject areas. The STEM-focused curriculum integrates learning tools and methods and prepares students for the adult world. On the other hand, we can consider STEM a basic content-based approach in education, in which the focus is on cognitive development, acquired competences and engagement in the learning process. [6]

The STREAM education model adds more layers to STEM - by including "Arts" and "Reading" in the equation. The Arts challenge and develop creative and innovative thinking and help teach morals, values, ethics and responsibility. Reading as the main element for discovering and interpreting new knowledge, adds another set of activities and tools to the learning process. The focus on Reading or Literacy provides conditions for a holistic learning experience and encourages critical thinking and creativity.

C. *PBL pedagogical practice*

1) *Practice design*

Awareness of the learning process implies that learners make independent decisions about choosing the method and environment that best suits them and the ability to adapt them according to their needs to achieve optimal results. Project-based learning is an example of metacognition and self-regulated learning. Pedagogical practice includes the following mandatory steps: clear preliminary formulation of the task; development and activation of a preliminary plan of both the student (team) and the teacher; teacher-student and student-team interactions; progress monitoring; structured assessment and self-assessment; conclusions. [7] The steps are given in figure 1.

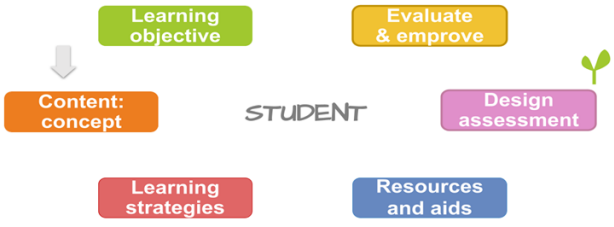


Figure 1. Design assessment and PBL

The established practices in the educational sphere are related to the placement of assessments at the end of the training of the relevant topic on a specific educational content. Evaluating a group student project implemented in the context of interdisciplinary PBL requires a different approach. [8]

In PBL, students perform a long-term set of tasks and it is logical that they are assessed with different methods and objectives that build a common assessment framework. In general, several types of assessments can be applied:

- Formative (to improve the individual activities in project implementation);
- Integrative (to reveal the direct connections between knowledge and learning);
- Summative (to measure the results achieved). [9]

According to Thuy Thu Vu [10], in order to be useful for student learning, the practical tasks subject to assessment should have the following strengths:

- The direct link between assessment and learning activities should be clearly visible;
- Students should be involved in the evaluation process;
- Teachers should provide clear, quality and timely feedback;
- On the other hand, the evaluation should be: clear, authentic, engaging. It should support students to draw their conclusions for further learning.

Teachers of various subjects require students to present as a result of a school project some digital product - video, presentation, text document, website, e-book, etc. The idea is good if the assignment is clearly formulated with details on the requirements for the form and content of the final product, and the evaluation criteria are clearly presented in advance, at the start of work on the project. Teachers without professional expertise in the field of IT cannot be armed with such knowledge, therefore working in a team of teachers is appropriate.

D. The outcomes in the context of IT

From the point of view of IT education, the digital products developed by the students as a result of the implementation of an educational project must have the following characteristics (Figure. 1):

Purpose /Aim/: The end product of PBL should check the level of understanding and provide opportunities to practise and apply knowledge and skills.

Effectiveness: The end product of PBL should increase opportunities for students to learn and demonstrate knowledge.

Personal attitude: The end product of PBL should offer students choice, be personable, and create a personal connection between students and the learning content.

Competencies: Students should be able to realise the end product of PBL independently and have a sense of improved skills after doing it.

Aesthetic aspect: Teachers should present such requirements that the students' task is attractive and motivates the students to complete it.

Quality criteria: Students need to be presented with a full set of feedback tools and standard requirements for digital products.

E. The outcomes in the context of ELT

From the point of view of English language learning, the products developed by the students as a result of the implementation of a learning project must have the following characteristics:

Purpose/Goal - reflect important aspects of acquiring language competences - working with written information and developing skills for analysis, synthesis and interpretation, as well as critical thinking.

Effectiveness - that the end product engages students' attention and directs them to creative expression and motivation to use aspects of applied arts to learn learning content.

Personal attitude - the project should reflect elements of the students' personal value system and approach the application of solutions in real-life situations.

Competencies - end products to demonstrate cultural competence and develop creative skills, abstract and non-standard thinking, ability to recognise and deal with moral dilemmas and conflicts.

Aesthetic aspect - the project should meet the needs of the students and take into account their interests and personal goals by improving the interaction between peers.

Quality criteria - the product must contribute to the improvement of language education and communication skills by providing students with opportunities for self-reflection, feedback and peer assessment tools.

When teachers work together to improve their practice, students learn more. Teaching must be transformed from a solitary activity into a process involving professionals sharing common attitudes and ideas for improvement. Steps in this direction make schools efficient, teachers satisfied and students successful. [7]

F. Professional Learning Communities (PLCs)

Professional Learning Communities (PLCs) have been entering education since recent decades and have nothing to do with teacher qualification courses. They provide an environment that encourages professional development, collaboration and innovation. The innovative concept in schools fundamentally changes the way educational services are provided to children and adolescents - through the dynamics in the goals it sets, through the mechanisms it works with, through the way the results are reported. [10]

Arguably, the difference between cooperation in professional learning communities and cooperation in other cooperative spheres is significant. [11]

School teams are created with a specific purpose, most often at the beginning of each school year. Teachers with different professional expertise work together to develop a space to improve student outcomes. Teacher effectiveness researchers share compelling evidence that when educators work as a team with their colleagues, they change school environments and policies, making improvements in their teaching that benefit both teachers and their students. According to Kathleen Fulton and Ted Britton, "great teaching is a team sport." [12]

According to researchers, school communities are effective because of a common system of interventions to support students with difficulties, regardless of the personal characteristics of the specific teacher.[13]

IV. STRATEGIES FOR ASSESSING PROGRESS

The model of perception of new knowledge, the way of learning and understanding and mastering skills are individual characteristics of the personality. According to the authors K. Shoylekova and B. Ivanova, teachers in each age group should choose teaching and assessment tools tailored to the student's characteristics. Teachers have the freedom and responsibility to apply critically and properly any form of teaching and assessment, to evaluate pedagogical practice and to revise their strategies. [15, 16]

In the conditions of STREAM education, the focus in determining the progress of students should be meaningfully redirected to transversal competences, to ensure the growing independence of the learner and to create conditions for the transfer of knowledge and skills from one learning area to another.

According to Rosemary Hipkins from New Zealand Council for Educational Research [18] they are traditionally comparative in nature and measure skills before and after training as well as actual achievements. They are authentic and meaningful if they are directly related to students' abilities and learning objectives. The evaluations prove the level of performance therefore they are necessarily related to previously announced criteria. It is best to have more than one author of tasks and criteria.

In cross-curricular learning, the tasks students perform are often interdisciplinary in nature. Learning challenges then require problem formulation skills and creativity and innovative ideas to solve it. Academic success is a process, and students must practice, persevere, and overcome obstacles. They should be challenged to use personal experiences, to transfer tools between subject areas, to develop their learning skills, leading them to meta knowledge. [17, 18]

Meta skills are universal and transversal skills that can help in professional and personal development. Rigorous professional expertise in a particular field of knowledge is no guarantee of success. It is often necessary to choose a new professional career, and meta skills expand the possibilities of realisation and open the door to progress.

According to Knight [19] it is essential to take into account the following features and characteristics of the evaluations:

- Estimates are provisional judgments based on current evidence;
- When designing PBL, the criteria for determining the quality of the achieved results should be specified;
- It is important for students to make sense of their development and what else they need to know and can do, to carry out critical self-reflection.

Knight also describes complex competencies, sometimes described as "soft" skills. These are competencies that are often difficult to define, cannot be measured in a quantitative way, and usually develop over a considerable period of time. These are often included in the qualities desired by an employer: teamwork skills, emotional intelligence, verbal communication, customer service skills, self-management, time management, confidence and efficiency.

V. A TRAINING SCENARIO IN ELS

The training scenario presented is an example of the common practice of three teachers of IT, English language and Philosophy. A total of 100 students in the 8th grade - first year of high school level, studying English for 18 hours per week, philosophy - 1 hour and IT - two hours per week are included.

The pedagogical scenario is realized in a total of 30 hours, and at least 5 of them are realized in the STEM centre of the school, in a room with workplaces for each student, equipped with a standard computer configuration.

Therefore, the focus of this study is the set of non-standard tools and methods for impact assessment developed by the team of teachers. They are under development and implementation. By the end of June, all classes will be held and the results of the practice will be summarized.

Big idea: Through teamwork to create electronic products (animations, presentations, collages, videos, etc.) to present ideas for solving global problems. Students should be encouraged to use their imaginations and create fictional characters with supernatural powers to help humanity.

Standards and intended outcomes

Students should acquire transversal skills: to argue and defend their personal position in various communication situations; to investigate a cognitive problem working in a small group; to make efforts to build a sense of empathy and trust in the community (group) through cooperation and constructive communication; to carry out individual and group studies, using a variety of information sources, achieve and present specific results; to present information to the public in a structured manner, following a predetermined regulation and using various visualisation tools and technologies.

Students should expand their digital literacy: the ability to use digital technologies, devices and social networks, find, organize, process, evaluate and information in an appropriate way; to observe and apply specific security rules, as well as norms related to the ethical and legal use of information; to choose and justify their choice of specialised software for creating web pages, digital images, sound and video information.

Students should expand their language skills and acquire a performance culture related to time management, English presentation, body language and confident public speaking. Students learn to work with diverse sources of information to analyse and synthesise in order to extract essential elements and linguistic units. The goal is to form a basic culture of the personality: moral, legal, aesthetic, physical, labour culture, of communication; developing creative skills and independent thinking of students.

Students to develop cultural competence and skills of expression through creativity by developing creativity and emphasising the distribution of tasks in a team and integrated construction of the four language skills: listening, speaking, reading and writing through understanding and composing oral and written messages, understanding and expressing thoughts, facts, opinions, feelings and attitudes both orally and in writing in

predictable and familiar situations; use of auxiliary materials and technique.

The roles of participants in educational activities

Teachers: mainly constructors/designers of the assignment (instructions, schedule of activities, evaluation criteria) and facilitators of the process.

Students: Researchers of the topics involved, active and engaged partners in the learning process, creators of creative and authentic final products.

Progress research and impact assessment is implemented with original and adapted tools - ongoing and at the end of the project, when the results are presented at the ELS open days in the last days of June.

A short feedback session at the end of an activity is an excellent way to emphasize the learning objectives and conclude that students are ready to move on to the next stage. This way, they understand what they did and why, and are better able to apply their knowledge and related skills to future activities. Time invested in feedback increases student understanding and increases their self-efficacy, which will lead to better learning outcomes.

After the lesson, teachers also do a short self-assessment. They comment orally on how far the lesson objectives have been achieved. If they are not satisfied with the goals achieved, the teachers determine what needs to be done differently. It is possible that one student needs more help or another needs a greater challenge. It is a quick, informal method of assessing the impact of teaching on learning.

Allowing students to comment on the work of others in class is likely to cause anxiety and lead to arguments. Students can be sometimes biased or superficial in their judgments. These problems can be avoided by careful planning and a teacher-developed assessment framework (rubric). Clear and understandable training objectives are needed at the very beginning of the activity. Students need to understand the criteria by which they are required to give feedback. It is good to create a pattern of how feedback should be provided. The teacher monitors the process carefully and mediates as needed until the students grasp the concept. Once students are aware of the ground rules, goals, and gain some experience, peer feedback can become a very effective tool.

The summative assessment includes a report on the work performed, an assessment by criteria for final products, rubrics, and a ticket for self-reflection.

It is important to understand that there are no easy routes to valuating results. However, this should not stop attempts to evaluate the impact of the teacher as a leader of the learning process.

VI. CONCLUSION

The activities presented in this study develop a variety of key transversal skills, use diverse tools, and aim to spark students' interest and increase students' motivation. Learning activities are implemented in a modern educational environment and by applying new methods of work: independent study, teamwork and general discussions and presentations.

Over the coming weeks, students will answer a series of surveys, participate in discussions, and provide feedback. The details are yet to be considered, but some conclusions can be drawn.

- The assessment of achievements in the implementation of a learning project includes feedback mechanisms - formal and informal, direct and indirect, individual and collective. Valuation can be conducted at different times and with different technological and methodological means.
- During the implementation of a learning project, a set of skills related to various key competences is developed. In interdisciplinary projects, teachers work together to develop checklists to reflect each student's skills in different areas. The teaching team develops joint criteria for the evaluation of a final product realised in the conditions of PBL.
- In the conditions of long-term implementation of a school project, students move from a one-time application of knowledge to using skills to perform significant tasks. Students make wider connections and learn multiple skills.

Schools like ELS bridge gaps in national policies by giving teachers the freedom to plan, implement and draw on innovative pedagogical practices across all subject areas. Most often, an innovative approach is thought about and planned in small groups of like-minded colleagues who design PBL together.

While we wait for policymakers to reform curricula, we can focus on learning in existing grades and build our lessons beyond the core concepts in middle schools. We can look for ways to provide training resources and engage teachers from different subject areas to apply STEM in their practice. In schools, STEM disciplines and methods can be used as a powerful educational tool to promote critical thinking, problem solving and creativity. Self-assessment skills are a key contributor to broadening one's worldview and deepening one's ability to learn, develop and perform ideas.

REFERENCES

[1] European Commission. European Skills Agenda, Retrieved from https://ec.europa.eu/social/main.jsp?catId=1223

[2] Definition of transversal skills, Northern Regional College, Retrieved from https://www.nrc.ac.uk/courses/what-are-transversal-skills#:~:text=Transversal%20skills%20are%20skills%20that,different%20 tasks%20and%20job%20roles)

[3] Barbara Hursh, Paul Haas, and Michael Moore, "An Interdisciplinary Model to Implement General Education", Journal of Higher Education, vol 54, №1, 1983, pp. 42-59.

[4] Holley, K. Interdisciplinary Curriculum and Learning in Higher Education. Oxford Research Encyclopedia of Education. Retrieved from https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-138

[5] Richard Bakken, James Dobbs, The Importance of Four Types of Knowledge for Leader Preparation in Radically Different Settings: Reflections on Data from a Case Study in Qatar and Teaching at the United States Military Academy, Creighton Journal of Interdisciplinary Leadership Vol. 2, No. 2, 2016, pp. 17 – 23

[6] Koes-H, S., Jewaru, A.A., Fuadah, F., Amalia, N.U., Sutaphan, S., Yuenyong, C., & Yuenyong, J. (2021). Developing the lesson plan of STEM education: designing Moringa leaf tea. Journal of Physics: Conference Series, 1835.

[7] Early Childhood and School Education Group, Department of Education and Training, Victoria (2016). A Reflective Guide to Mentoring and being a teacher-mentor, Retrieved from https://www.education.vic.gov.au/Documents/school/teachers/profdev/reflectiveguidetomentoringschools.pdf

[8] K. Dyulgerova, D. Atanasova (2022) Assessment Teaching and Learning in an Interdisciplinary High School Project, INTED2022 Proceedings, pp. 4497-4506.

[9] Geoffrey T. Crisp, Integrative assessment: reframing assessment practice for current and future learning, Assessment & Evaluation in Higher Education, Vol. 37, No. 1, February 2012, 33–43

[10] Thuy Thu Vu, Designing Integrative Assessment to Enhance Student Learning for Their Future, Retrieved from https://files.eric.ed.gov/fulltext/ED596772.pdf

[11] Roger Crawford (2000) Information technology in secondary schools and its impact on training information technology teachers, Journal of Information Technology for Teacher Education, 9:2, 183-198, DOI: 10.1080/14759390000200082

[12] Brown, B. D., Horn, R. S., & King, G. (2018) The Effective Implementation of Professional Learning Communities. Alabama Journal of Educational Leadership, 5, 53-59.

[13] Kappan, P. D. (2016). The futility of PLC lite. 97(6), 69-71. DOI:10.1177/0031721716636878

[14] Fulton, K., & Ted, B. (2016). High Impact Teaching Strategies: Excellence in Teaching and Learning. Victoria. Department of Education and Training.

[15] K. Shoylekova, B. Ivanova (2022) Analysis of the Results Achieved by Different Forms of Teaching Computer Science at Secondary Schools, EDULEARN22 Proceedings, pp. 1986-1991.

[16] B. Ivanova, K. Shoylekova, R. Rusev (2022). Identifying the Challenges in the Field of Computer Science in Distance Learning in Secondary Education Proceedings of INTED2022 Conference 2022, pp.5824-5833, doi: 10.21125/inted.2022.1489

[17] Dimitrov,M,N.Venelinova (2019). Smart, Secure and Safe Energy Management Approach – An Education Framework Improving the Competence Grid of the Professionals in the Energy Sector 25th International conference KNOWLEDGE-BASED ORGANIZATION, vol. XXV, No 1, 213-218,

[18] Hipkins, Rosemary. (2007). Assessing Key Competencies: Why Would We? How Could We?. Retrieved from https://www.nzcer.org.nz/system/files/Key_Competencies.pdf

[19] Knight, Peter and Page, Anna (2007). The assessment of 'wicked' competences. Practice-based Professional Learning Centre, Centre for Excellence in Teaching and Learning, The Open University. Retrieved from http://oro.open.ac.uk/76353/1/Knight-and-Page-The-assessment-of-wicked-competences-2007.pdf

When User Experience Design Meets LMS Desktop Users: How to Mitigate Mobile-Oriented Interface?

M. Banek Zorica*, J. Klindžić **
* Faculty of Humanities and Social Sciences, University of Zagreb/Department for Information and Communication Sciences, Zagreb, Croatia
** Faculty of Humanities and Social Sciences, University of Zagreb/E-learning Support Center, Zagreb, Croatia
mbanek@ffzg.hr, jasmin@ffzg.hr

User experience design (UX design) of Learning Management Systems (LMS) is recently becoming more mobile device oriented, since large numbers of users are using mobile devices (smartphones primarily) to access learning content and activities. In this paper we will give an overview of device and platform use at the large HEI in Croatia (Faculty of Humanities and Social Sciences, University of Zagreb), based on 3-year survey of first year undergraduate students (over 2100 respondents) as well as the possible steps to mitigate the issues with the mobile-oriented LMS interface (Moodle 4.x) when used on desktop devices. While move to the mobile-oriented interface design is helping most of the users, some of the teachers and content creators, who by default need to use desktop devices (PCs and laptops) have experienced certain issues with LMS usage: too large fonts and icons (designed for touchscreen and small screen devices), several features and navigation elements no longer initially visible (needs for additional interface clicks and new menus) and so on. Most of those issues could be mitigated using customized interface themes and modifications, but some require organizational and methodical changes that we will discuss in detail.

Keywords – UX design; LMS; mobile access; MOODLE;

I. INTRODUCTION

User experience design (UX design) of Learning Management Systems (LMS) is recently becoming more mobile device oriented, since large numbers of users are using mobile devices (smartphones primarily) to access online content, from regular browsing to learning content and online activities. The available statistics for the past 12 months (January 2022 - January 2023) are showing that the global mobile market share has reached 59.76% and desktop market share stands at 40.24% [1]. Both the screen size and orientation of the screens are playing an important role when considering the mobile end-user experience (both teachers/content creators and students/learners) - mobile screens are smaller than desktop screens and their orientation is vertical (most of the time, especially for the mobile phone users). Tablet screens and their users face additional issues, since they are usually larger than mobile phone screens, but still closer to some desktop/laptop screens size and part of the time they are also used in horizontal/landscape mode.

There are numerous UX definitions, and they vary from describing what UX is not - guessing what your user need, web development, interface aesthetics or information technology itself [2], visual design or service design [3] or what it should not be confined to: pure usability [4], and the International Organization for Standardization defines the user experience as a "person's perceptions and responses resulting from the use and or expected use of a product, system or service." [5], and that, according to Jusoh, Almajali and Abu Albasal includes "the user's emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours [...] that occur before, during and after use" [6] of a certain system, interface, service. Of course, UX usually includes most of the things that are usually dismissed in those definitions (usability, web development, interaction design, visual design…), but it still should not be limited or perceived like just one of those facets and parts, while excluding or omitting others.

Morville's model describes that notion best by using the famous UX honeycomb [4]:



Figure I. UX honeycomb [4]

in which different facets: useful, usable, desirable, valuable, findable, accessible and credible can help UX designers and other people involved in the process to define their priorities. While all facets cannot always be balanced and prioritized and equally developed, they should all be taken into consideration, to a varying degree (depending on the organization / software needs/priorities).

Learning management systems (LMS) are web-based software platforms that provide the access to and delivery of online content and activities for learning purposes, facilitating the delivery of online, face-to-face or blended

courses [7], [8]. They accomplish that by organizing content and activities into separate courses with their respective teachers and students enrolled in them, enabling communication, tracking of student progress, activity completion and grades, providing access to content in a structured way combined with almost 24/7/365 availability (but there are caveats to that availability such as: motivation, network and infrastructure readiness as well as time available to students and teachers). Their usage has grown considerably in the past 20+ years, and one of the LMS solutions with the largest user base, and open source LMS MOODLE [9], has been installed on over 165.000 sites with over 43.300.000 courses and more than 350.000.000 users globally [10].

Faculty of Humanities and Social Sciences University of Zagreb (FHSS) is the largest higher education institution in Croatia which carries out research activities and has over 100 university programs in the field of humanities and social sciences (undergraduate, graduate and postgraduate levels). The FHSS is a part of the University of Zagreb, the oldest university in Croatia and one of the oldest universities in Europe, and has over 7000 students, more than 800 teachers and 700+ external associates.

E-learning and educational technology implementation at FHSS began in 2002. as a part of the Ministry of Science and Technology project "Organization of Information and Knowledge in the Electronic Learning Environment" (OIZEOO) [11], and it resulted in the faculty-wide implementation of MOODLE based LMS called Omega (https://omega.ffzg.hr). In January 2023, LMS Omega has over 9.800 users, out of which there are 700 teachers and over 600 external associates teaching on 4.500+ online courses. All the support, internal development and implementation of the educational technologies and services at FHSS is done by FHSS E-learning Support Center (ESC). Since August 2022, LMS Omega has been running MOODLE version 4.x, which has undergone long-term and serious UX design and development by the MOODLE HQ (creators of the software) in cooperation with the wider global community, with improvements and changes in new primary navigation, new secondary menus and course components, improved student and mobile experiences [12].

After the upgrade of LMS Omega to MOODLE 4.x in August 2022, and the basic introduction of users to the new interface and UX through webinars and workshops (both teachers and students), a number of complaints and comments were received regarding the changes. As usual, with the significant changes to the organization and look of the interface, some of those complaints were resolved quickly with creation of additional instructions (text, video and animations). But some of those complaints persisted, and ESC was not able to resolve all of them with additional education and instruction for users. Most of the complaints came from teachers and content creators using the desktop or laptop devices (computers), because the UX / UI in Moodle 4.x is more oriented on mobile device users (largely student population), and it ranged from too large fonts and icons (which can easily be solved using internet browser font size commands), several important features and navigation elements no longer initially visible to the user (needs for additional interface clicks and introduction of new drop-down menus instead of navigational blocks on the left or right side of the screen) and course content being "too tall" - because the approximate 30% increase in height of the resource and activity names visible on the homepage of the course. Students and other users who accessed the LMS using mobile devices had, on the other hand, commented that the interface is more usable, touch-screen friendly, simpler and easier to use. Some of the teachers and content creators have also had positive feedback (while using the LMS on mobile devices for grading, giving feedback and using previously created activities and resources).

So keeping a balance between desktop and mobile users is our main challenge. Still, the majority of the students are moving towards mobile device use. Although pandemics was the main motivator and trigger towards massive transfer to mobile devices, users have also discovered the benefits and decided to use it even further in the future. So, our chosen approach is to raise awareness among the desktop users, majority of them being the professors and teaching staff, that there is a necessary shift ahead to mobile environment and adaptation of the learning objects / content. In order to scientifically prove this position, we have conducted a research that showcases our statements with numbers.

II. MATERIAL AND METHOD

ESC had previously ran a 3-year survey on first year undergraduate students enrolled at FHSS during the initial LMS training in academic years 2020 to 2022 (held in September each year, before the classes start) regarding the devices and other equipment that those students possess in order to try to prepare the teachers, educational software solutions and infrastructure to students' needs. LMS initial mass-training for undergraduate students has been running since 2013 and has proven to be an effective way to quickly introduce freshmen to LMS and educational technologies present at the FHSS. From 2013 it was run in 45 30-minute workshops in PC labs, and since 2019 it was moved completely online, and now it lasts 10 working days (about 3 hours workload total). Survey was done using MOODLE Feedback activity, and over 90% of first-year students each year have answered. It was an 11-question anonymous survey, with questions ranging from device types they are using (desktop, laptop, smartphone, tablet - multiple possible devices), age of those devices, operating systems (on all devices that they possess), and what type of internet connection are they going to use to access online services at FHSS (if they had that information). Primary motivation was COVID-19 measures and online classes in 2020, but it continued later on because it provided valuable information for both teachers, ESC and FHSS management. The number of surveys submitted per year: 2020 - N = 721, 2021 - N = 709 and 2022 - N = 698 (total of 2128).

Additional survey was also run for 3 years in academic years 2020 to 2022 regarding the speed of the students' internet connection, where first-year students had to use OOKLA Speedtest website [13] to measure their current

speed and submit the result, also using MOODLE Feedback anonymous activity. The number of surveys per year is similar to the device type survey, but we are only using the sample of 100 submitted surveys per year in this article (total of 300).

In 2023, both of those surveys will be integrated and additional questions will be added, regarding the UX / UI and development and implementation of new LMS features, as well as accessibility issues/customizations. Additional survey for teachers and content creators on similar topics will also be implemented later this year as a longitudinal study.

III. RESULTS

As mentioned before, the number of surveys submitted per year was: 2020 - N = 721, 2021 - N = 709 and 2022 - N = 698, with total of 2128. When it comes to device distribution among the student population as presented in Figure 1, the least number of students were using tablets (from 24 in 2020 to 33 in 2022). The limiting factor of those devices could be their price and battery duration, as well as limited connectivity options (the low-cost models are limited to Wi-Fi only). Desktop users were following, with a visible decline in numbers (from 139 in 2020 to 96 in 2022), which could be explained with their limited mobility (for students originally living outside Zagreb, the university center), size and price, when compared to laptops and mobile phones. Laptops number is also declining, but not as rapidly (from 613 in 2020 to 581 in 2022), especially given their large initial numbers. They are mobile, can be carried to classes, they have multiple connectivity options and can be considered cost effective. And finally, smartphones are also showing a slight decrease in numbers (from 623 to 602). They are portable, can be carried to classes, they also have multiple connectivity options, but some models are rather costly.

Number of devices with smaller screens (smartphones, tablets, and to some extent laptops, especially ultra-portable models) are overwhelmingly present in the student population. Designing UI and UX in general with the "mobile-first, every other platform later" idea seems to be prudent, giving the students opportunity to enjoy almost all of the UX facets mentioned before: LMS becomes more useful, usable, desirable, valuable, findable, accessible and credible. The devices they are using are also shaping the content the teachers, content providers and institutions are creating and customizing for their consumption. It does also mean that there is still a need to design / customize UI also for students / learners who are still dominantly using desktop or large screen devices (be it because of the financial or other reasons).

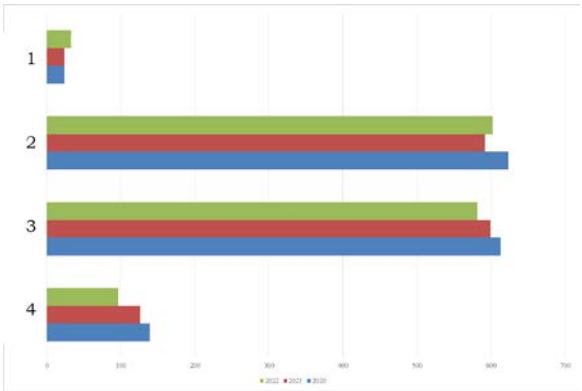


Figure II. What type of device are students using to access educational technologies at FHSS. Legend: 1 – tablet, 2 – smartphone, 3 – laptop, 4 - desktop

Internet access speed was also measured by the students, using OOKLA Speedtest, preferably on the devices and on the location where they will be living while studying at the University of Zagreb (private accommodation, dorms). Although the number of surveys per year is approximately the same as the device used survey (around 2100), we have taken a sample of 100 surveys per year. All the results presented in Table I. were submitted as the OOKLA Speedtest URL, so there are no ambiguities or mistakes that could be present in manual result entry. When compared, average download and upload speed in 2022 (47.29 Mbps / 19.16 Mbps) is almost twice as high as the speeds in 2020 (26.72 Mbps / 11.41 Mbps). Maximum download and upload speeds have risen significantly (92.98 Mpbs / 53.83 Mpbs in 2020 to 240 Mpbs / 124.04 Mbps in 2022), probably as a result of wider introduction of fiber optics home connectivity in Croatia during that period. Minimum download and upload speeds measured are really low, barring any meaningful use of the internet, let alone online educational technologies (from 1.56 Mbps / 0.27 Mbps in 2020 to 4.54 Mbps / 0.32 Mbps in 2022). Upload speeds in general need to be much higher, especially if students and teachers are to use video conferencing or any other modern communication two-way tools that have become more and more a standard recently.

TABLE I. STUDENT INTERNET CONNECTIVITY SPEED RESULTS, SEPTEMBER 2020 – SEPTEMBER 2022

	September 2020	September 2021	September 2022
	Mbps	Mbps	Mbps
Minimum download speed	1.56	2.08	4.54
Minimum upload speed	0.27	0.34	0.27
Maximum download speed	92.98	211.12	92.98
Maximum upload speed	53.83	124.81	53.83
Average download speed	26.72	41.39	26.72
Average upload speed	11.41	20.09	11.41

Some of the measured students' internet connectivity speeds paint a clear picture on why certain parts of the student population are not willing or able to use online educational tools, and why they are having issues with teachers and institutions asking them to use those tools on a daily basis. When it comes to UX / UI issues and their slow internet connections, some students are reporting (through support tickets to ESC) that their older desktop and laptop computers are having issues rendering certain types of audio-visual content (both real time AV content or pre-recorded material).

IV. DISCUSSION

One of the main LMS usage goals is teaching and learning, in a structured way, using modern pedagogy and available technologies (and emphasis should be on modern pedagogy and less on "cool" technological solutions). In order to teach someone, it is important to make them feel at ease with the tools used in the classroom or learning environment, be it digital or physical room. UX design, adaptation and analysis of current LMS could help teachers and students alleviate certain issues regarding usability, usefulness, accessibility and other UX facets when it comes to using LMS in everyday educational contexts. As stated before, since the upgrade to MOODLE 4.x, a number of FHSS users had expressed their comments and complaints, most of those negative comments / complaints coming from teachers and content creators using desktop computers, while students mostly had positive feedback. Since UX design in MOODLE 4.x was oriented towards mobile device users (and they are in our case mostly students), issues arising for the desktop users were to be expected. Some of the issues could be solved using instructional design principles when creating activities and resources in online courses, ie. a number of teachers were having issues with "too tall course content, because several pages long/tall course became much longer" - meaning, for each link and icon combination on the previous MOODLE interface that was in 1 row, in MOODLE 4.x it was presented in 2.5 to 3 rows space. Needless to say that those changes were made both for touchscreen usability and screen reader software for students with vision impairment issues. Upon inspection of those courses, it was obvious that even before the software upgrade of the LMS, there were serious issues regarding the large number of activities and resources present (with 30-50% of those activities and resources being from previous semesters and hidden from students, so only teachers had issues with scrolling pages and pages of activities and resources, while students only saw actual, not hidden materials). In the MOODLE community, very long courses, rich with numerous resources and activities that keep students and teachers scrolling for a long time in order to find something is called "the scroll of death" [14] and it should be avoided because of the information overload it presents on the users. So, part of the issues that teachers are having could be resolved by changing the course design / structure to better suite their needs, while also removing the unnecessary activities / resources from the previous academic years (decluttering the course home page in the process).

V. CONCLUSIONS

Large number of FHSS students are using mobile devices and small screen devices to access our MOODLE based LMS Omega, and the implementation of MOODLE 4.x version oriented towards mobile touchscreen users has helped them to use the system in a more efficient and easy way. Most of their previous UX / UI issues that students have faced are not gone completely, but are mitigated to a large extent, with possible improvements in announced MOODLE releases with UX issues in mind.

On the other hand, since most FHSS teachers and content creators are using desktop computers for their content developing and creation purposes, ESC will have to find possible solutions for their UX issues.

Some of the solutions include: further webinars and education events for teachers on UX changes in MOODLE 4.x onward, tips and tricks for desktop users on internet browser / operating system solutions and tools, information on reasons why those changes were needed (majority of students using mobile devices, accessibility improvements for visually impaired users and so on), also education on using instructional design principles that could help alleviate some of the "the scroll of death" situations, and probably selection and customization of a new MOODLE 4.x theme for desktop users.

Our future research in UX/UI design and LMS Moodle will definitely include standardized UX satisfaction surveys and cover the device types used (models and year of production).

REFERENCES

[1] Desktop vs Mobile Market Share Worldwide | Statcounter Global Stats. (n.d.). StatCounter Global Stats. https://gs.statcounter.com/platform-market-share/desktop-mobile/worldwide Retrieved 19 Jan. 2023.

[2] What UX is NOT According to The Experts – UX Design Recruiting and Staffing Agency | UX Hires. (n.d.). https://uxhires.com/what-user-experience-is-not-according-to-experts/ Retrieved 21 Jan. 2023

[3] Iorgulescu, Costin. "This Is NOT User Experience." Medium, 29 Dec. 2018, https://uxdesign.cc/this-is-not-user-experience-d1c763224038 Accessed 19 Jan. 2023.

[4] Morville, P. (2004, June 21). User Experience Design. Semantic Studios. Retrieved January 10, 2023, from http://semanticstudios.com/user_experience_design/

[5] International Organization for Standardization. (2010). ISO 9241-210:2010(en) Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. Retrieved January 8, 2023, from https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en

[6] Jusoh, Shaidah & Almajali, Sufyan & Basal, Abdelraheem. (2019). A Study Of User Experience For E-Learning Using Interactive Online Technologies. Journal of Theoretical and Applied Information Technology. 97. 4036-4047.

[7] Barreto, D., Rottmann, A., Rabidoux, S., Daisyane Barreto. Learning management systems. EdTech Books. Retrieved February 19, 2023, from https://edtechbooks.org/learning_management_systems

[8] Yueh, H.-P., & Hsu, S. (2008). Designing a learning management system to support instruction. Communications of the ACM, 51(4), 59–63. doi:10.1145/1330311.1330324

[9] Poulová, Petra & Šimonová, Ivana & Maněnová, Martina. (2015). Which One, or Another? Comparative Analysis of Selected LMS. Procedia - Social and Behavioral Sciences. 186. 1302-1308. 10.1016/j.sbspro.2015.04.052.

[10] Moodle statistics. (2023, January 30). Retrieved February 5, 2023, from https://stats.moodle.org/

[11] Ministry of Science, Education and Sport. (2002). Organizacija, upravljanje i razmjena znanja u elektroničkom obrazovnom okruženju. MSES. Retrieved January 11, 2023, from http://zprojekti.mzos.hr/public/c-prikaz_det.asp?psid=0&ID=1265

[12] Fry, A. (2022, April 19). Welcome to a new user experience – Moodle 4.0 is here! Moodle.com. Retrieved January 4, 2023, from https://moodle.com/news/moodle-4-is-here/

[13] Speedtest by OOKLA. (n.d.). OOKLA. Retrieved January 10, 2023, from https://www.speedtest.net/

[14] Ballantyne, Emily. "Small Optimizations: Five Simple Tweaks to Maximize Student Engagement in Online Courses." Proceedings of the Atlantic Universities' Teaching Showcase 21 (2017): 33-42.

Survey of recent research topics on effective use of communication tools in higher education

M. Nurminen*, H-M. Järvinen†, J. Viteli† M. Saari*, P. Rantanen*
* Tampere University/Faculty of Information Technology and Communication Sciences, Pori, Finland
† Tampere University/Faculty of Information Technology and Communication Sciences, Tampere, Finland
mikko.nurminen@tuni.fi

***Abstract*—Effective communication is crucial for the success and well being of students and faculty members in higher education. Communication can improve student's learning outcomes by clarifying the expectations and learning objectives on both curriculum and course levels. Further inclusive communication can enable creating a sense of community and belonging among students and the faculty. In this article, we present the results of a literary survey conducted on recent research discussing communication practices and software communication tools in higher education. Both linear and interactive communication occur in variable contexts and between different participating roles. During campus lock downs caused by COVID-19 pandemic, teaching and related communication were moved into fully distance mode. For teaching and learning higher education institutions needed to rely on existing and newly created software platforms. Synchronous teaching and communication could be conducted using video and voice conferencing tools, like Zoom or Teams. Similarly asynchronous communication could use tools including forums or emails. The tools adopted during the pandemic largely remain in use after the comeback of blended learning. For future research one goal could be the creation of communication models that can integrate the roles, practices, and tools currently interacting in higher education.**

Keywords—higher education, communication, communication tools, COVID-19

I. INTRODUCTION

Global studies conducted by the International Association of Universities (IAU) reveal how the COVID-19 pandemic had a strong impact on teaching and learning in higher education institutions (HEIs). [1] [2] The pandemic effectively forced most universities and other higher education institutions to extensively adapt online distance teaching methods and tools, supplanting the more traditional contact teaching on campuses. The comparison of results of the two global studies shows that the move to online distance teaching was intensified with as the pandemic wore on, with the number of surveyed institutions offering online distance teaching increasing from 67% in 2020 to 89% in 2022, while even as of 2022 11% of these institutions did not offer remote teaching.

The move towards distance teaching and learning implies that appropriate modifications were required in both the teaching methods and also in the online teaching and communication tools that the faculty and students use, or that even new methods or tools needed be utilized. In this study the modifications HEIs undertook especially in their use of teaching related communication tools are reviewed, guided by the following two research questions:

1) Were there changes in the use of communication tools in distance teaching and learning in higher education as the result of COVID-19? What were these changes?
2) How did the students and faculty experience these possible changes?

The method applied in seeking answers to these research questions was the literary survey method. [3] Literary survey method was used to discover if the recently published scientific literary covered matters related to the research questions.

The hypothesis of this work is that COVID-19 pandemic's effects were so severe also to higher education, that the literature survey will reveal distinct ways in which higher education institutions chose to cope with these effects. It is also hypothesized that the survey enables exploring the various ways students and faculty of these institutions experienced these changes, even though these experiences could be somewhat intermixed with their experiences during the outset years of the pandemic.

The structure of this paper is as follows: In Section II, we review the related research about communication in higher education. In Section III, we present systematic mapping survey research method. Section IV introduces the findings of the survey. In discussion Section V we expand on the findings of the survey and the study is summarized in Section VI.

II. BACKGROUND

The increased use of ICT technology in higher education distance learning and teaching has been an on-going trend worldwide, aided by the goals set by governing bodies, like EU's Digital Education Action Plans [4], and higher education institutions themselves, with COVID-19 pandemic giving urgency to this trend. [5] While distance learning must occur online, the teaching methods, communication tools and communication modes (synchronous and/or asynchronous) need to be carefully considered to ensure positive learning outcomes and technically efficient teaching. Xie et al. [6] analyzed synchronous and asynchronous online distance learning environments and their communication tools. Synchronous tools offer opportunities for real-time communication, which for example

enables students cooperatively work on their group work. Asynchronous communication implies course implementations where the students and teachers send messages, to which the their recipient will answer at the time of their own choosing. [7] In the theorical framework by Xie et al. asynchronous communication tools include discussion forums, blogs, emails and social media messages improve students reading and writing skills. Teachers give feedback for student's activities using asynchronous tools, and provide materials as pre-recorded lecture videos, material as files (PDFs, PowerPoint presentations, etc.). Asynchronous learning demands more rigorous self-regulation from the learner. These are contrasted with synchronous e-learning tools for communication and feedback, and instruction material distribution. In synchronous e-learning communication tools video, audio and/or web conferencing, live chats, white boarding, and application sharing. Students receive immediate direct input and feedback from the teachers. Surveyed students felt more comfortable using asynchronous communication tools, for they can be easily accessed offline. Students preferred to have an instructor present to promptly receive feedback and thus improve their learning. Doubtful and shy students preferred asynchronous mode. Asynchronous mode might be uncomfortable when engaging in computer mediated discussions. As both synchronous and asynchronous e-learning and communication tools are shown to have advantages not available in the other mode, Xie et al. come to conclusion, that best mode is a blended mode where synchronous and asynchronous modes are applied when appropriate.

Santos et al. presented a taxonomy for online based communication technologies [8], which is applied when we categorize the communication tools discovered in the mapping study presented in this article. The taxonomy is shown in Table I.

Taxonomy includes both synchronous and asynchronous tools, divided into categories based on their communication capabilities. In the Publishing and sharing technologies category are synchronous tools which enable users to publish instructions or other materials to other students or teachers. These tools often have communication capabilities beyond publishing and sharing, like the discussion forums in Moodle or viewer comments in Youtube. Whether these communication capabilities are utilized is up to the teachers to decide. Tools in the Collaborative technologies category provide platforms for collaboration where discussions and files can be shared with others. Collaborative tools are mostly synchronous, with the exception of tools like Slack, which provides instant messaging. Interpersonal communication technologies category has three sub-categories: email, instant messaging, and video conferencing and voice systems. With the tools in Social networks category, Interpersonal communication technologies which students can use in the on and off their studies. Social media usage in education has been extensively researched, with some studies showing that teachers use social media concentrating in teaching, while students, especially younger university students, use social media to build communities. [9] [10]

III. LITERARY SURVEY

Systematic mapping study was selected as the research method for the conducted research. In their article Petersen et al. [3] describe the process involved in this method and contrast it with systematic literary review. The goal of a systematic mapping study is to "give an overview of a research area through classification and counting contributions in relation to the categories of that classification".

Searching from knowledge databases is the most common way of finding the articles for survey in systematic mapping studies, and our research followed this path, too. We used the ACM Digital Library service, which provides a web page for searching and accessing the publisher's article database. The selection of the search terms to scope the materials for our systematic mapping study was straight-forward and framed by the research questions. After evaluating some search terms and their combinations, search terms that produced relevant search results were "higher education", "distance learning", and "communication tools". Suitable combination of search terms was required, as for example the search for research articles published between 2020 and 2023 using combination of search terms "higher education", "covid-19", and "communication" yielded 704 results, excessive amount to evaluate. However, the need to find a narrower result set left out some potentially interesting articles that didn't match the search terms, such as [11] whose findings remind that some students or educators prefer face-to-face teaching due the benefits it provides to the students with certain learning styles.

The final selection criteria for a publication to be included in the mapping study were the following:

1) The search string we used combined the search terms: "higher education" AND "communication tools" AND "distance learning". ACM Digital Library search with this search string yielded 79 results, which various content types (peer-reviewed research articles, abstracts, etc.).
2) Other content types than research articles were removed from the result set, reducing it to 51 entries.
3) As the effects of COVID-19 were of interest to us, article published prior to 2020 were filtered from the results, which produced 23 results.
4) These 23 research articles were evaluated, and if they held no information related to our research questions, they were removed from the mapping study. However, if education related communication methods or practices, or communication tools was mentioned in the study, that study was included in the mapping study.

Studies that were excluded from the mapping study included [12], which focused on Learning Analytics, [13] about university opening its technical support to local community, and [14] which describes technologies applicable for project-based approach training teachers, study

TABLE I: Taxonomy of online based communication technologies (from [8])

Categories	Sub-categories	Examples
Publishing and sharing technologies		Youtube, Moodle, Flickr, Blogs, etc.
Collaborative technologies		Google Drive, Slack, Wiki, etc
Interpersonal communication technologies	Electronic mail	Gmail, Hotmail, etc.
	Instant messaging	Messenger, WhatsApp, SMS, etc.
	Videoconferencing and voice systems	Skype, Google Hangouts, et
Social networks		Facebook, Twitter, Linkedin, etc

of applying blended synchronous and asynchronous online learning in Indonesia [15] , a study comparing two LMSes [16], and a study of adoption of a student mobility program [17].

IV. Results of survey

Table II shows the articles that were included in the mapping study. Where an X is entered on the row of an article for any of the categories on the columns, the article was evaluated to include descriptions or discussion related to that category.

Depending on the article, the communication and communication tools were examined in various contexts. In some articles the tools were discussed in the context of course development and teaching, mentioning specific tools and practices. [19] [23] In contrast to those, some studies examined the evolution and properties of distance teaching and related communication tool use on more general level, like in [26], [21], [20] and [22]. The most closely scoped study examined Remote Pair Programming tasks as part of distance learning, highlighting how communication capabilities must be integrated to any online environment for it to be usable. [25]

Our first research question was: "Were there changes in the use of communication methods and tools in distance teaching and learning in higher education as the result of COVID-19? What were these changes?" When we look at the articles, we can see that there were changes, but changes in communication were linked to holistic move to distance learning pedagogical methods, whether it was e-learning [18], blended learning [30], or even "emergency remote teaching" [31].

[25] suggested that providing guidance to teachers and students on how they can best to use existing communication tools was preferable, instead of deploying or developing new tools. Other articles gave more weight to this suggestion, as most communication tools were already publicly available before the pandemic. These tools include Teams [31] [25], Zoom [18] [25] and tools used in Chinese higher education institutions, including QQ and WeChat [23] [24].

The communication methods and tools introduced during the pandemic seem to have had favorable effects on teaching and learning, as teachers reportedly are going to persist in using them after the pandemic, too. According to [31] 82% of the teachers had used formal communication tools during the pandemic and 78% will continue with them. These formal communication tools included forums, Piazza, and Teams.

The second research question was: "How did the students and faculty experience these possible changes [in communication methods and tools]?" [18] findings included that students were negatively impacted by the limitations and incorrect choices of communication tools in e-learning courses. These partly led to weaker learning results. The shift from on campus face-to-face teaching to fully online classes caused stress and adjustment difficulties for some students. [19]. [30] found that on a course which used blended learning students were overall satisfied interaction during the course, be it student-student interaction, student-lecturer interaction, student-blended learning system.

[14] survey findings included that teachers' digital competencies before pandemic correlated with their experience in e-learning, and teacher's implemented their online teaching. Inexperienced teachers had more negative attitudes towards e-learning, and they could fail to see how quality e-learning can benefit students. When teacher's opinions about increased use of ICT and moving to online teaching were studied in [20], practical problems surfaced, such as unspecific guidance in application of ICT, financial burdens brought on by acquiring the needed hardware and software, inadequate support staff, and lacking supervision and inspection of the application of ICT in teaching. These challenges could hinder effective communication in many countries and HEIs.

V. Discussion

This article's goal was to evaluate how COVID-19 affected higher education communication in teaching and how participants of academic life experienced those effects. A survey of recent research articles was used to gather findings from globally diverse authors covering local and global aspects of the changes.

The effects presented in the surveyed papers were comparable to the changes explored in larger scale reports, such as [5] and [31]. Increase in innovative online teaching practices and tools and their impacts on students are seen to have longer term implications on teaching. As [5] states it, COVID-19 worked as a catalyst for these changes.

TABLE II: Articles included in the mapping study and categories they discussed

Cited article	Communication tool(s)	Communication method(s) or practice(s)	Student perceptions	Faculty perceptions
[18]	X	X	X	
[19]	X	X	X	X
[20]				X
[21]	X	X	X	
[22]	X	X	X	X
[23]	X		X	X
[24]		X		X
[25]	X			X
[26]		X	X	X
[27]	X	X		
[28]		X		X
[29]	X			X
[30]	X		X	
[31]		X		X
[14]				X

Both [21] and [22] discuss Massive Open Online Course (MOOC) providers, which offer their online courses and materials to everyone, often at no cost to their students. Of MOOC providers Coursera, EdX, Khan academy, Udemy, and Open Yale courses were mentioned in the articles. These MOOCs are sometimes seen as direct competitor to traditional higher education, the quality and availability of the MOOCs has increased in last ten years. [32] However, established, traditional education is seen as having advantages for career advancement and rigorous study structures.

A noteworthy matter is the implicit requirement of integrating communication capabilities to MOOCs and LMS's. These online learning systems do not merely offer students material and exercise feedback and results, but also a wide range of peer-to-peer and student-teacher communication and feedback needs occur for the course to successfully proceed. Moodle LMS was discussed in [23] and [27]. It is also noteworthy from the viewpoint of communications that Neto et al. describe the creation of a new "Planning and Developing Courses in Distance Learning Environments" course, with communication taking place on Moodle's general and module-specific forums. This combined with the number of articles where Moodle is mentioned as the learning platform in use or being deployed indicate that Moodle's communication affordances are adequate for higher education courses. However, in the writers' experience, we find that Moodle's forums are somewhat cumbersome, and only lend themselves to asynchronous communication. Therefore, if synchronous communication is required in the teaching, another communication tool should be used to complement discussion forums.

Future work could be undertaken to formulate a communication model for higher education. This model would present the roles and tools involved in communication and could provide a clear structure for designing communication.

VI. Conclusion

This paper presented a survey of recent research on communication in higher education. The focus of the survey was the pandemic-related changes in the use of communication tools and methods, and the experiences students and faculty had with these changes. Based on the survey, communication tools used have changed globally to be in line with distance and blended learning models, and online communication tools that were already in use prior to the pandemic saw increased usage. As some of the published research highlighted, the changes in communication tool use could be sustained as long-term solutions, as teachers indicated they were satisfied with their tools and the communication they enabled. However, in surveyed articles, the teachers' and students' experiences with this "new normal" of communication and teaching methods varied. While some students felt stress in adapting to the new education landscape, overall students were at least somewhat positive about the opportunities for learning provided. In this new landscape, teachers were required to mix modes in their online courses, for example, by requiring asynchronous self-paced learning from students while applying synchronous teaching, such as live-streamed lectures. In conclusion, teachers have adapted or launched new innovative ways of teaching and learning as a result of the pandemic, and this innovative development is likely to endure into the future.

References

[1] G. Marinoni, H. Vant Land, and T. Jensen, "The impact of covid-19 on higher education around the world," *IAU Global Survey Report*, 2020.

[2] T. Jensen, G. Marinoni, and H. van't Land, "Higher education one year into the covid-19 pandemic, second iau global survey report," *International Association of Universities (IAU), Paris, France*, 2022.

[3] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and Software Technology*, vol. 64, pp. 1–18, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950584915000646

[4] European Union, "Digital education action plan 2021-2027," https://ec.europa.eu/education/sites/education/files/digital-education-action-plan-2021-2027_en.pdf, 2021, accessed: 2023-02-05.

[5] R. Y. Chan, K. Bista, and R. M. Allen, *Online teaching and learning in higher education during Covid-19 : international perspectives and experiences*, ser. Routledge studies in global student mobility. New York, NY ;: Routledge, Taylor & Francis Group, 2022.

[6] H. Xie, W. Liu, J. Bhairma, and E. shim, "Analysis of synchronous and asynchronous e-learning environments," in *Proceedings of the 2018 3rd Joint International Information Technology, Mechanical and Electronic Engineering Conference (JIMEC 2018)*. Atlantis Press, 2018/12, pp. 270–274. [Online]. Available: https://doi.org/10.2991/jimec-18.2018.58

[7] S. Hrastinski, "Asynchronous and synchronous e-learning," *Educause quarterly*, vol. 31, no. 4, pp. 51–55, 2008.

[8] H. Santos, J. Batista, and R. P. Marques, "Digital transformation in higher education: the use of communication technologies by students," *Procedia Computer Science*, vol. 164, pp. 123–130, Jan. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050919322021

[9] A. Z. Klein, J. C. d. S. F. Junior, J. V. V. M. M. d. Silva, J. L. V. Barbosa, and L. Baldasso, "The educational affordances of mobile instant messaging (mim): Results of whatsapp used in higher education," *International Journal of Distance Education Technologies (IJDET)*, vol. 16, no. 2, pp. 51–64, Apr. 2018, publisher: IGI Global. [Online]. Available: https://www.igi-global.com/article/the-educational-affordances-of-mobile-instant-messaging-mim/201861

[10] O. Castro-Romero, "Social media as learning tool in higher education: The case of mexico and south korea," *Sinéctica*, no. 44, pp. 1–16, 2015.

[11] C. Chhetri, ""i lost track of things": Student experiences of remote learning in the covid-19 pandemic," in *Proceedings of the 21st Annual Conference on Information Technology Education*, ser. SIGITE '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 314–319. [Online]. Available: https://doi.org/10.1145/3368308.3415413

[12] T. P. Falcão, R. F. Mello, R. L. Rodrigues, J. R. B. Diniz, Y.-S. Tsai, and D. Gašević, "Perceptions and expectations about learning analytics from a brazilian higher education institution," in *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, ser. LAK '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 240–249. [Online]. Available: https://doi.org/10.1145/3375462.3375478

[13] J. T. Biehl, R. Farzan, and Y. Zhou, "Can Anybody Help Me?: Using Community Help Desk Call Records to Examine the Impact of Digital Divides During a Global Pandemic," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–13. [Online]. Available: https://doi.org/10.1145/3491102.3517693

[14] P. S. Bazhina, T. G. Sepik, O. P. Zhigalova, and T. L. Kopus, "Technologies for project-based approach to the system of teachers training," in *Proceedings of the III International Scientific and Practical Conference*, ser. DEFIN '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–5. [Online]. Available: https://doi.org/10.1145/3388984.3390619

[15] H. Wibawanto and Roemintoyo, "The Learning Method of Society 5.0 During New Normal in Indonesia: Case Study: Vocational Highschool in Surakarta, Indonesia," in *Proceedings of the 4th International Conference on Learning Innovation and Quality Education*, ser. ICLIQE 2020. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–6. [Online]. Available: https://doi.org/10.1145/3452144.3452196

[16] M. J. J. Gumasing, A. B. Vasquez, A. L. S. Doctora, and W. D. D. Perez, "Usability Evaluation of Online Learning Management System: Comparison between Blackboard and Canvas," in *2022 The 9th International Conference on Industrial Engineering and Applications (Europe)*, ser. ICIEA-2022-Europe. New York, NY, USA: Association for Computing Machinery, 2022, pp. 25–31. [Online]. Available: https://doi.org/10.1145/3523132.3523137

[17] H. Pham, N. Huynh, and H. Nguyen, "Virtual Mobility Adoption: a Study of Factors Affecting Students’ Satisfaction toward an Online English Program at a Young Vietnamese Higher Education Institution: Virtual Mobility Adoption," in *2022 13th International Conference on E-Education, E-Business, E-Management, and E-Learning (IC4E)*, ser. IC4E 2022. New York, NY, USA: Association for Computing Machinery, 2022, pp. 45–52. [Online]. Available: https://doi.org/10.1145/3514262.3514282

[18] H. Dao Thi Thu and N. Duong Hong, "A Survey on Students’ Satisfaction with Synchronous E-learning at Public Universities in Vietnam during the COVID-19," in *Proceedings of the 5th International Conference on Education and Multimedia Technology*, ser. ICEMT '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 196–202. [Online]. Available: https://doi.org/10.1145/3481056.3481069

[19] M. M. B. De Gracia, J. C. C. Cinches, A. C. Galang, and C. J. Villanueva, "An Analysis of the Perception of Students and Teachers towards the Home-Based Online Learning Modality and the Students’ Academic Achievement of a Private School in the National Capital Region (NCR) of the Philippines," in *2022 13th International Conference on E-Education, E-Business, E-Management, and E-Learning (IC4E)*, ser. IC4E 2022. New York, NY, USA: Association for Computing Machinery, 2022, pp. 216–223. [Online]. Available: https://doi.org/10.1145/3514262.3514275

[20] N. Ta Trung, "Difficulties in Applying Information and Communication Technologies in Teaching at Public Universities of Vietnam," in *Proceedings of the 8th International Conference on Frontiers of Educational Technologies*, ser. ICFET '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 29–35. [Online]. Available: https://doi.org/10.1145/3545862.3545867

[21] B. U. Zholdasbekova, K. N. Galay, S. K. Mizanbekov, Z. A. Bayanbayeva, and J. B. Dzhalamova, "E-learning as a factor of quality assurance of professional training of future teachers of Russian language and literature," in *Proceedings of the 6th International Conference on Engineering & MIS 2020*, ser. ICEMIS'20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–6. [Online]. Available: https://doi.org/10.1145/3410352.3410839

[22] P. Huang and H. C. Lucas, "Early Exploration of MOOCs in the U.S. Higher Education: An Absorptive Capacity Perspective," *ACM Transactions on Management Information Systems*, vol. 12, no. 3, pp. 22:1–22:28, 2021. [Online]. Available: https://doi.org/10.1145/3456295

[23] F. Shu, C. Zhao, Q. Liu, H. Li, and Y. Huang, "Enhancing Adults' Online Course Learning Behavior Performance through Live Class in Distance and Open Education: Learning Analysis Based on a Course," in *Proceedings of the 6th International Conference on Education and Multimedia Technology*, ser. ICEMT '22. New York, NY, USA: Association for Computing Machinery, Mar. 2022, pp. 65–69. [Online]. Available: https://doi.org/10.1145/3551708.3556205

[24] X. Chen, S. Chen, X. Wang, and Y. Huang, ""I was afraid, but now I enjoy being a streamer!": Understanding the Challenges and Prospects of Using Live Streaming for Online Education," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW3, pp. 237:1–237:32, 2021. [Online]. Available: https://doi.org/10.1145/3432936

[25] A. Adeliyi, M. Wermelinger, K. Kear, and J. Rosewell, "Investigating Remote Pair Programming In Part-Time Distance Education," in *Proceedings of the 2021 Conference on United Kingdom & Ireland Computing Education Research*, ser. UKICER '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–7. [Online]. Available: https://doi.org/10.1145/3481282.3481290

[26] Z. Chen, H. Cao, Y. Deng, X. Gao, J. Piao, F. Xu, Y. Zhang, and Y. Li, "Learning from Home: A Mixed-Methods Analysis of Live Streaming Based Remote Education Experience in Chinese Colleges during the COVID-19 Pandemic," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–16. [Online]. Available: https://doi.org/10.1145/3411764.3445428

[27] J. Neto, S. Nolan, and A. Mendes, "Planning and Developing Courses in Distance Learning Environments: a training course for HiEdTec Project," in *Proceedings of the 22nd International Conference on Computer Systems and Technologies*, ser. CompSysTech '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 201–206. [Online]. Available: https://doi.org/10.1145/3472410.3472440

[28] T. D. Palaoag, J. G. Catanes, R. Austria, and J. S. Ingosan, "Prepping the New Normal: The Readiness of Higher Education Institution in Cordillera on a Flexible Learning," in *2020 The 4th International Conference on Education and Multimedia Technology*, ser. ICEMT 2020. New York, NY, USA: Association for Computing Machinery, 2020, pp. 178–182. [Online]. Available: https://doi.org/10.1145/3416797.3416829

[29] A. T. V. Pham and N. T. T. Ho, "Strengths and Weaknesses of MOOC-based Blended Learning as Perceived by Lecturers and Administrators," in *Proceedings of the 4th International Conference*

on Modern Educational Technology, ser. ICMET '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 68–72. [Online]. Available: https://doi.org/10.1145/3543407.3543419

[30] T. Thi Kim Le and B. Thi Tran, "Student Satisfaction of the Blended Learning Implementation in Email Marketing Course: A Case Study at a Vocational College in Vietnam," in *Proceedings of the 5th International Conference on Education and Multimedia Technology*, ser. ICEMT '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 44–50. [Online]. Available: https://doi.org/10.1145/3481056.3481067

[31] A. A. Siegel, M. Zarb, B. Alshaigy, J. Blanchard, T. Crick, R. Glassey, J. R. Hott, C. Latulipe, C. Riedesel, M. Senapathi, Simon, and D. Williams, "Teaching through a Global Pandemic: Educational Landscapes Before, During and After COVID-19," in *Proceedings of the 2021 Working Group Reports on Innovation and Technology in Computer Science Education*, ser. ITiCSE-WGR '21. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–25. [Online]. Available: https://doi.org/10.1145/3502870.3506565

[32] R. Borasi, R. DeMartino, N. Harris, and D. Miller, "Could covid-19 be a catalyst for disruption in higher education?" in *Online Teaching and Learning in Higher Education during COVID-19*. Routledge, 2021, pp. 153–166.

Does Daily Use of Digital Technologies Influence the Reading and Information Literacy of 15-year-old Students?

V. Lang*, A. Špernjak* and A. Šorgo*, **
* Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor, Slovenia
** Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia
andreja.spernjak@um.si

***Abstract* - This paper focuses on the impact of out-of-school and in-school use of digital technologies on the reading and information literacy skills of 15-year-old Slovenian students who participated in the PISA 2018 study. For our study, we selected PISA 2018 variables related to the use of digital technologies in school or outside school. The sample (N = 2612) consists of 50.8% girls and 49.2% boys. Part of the questionnaire refers to the use of digital technologies outside school: for entertainment and practical information purposes, for school purposes, and time spent with digital technologies. Part of the questionnaire refers to the use of digital technologies in school: for various purposes, for learning or teaching, and the time spent using digital technologies in class. The results show that the constructs of using digital technologies in school or out of school are strongly related but have a very low/insignificant impact on reading literacy and information literacy.**

Keywords* - *digital technologies, secondary education, information literacy, reading information literacy, PISA 2018.

I. INTRODUCTION

The development of information and communication technology (ICT) has brought many changes and challenges to our lives and has a great impact on education. Many believe that information literacy is an important skill in the modern world. Those who have information literacy, must be able to obtain, evaluate, and use information from a variety of sources. The OECD [1] stated that digital technologies and the Internet enhance children's learning experiences and promote different learning methods. Digital technologies are also present in individuals' daily lives and leisure time. However, new services and knowledge requirements are emerging every day that demand digital and information literacy from users [1].

Most of the revised elementary school curricula promote students' digital literacy by researching databases and information bases, searching, selecting, processing, analysing and critically evaluating the acquired information. It is important to view digital technologies as a tool that students can use to help themselves fulfil their academic and many others responsibilities. Modern educational reforms expect teachers to support students in formal and non-formal learning with digital technologies, in school and at home [2]. Learning science and biology using digital technology as a learning tool can be as motivating as teaching with active teaching methods [3]. The teacher's task is to stimulate students' curiosity for independent and lifelong learning using modern digital technologies [4].

Even though digital technologies are mostly seen as something positive [5], students' recreational use of digital technologies also poses risks and is a cause of concern for parents and policy makers [6]. Inappropriate or unsafe internet use can expose students to harmful content or cyber-bullying. Students are also exposed to an enormous amount of information online, which can help them develop online reading skills, but can also have negative effects if students are unable to distinguish facts from fake news and check their sources. Additional risks, such as excessive use of video games and compulsive use of social media, can have serious physical, social, psychological, and cognitive consequences [1,7].

We already know that time spent in front of screens outside of school has a small impact on better information literacy. We also know that the information literacy of Slovenian students is far from excellent [8, 9, 10]. Much less is known about how the time spent in front of screens in a school for learning and teaching affects information literacy. Even less is known about how much time students spend using digital technologies in class in different subjects and how much time is spent outside of class on the requirements of school work in schools. Therefore, we are interested in whether time spent in front of screens for schoolwork has a greater impact on information literacy of 15-years old adolescents than time spent in front of screens for entertainment.

The aim of the research was to answer the six research questions, which can be considered as hypotheses H1 - H6:

- Does the frequency of using digital technologies outside of school for entertainment and practical information purposes have an impact on reading and information literacy? (H1)
- Does the frequency of use of digital technologies outside school for school purposes has an impact on reading and information literacy? (H2)
- Does the amount of time digital technologies are used outside of class (at home or at school) have an impact on literacy and information literacy? (H3)

- Does the frequency of using digital technologies in school for different purposes have an impact on reading and information literacy? (H4)
- Does the timing of digital technologies use in the classroom has an impact on reading and information literacy? (H5)
- Does the use of digital technologies for learning or teaching in the classroom have an impact on literacy and information literacy? (H6)

At this point in the research, we did not look for differences based on personal characteristics and traits and socio-demographic factors. The reason for this was not that such information was unimportant, but that we were looking for robust models in the first phase.

II. MATERIALS AND METHODS

To obtain answers to these research questions, the authors analyse data of interest collected in the PISA 2018 survey to find patterns that were not included in the final PISA 2018 report.

The research is based on publicly available databases provided by the OECD, PISA [11]. The data were filtered to the Slovenian sample and the items of interest were identified from the manuals and code tables. The data were transferred to Excel files and cleaned of missing and incomplete data that were not suitable for the planned analyses. Statistical analyses were performed using IBM SPSS 27 and AMOS 27 software.

A. Sample and sampling

The PISA 2018 survey was conducted in Slovenia in March and April 2018 and followed international protocols. For the current study, Slovenian data from PISA 2018 were transferred to the local drive. The target group of the survey was 15- and 16-year-old students participating in various educational programs. The interested reader can find information about the Slovenian school system on the websites of (Eurydice, Ministry). 6401 pupils/students from 302 secondary education programs, 43 elementary schools, and two adult education institutions participated in the survey. More than 90% were enrolled in the 1st year of lower-secondary school programs at the time of the survey. The sample was representative of a population of participants in all Slovenian formal education programs. The total number of births in this calendar year is 17501 (Statistical Office). Only records without missing data were used for further analysis. This sample (N = 2612) consists of 50.8% girls and 49.2% boys.

B. Instrument

Question ST 166 (Student Questionnaire) from the ST knowledge test and six questions from the IC questionnaire - ICT Familiarity Questionnaire (IC008, IC010, IC011, IC150, IC151, IC152) from the Slovenian Pisa 2018 study [12] were selected for the study.

ST166 [12] was considered a reading and information literacy task as an outcome (latent) variable. The instructions were as follows:

"*You have received a message in your inbox from a well-known mobile phone operator telling you that you are one of the winners of a smartphone. The sender asks you to click on the link to fill out a form with your data so they can send you the smartphone. In your opinion, how appropriate are the following strategies in reaction to this email?"*

Students could indicate how much they agreed with the opinion on the 6-point Likert scale in each of the 5 rows (Table 1), where 1 means – not appropriate and 6 – very appropriate.

We selected tasks IC008, IC010, IC011, IC150, IC151, and IC152 [12] as predictors of reading and information literacy. Part of the questionnaire relates to collecting data on the frequency and time spent using electronic devices and digital technologies outside school for entertainment and practical information purposes (IC008), using digital technologies outside school for school purposes (IC010), and using digital technologies in school for various purposes (IC011). We also selected the question where students indicate how much time they spend using digital devices and digital technologies in class during a normal school week (IC150), for 9 school subjects. This is followed by questions in which participants indicate how much time they spend using digital technologies outside of class (at home or at school) for 9 school subjects (IC151) and how much time they spend using digital technologies for learning or teaching in class for 9 school subjects (IC152).

Items on frequency and time spent using electronic devices served as predictors of reading literacy. We selected questions IC008, IC010, and IC011, which asked respondents to indicate how often they used electronic devices outside of school (IC008 and IC010) or in school (IC011) for various purposes. In each row, students choose a response on a 5-point scale indicating how much they agree with the opinion, where 1 means – never or almost never and 6– every day.

Question IC150 asks students how much time they spend using digital devices during class in a typical school week. Question IC151 asks students how much time they spend using digital devices outside of class in each of the nine school subjects listed. Students were asked to choose one answer on the 5-point scale in each row, where 1 means – no time, 2 means – 1-30 minutes per week, 3 means – 31-60 minutes per week, 4 means – more than 60 minutes per week and 5 means – not in this subject. The school subjects listed in the questionnaire were: Slovenian language, Mathematics, Science, Foreign language, Social sciences, Music, Physical education (Sport), Performing arts, and Visual arts.

Question IC152 asks students whether they have used electronic devices for learning or used in teaching during lessons in different subjects in the last month. The

questionnaire covers the same school subjects as question IC151 previous item Students choose one answer on the 5-point Likert scale in each row, where 1 means – Yes, both the teacher and the students used it, 2 means - Yes, but only the students used it, 3 means - Yes, but only the teacher used it, 4 means – No and 5 means - I do not in this subject.

C. Statistical analyses

The statistical procedure was carried out in several steps. In the first step, data from two databases (IC - ICT Familiarity Questionnaire and ST - Student Questionnaire) were merged into a single SPSS file, based on the students' code ID. This procedure resulted in a sample of 6410 students. The data were reviewed, and cases with missing data were deleted from the file. This left 2612 cases that could be included in the modelling.

Each selected variable was analysed for frequency of responses (F%), mean, median, mode, and standard deviation (SD).

Each construct (latent variable) was examined. Cronbach's alpha was calculated to assess the reliability of the construct. To assess the unidimensionality of the constructs, Principal Component Analysis (PCA) was conducted with Direct Oblimin rotation. Previously, KMO and Bartlett's tests for sphericity were performed to assess the fit of the matrices.

It was found that the construct ST has two components, that explain 68.9% of the variance. Since knowledge can be considered summative, the sums of all items forming the ST construct were later used in the models [13].

III. RESULTS

A. Task to check reading and information literacy

Examination of the unidimensionality of the construct revealed that two components were extracted, explaining 68.9% of the variance. All items loaded above the 0.4 level, so they were included in the analyses. We note that the task of assessing reading and information literacy can be divided into two components: Curiosity (PC1) and Caution (PC2). We can explain 40.7% of the variance with the first component (Cronbach's alpha = 0.67). The first component consists mainly of curiosity attitudes: *"Reply to the email and ask for more information about the smartphone."* and *"Click on the link to fill out the form as soon as possible."* The second component (Cronbach's alpha = 0.53) explained 28.2% of the variance and included three items of caution: *"Check the sender's email address," "Check the mobile operator's website to see if the smartphone offer is mentioned"* and *"Delete the email without clicking the link."* According to the results respondents think that the least appropriate response to a message in their inbox from a well-known mobile phone provider telling them that they are one of the winners of a smartphone is to click on the link and fill out the form as soon as possible. However, most agree that it is appropriate to check the sending email address.

B. Frequency of use of digital technologies outside school for entertainment and practical information purposes (IC008) and for school purposes (IC010)

Response frequencies and descriptive statistics for questions IC008 and IC010 (*How often do you use digital devices for the following activities outside of school?* N = 2612). Respondents outside school never or once play online games via social networks (e.g., Farmville®, The Sims Social) and never or once using social networks for communication with teachers (e.g., Facebook, MySpace). These results were relevant before the SARS-CoV-2 pandemic. It would be good to know the current situation, because majority Slovenian teachers and students in period SARS-CoV-2 use social networks even daily for communication. Students still almost daily they use social networks to communicate with other students about schoolwork (e.g., Facebook, MySpace), to chat online (e.g., MSN®), to participate in social networks (e.g., Facebook, MySpace), and to surf the Internet for fun (e.g., to watch videos, e.g., YouTube™). A review of the use of digital technologies outside of school for cutting-edge entertainment and practical purposes helps us to show that technologies are rapidly being integrated into everyday life and the available evidence [14] points to an increase in the use of digital technologies particularly during adolescence. 86.7% of young people in this survey agree that they really enjoy using digital devices. The results of this research show that Slovenian 15 - year - old outside the school most often use digital technology for online chatting, for participating in social networks and for browsing the internet for fun, but not for playing online games via social networks. For school purposes outside of class, they most often use social networking sites to communicate with other students about schoolwork (e.g., Facebook). It has already been noted that the impact of using digital technologies based on social media outside school varies depending on the use and type of social networks. The results and models show that out-of-school digital technology use is related to other constructs. It is most strongly related to the construct IC010 (Using digital technologies outside school for school purposes). However, it has a very small statistically insignificant impact (0.06) on reading and information literacy.

When asked how much time they spend on digital technology outside of school per week for different subjects, they answered that they spend most of their time, i.e., up to 60 minutes per week, on music, art and foreign language. In other subjects, they do not use digital technology outside class or spend at most 30 minutes per week. They spend the least time in front of screens for sport and mathematics. The results and models show that time to using digital technologies outside of class is related to another construct. It is most strongly (0.40) related to the construct IC150 (Time of using digital technologies in the class). However, it has a very small statistically insignificant impact (0.05) on reading and information literacy. In particular, recreational use of email, reading online news for pleasure, and surfing the Internet for pleasure (e.g., watching YouTube) were positively correlated with students' digital reading achievement from PISA 2009 to PISA 2018, which partially confirmed [15] findings that online social activities play a role in explaining adolescents' digital

reading achievement, but contradicted Lee and Wu's [16] findings that social entertainment activities negatively influence adolescents' reading achievement.

The various effects of out-of-school digital technology use have shown either positive effects [17] or negative effects [18, 19, 20] on students' reading achievement.

C. Frequency of use of digital technologies in school for different purposes

The reliability of the questionnaire about use of digital technologies in school for different purposes is very good, the Cronbach's alpha 0.93 determined. Students in school never or almost never use them to post work on school websites, to do homework on school computers, they almost never use computer simulations in school. However, they are more often used to surf the Internet for schoolwork and to chat online at school.

With the two remaining components, 76.3% of the variance can be explained. All items had a loading value above 0.4 and were therefore included in the analyses.

With the first component, we can explain 53.2% of the variance (Cronbach's alpha 0.95). The items included in this component were student's minority use ICT for school work, such as use of ICT to post work on school websites, to do homework on school computers, or use computer simulations in school: *"Downloading, uploading or browsing material from the school's website (e.g.)."*, *"Using school computers for group work and communication with other students."*, *"Practicing and drilling, such as for foreign language learning or mathematics."*, *"Using learning apps or learning websites."*, *" Posting my work on the school's website."*, *"Playing simulations at school."*, *"Doing homework on a school computer."*

Three items of second component explains 23.1% of the variance (Cronbach's alpha 0.76), where students ICT mostly use for communication via email or online chatting and *"Browsing the Internet for schoolwork."*, *"Using email at school."* and *"Chatting online at school."*, fall into the second component, which explains 23.1% of the variance (Cronbach's alpha 0.76).

D. Time of using digital technologies in the class

The reliability of the questionnaire is very good, the Cronbach's alpha 0.86 determined. The results show that they do not use electronic devices in physical education (sports) and mathematics classes at school. More frequently i.e., 1-30 minutes in a typical school week in school, they are used in visual arts, music and performing arts. With the two remaining components, 65.4% of the variance can be explained. All items had a loading value above 0.4 and were therefore included in the analysis. With the first component, we can explain 40.7% of the variance (Cronbach's alpha = 0.88). The first component includes school subjects where electronic devices are not used. There are Sport, Social sciences, Foreign language, Mathematics, Science, and Slovene. The second component (Cronbach's alpha= 0.79) explains 25.0% of the variance and includes three school subjects' Visual arts, Music, and Performing arts. In these school subjects, electronic devices are used up to 30 minutes per week.

E. Use of digital technologies for learning or teaching in the class

The reliability of the questionnaire is very good, which determines Cronbach's alpha 0.93. The results show that in the last month only teachers in the following subjects used electronic devices for teaching: Foreign language, Slovene, Science, Social sciences, Mathematics, Visual arts, and Sport. No electronic devices were used in the subjects of music and Performing art in the last month.

With the two remaining components, 75.7% of the variance can be explained. All items loaded above the 0,4 level therefore they were included for analyses.

We can explain 42.4% of the variance with the first component (Cronbach's alpha = 0.92). The first component includes the school subjects Mathematics, Slovene, Social sciences, Foreign language, and Science. The second component (Cronbach's alpha = 0.87) explains 33.3% of the variance and includes the four school subjects Visual arts, Sport, Music, and Performing arts.

The results show that young people are also use social networks and digital technologies very frequently at school, especially for online chats, using emails and surfing the internet. However, digital technology is almost never used at school for completing homework on the school computer. The results and models show that time spent using digital technology in class is related to other constructs. It is most strongly related (0.40) to the construct IC151 (Time of using digital technologies outside of school). However, it has a very small statistically insignificant impact (0.05) on reading and information literacy. The use of digital technology for learning or teaching in different subjects also has a very small and statistically non-significant impact on digital and information literacy (0.13). Other authors [17, 18, 19, 20] found that the use of digital technology in school for extracurricular purposes has a negative impact on literacy and information literacy. These findings were also confirmed by a secondary longitudinal study with the results from PISA [21]. The statistically insignificant effects in our study can be explained by the fact that the use of digital technologies for extracurricular purposes robs students of a lot of time that could be spent focusing on school work. It is also known that the use of digital technologies in school quickly diverts students' attention from school work [22]. It is believed that the use of digital technologies in school under the supervision and guidance of a teacher would reduce disruption and overuse of social networks in the classroom [23] and potentially have a greater impact on information literacy.

F. Hypothetical and final model

Although there are correlations within the constructs, in Figure 1 we only show the effects of the constructs on information literacy to test the hypotheses. With Hypothesis 1, we tested whether there is a statistically

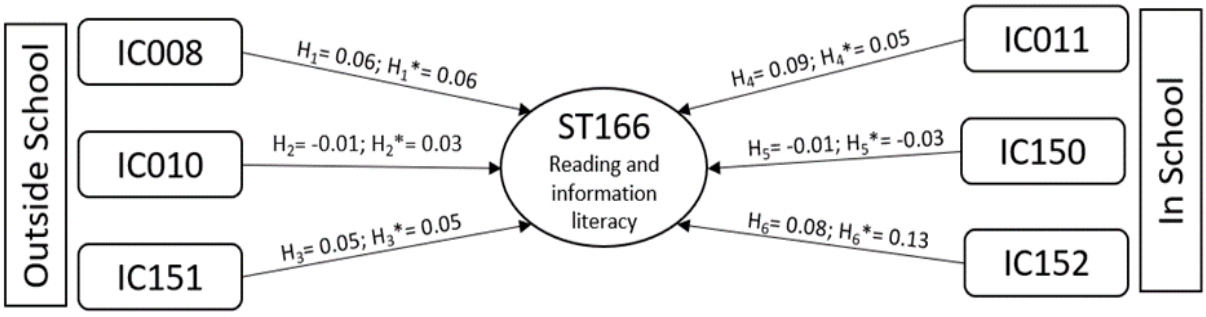


Figure 1. Hypothesis testing model (*Result for Final model).

significant relationship between the frequency of using digital technologies outside school for entertainment and practical information purposes to reading and information literacy. We checked the correlation between the variable and found that there was no statistically significant correlation.

There was also no statistically significant difference in the testing of Hypothesis 2, in which we examined the correlation between the frequency of using digital technologies outside of school for school related purposes to reading and information literacy. In Hypothesis 3, we were interested in the correlation between time to using digital technologies outside of class to reading and information literacy. We found that there was no statistically significant correlation. With Hypothesis 4, we wanted to test the correlation between frequency of using digital technologies in school for different purposes to reading and information literacy. With Hypothesis 5, we tested whether there is a statistically significant relationship between the time of using digital technologies in the class to reading and information literacy. Also, there was no statistically significant correlation. There was also no statistically significant difference in the testing of Hypothesis 6, where we chose to look at the relationship between using of digital technologies for learning or teaching in the class. We were also unable to find a statistically significant relationship among these hypotheses. We conclude that no construct of using digital technologies for school work or leisure purposes has a statistically significant impact on reading and information literacy.

IV. CONLUSION

Using a step-by-step analysis of data from 2612 students from 347 Slovenian educational programs, the current cycle of the 2018 PISA Digital Reading Study reveals the impact and timing of digital technology use on 15-year-olds' reading achievement. The factors that impact the use of digital technologies on digital literacy and information literacy have been poorly explored in previous literature. The results of this study are based on 6 different aspects of the use of digital technologies for in-school or out-of-school purposes, and results reveal whether the time spent in front of digital technology screens for school purposes influence students' information literacy. Based on the results of the secondary analysis of the data from questionnaire of PISA 2018, we can answer the research questions about the influence on reading and information literacy. To test reading and information literacy, we selected a task from the PISA 2018 questionnaire that tests students' confidence in receiving unsolicited text messages designed to motivate a person to enter sweepstakes and provide personal information in order to receive a prize - a smartphone.

Most 15-year-olds agree that it is inappropriate to click on a link to such a message and to reply to messages containing personal information. According to Cheshire, Antin and Churchill [24], in interaction with internet systems, individuals develop implicit or explicit attitudes about the risks and uncertainties in online environments, a finding which relates to our results; most students agree that it is appropriate to check the sender's email address. As expected, the students' responses split into two components. One group is curious, and the other is cautious. Both groups of 15-year-old students mostly agree (86.7%) or strongly agree that they enjoy using digital devices.

REFERENCES

[1] OECD. (2015). *Students, computers and learning: Making the connection, Programme for international student assessment.* Paris: OECD Publishing. Available at http://www.oecd.org/publi

[2] L. Johnson, S. A. Becker, V. Estrada, A. Freeman, P. Kampylis, R. Vuorikari, & Y. Punie. (2014). *Nmc horizon report europe: 2014 schools edition*. The New Media Consortium.

[3] M. Bulić, & I. Blažević. (2020). The impact of online learning on student motivation in science and biology classes. *Journal of Elementary Education, 13*(1), 73-87. https://doi.org/10.18690/rei.13.1.73-87.2020

[4] C. Dede, J. A. Bellanca, & R. S. Brandt. (2010). 21st century skills: Rethinking how students learn. J. Bellanca & R. Brandt.

[5] C. Buzzard, V. L. Crittenden, W. F. Crittenden, & P. McCarty. (2011). The use of digital technologies in the classroom: A teaching and learning perspective. *Journal of Marketing Education, 33*(2), 131-139.

[6] J. H. Graafland, (2018). New technologies and 21st century children: Recent trends and outcomes. https://www.oecd-ilibrary.org/docserver/e071a505-en.pdf?expires=1675504659&id=id&accname=guest&checksum=66021BAA93DFB5E6483EAD1B9E038605

[7] J. Meates. (2020). Problematic Digital Technology Use of Children and Adolescents: Psychological Impact. *Teachers and Curriculum, 20*(1), 51-62.

[8] K. Dolenc, & A. Šorgo. (2020). Information literacy capabilities of lower secondary school students in Slovenia. *The Journal of Educational Research, 113*(5), 335-342. https://doi.org/10.1080/00220671.2020.1825209

[9] D. Dolničar, B. Boh Podgornik, T. Bartol, & A. Šorgo. (2020). Added value of secondary school education toward development of information literacy of adolescents. *Library & Information Science Research*, *42*(2), 101016. https://doi.org/10.1016/j.lisr.2020.101016

[10] A. Šorgo, T. Bartol, D. Dolničar, & B. Boh Podgornik. (2017). Attributes of digital natives as predictors of information literacy in higher education. *British Journal of Educational Technology*, *48*, pp. 749-767. https://doi.org/10.1111/bjet.12451

[11] OECD. PISA 2018 Database. https://www.oecd.org/pisa/data/2018database

[12] OECD. (2019). PISA 2018 results (volume I): What students know and can do. PISA, OECD Publishing. https://doi.org/10.1787/5f07c754-en

[13] R. B., Kline. (2016). Principles and Practice of Structural Equation Modeling (4th ed.). New York, NY: The Guilford Press.

[14] G. Merchant. (2012). Mobile practices in everyday life: Popular digital technologies and schooling revisited. *British Journal of Educational Technology*, *43*(5), 770-782.

[15] J. Gil-Flores, J.-J., Torres-Gordillo, & V.-H., Perera-Rodríguez. (2012). The role of online reader experience in explaining students' performance in digital reading. *Computers & Education*, *59*(2), 653–660. https://doi.org/10.1016/j.compedu.2012.03.014.

[16] Y.-H. Lee, & J.-Y. Wu. (2013). The indirect effects of online social entertainment and information seeking activities on reading literacy. *Computers & Education*, *67*, 168–177. https://doi.org/10.1016/j.compedu.2013.03.001

[17] M. Skryabin, J. Zhang, L. Liu, & D. Zhang. (2015). How the ICT development level and usage influence student achievement in reading, mathematics, and science. *Computers & Education*, *85*, 49–58. https://doi.org/10.1016/j.compedu.2015.02.004

[18] D. Petko, A. Cantieni, & D. Prasse. (2017). Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 scores. *Journal of Educational Computing Research*, *54*(8), 1070–1091. https://doi.org/10.1177/0735633116649373

[19] Y. Xiao, Y. Liu, & J. Hu, J. (2019). Regression analysis of ICT impact factors on early adolescents' reading proficiency in five high-performing countries. *Frontiers in Psychology*, *10*, 1646. https://doi.org/10.3389/fpsyg.2019.01646

[20] X. Hu, Y. Gong, C. Lai, & F. K. S. Leung, (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Computers & Education*, *125*, 1–13. https://doi.org/10.1016/j.compedu.2018.05.021

[21] J. Hu, & R. Yu. (2021). The effects of ICT-based social media on adolescents' digital reading performance: A longitudinal study of PISA 2009, PISA 2012, PISA 2015 and PISA 2018. *Computers & Education*, *175*, 104342.

[22] L. E. Levine, B. M. Waite, & L. L. Bowman, (2007). Electronic media use, reading, and academic distractibility in college youth. *CyberPsychology and Behavior*, *10*(4), 560–566. https://doi.org/10.1089/cpb.2007.9990

[23] A. Andersson, M. Hatakka, A. Gronlund, & M. Wiklund. (2014). Reclaiming the students - coping with social media in 1:1 schools. *Learning, Media and Technology*, *39*(1), 37–52. https://doi.org/10.1080/17439884.2012.756518

[24] C. Cheshire, J., Antin, & E. Churchill. (2010). Behaviors, adverse events, and dispositions: An empirical study of online discretion and information control. *Journal of the American Society for Information Science and Technology*, *61*(7), 1487-1501. https://doi.org/10.1002/asi.21346

Self-Assessment of Pre-service English Teachers' Digital Competence

T. K. Harakchiyska
University of Ruse / Department of Bulgarian Language, Literature, History and Arts, Ruse, Bulgaria
tharakchiyska@uni-ruse.bg

***Abstract* – Digital competence is considered an integral component of the skills and professional expertise of teachers nowadays. The aim of this study was to explore the perceptions of 82 pre-service English language bachelor and master level teachers from a state university in Bulgaria on their digital competence based on the interplay of five variables – competence for using and integrating ICT in the foreign language classroom, technological pedagogical knowledge, support received during training, capacity to promote the digital skills of their learners and attitude towards integrating digital technologies in their future professional practice. Based on the performed descriptive, correlational and inferential analysis of the data obtained, it was established that the most strongly related variables were the self-perceived competence of the pre-service teachers for using and integrating digital technologies in the teaching and learning of foreign languages and their capacity to promote the digital skills of their learners.**

Keywords – ***digital competence; pre-service teacher training; assessment.***

I. INTRODUCTION

Digital technologies constitute an essential part of the 21st century classrooms and require specific knowledge, competences and skills from teachers (and subsequently from learners) in order to function successfully in an interactive, dynamic and ICT-based educational context. Teachers nowadays are expected to be not only technically competent in managing the rich spectrum of devices, tools and applications to communicate with students and colleagues, to collaborate professionally, to create and share content, but also possess an ability to critically approach technology from a pedagogical perspective.

Given this, teachers' digital literacy has become a topic of interest in a vast number of theoretical frameworks which approach the concept from different perspectives. While some authors stress the fact that it entails elements from many domains (e.g. literacy, cultural studies, communication, technology studies) [1], others directly state that it is a "pluralist concept" [2] due to its multidimensional nature [3]. Despite the variety of interpretations, most researchers agree that educators' digital competence is an umbrella term which entails the knowledge, skills and abilities of teachers to implement effectively digital technologies in their teaching practices and in supporting their students' active learning. Training future teachers to deploy productively and creatively digital tools into education is thus a key priority on the European agenda and a target in the educational policies of many countries.

Bulgaria is not an exception in this case. In 2019 the Government introduced the "Digital Bulgaria 2025 National Programme" [4] which promotes the strategic development of digital competence of all citizens aiming to increase their competitiveness and strengthen their employability. Central to the envisaged reforms is the enrichment of teachers' pedagogical repertoire with new formats and methods for digital engagement of learners. In line with this, the Ministry of Education and Science designed the "Digital Qualification" National Programme [5] in 2021 and has already started upskilling school teachers and university academic staff for using ICT, digital content and online teaching tools in their classes. These steps, however, are a follow-up of the profound changes made in the state regulated requirements concerning the training of pre-service teachers. According to these regulations introduced in 2018, all pre-service teacher training programmes offered by higher educational institutions on a national level should include the following two courses "Information and Communication Technologies and Work in Digital Environment" (as a compulsory component) and "Digital Competence and Digital Creativity" (as an elective component) each comprising of at least 30 academic hours. Apart from these two officially required courses, universities have the freedom to enrich their educational offer with study courses designed to upgrade pre-service teachers' digital competence. In fact, higher educational institutions are supposed to do so since the digital capacity of teacher trainees to integrate ICT-based technologies in an innovative and sustainable way in their teaching and in the learning experiences of their pupils are among the core outcomes of the pedagogical training offered at university level.

Since these courses are part of the training of pedagogical specialists in the last 3 years, it is expected that each higher educational institution in the country has not only offered a blend of courses fostering the digital competence of student teachers, but has also assessed their relevance, usefulness and efficiency. Contrary to the expectations, data on the quality of this training on the digital competence of teacher trainees in Bulgaria is limited. This study, therefore, is an attempt to address the existing gap in research literature by providing insights into the digital competence of pre-service teachers in the University of Ruse based on their self-assessment. It

examines the perceptions of pre-service teachers about their skills to employ digital technologies in the teaching of English as a foreign language (L2) and in empowering learners to use digital tools and resources for self- or collaborative L2 learning and assessment. The objective of the provided analysis is to determine also the attitude of future English language teachers towards the integration of ICT-based educational technologies in their L2 classrooms.

II. LITERATURE REVIEW

The advancements in digital infrastructure in all spheres of our daily lives along with the availability and vast use of digital devices have shaped the ways in which the generations born and raised in a digitally saturated world, i.e. the so called "digital natives" [6], communicate, learn and socialize. Feeling comfortable with technologies from an early age, young people nowadays are good at multitasking, have advanced technical skills, rely heavily on technology when they need or want to share information and learn best when actively engaged. These characteristics of the digital natives pose new requirements on the training of pre-service teachers, who although being digital natives themselves, need formal instruction for integrating digital technologies in a pedagogically sound manner in their classrooms. That applies in particular to the future L2 teachers since digitization has reshaped the teaching and learning of foreign languages by enriching them with technology-based applications and tools that foster meaningful communication and language skills refinement, on the one hand, and which trigger learners' autonomy, collaboration and creativity, on the other hand.

The integration of digital technologies in the training of teachers at university level requires from higher educational institutions to embrace new paradigms for knowledge co-creation through digital inclusion. To achieve this, universities are guided by the methodological and assessment frameworks which define the knowledge, skills and competences which future educators need to possess in order to create, manage and support digital learning experiences.

One of these frameworks is the European Framework for the Digital Competence of Educators (DigCompEdu) [7]. It specifies 22 digital competences distributed among six areas of expertise which cover teachers' professional and pedagogical abilities to utilize digital applications and tools in education, on the one hand, and their skills to facilitate learners' digital literacy, on the other hand. Another framework, which also organizes the competences of teachers in six main thematic areas, is the DC4LT Assessment Framework [8]. Although it is targeted at foreign language teachers, its categories correspond to a significant extent to the areas of teachers' digital expertise in DigCompEdu (Table I).

The thematic areas of these two frameworks and the digital competences identified in them suggest that teachers' skills to engage with digital technologies in a competent, critical and creative way are an essential prerequisite for incorporating ICT-based technologies in the classroom, but only technological knowledge and skills are not enough for efficient teaching and learning.

TABLE I. THEMATIC AREAS OF DIGCOMPEDU AND DC4LT

Digital competence assessment frameworks	
DigCompEdu	***DC4LT***
Pedagogical engagement	Professional development
Digital Resources	Technology
Teaching and learning	Pedagogy
	Content
Assessment	Assessment
Empowering learners	Learner's support
Facilitating learners' digital competence	

Recent research studies [9, 10] demonstrate that pre-service teachers need to be equipped with knowledge and competences to approach technology as a pedagogical tool. In this respect, the TPACK model proposed by Mishra and Koehler [11] puts forward the idea that technology, pedagogy and content are in constant interplay so digital teaching involves a more complex type of knowledge – technological pedagogical content knowledge. This type of knowledge allows teachers to gain understanding of how specific subject-related content can be represented and taught through technology, how a diverse scope of pedagogical approaches, strategies and techniques can be used to engage students with the subject matter and with each other in digitally mediated forms of instruction and how students' previous knowledge and skills (incl. digital literacy) can be used to boost learners' performance and development.

It has been established that prospective teachers' attitudes have a profound influence on their willingness to use digital tools and applications [12] and are a predictor for both the integration of TPACK and technology in the education process [13]. These findings correlate with the theoretical underpinnings of the Technology Acceptance Model (TAM) [14] and its extensions which postulate that the adoption of new technology depends on the attitude of the user to that technology, which in turn, depends on two factors: perceived usefulness of the technology and perceived ease of use. Technology acceptance, however, is not only linked to ICT attitudes. Rather, it is a follow-up of rigorous training which concentrates on the pedagogical aspects of embedding digital technologies in the teaching and learning process. Such an assumption is also shared by [15] who designed the Synthesis of Qualitative Data (SDQ) model for training pre-service teachers to implement technology. It comprises three levels which have a direct reference to the development of pre-service teachers' skills for technology integration and offers six strategies on the micro level: Role models, Collaboration, Instructional design, Authentic experiences, Reflection, and Feedback. When operationalized, these six strategies enhance the intersection of content, pedagogy and technology in the training of pre-service educators (Table II).

In addition to the SQD model, teacher candidates' self-efficacy with digital tools and applications is considered an indicator for their aptitude to create and manage a technology-driven educational process. Empirical data confirm that trainee teachers who conduct classes in

which they frequently integrate digital technologies, have higher self-efficacy [16]. Similarly, [17] and [18] acknowledge that pre-service training programmes that provide digital training courses boost candidate teachers' abilities and readiness to include ICT-based tools and resources in their classrooms, thus increasing their technological self-efficacy. Consequently, the assessment of future teachers' technology self-efficacy can provide valuable data on the effectiveness of such a programme.

TABLE II. MICRO LEVEL STRATEGIES OF THE SDQ MODEL

Micro level strategy	Manifestation in the pre-service teacher training
Role model	Teacher trainees observe their tutors or school teachers use technology in a meaningful way in the classroom
Reflection	Pre-service teachers reflect upon and discuss with their mentors about the use of technology in the teaching and learning process
Instructional design	Pre-service teachers get help in designing lessons that include digital technologies
Collaboration	Teacher candidates share experience, work together and support each other in their technology learning endeavours
Authentic experiences	Teacher trainees check their digital competences in real classrooms (e.g. during teaching practice, internships, etc.)

Adapted from [15]

Almost all related literature works suggest that new studies on pre-service teachers' digital competence need to be conducted in order to extend the bulk of research with new perspectives and relevant findings on the effect of factors such as attitudes towards technology on the digital capacities of beginning teachers.

III. METHODOLOGY

A. Study sample

The participants in the study were 82 pre-service teachers from the University of Ruse (Bulgaria) out of which 62 (75.6 %) enrolled in the bachelor level programme preparing teachers of English to young learners and 20 (24.39 %) enrolled in the two master level programmes, respectively for training future teachers of English to young learners and for preparing teachers of English to lower secondary school learners. The mean age was 24.85 (SD = 5.114) and the gender ratio was typical for the educational strand of programmes at the university – 92.68 % (N = 76) female and 7.32 % (N = 6) male students. The participants were in their second (66.13 %; N = 41), third (32.26 %; N = 20) and fourth year (33.87 %; N = 21) of training.

B. The instrument

A self-evaluation questionnaire specifically designed for the purposes of the study was used. It comprised six sections and was written in Bulgarian. The first section collected demographic data (age, gender, year and level of study). The other sections assessed the responses of the study sample on each of the items in the variables along a 5 point Likert scale ranging from 1 (Strongly disagree) to 5 (Strongly agree).

The Self-perceived pre-service teachers' competence for integrating ICT in L2 instruction (SP CUT) was measured in the second section with 8 items focused on assessing the ability of the study participants to use technology in L2 teaching and learning and 6 items assessing their skills to organize and manage digital assessment. Each of these items is developed on the basis of descriptors from two of the assessment frameworks – DigCompEdu and DC4LT. From DigCompEdu we have utilized and adapted competences from Area 2: Digital resources and Area 4: Assessment, while from the second framework – descriptors from the thematic categories "Technology" and "Assessment". The Cronbach's alpha value of the 14 items was 0.843.

The Self-perceived technological pedagogical knowledge (SP TPK) of the study sample was measured with 5 items resulting from the adaptation of descriptors from the thematic categories "Pedagogy" and "Content" of DC4LT and the domains of the TPACK model [17]. The Cronbach's alpha of these 5 items was 0.702.

The Self-perceived support received during training (SP SRD) of the pre-service study participants was measured by 6 items reflecting each of the strategies in [15]. The Cronbach's alpha of these 6 items was 0.780.

The Self-perceived capacity to promote the digital skills of L2 learners (SP CDL) of the future English language teachers was assessed by 6 items designed following the descriptors of the DigEduComp in Area 6: Facilitating learners' digital competence and the DC4LT thematic category "Learner support". The Cronbach's alpha of these 6 items was 0.786.

The last section – Self-perceived attitude towards the use of digital technology in L2 instruction (SP ATU) contained 4 items providing data on the individual beliefs of the study subjects towards ICT and their intention towards incorporating it in English language teaching contexts. The Cronbach's alpha of these 4 items was 0.831.

C. Research questions

Two main questions were considered in the research study:

1. What is the interplay among the five variables constructing the digital competence of the pre-service bachelor and master level teachers?
2. Which of the variables constructing the digital competence of pre-service teachers (SP CUT, SP TPK, SP SRD, SP CDL) has the strongest influence on their attitude to using technology (SP ATU) in the future?

D. Data collection procedure and analysis

The questionnaire was administered in paper to the study participants in the winter semester of the 2022 / 2023 academic year during the classes of the students with the researcher. Participation in the data collection procedure was voluntary. The questionnaire was anonymous and the time for its completion was about 15 minutes.

The collected data was statistically processed using the SSPS software (version 29). Firstly, in order to determine the reliability and validity of the instrument, the Cronbach's alpha was calculated for each of the variables. Secondly, we have used descriptive, inferential and correlational statistics to analyse the demographic data and the variables. We have computed the Skewness and Kurtosis of each deliverable to analyse the assumption of normality of the collected data. Along with that a multiple regression analysis was made to examine the relations between the different variables. Lastly, we applied a parametric (one-way ANOVA test) to determine the mean values of the bachelor and master level students on each of the variables and a homogeneity test on the variances (an ANOVA Levene's test for equality of variances) to test if there is equal variance for each variable calculated among the two groups of respondents (bachelor and master level students).

IV. RESULTS

A. Statistical descriptive analysis

To get a better understanding of the data obtained we calculated the mean (M), standard deviation (SD), Skewness and Kurtosis plus the Cronbach's alpha (α) of each of the variables (Table III).

TABLE III. DESCRIPTIVE STATISTICS OF THE VARIABLES

Variable	M	SD	Skew-ness	Kurtosis	Cronbach's alpha (α)
SP CUT	3.85	0.52	-0.95	0.213	0.834
SP TPK	3.57	0.69	-0.74	-0.375	0.702
SP SRD	3.46	0.34	-0.33	-0.349	0.780
SP CDL	3.74	0.66	-0.80	-0.495	0.786
SP ATU	3.54	0.80	-0.79	-0.361	0.831

The descriptive analysis of the variables showed that the mean (M) ranges from 3.46 to 3.85 and the standard deviation (SD) was from 0.34 to 0.80, which is suggestive of a relatively narrow distribution of the responses of the study participants. The Kurtosis (ranging from -0.495 to 0.213) and Skewness (with a range from -0.95 to -0.33) also indicated a normal univariate distribution of the data since they were both within the acceptable range, i.e. the Kurtosis falls between the absolute values of 3 and 10 [18]. The reliability coefficient of each of the variables was also adequate as the Cronbach's alpha (α) was from 0.702 to 0.834 [18].

B. Correlational analysis

A correlational analysis was performed in order to give an answer to our first research question related to identifying the relations among the five variables based on the self-perceptions of the study sample of their digital competence.

The calculations of the Pearson correlation coefficient (r) for each pair of variables, showed that the variables with the strongest relationship were SP CUT and SP SRD (r = 0.896) followed by SP CUT and SP ATU (r = 0.867), while the variables least dependent on each other were SP SRD and SP CDL (r = 0.679) (Table IV).

TABLE IV. CORRELATION COEFFICIENTS AMONG THE VARIABLES

	SP CUT	SP TPK	SP SRD	SP CDL	SP ATU
SP CUT	1	0.826**	0.896**	0.861**	0.867**
SP TPK	0.826**	1	0.734**	0.778**	0.699**
SP SRD	0.896**	0.734**	1	0.679**	0.673**
SP CDL	0.861**	0.778**	0.679**	1	0.824**
SP ATU	0.867**	0.699**	0.673**	0.824**	1

** Correlation is significant at the 0.01 level (2-tailed)

Since we wanted to check whether there were differences between the bachelor and master level students in their self-assessment on the different variables comprising their digital competence, we performed a one-way ANOVA test (Table V).

TABLE V. ONE-WAY ANOVA TEST (MULTIPLE COMPARISONS AMONG THE BACHELOR AND MASTER LEVEL STUDY SUBJECTS)

Variables	Level	N	M	SD	Std. Error	F (between groups)
SP CUT	Bachelor	62	3.76	0.56	0.07	9.238
	Master	20	4.15	0.17	0.04	
SP TPK	Bachelor	62	3.41	0.70	0.09	16.156
	Master	20	4.07	0.39	0.09	
SP SQD	Bachelor	62	4.10	0.30	0.04	2.920
	Master	20	3.95	0.42	0.09	
SP CDL	Bachelor	62	3.59	0.68	0.09	15.478
	Master	20	4.20	0.29	0.06	
SP ATU	Bachelor	62	3.41	0.86	0.11	6.630
	Master	20	3.93	0.41	0.09	

* $p < 0.001$

The comparison of the means of the two groups of study respondents on each of the variables revealed that the master students reported higher mean values in all deliverables compared to the bachelor students with the exception of SP SRD, where the undergraduate study subjects had slightly higher values (M = 4.10, SD = 0.30) than the postgraduate study subjects (M = 3.95, SD = 3.95). The variables with the highest means for the bachelor level students were SP SRD (M= 4.10, SD = 0.30), SP CUT (M = 3.76, SD = 0.56) and SP CDL (M = 3.59, SD = 0.68), while the master students self-assessed themselves higher in terms of their SP CDL (M = 4.20, SD = 0.29), SP CUT (M = 4.15, SD = 0.17) and SP TPK (M = 4.07, SD = 0.39). Both groups self-assessed their attitude towards the use of digital technologies (SP ATU) lower compared to all other variables.

A subsequent Levene's test was used to test the homogeneity of variances between the two groups of study subjects. Since $p = 0.021$ ($p < 0.05$) we concluded that the data spread was equal.

C. Inferrential analysis

The performed inferential statistical analysis provided an answer to the second research question related to identifying which of the variables constructing the digital competence of the study subjects (SP CUT, SP TPK, SP SRD, SP CPL) had the strongest influence on their attitude to use technology (SP ATU) in their future careers as L2 teachers (Table VI).

TABLE VI. MULTIPLE REGRESSION ANALYSIS OF THE DEPENDENT VARIABLE SP ATU

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	*Std. Error*	*Beta (β)*		
1	(Constant)	2.188	0.581		3.751	< 0.001
	SP CUT	1.480	0.186	0.675	5.644	< 0.001
	SP TPK	0.152	0.112	0.131	1.350	< 0.001
	SP SRD	0.169	0.127	0.071	1.328	< 0.001
	SP CDL	0.410	0.131	0.337	3.141	< 0.001
Information about the Model		F (4.77) = 69.854, R^2 = 0.784, p < 0.001				

Dependent variable: SP ATU

The results from the multiple regression analysis revealed that the variable with the highest unstandardized beta value was SP CUT (B = 1.480, p < 0.001) followed by SP CDL (B = 0.410, p < 0.001) and SP SRD (B = 0.169, p < 0.001). This suggested that the strongest predictor for the possible future integration of ICT technologies and tools in the L2 classroom of the study participants was their Self-assessed competence for using and integrating digital technologies in the L2 teaching and learning process (SP CUT). The Self-perceived attitude of the study sample towards the use of digital technologies (SP ATU) was least dependent on their Self-perceived technological pedagogical knowledge (SP TPK).

V. DISCUSSION

The results of the study regarding the first question indicate that the variables which are most strongly associated are the self-perceived competence of pre-service L2 teachers for integrating ICT in L2 instruction (SP CUT) and the self-perceived support received during training (SP SRD). These results correspond to research data of previous studies on the positive interrelation between pre-service teachers' ICT capacity and the demonstrations on how to integrate digital tools in the L2 classroom, on the hands-on experience and support received from course tutors and peers during their university training [9, 10, 19].

The second pair of variables that are found to be strongly related are the self-perceived competence of the study subjects for integrating ICT in L2 instruction (SP CUT) and their self-perceived attitude towards the use of digital technologies in their future teaching career (SP ATU). These results comply with the data of recent research studies on the relation between these two variables [19, 20].

Contrary to our initial expectations, the results reveal that the variables with the lowest association are the self-perceived support received during training (SP SRD) and the self-perceived capacity to promote the digital skills of L2 learners (SP CDL). It is difficult to give a logical explanation of that because research data support the positive relation between candidate teachers' self-efficacy in using technology (gained during their studies at university) and their skills in developing the digital competence of their learners [10, 20]. One possible interpretation is the assumed low level of knowledge and skills which some of the L2 pre-service teachers might have with regard to incorporating digital learning activities and resources in the English language instruction. But it is also possible to approach the result from a different perspective. Given the fact that all study subjects are digital natives (their mean age is 23) who are trained to become teachers of another generation of digital natives, it could be supposed that the study respondents consider their young L2 learners as digitally literate individuals who do not need any additional training on how to use technology. Although this seems quite an unreasonable claim since digital natives are by no means sophisticated users of digital technology, it could be the reason behind a small percentage of the answers.

A detailed consideration of the interplay of the variables comprising the digital competence of the bachelor and master level study participants shows that the master students have higher means on the different variables than the bachelor students. This is indicative of the fact that they self-assess their mastery with digital technologies higher compared to the bachelors. The means of the master level students on the examined variables create the perception that they feel themselves better equipped in promoting the digital skills of L2 learners (SP CDL) and integrating ICT in L2 instruction (SP CUT) than the bachelor level students, who in turn value high the support received during training (SP SRD) and their competence in integrating ICT in L2 instruction (SP CUT). It is difficult to comment on these data in the light of available research as most of the recent studies examine the levels of digital competence mostly of undergraduate pre-service teachers [10, 19, 21]. The experience which the master students have gained with digital technologies during their undergraduate pedagogical training is most probably the main reason prompting them to consider themselves better skilled than the bachelors.

Interestingly, the self-perceived attitude of the study respondents towards the integration of digital technologies (SP ATU) is the variable that is ranked lowest by both the bachelor and the master level students. A plausible explanation for that could be the fact that the two groups of study participants perceived their abilities in engaging with digital technologies for professional purposes more important than their attitude towards implementing digital tools and resources in the educational process. This does not indicate that the study participants will not adopt new technology in their pedagogical encounters, but that they have valued it less than all other variables comprising their digital competence.

Regarding our second research question, the data prove that the pre-service teachers' self-perceived competence for integrating ICT in L2 instruction (SP CUT) is the strongest predictor for the positive attitude and intention of the study respondents towards implementing ICT-based technologies in their L2 classrooms. This complies with results of a number of studies which have established the strong relation between these two variables [19, 20]. It is important to note that although we have not examined the interrelation between the perceived usefulness of the technology and perceived ease of use [15], we cannot state whether they have any influence of the attitude of the study participants and on their intentions on using digital technologies in their

profession. This will be an issue under analysis in our future studies on the topic.

Although some prior studies [20, 22] have demonstrated that the adoption of TPACK training modules or courses in language teacher preparation programmes increase future L2 teachers' inclination to engage with digital technologies and facilitate their actual incorporation in class, our results contradict them. It is established that L2 pre-service teachers' attitude towards integrating digital technologies in language instruction is least dependent on their Self-perceived technological pedagogical knowledge (SP TPK). This implies that the attitude of pre-service teachers towards digital technologies depends more on their capabilities of using digital tools and applications, rather than on the implementation of the four components of the TPACK model in the pedagogical training of candidate teachers. In any case, these results do not imply that the pre-service teachers do not need consistent technological pedagogical training.

VI. Limitations of the Study

The use a self-assessment instrument creates conditions for subjectivity in the self-evaluation performed by the study subjects which could have a negative effect on the reliability and validity of the data. Therefore, future research could include focus group interviews with the study subjects, open-ended questions or an observation and assessment of their actual digital competences in the L2 teaching and learning process.

A second limitation of the reported study is a result from the methodological decision of the researchers to exclude gender as a variable determining the digital competence of the study subjects due to small percentage of male respondents (7.32 %). This limitation could hardly be addressed in our future research since male students constitute a small proportion of all students in the pedagogical programmes.

VII. Conclusion

The results of the study shed light on the interplay of the five variables on the digital competence of L2 pre-service teachers and their attitude towards the use of digital technologies in their L2 classrooms. The major implication that arises from this research is that self-efficacy plays a crucial role in the development of L2 teacher trainees' digital competence and positive intention to the implementation of ICT tools and applications in their future pedagogical practice.

We hope that this study would serve as a benchmark for future research on the topic and a starting point for rethinking the quality of digital training of future language teachers on a national level.

References

[1] Ilomäki, L., S. Paavola, M. Lakkala and A. Kantosalo, "Digital competence: An emergent boundary concept for policy and educational research," Education and Information Technologies, 21, 2016, pp. 655-679, doi: 10.1007/s10639-014-9346-4.

[2] F. Caena and C. Redecker, C. "Aligning teacher competence frameworks to 21st century challenges: The case for the European Digital Competence Framework for Educators (DigCompEdu),". European Journal of Education, 54(3), 2019, pp. 1-14.

[3] J. Janssen, S. Stoyanov, A. Ferrari, Y. Punie, K. Pannekeet and P. Sloep, "Experts' views on digital competence: Commonalities and differences," Computers & Education, 68, 2013, pp. 473-481.

[4] Digital Bulgaria 2025 National Programme (2019). [Online]. Available: https://www.mtc.government.bg/sites/default/files/uploads/it/09-12-2019_programa_-cifrova_bulgariya_2025.pdf (in Bulgarian)

[5] "Digital Qualification" Nationa Programme, Ministry of Education and Science (Bulgaria), 2021. [Online]. Available: https://web.mon.bg/bg/101029 (in Bulgarian).

[6] Prenksy, M. Digital natives, digital immigrants. *On the Horizon*, 9, 5, pp.1–6, 2021.

[7] C. Redecker and Y. Punie, European Framework for the Digital Competence of Educators (DigCompEdu), Luxembourg: Publications Office of the European Union, 2017.

[8] M. Perifanou, Digital Competence Assessment Framework & Tool for Language Teachers. DC4LT Consortium, 2021. [Online}. Available: https://www.dc4lt.eu/

[9] J. Tondeur, R. Scherer, F. Siddiq and E. Baran. "A comprehensive investigation of TPACK within pre-service teachers' ICT profiles: Mind the gap!," Australasian Journal of Educational Technology, 33(3), 2017, pp. 46-60.

[10] V. Milutinovic, "Examining the digital competencies of pre-service teachers," In Professional Competences for Teaching in the 21st Century (Proceedings of the International Conference Organised by the Faculty of Education in Jagodina on May 23−25, 2019), Ed. by V. Savić and O. Cekić-Jovanović, Section 5: Teaching Competences in ICT Education, 2019, pp. 373-391.

[11] P. Mishra and M. J. Koehler, "Technological pedagogical content knowledge: A framework for integrating technology in teachers' knowledge," Teachers College Record, 108(6), 2006, pp. 1017-1054.

[12] R. Scherer and T. Teo, "Unpacking teachers' intentions to integrate technology: A meta-analysis," Educational Research Review, 27, 2019, pp. 90-109.

[13] R. Scherer, J. Tondeur, F. Siddiq, E. Baran, "The importance of attitudes toward technology for pre-service teachers' technological, pedagogical, and content knowledge: Comparing structural equation modeling approaches," Computers in Human Behavior, 80, 2018, pp. 67-80.

[14] F. D. Davis. A technology acceptance model for empirically testing new end-user information systems: Theory and results. Massachusetts, United States: Sloan School of Management, Massachusetts Institute of Technology, 1986.

[15] J. Tondeur, et. al., "Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence," Computers & Education, 59(1), 2012, pp. 134-144.

[16] S. Kavanoz, H. Yuksel and E. Ozcan, "Pre-service teachers' self-efficacy perceptions on WEB pedagogical content knowledge," Computers & Education, 85, 2015, pp. 94-101.

[17] D. A. Schmidt et al., "Technological Pedagogical Content Knowledge (TPACK)," Journal of Research on Technology in Education, 42(2), 2009, pp. 123-149.

[18] J. Cohen, P. Cohen, S. G. West and L. S. Aiken, Applied multiple regression/correlation analysis for the behavioral sciences, 3rd ed., Mahwah, NJ: Lawrence Erlbaum, 2003.

[19] G. Kessler, "Formal and informal CALL preparation and teacher attitude toward technology," Computer Assisted Language Learning, 20, 2007, pp. 173–188.

[20] A. Lachner at al.,"Fostering pre-service teachers' technological pedagogical content knowledge (TPACK): A quasi-experimental field study," Computers & Education, vol. 174, 2021, 104304.

[21] H.-J. Cheng, "A TL-TPACK model on CSL pre-service teachers' competencies of online instruction," in Preparing foreign language teachers for next-generation education, C.-H. Lin, D. Zhang and B. Zheng, Eds., Hershey: IGI Global, 2017, pp. 198-225.

[22] K.-O. Jeong, "Preparing EFL student teachers with new technologies in the Korean context," Computer Assisted Language Learning, 30, 2017, pp. 488–509.

Examining the Relationship Between Students' Social Media Usage, Smartphone Checking During Lessons/Studying, and Academic Achievement (GPA)

M. Kućar*, V. Vidaček Hainš**, R. Kovačić***
*Institute of Social Sciences Ivo Pilar, Zagreb, Croatia
**Faculty of Organization and Informatics Varaždin, University of Zagreb, Croatia
***Electromechanical school Varaždin, Croatia
*mkucar@pilar.hr, **vvidacek@foi.hr, ***ratko.kovacic@skole.hr

***Abstract* - The interaction of young people with modern technology, specifically social media, is being studied in various contexts. This study aimed to examine the relationship between the students' social media usage (SMU), smartphone checking during lessons or while studying, and their academic achievement (GPA). In total, 225 participants took part in the study, with a mean age of 17.22 years (SD=1.57). To examine the role of time spent on Facebook (FB) and Instagram (IG), frequency of checking FB and IG, and smartphone checking during lessons/while studying, a hierarchical regression analysis was made. In the first step of the analysis, the participants' age and gender were controlled for, and both were significant predictors of the students' GPA. According to the findings, younger age, and female gender both predict a higher GPA. The amount of time spent on IG was a negative predictor of the student's GPA. Furthermore, the frequency of FB checking was also a predictor of the GPA, students who checked FB more frequently had higher GPAs. The two different measures of SMU (time spent and frequency of checking) were moderately congruent. The limitations and contributions of this study were discussed.**

Keywords – academic achievement; GPA; screen time; smartphone use; social media usage

I. INTRODUCTION

A. Social Media Usage (SMU)

Social media are internet-based, disentranced channels enabling mass personal communication through user-generated content [1], primarily in the form of text, audio, and video [2]. The way individuals spend their time, present themselves, and socialize has been fundamentally changed because of it. Questions have been raised about the potentially harmful effects of increased SMU. The amount of time that young people spend online has doubled over the past ten years and there is a growing debate about whether this change has a harmful effect on children and adolescents [3].

Social media is commonly used for social interactions in groups with similar interests, photo sharing, news reading, and business purposes. [3]. Whiting and Williams [4] emphasize that motivations vary depending on the platform in question. Cross-cultural research [5] showed a plethora of social media platforms used by adolescents – the most-popular were Instagram (IG), Snapchat, Facebook (FB), and Twitter. Social media platforms' popularity is continually changing throughout time. The change in popularity of various social media platforms is influenced by different factors, such as age, culture, or point in time. For example, FB's popularity is declining among younger generations, but it remains one of the world's most-used social media sites with over a billion active daily users. Users primarily access Facebook through mobile devices for personal information sharing and social connections [6].

Another example is Instagram (IG), a mobile application mostly used for photo and video-sharing, predominantly by young adults, with over 1.3 billion active users [5]. Research indicates that IG usage can be motivated by the need for self-presentation, social documentation (e.g., travel), and may be associated with certain aspects of narcissism [7], [8].

There is a growing body of studies exploring the effects of SMU in various contexts, but the studies are still mostly cross-sectional and use self-assessment measures of SMU [9]. Social media is widespread in both personal and professional settings, which has increased the significance of these technologies for education. College students reported that they use FB for social networking, work related activities and daily activities. Students primarily use FB for communication in academic groups, exchanging learning materials, and information regarding lessons, according to the study that aimed to identify the academic purposes for which students use the social media platform [10]. Research indicates that students

have positive attitudes towards using FB for communication and online interaction with peers, reporting that it aids in a better understanding of course content [11].

B. Smartphone Usage

Evolving research focused on the potential influence of modern technology on various psychological functions is not limited to specific social media sites. Some researchers focus on specific devices like television, tablet [9], or smartphone – the most used digital device for many individuals, with over 90% of the population owning one [12]. For example, a study found that college students reported that smartphones can have a positive impact on students' educational and psychological well-being [13]. On the other hand, a systematic report has linked problematic smartphone use to decreased health-related and overall quality of life as well as life satisfaction, and subjective well-being in adolescents [14]. The research connecting smartphone usage and academic achievement typically reports a negative relationship between the variables, although not necessarily causal. Such reports usually indicate that spending time on social media affects academic performance, not vice versa [15], [16]. Regarding the measurement of smartphone usage, research was first largely dominated by self-assessment measures. They were proven to be imprecise in most cases, so researchers are advocating for different ways of measuring smartphone usage and screen time [9]. However, self-assessment measures are still often present in this field because they are an effective and quick way of gathering wanted information. Possible alternatives include direct measures of smartphone usage. This information is available for most smartphones on both commonly operating systems (Android and iOS). Depending on the device and OS, information about the amount of time spent interacting with smartphone last week or on a specific app is also given [17]. The frequency of smartphone checking refers to checking the device for possible new notifications, even when no incoming notification is signaled. This measure can fall in the category of self-assessment, although information about the number of times a phone was unlocked is given for some mobile devices. Interestingly, this type of measurement is not necessarily highly correlated to smartphone screen time [9]. Three types of self-assessment measures are included in this study. The self-assessment of time spent daily on FB and IG, the frequency of smartphone checking while studying and during lessons, and the frequency of checking FB and IG.

C. Academic Achievement

Academic achievement refers to the performance outcomes that indicate how far a student has progressed toward specific goals that were the focus of activities in instructional settings. School systems typically define cognitive goals that apply across multiple subject areas or include knowledge and understanding in a specific intellectual domain (e.g., literacy) [18]. As a result, academic achievement should be viewed as a multifaceted construct encompassing various domains of learning. Because the field of academic achievement is so broad and encompasses such a wide range of educational outcomes, the definition of academic achievement is dependent on the indicators used to measure it. Academic achievement measures usually include standardized tests, course exam scores, course grades, and a cumulative or current grade point average (GPA). Test scores and GPA are the most used measures of educational or academic achievement, mostly because of their objectivity, simplicity, and possibility of comparison between different samples [19], [20]. Some researchers found empirical evidence implicating a link between academic achievement and SMU, while others reported the absence of such a link. For example, Alaslani and Alandejani [21] found that undergraduate and graduate students who use social networks more often have significantly better interactions with colleagues and instructors, higher engagement, and better, cooperative learning and student achievement. Sharing knowledge among colleagues improves the learning process and has an impact on improving students' achievement.

However, there is a growing body of evidence that unmonitored, excessive social media use which is not related to school/academic activities leads to lower academic achievement. One explanation is the fact that social media often presents a distraction from academic activities and therefore leads to lower educational success [22]. It was found that university students who spend more time on social media have lower GPAs than students who spend less time on social media or do not use social media. Although, their SMU for academic purposes was positively related to their academic achievement [23],[24]. Similar evidence was found by Junco [25], who reported that senior students spent significantly less time on FB compared to freshmen. His research explored the negative correlation between the time that freshman students spend on FB and their GPA. Multitasking with FB was negatively correlated with GPA for younger students, but there was no negative correlation for seniors.

II. METHOD

A. Research Problem

The overall objective of this study is to explore the role of students' SMU (FB and IG) and smartphone checking in explaining their academic achievement (GPA).

The hypothesis is that students' FB/IG usage, frequency of checking FB/IG, smartphone checking while studying and during lessons will explain a significant amount of variance of the GPA in the last academic year. It is expected that all predictors regarding the amount of usage will be negative, meaning that higher usage is related to lower GPA. Frequency of checking FB/IG will be positively related to GPA, because higher values in checking frequency mean that more time passes between two checking occasions.

The second objective is to check the congruence between the two different measures of SMU (time spent and frequency of checking) for both FB and IG. It is expected that the two measures of SMU will be moderately correlated.

B. Participants

There were 225 participants in total, with a mean age of 17.22 years (SD=1.57). There were more male participants in the study (72%), than female (26%). Five participants (2%) have not answered the question related to gender or stated "other/undefined". Participants were largely students from two vocational high schools in Croatia (84%), from programs focused on computer science and engineering. The rest of the sample (16%) consisted of public university students, all of which were first-year students in applied computer science program. There were 37 first-year, 37 second-year, 36 third-year, 80 fourth-year high school students, and 35 first-year university students.

C. Data Collection

The participants were told that the research focuses on examining students' technology usage habits and that their participation is voluntary. They were given a questionnaire (Social Media Questionnaire – SMQ) and were told that their answers were anonymous and interpreted only on a group level. Initially, students answered socio-demographic questions related to their gender, age, school/university program, and their previous year's GPA. Participants answered questions separately for each of the popular social media networks (FB and IG). Their participation was not followed by compensation of any kind. All data was collected via the paper-pen method. Prior to data collection, research approval from the Ethics committee of the Faculty of Organization and Informatics was obtained.

D. Measures

FB & IG usage - participants were first asked if they had a FB/IG profile. Out of 225, only 3 participants (1.3%) answered that they do not have a FB profile, and 21 participants (8.5%) answered they do not have an IG profile. If the answer was "yes", they were asked to estimate how much time they typically spend on a certain media site daily. They were provided four possible answers with gradually increased time spent on the certain platform (1= ''30 minutes or less'', 2= "around 1 hour or less", 3= "around 2 hours or less", 4= "more than two hours'').

Frequency of FB & IG checking - participants were asked to estimate how often they check or turn on the mobile application FB/IG. This was a single item question in an open response format.

Smartphone checking while studying and during lessons - these constructs were assessed as two separate single item questions *("How often do you check your smartphone while you are studying/during lessons?"*). The items were rated on a 5-point Likert-type scale ranging from 1 to 5 (1= almost never, 2= rarely, 3= sometimes, 4= often, 5= almost all the time).

Academic achievement - participants were simply asked to write down their GPA (with two decimals) from last academic year which means that academic achievement was evaluated with a single item. For the first-year high school students, it was their average grade in the 8th grade of elementary school. For first-year university students, it was their grade point average from the fourth year of high school education.

III. RESULTS

A. Descriptive Statistics

Descriptive statistics for all study variables are displayed in Table 1. Variables regarding time spent

on and frequency of checking FB/IG are expressed in minutes. The distribution of students' estimations of their FB/IG usage is positively skewed. Participants' estimations of the frequency of checking FB/IG have high variability. Answers which implied that a person is, on average, checking FB/IG less often than once every 12 hours (720 minutes) were removed from the study because they arguably do not qualify as regular users of the application. Average checking times for both FB and IG are slightly over 1 hour, every 71.10 minutes for FB and 65.24 for IG. Overall, the participants are checking their smartphones while studying and during lessons moderately often (studying-M=3.19, SD=1.01, lessons-M=3.01, SD=1.17). The distribution of students' GPA is negatively skewed with a small standard deviation (M=4.18, SD=0.51). The correlation matrix showed a moderate correlation between time spent on FB and IG ($r=0.44$, $p<0.01$) and between frequency of FB and IG checking ($r=0.35$, $p<0.01$). Two measures of smartphone checking are also moderately correlated ($r=0.35$, $p<0.01$). Students' GPA is significantly correlated with the participants' age ($r=-0.20$, $p<0.01$), time spent on both FB ($r= -0.19$, $p<0.01$) and IG ($r=-0.32$, $p<0.01$), with checking smartphone during studying ($r=-0.23$, $p<0.01$), and during lessons ($r=-0.18$, $p<0.01$).The correlations between the two different measures of using FB and IG are both negatively moderately correlated, which confirms the stated hypothesis (FB =-0.37, $p<0.01$, IG=-0.44, $p<0.01$).

TABLE 1. DESCRIPTIVE STATISTICS AND FREQUENCIES FOR ALL STUDY VARIABLES

	N	**Min.**	**Max.**	**Mean**	**SD**
age	220	14	21	17.22	1.57
gender	220	0	1	-	-
FB –time spent	221	1	4	2.14	1.08
	N	%	c. %		
30min or less	77	34.8	34.8		
Around 1h or <	74	33.5	68.3		
Around 2h or <	32	14.5	82.8		
More than 2h	38	17.2	100		
FB – checking	217	2	720	71.10	108.69
IG –time spent	201	1	4	2.28	0.97
	N	%	c. %		
30min or less	47	23.4	23.4		
Around 1h or <	78	38.8	62.2		
Around 2h or <	49	24.4	86.6		
More than 2h	27	13.4	100		
IG – checking	201	2	600	65.24	100.77
Phone-studying	225	1	5	3.19	1.01
Phone- lessons	225	1	5	3.01	1.17
GPA	221	2.59	5.00	4.18	0.51

Note: C.% - cumulative percent, Phone - Smartphone

B. Hierarchical regression analysis

A hierarchical regression analysis was conducted, with GPA as the dependent variable and FB-time spent, FB-checking, IG-time spent, IG-checking, smartphone-studying, and smartphone-lessons as predictors. To account for the possible effects of socio-demographic factors, age and gender were added in the first step as control variables. The predictors were added in the second step of the analysis. Before conducting the regression analysis, certain underlying assumptions were checked, including the multicollinearity, linearity, normality, independence, homoscedasticity, and absence of outliers [26]. Absence of outliers was checked with Cook's distance value which shows the influence of each individual observation on the fitted response values. The highest Cook's distance value in the dataset was 0.1 which would imply the absence of outliers. The multicollinearity assumption was tested with intercorrelations among the independent variables and VIF values. First, there was no correlation between any study variables higher than $r=0.5$. Second, the highest VIF value in the analysis was 1.46 which is acceptable and implies a specific contribution of every independent variable. The residual plot was used to test linearity and homoscedasticity assumptions. The plot shows the relationship between standardized residuals and standardized predicted values, in this case the residuals were scattered randomly around the horizontal and vertical lines representing residuals equal to zero. The normality assumption was tested using the Normal P-P Plot of the standardized residuals. The observations are falling on the straight diagonal line which means the assumption is met. Lastly, Durbin-Watson value was used to test the independence assumption and test for auto correlation in the residuals of a statistical regression analysis. The observed value was 2.20, which falls in the acceptable range. Including socio-demographic variables in the first step of the analysis explained only 5% of the variance in the criteria but reached the statistical significance ($F_{(2,221)} = 4.57$, $R^2 = 0.05$, $p<0.05$). Adding the predictor variables increased the variance explained by 16%, making total variance explained 20% ($F_{(8,213)} = 5.81$, $R^2 = 0.20$, $p<0.05$). The regression model is predicted by the sociodemographic factors and independent variables related to social media/smartphone usage. As seen in Table 2, significant predictors of the students' GPA were age ($\beta= -0.24$, $p<0.01$), female gender ($\beta= 0.16$, $p<0.05$), time spent on IG ($\beta= -0.35$, $p<0.01$), and FB checking ($\beta= -0.20$, $p<0.01$). The main hypothesis is partially supported.

TABLE 2. HIERARCHICAL REGRESSION ANALYSIS OF THE CONTROL AND INDEPENDENT VARIABLES ON GPA

	GPA	
	Step 1	***Step 2***
Predictor	**β**	**β**
Age	-0.21**	-0.24**
Gender	0.11	0.16*
FB - time spent		-0.06
FB - checking		-0.20**
IG - time spent		-0.35**
IG - checking		-0.01
Smartphone – studying		-0.08
Smartphone - lessons		0.03
F	4.57*	5.81**
R^2	0.05	0.20
ΔR^2	-	0.16

Note: * p<0.05, **p<0.01

IV. DISCUSSION

The main hypothesis of this study is partially confirmed. According to the results, only the frequency of FB checking, and the time spent on IG significantly predict the criteria. Surprisingly, the frequency of FB checking is negatively related to the GPA, which means that students who check FB more frequently have a higher GPA. Conversely, the more time spent on IG, the lower the GPA. One possible explanation for this finding is the fact that FB is often used for educational activities, such as exchanging course materials, or forming lesson groups, while IG is not. Typically, IG is quite rarely researched or mentioned in the educational context. This result replicates a finding of a study conducted by Junco [25] who reported that there is no negative correlation between spending time on FB and GPA. If this explanation is true, it still does not explain why time spent on FB does not predict the GPA, while the frequency of checking does. It is possible that students who are highly conscientious, motivated, or even anxious about their academic results, check FB more often in case new course material or test results in their group appear.

The amount of time spent on IG is found to be a significant negative predictor of GPA. As a rapidly evolving social media platform, IG offers plenty of content and has advanced algorithms that predict users' interests, resulting in excessive scrolling and spending time in the application. It is reasonable to presume that more time spent on IG means less time left for productive work (learning), which is reflected in the grades (GPA). Although the correlational matrix showed a negative correlation between smartphone checking (while studying and during lessons) and GPA, the two variables were not significant predictors in the hierarchical model. It needs to be considered that oftentimes participants will purposefully decrease their frequency of checking their smartphone during lessons or while studying. That can especially happen in the educational context because of social desirability. The control variables were both significant predictors of the outcome. Age was a negative predictor of the GPA, which is expected considering that educational challenges increase as a person's academic pathway progresses. On average, female participants had a higher GPA in this sample. However, the study sample is not balanced by gender, so this difference needs to be interpreted with caution. There are several important limitations to this study that need to be addressed. First, the cross-sectional nature of this study makes causal explanations impossible. In the future, longitudinal research should be considered. Second, the sample size is acceptable, but a larger sample with a better gender balance would make the study results more generalizable to the broader population. The sampling method may also limit the generalizability of the findings. This sample is not probabilistic, it is convenient. Arguably, it is possible that the high school and university students involved in this study do not differ significantly from the average Croatian student. One reason for that claim could be the lack of untypical or extreme characteristics related to this group. For example, they are not students at art schools, academies, or private institutions, and high school students come from vocational schools with, typically, average state exam results. Their average GPA is also in accordance with the GPA of vocational school students in Croatia. Many of them continue their education in the university program included in this study [27]. Future studies could consider gathering information from different institutions in a certain municipality or in the whole country. Also, future studies could include possible mediator or moderator variables which could explain the link between social media and academic achievement. Such factors may include self-efficacy, learning-related habits, personality factors, or intellectual ability. Also, future research could make a comparison between different types of students (e.g., public vs. private institutions, STEM vs. the non-STEM field). Different ways of measuring SMU or the usage of a certain device are discussed in the introduction. Future studies could incorporate various measures, it would especially be useful to include a direct measure of SMU. The contribution of this study is the finding that using different social media platforms can have indigenous relationships with academic achievement. Another contribution is the comparison of the two measures of SMU and the finding that they are only moderately congruent. This field of study has significant challenges regarding the measurement of SMU, so studies that include and compare various types of

screen time measurement are possibly of value for future researchers.

V. CONCLUSION

This study explored the link between SMU (FB and IG), smartphone checking, and the GPA of students in secondary and tertiary education. Control variables (age and gender) were both significant predictors of the GPA; older students had lower GPAs, while female students had higher GPAs. The results indicate that the amount of time spent on IG negatively predicts the student's GPA score. The more time a student spends on IG, the lower his GPA. Students' frequency of FB checking is also a predictor of the GPA, those who check FB more often have higher GPAs. In total, the regression model explained 20% of the variance in the criterion (GPA).

VI. REFERENCES

[1] N. Akhtar, "Relationship between Internet Addiction and Academic Performance among University Undergraduates," Educational Research and Reviews, vol. 8, no. 19, pp. 1793-1796, 2013.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[2] A. Orben, "Teenagers, screens and social media: a narrative review of reviews and key studies," Social psychiatry and psychiatric epidemiology, vol. 55, no. 4, pp. 407-414, 2020.P. Perić, "Naslov rada," naziv časopisa, prihvaćen za tisak

[3] G. Seidman, "Self-presentation and belonging on Facebook: How personality influences social media use and motivations," Personality and Individual Differences, vol. 54, no. 3, pp. 402–407, 2013.

[4] A. Whiting, D. Williams, "Why people use social media: a uses and gratifications approach," Qualitative market research: an international journal, vol.16, no. 4, pp. 362-269, 2013.

[5] S. Alhabash and M. Ma, "A Tale of Four Platforms: Motivations and Uses of Facebook, Twitter, Instagram, and Snapchat Among College Students?" Social Media and Society, vol. 3, no. 1, pp. 1-34, 2017.

[6] E. Larson and C. Vieregger, "Strategic actions in a platform context: what should Facebook do next?" Journal of Information Systems Education, vol. 30, no. 2, pp. 97-105, 2019.

[7] T. T. Cheung, "A study on motives, usage, self-presentation and number of followers on Instagram," Outstanding Academic Papers by Students (OAPS), City University of Hong Kong, CityU Institutional Repository, 2014.

[8] P. Sheldon and K. Bryant, "Instagram: Motives for its use and relationship to narcissism and contextual age," Computers in human Behavior, vol. 58, pp. 89-97, 2016.

[9] L. K. Kaye, A. Orben, D. A. Ellis, S. C. Hunter and S. Houghton, "The conceptual and methodological mayhem of "screen time"," International Journal of Environmental Research and Public Health, vol. 17, no. 10, pp. 3661, 2020.

[10] Y. M. Al-Dheleai and Z. Tasir, "Using Facebook for the purpose of students' interaction and its correlation with students' academic performance," The Turkish Online Journal of Educational Technology, vol. 16, no. 4, pp. 170–178, 2017.

[11] N. Đurica, D. Soleša, V. Šimović and M. Đurica, "The Students' Perception of the Importance of Using Facebook for Academic Purposes," Croatian Journal of Education, vol. 20, no. 4, pp. 1059–1087, 2018.

[12] "Mobile Fact Sheet," Pew Research Center. Available:https://www.pewresearch.org/internet/fact-sheet/mobile/ (accessed on Jan 19, 2023).

[13] G.R. Jesse, "Smartphone and App Usage Among College Students: Using Smartphones Effectively for Social and Educational Needs," Proceedings of the EDSIG Conference; Wilmington, North Carolina, USA, pp. 1-13, 2015.

[14] L. Fischer-Grote, O. D. Kothgassner, O. D. and A. Felnhofer, "The impact of problematic smartphone use on children's and adolescents' quality of life: A systematic review." Acta Paediatrica, vol. 110, no. 5, pp. 1417-1424, 2021.

[15] A. Lepp, J.E. Barkley and A. C. Karpinski, "The relationship between cell phone use and academic performance in a sample of US college students," Sage Open, vol. 5, no. 1, 2015.

[16] O. Sapci, J. D. Elhai, A. Amialchuk i C. Montag, "The relationship between smartphone use and students' academic performance," Learning and Individual Differences, vol. 89, 2021.

[17] T. S. Van Endert and P. N. Mohr, "Likes and impulsivity: Investigating the relationship between actual smartphone use and delay discounting," PloS one, vol. 15, no. 1, 2020.

[18] R. Steinmayr, A. Meiŋer, A. F. Weideinger and L Wirthwein, "Academic achievement," Oxford University Press, 2014.

[19] T. T. York, C. Gibson and S. Rankin, "Defining and measuring academic success," Practical assessment, research, and evaluation, vol. 20, no. 1, 2015.

[20] E. S. Han and J. Keefe, "The impact of charter school competition on student achievement of traditional public schools after 25 years: Evidence from national district-level panel data" Journal of School Choice, vol. 14, no. 3, pp. 429-467, 2020.

[21] K. Alaslani and M. Alandejani, "Identifying factors that influence students performance through social networking sites: An exploratory case study," Heliyon, vol. 6, no. 4, 2020.

[22] A. G. Kolo, W.M.W. Jaafar and N. Ahmad, "Influence of psychosocial factors on student's academic performance in one of Nigerian colleges of education," Malaysian Journal of Social Sciences and Humanities (MJSSH), vol. 2, no. 1, pp. 1-10, 2017.

[23] W. W. Lau, "Effects of social media usage and social media multitasking on the academic performance of university students," Computers in human behavior, vol. 68, pp. 286-291, 2017.

[24] M. Muslikah and A. Andriyani, "Social media user students' academic procrastination" Psikopedagogia Jurnal Bimbingan Dan Konseling, vol. 7, no. 2, pp. 53-57, 2018.

[25] R. Junco, "Student class standing, Facebook use, and academic performance," Journal of Applied Developmental Psychology, vol. 36, pp. 18–29, 2015.

[26] B. G. Tabachnick and L. S. Fidell, "Using Multivariate Statistics," 5th edn Allyn & Bacon.

[27] Školski e-Rudnik ŠeR - Školski e-Rudnik (Vol. 3), https://mzo.gov.hr/istaknute-teme/ser-skolski-e-rudnik-3419/3419/ (accessed Jan. 19, 20).

Face-to-Face and Online Communication and Etiquette in Secondary Education

Ena Martinčević* and Violeta Vidaček-Hainš **
* Faculty of Organization and Informatics, Varaždin, Croatia
Email address:
martincevic.ena@gmail.com
vvidacek@foi.unizg.hr

***Abstract* - Good communication is the key for succeeding in any aspect of our lives, meaning it is essential for education as well. Besides communication, the significance of etiquette plays an important part in providing a sense of belonging to a certain community. The aim of this research paper is to investigate and shine new light on communication and etiquette among high school students and teaching staff in Varaždin County. This paper has been divided into theoretical and empirical part. In the theoretical part of the paper, the emphasis is on communication and etiquette and their significance in secondary education. For the empirical part of the work, research has been carried out and obtained data have been statistically analyzed and interpreted. Results showed that there are no gender differences in etiquette behavior in face to face student communication, but in online environment female students know more etiquette rules than male. Students find professors' nonverbal expression as important part of communication, particularly those who are attending more complex programs. Students' disagreements in per communication and the frequency of disagreements are not dependent.**

The findings of this research will serve as a base for determining the quality of communication and etiquette in the aforementioned high school.

Keywords* - *component; communication, etiquette, education

I. INTRODUCTION

Face to face and online communication are becoming important elements in teaching and learning process. According to research results [1], there are authoritarian and authoritative teachers. Authoritarian teachers usually take the floor themselves and do not involve students in discussions, while authoritative teachers regularly involve students in the very process of teaching and setting rules. An important aspect of teaching is active listening [1]. Students can enrich their vocabulary through active listening and this can ultimately have an impact on the development of literacy, which increases success in school and raises the students' self-confidence. Active listening enables the teacher to create better relationships with students, encourages innovation, and helps in decision-making [2].

Non-verbal communication is also important part of everyday communication. Special attention should be paid to the facial expressions of teachers and students, the movement of the teacher through the classroom, the tone of voice, as well as a dose of humor that can create a relaxed atmosphere, although it should be used only within certain limits. Already upon entering the classroom, the teacher can evoke various feelings in the students with his or her non-verbal cues [3]. For example, if students do not look the teacher in the eye, show signs of nervousness or have a shaky voice, they can show fear and discomfort. [4].

With the constant development of technology, teachers and students are faced with increasing and modernizing school standards. Social networks can help create greater student motivation, enable easier access to school materials and information as well as offer help with homework [5]. According to the research conducted by [6], students use Instagram most often while they most often communicate via the mobile application Viber. Distance learning has been made interesting and interactive by the Kahoot platform designed for learning in a fun and more engaging way where students are also involved in the teaching process. The platform allows repeating the material, but also checking students' knowledge, where students answer the questions individually or in teams, and those who answer correctly the fastest get the most points [7].

Information communication technology (ICT) refers to all technical means that help in handling information. E-learning, with the help of new technologies, enables students to spread and transfer knowledge outside their educational institution [8]. Usage of ICT is helpful for teachers and their lecturing and enables students to have access to teaching materials in one place and to search for them faster. Research conducted by [9] showed that ethical issues are becoming progressively more important by increasing usage of social media in education. The new information and communication technologies, particularly social networks such as Facebook, have had a significant impact on psycho-social adjustment of young people [10]. Teachers can provide assignments and instructions to students in various ways, such as recording sound on an existing PowerPoint presentation, sending students useful

links, creating quizzes, compiling instructions in PDF format, and the like [11].

Children first learn about the rules of good behavior from their parents, where parents are obliged to devote themselves individually to the child's needs and prepare them for social life so that they acquire the necessary social skills. Enrolling children into school is important for their social development, as children get to know their peers and are confronted with new situations or challenges. It is teacher's duty to encourage children to cooperate with each other in a supportive manner and guide them towards acceptable behavior [12]. For example, environmental topics could be presented in the curriculum, which would awaken student's awareness of the environment and influence their perspective on the environment [13]. Teachers reported that students' misbehavior has a negative impact on their job satisfaction and student achievement. Specific problems with students' misbehavior are situations where students are not following teachers' instructions and are not paying attention during lessons [14]. In order for both parts, teaching staff and students, to feel comfortable during class, it is essential to achieve a positive working atmosphere that enables a pleasant stay at school. Teachers are the ones who can establish the same with their skills, and encourage students to accept diversity and mutual respect so that they can successfully overcome all the changes of the modern school [15]. Authors [16] reported that school misbehavior often leads to peer violence. Their research results showed that nearly half of teenagers were victims during schooling and that bullying happens increasingly in the classroom (58,1 %). Additional prevention programs need to be created and implemented in schools which aim to raise awareness of this problem and to help students develop communication competences and appropriate behavior patterns [16]. Sometimes a lack of etiquette appears in communication at appears on social networks communication as individuals unnecessarily express their opinion towards others, say things they would never say in face-to-face communication and forget about good manners.

In the online environment, each individual must know how to publish personal information, and be careful what kind of comments they leave on other people's posts. Every post in the online environment leaves a digital trail; it is preferable to avoid unkind posts because this can eventually affect the reputation of individuals [17]. Some of the advantages of using the Internet in education are certainly the increase in student motivation, engagement and cooperation, which also increases students' technological skills. However, there are also negative aspects of the Internet such as violent behavior that can consequently lead to the psychological and emotional decline of individuals [18]. It is important to teach students how to communicate in the virtual world, how to recognize the electronic violence, which is most often associated with sending offensive messages to peers [19]. There is a gender difference on social network behavior among high school students; girls are more exposed to virtual violence and boys participate more often in virtual violence. Virtual addiction is more often among male students [20].

The general goal of this research is to examine the quality of communication between students and teachers in the selected secondary school. In addition, the aim is to explore how students are familiar with school etiquette, i.e. whether students behave politely during school face to face communication and in the online environment.

Specific goals are to explore if there are gender differences in terms of online an face to face communication, as well as to explore if there are potential difference between students attending three and four years of high school programs.

II. RESEARCH METHODOLOGY:

A. *Respondents*

Respondents are students attending one high school in Varaždin region. Total number of respondents are N=90 students, 45,56% are male students, 48,89% are female students and 5,56% of students did not want to reveal their gender. Students are between 17 and 18 years of age.

Respondents are attending three and four years programs, such as general high school program (26.67%), computerized numerical control operators (21.11%), economists (20.00%), sales (15.56%) and others (16.67%).

B. *Research*

For the purpose of this research, the questionnaire Communication skills (CS) has been designed. The questionnaire consists of 30 items divided in demographic part and Likert scale items. Likert scale items consist 1 to 5 and express agreements with the statements form strongly disagree to strongly agree.

The research was conducted in March 2022. The written approval from school authorities and parents was issued. The questionnaire was completed anonymously and voluntarily.

III. RESEARCH RESULTS AND INTERPRETATION

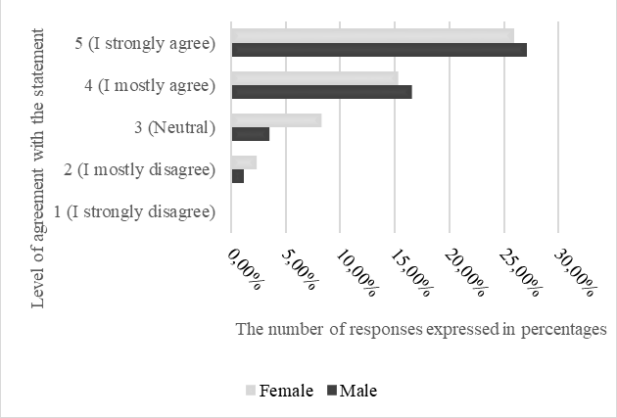
The results for the conducted research are shown below. Respondents were asked about the quality of communication between them and teachers.

A. *Etiquette at school and online*

Students were asked about the etiquette in communication and gender differences were explored. The gender difference for the item: "I always greet

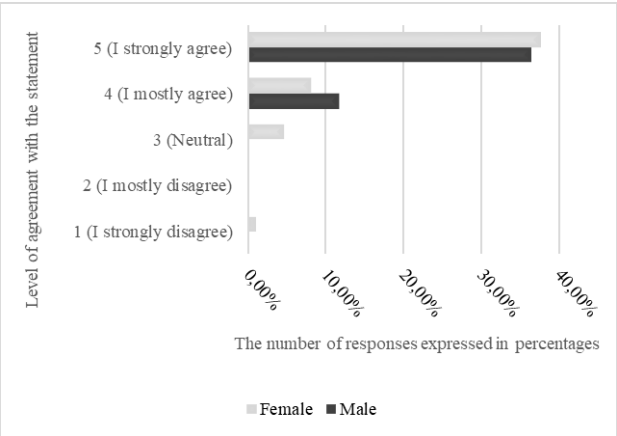
teachers and other school employees when I meet them in the corridor or outside the school and I do not wait for them to greet me first" was explored. Results showed that there is no statistically significant difference in values regarding greeting teachers in school or outside of school between male and female students. By calculating the t-test (t =1.630), the degrees of freedom (df=83) and p=0.1069, as well as p>0.05 it can be confirmed that there is no statistically significant difference in greeting teachers at school or outside of school between male and female students.

This result showed that high school students pay attention on some elements of etiquette such as greetings their teachers.



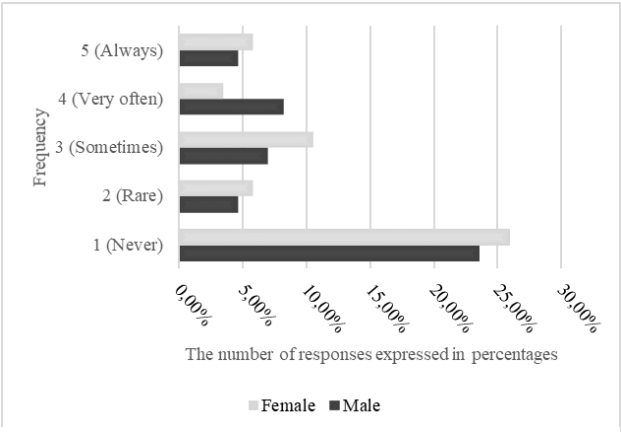
Graph 1. The difference between male and female students in greeting the teacher and other school employees

The following statement: "When I address a teacher via email or some other online tool, I address them with respect (e.g. I start the email with Dear Sir or Madam, when I address them I use capital letters for words such as You, Yours, etc.)" was explored in terms of gender differences. It was confirmed that there is no statistically significant difference in mean values in expressing respect for teachers via e-mail between male and female students (t=0.7998, df=83, p=0.4261).



Graph 2. The difference between male and female students in showing respect to teachers via email or other online communication tools

The next statement that was explored in terms of gender difference is related to the online school etiquette: "When I address a teacher or another student via email or another online tool, I use capital letters (e.g. *ETIQUETTE IN SCHOOL*)". Results showed that there is statistically significant difference in means of values for usage of capital letters in emails or other online tools between male and female students. Statistical method of t-test (t=2.355, df=83, p=0.0287) confirms that female students pay more attention to etiquette in online communication and use capital letters for the whole sentences and words less often than male students.



Graph 3. The difference between male and female students in the use of capital letters during online communication

B. Satisfaction with school and online communication

The second research question is: "Are students more satisfied with mutual communication at school or in an online environment?". Statements from the questionnaire related to satisfaction with mutual communication at school and online environment were taken into account, with N=90 students responding to both statements. According to the value of the t-test (t= 0.724, df=178, p=0.4700), it could be concluded that there is no difference in mean of values related to satisfaction with communication between students in face-to-face communication comparing with online communication.

C. The importance of non-verbal communication from teachers towards students

The third research question is: "Does teachers' non-verbal communication affect the students more than teachers' verbal communication?". In order to find out the answer to the mentioned research question, t-test was calculated for two statements from the questionnaire.

According to the value of the t-test (t= 2.514, df=88, p=0.0138), it can be concluded that there is a difference in the mean values related to the influence of the teachers' tone of voice as the element of non-verbal communication on students between three- and four-year courses. More precisely, the teachers' tone of voice has less influence on

students who are attending three-years of high school compared to students who are attending the four years high school program.

Additional item "The teachers' facial expression during the lecturing is important to me" was analyzed in terms of potential differences between students attending three and four years of high school program. Results showed that there is a difference in the mean values related to the influence of the teachers' facial expression on students who attended three years of high school compared to students enrolled in four years' education. T-test (t=2.604, df=88, p=0.0108) confirms that there is a difference in means of values related to the influence of the teachers' facial expressions during lecturing/ presentation of the material on the students between students enrolled to the three year and four-year high school programs. More precisely, the influence of the teachers' facial expression is lower on students of three-year courses compared to students of four-year courses. It could be concluded that non-verbal communication is more important for students attending more complex, four years of high school programs.

D. Dependence of disagreements resolution on the frequency of disagreements

The additional research question is: "Do students' ways of resolving disagreements depend on the frequency of encountering disagreements?". In order to get an answer to the research question, two variables were taken into account. A critical value with a significance level of 5% and df=12 was found in the table of degrees of freedom for Chi-square test (X^2 =0.2021). The critical value for Chi-square is 21.0261. As the value obtained for the Chi-square is 0.2021 and Chi-square value in this research is significantly smaller than the critical value, we concluded that the variables linked to the different ways of resolving disagreements and the frequency of encountering disagreements are not dependent.

Although, the obtained results are interesting, it needs to be mentioned certain limitations appear during the implementation of the research itself. It is important to mention that relatively small number of respondents, only the graduates of one high school, were considered and have participated in research. In order to get the entire picture of the development of communication and adherence to etiquette in the selected high school, bigger research sample could be considered. Likewise, if one would like to get an overall picture of communication and etiquette in high school education, students from other high schools should also be included in the research. Another limitation could be the objectivity and honesty of the students when giving answers. A relatively small number of respondents can be mentioned as a limitation of the research. Likewise, the research was conducted on a convenient sample of respondents, i.e. available students or graduates of the selected high school who were present at the time of the research. Therefore, the representativeness of the sample is limited.

IV. CONCLUSION

Etiquette is important part of everyday communication between students and teachers in high school.

The research results show that there are no statistically significant differences in the answers between male and female students in terms of school etiquette. The only difference occurred in the statement related to the usage of capital letters in online communication. This has shown that female students are more familiar with online etiquette and the correct use of capital letters compared to their male peers. It was found that female students pay better attention to the use of capital letters in an online communication, compared to male students, meaning that female students pay slightly better attention to online etiquette.

It was also shown that students are equally satisfied with face-to-face communication and online communication, and that the same is considered for teaching. Likewise, students attending three-year high school program are less influenced by non-verbal communication from the teacher compared to students attending four-year high school program. In addition to the above, according to the obtained results, it was concluded that the variables ways of resolving disagreements and the frequency of encountering disagreements do not depend on each other.

Results of this research could be used for creating the communication training programs for online and face-to-face communication in high schools. Further research could be extended to other high schools and different student body.

REFERENCES

[1] K. Vanek and P. Karabin, "Tko su dobri učitelji? Školski vjesnik, Journal for Pedagogic Theory and Practice, vol. 70, no. 2, 349-370, 2021.

[2] R. Varga and I. Somolanji Tokić, "Is a Systematic Approach to Teaching the Active Listening Skill Necessary?", Školski vjesnik: Journal for Pedagogic Theory and Practice, vol. 64, no. 4, pp. 647-660, 2015.

[3] D. Kožić, A. Globočnik Žunac and Lj. Bakić-Tomić, "Use of Non-verbal Communication Channels in the Classroom", Croatian Journal of Education, vol.15, no.. Sp.Ed.1, pp. 141-153, 2013.

[4] S. Bangoji, "Communication in Education", Knowledge-International Journal, vol. 46, no. 6, pp. 1159-1163, 2021.

[5] J. Diklić, M. Nakić and D. Šošić, "The Educational Role of Social Networks in the Media Education of Children", Communication Management Review, vol. 04, no. 01, pp. 178-195, 2019.

[6] G. Buljan Flander, E. Selak Bagarić, K. Prijatelj and M. Čagalj Farkas, "Analysis of Current Trends in Use of Social Networks among First and Third Grade Students of Secondary Schools in Croatia", Criminology & Social Integration: Journal for Criminology, Penology and Behaviour Problems, vol. 28, no. 2, pp. 277-294, 2020.

[7] B. Groznik, “Teaching Using the Kahoot Learning Platform”, Varaždin Teacher: Digital Professional Magazine for Education, vol. 4, no. 6, pp.190-204, 2021.

[8] I. Smiljčić, I. Livaja and J. Acalin, “ICT in Education”, Collected papers of Šibenik College, vol. 1, no. 3-4/2017, pp. 157-170, 2017.

[9] O. A. Ige, “What We do on Social Media! Social Representations of Schoolchildren's Activities on Electronic Communication Platforms”, Heliyon, vol. 6, no. 8, 2020.

[10] S. Rogošić, B. Baranović and J. Šabić, “Application of ICT in the Process of Learning, Teaching and Evaluation in Upper Secondary Vocational Schools: Qualitative Analysis”, Methodical Review : Journal of Philosophy of Education, vol. 28, no. 1, pp. 63-88, 2021.

[11] K. Marciuš Logožar, “Distance Education (Online Education) Focused on the Student”, Progress - *Napredak:* Journal for Interdisciplinary Research in Education, vol. 162, no. 3-4, pp.345-369, 2021.

[12] Z. Hercigonja, “Parents and Teachers - Educational Partners”, Varaždin Teacher: Digital Professional Magazine for Education, vol. 3, no. 3, pp.44-55, 2020.

[13] R. Jukić, “Ecological Question as an Educational Need”, Social Ecology: Journal for Environmental Thought and Sociological Research, vol. 20, no. 3, pp. 267-286, 2011.

[14] T. Vidić, “Student Classroom Misbehavior: Teachers' Perspective”, Croatian Journal of Education, vol. 24, no. 2, 599-639, 2022.

[15] D. Markulin, “The Teacher's Skills in Encouraging a Successful Teaching Climate”, Journal of Educational Sciences Foo2rama, vol.4, no. 4, pp. 45-60, 2020

[16] N. Kosić Bibić and J. Kovačević “Peer Violence in Schools - Violence or Style of Communication?” Food in Health and Disease , vol. Special Issue , no. 10. Štamparovi dani, 79-85, 2018.

[17] C. Kryder, “Online etiquette in the Digital Age”, AMWA Journal, vol. 28, no. 03, pp.130-131, 2013.

[18] E. Hasmujaj and J. Bushati, "Internet in Education System", Pannoniana, vol.1, no. 2, pp. 71-79, 2017.

[19] L. Vejmelka, N. Strabić and M. Jazvo, “Online Activities and Risky Behaviors of Adolescents in a Virtual Environment”, Društvena istraživanja, Journal for General Social Issues, vol. 26, no. 1, pp. 59-78, 2017.

[20] A. Čelk, “Dimensions of Virtual Behaviour of Adolescents”; Školski vjesnik , School News, vol. 68., no. 2., 413-425, 2019.

Non-Formal Education: Dynamics of Changes in the Popularity of Mass Open Online Courses

N. Dehtiarova*, O. Zhmud**, V. Makarova*, O. Hontar*, M. Zakharevych**
* Sumy State Pedagogical University named after A.C. Makarenko, Sumy, Ukraine
**Pavlo Tychyna Uman State Pedagogical University, Uman, Ukraine
degtyarevanv@fizmatsspu.sumy.ua, o.v.zhmud@gmail.com, vikaodinstova@gmail.com,
helenhontar18@gmail.com, zakharuman@gmail.com

***Abstract* - Urgency and peculiarities of the integration of formal and non-formal education forms were considered in the article. Conscious assimilation of knowledge when integrating non-formal courses into the educational process of the university on the basis of experimental research was proven.**

Results of a survey of more than 200 respondents from different cities of Ukraine regarding non-formal education in 2021 and 2022 were analyzed. Dynamics of changes in relation to non-formal education in peacetime and during military operations were presented. Recommendations for taking into account the student's learning results during non-formal education were summarized.

Keywords - non-formal education, mass open online courses, computer in education, self-education, lifelong studying.

I. INTRODUCTION

Education in Ukraine and in the whole world has begun to change. The pandemic has caused the transition to distance form of education for pupils/students. New professions appear on the labor market. Those professions that already exist are updated in such a way that the skills or knowledge acquired by a specialist become out-dated. Retraining should be quick and meaningful. Now the specialist needs clear knowledge, which he needs just today. In this way, demand and supply in the field of non-formal education were formulated.

Pupils, students and adults combine non-formal and formal education today. Massive open online courses (MOOC) are among the most in-demand forms of non-formal education. The first courses were created on the basis of educational institutions - the University of Prince Edward Island, Stanford, and others. Therefore, in the future, MOOC will be created with the support of universities or public organizations. Courses are becoming increasingly popular in Ukraine as well. The authors of this paper conducted research in 2021 on the awareness of lecturers and students about various courses available in the country, planning to further analyze the demand for courses by areas and offer students to develop a course according to certain criteria. However, the war has changed the direction of research. In the conditions of military operations, studying has become more difficult. Therefore, a survey was conducted in December 2022 - at the beginning of January 2023 in the same institutions that were used as a base for the questionnaire in 2021. And the results were compared.

A large number of studies and scientific publications are dedicated to the problem of MOOC. S.J. Blackmon, C.H. Major considered the issues of the essence of MOOC, their variations and offered their own vision of their typology [1]. U. Nazir, H. Davis, and L. Harris studied motivation for studying and indicators of MOOC courses completion by the participants [2]. Quantitative analysis of IT courses was developed by O. Semenikhina, A. Yurchenko, A. Sbruieva, A. Kuzminskyi, et al. [3;4].

The purpose of this article was to identify the current state of MOOC demand and analyze its efficiency in difficult conditions.

Research tasks:

- publication of the questionnaire results of 2021 and December 2022 - January 2023;
- identifying the dynamics of changing attitudes towards MOOC;
- experimental checking of the efficiency of MOOC courses completing in difficult conditions.

The authors conducted a survey at the beginning of the experiment. This provided information on attitude towards non-formal education and MOOC courses. The next step was a knowledge check in the discipline “Digital Technologies”. After that, students were recommended to take a course in which they acquired skills in working with text, electronic tables, and algorithmic structures. These skills extended studying on the university's course “Digital Technologies”. Students chose the course independently and agreed with the teacher. This was necessary so that the skills that were formed and subsequently evaluated would be approximately the same. If the student chose a course that did not have all the necessary materials and tasks, the teacher recommended several courses to choose from. The following stage was a control check and comparison with the results at the beginning of the experiment.

Therefore, this publication presents the results of the questionnaire, which demonstrate the awareness and attitude of teachers and students to MOOC. The questionnaire was conducted in 2021 and in December 2022 – January 2023. Changes taken place over the past 1.5 years were analyzed in the article. In addition, it should be taken into account that in 2022 the war started in Ukraine and the conditions of education became

difficult. Then the results of the experiment conducted between these two surveys were presented. Such an experiment demonstrated that MOOC courses are support of formal education in various conditions.

II. DYNAMICS OF ATTITUDES TO MASS OPEN COURSES ACCORDING TO QUESTIONNAIRE RESULTS FROM 2021 AND DECEMBER 2022 - JANUARY 2023

The questionnaire was conducted on the basis of Sumy State Pedagogical University named after A.S. Makarenko, Pavlo Tychyna Uman State Pedagogical University, Kremenets Regional Humanitarian and Pedagogical Academy named after Taras Shevchenko, schools of Sumy, Uman, Kremenets, Kyiv and urban villages. The survey was conducted using Google Forms. 205 respondents took part in 2021; there were 215 respondents in December 2022 - January 2023. Each questionnaire contained 15 questions. The number of students, lecturers and other specialists among them was distributed in the way shown in the diagram (Fig. 1).

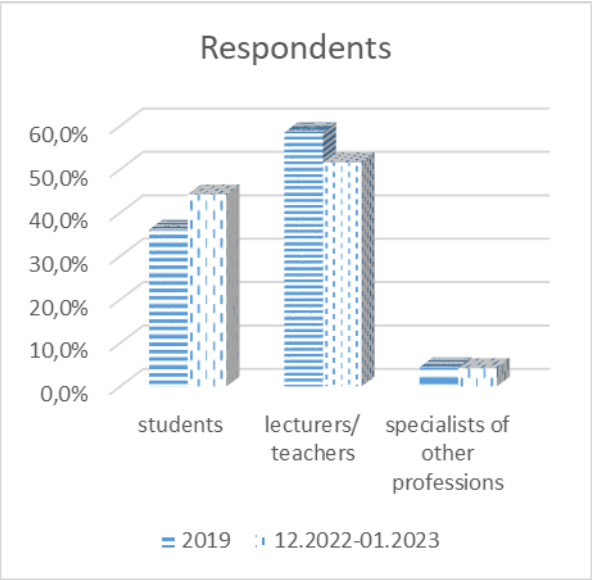


Figure 1. Distribution of the number of students, lecturers and other specialists taken part in the survey

The following answers to the question “Do you know what non-formal education is?” were offered:

- “yes, I completely understand”;
- “I have a rough idea”;
- “no, I don't know what it is”.

The number of respondents who fully understood the concept of non-formal education increased from 42.9% to 44.2%. The number of respondents who had a rough idea also increased by 47.9% (data of 2022, January 2023) compared to 44.9% (2021). The number of respondents who were not interested in non-formal education decreased from 12.2% to 7.9%.

One of the questions of this questionnaire was the following: “Has your attitude to studying changed since the start of Russia's war against Ukraine?” And 68.9% of the respondents (which were 144 people out of 215) answered that “now it is necessary to value the opportunity to study more”. 21.5% of the respondents did not change their attitude, 7.2% chose the answer that the main thing was “now it is not time for studying, you can study later”. The results of the answer to the question “Did you take courses in 2022 or are you taking them now?” were also indicative. 78.6% answered “Yes”, which means that 169 people out of 215 are engaged in self-development in difficult conditions.

The next block of questions revealed the opinion of the interviewees about massive open online courses. Prometheus, Diia. Digital Education, Coursera turned out to be the most famous for this category of respondents. These categories were chosen by the most respondents at both stages of the survey. In 2022-2023 survey, the number of respondents who learned about EdEra increased. Prometheus, Diia. Digital education, EdEra are Ukrainian platforms where mostly free courses in various fields are located. Paid courses are also offered among Prometheus MOOC.

Distribution of answers to the question “Which device do you use most often for studying and taking courses?” is shown below (Fig.2)

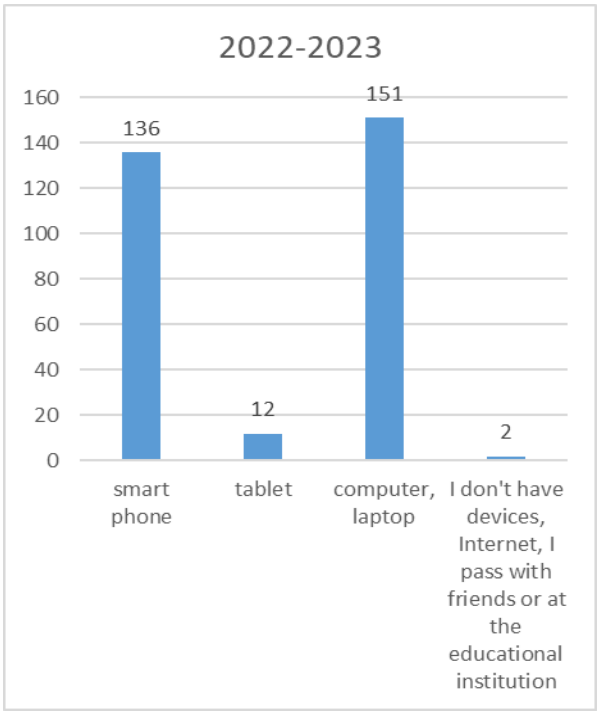


Figure 2. Use of different devices

Various cities and urban villages were chosen for the survey:

- Kyiv which is the capital of Ukraine with a population of more than 2.8 million,
- Sumy which is a regional city with a population of over 260 000,
- Uman which is an administrative center of Uman district with a population of about 82 thousand,

- Kremenets which is a small town in Ternopil region with a population of less than 21 000,
- small towns and urban villages of Sumy region and Uman district, the population of which could be no more than 3-5 thousand people.

Only 2 respondents out of 215 had problems with access to the device or network, and over 50% had several devices with the help of which they had the possibility to study. This fact gave grounds for asserting that the problem of access to courses in cities was not significant.

The vast majority chose “their own desire” and “self-education” as the reasons for taking the course: 54.5% (2022-January 2023) and 46.4% (2021). Other reasons included “a requirement/recommendation by the place of study or work”, “curiosity” and other reasons.

The second stage of the survey (2022 - January 2023) included questions about the form of education. 111 respondents (53%) considered it appropriate to return exclusively to full-time education under the right conditions, 77 participants (37%) considered it appropriate to remain blended learning, the rest could not decide. A total of 209 people answered this question. The question and the distribution of answers are shown in the figure (Fig. 3).

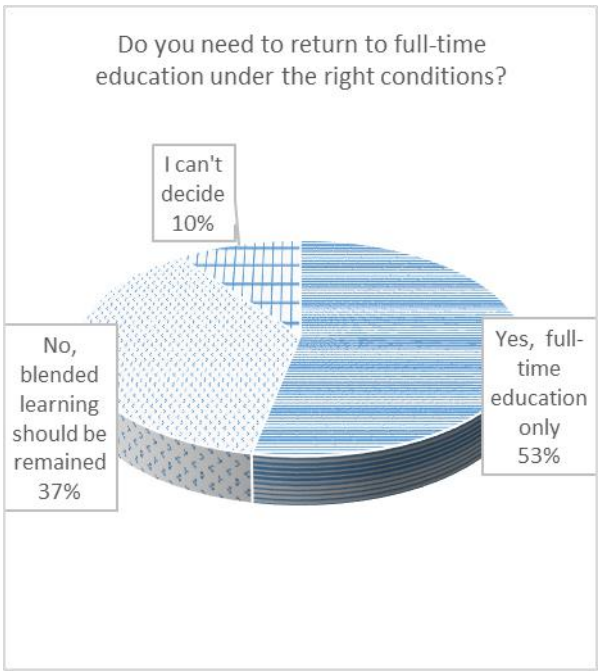


Figure 3. Choosing the form of studying by respondents

At the same time, 55.3% of respondents considered full-time studying to be the most effective, and 32.7% of respondents considered blended learning to be the most effective one. Others mentioned exclusively distance studying and self-education. No one chose “family education” option.

Thus, non-formal education is gradually taking its place in continuous lifelong education. Everything goes quite harmoniously when an adult expands his competences with the help of MOOC. It appears inconsistency in the case when a student has completed a certain course, acquired knowledge and skills, and has no desire to take the same course in the educational institution. . Differences between courses by rating were studied by Gamage D., Perera I., Fernando S. The authors concluded that courses with a higher rating do not differ in content, but such courses are easy to understand, with a good material presentation. Students choose courses based on their intuition. The authors note that some developers do not pay much attention to design or methodological issues when creating, but only improve courses in the future according to comments [5]. The lecturer must determine whether there is a correspondence in the theoretical material, or whether the skills that were expected as a result of studying have been formed. The lecturer must recognize or not recognize the results and justify it. And here the problems that need to be solved now are formulated:

- a clear coordinated process of integrating non-formal education into formal education is needed;
- a procedure for determining the criteria for taking into account knowledge and skills acquired in non-formal education when evaluating the results of studying is needed;
- the fact that determining the features of the choice of platforms and courses depends on the student is important, but the lecturer is not obliged to know the level of teaching in all courses of all platforms;
- there is a problem of psychological readiness of lecturers themselves to accept non-formal education of students;
- the student must be responsible for the quality of knowledge acquisition.

Everything that has been mentioned is also supported by the studies of other scientists. Lazarus Fc. and Suryasen. R. studied the problem of implementing MOOC in the education of developing countries. The authors, based on an in-depth analysis, note that the higher the penetration of MOOC, the greater will be the improvement in the quality of higher education [6]. Thus, the fact that the introduction of courses, in particular, and non-formal education in general is relevant in Ukraine, is an important factor in improving the quality of education.

III. DESCRIPTION OF THE RESEARCH AND ITS RESULTS

In this context, students - future computer science teachers were offered to take courses of their choice. Two groups at Sumy State Pedagogical University named after A.S. Makarenko and Pavlo Tychyna Uman State Pedagogical University to evaluate the efficiency of MOOC integration into formal activities were formed. Students of both groups had statistically the same studying results. The experimental group (EG) included 25 students. It consisted of 12 students of the 4th year of Bachelor training and 13 students of the 1st year of Master's degree. The control group (CG) consisted of 24 students, including 12 students of the 4th year of Bachelor

training and 12 students of the 1st year of Master's degree. Determination of efficiency was carried out based on the evaluation of theoretical knowledge, practical skills and reflective perception of the student's own result. Students chose a platform and a course by their own will that corresponded to the acquisition of skills in a specific discipline. The ability to work with text documents, tabular data, and algorithmic structures was evaluated. Control was carried out at the beginning and at the end of the course. It included testing from the theoretical part, performing of practical tasks. This part was evaluated by the lecturer. In addition, reflection was a separate criterion: a conversation and a survey of students regarding the level of awareness of their own knowledge. It should be noted that reflection indicators were a subjective evaluation of students. Surveys for Bachelor's and Master's students were different, according to their disciplines.

The results were expressed in the levels of high, sufficient, medium, and low. Obtained results were presented in Table 1,2.

TABLE I. RESULTS OF THE TASKS AT THE BEGINNING AND AT THE END OF THE EXPERIMENT (Experimental group)

Criterion	Levels	Initial evaluation		Final evaluation		Changes	
		Number	in %	Number	in %	Number	in %
Theoretical knowledge	High	2	8	5	20	3	12
	Sufficient	8	32	11	44	3	12
	Medium	10	40	8	32	-2	-8
	Low	5	20	1	4	-4	-16
Total:		***25***	***100***	***25***	***100***		
Practical skills	High	3	12	6	24	3	12
	Sufficient	8	32	12	48	4	16
	Medium	9	36	7	28	-2	-8
	Low	5	20	0	0	-5	-20
Total:		***25***	***100***	***25***	***100***		
Reflection	High	0	0	5	20	5	20
	Sufficient	10	40	12	48	2	8
	Medium	12	48	8	32	-4	-16
	Low	3	12	0	0	-3	-12
Total:		**25**	**100**	**25**	**100**		

TABLE II. RESULTS OF THE TASKS AT THE BEGINNING AND AT THE END OF THE EXPERIMENT (Control group)

Criterion	Levels	Initial evaluation		Final evaluation		Changes	
		Number	in %	Number	in %	Number	in %
Theoretical knowled	High	1	4,2	3	12,5	2	8,3
	Sufficient	8	33,3	9	37,5	1	4,2
	Medium	11	45,8	10	41,7	-1	-4,2
	Low	4	16,7	2	8,3	-2	-8,3
Total:		***24***	***100***	***24***	***100***		
Practical skills	High	2	8,3	3	12,5	1	4,2
	Sufficient	8	33,3	10	41,7	2	8,3
	Medium	10	41,7	10	41,7	0	0,0
	Low	4	16,7	1	4,2	-3	-12,5
Total:		***24***	***100, 0***	***24***	***100***		
Reflection	High	1	4,2	3	12,5	2	8,3
	Sufficient	10	41,7	10	41,7	0	0,0
	Medium	10	41,7	10	41,7	0	0,0
	Low	3	12,5	1	4,2	-2	-8,3
Total:		**24**	**100**	**24**	**100**		

Given data showed that the increase in the number of students in the experimental group who passed to a high and sufficient level and a decrease in their number in the medium and low level indicated the efficiency of the integration of non-formal education into formal one. Comparing changes that occurred in EG and CG also confirmed this fact. Such changes were more obvious for the experimental group.

We believe that experimental confirmation of the efficiency of integrating of non-formal education, namely MOOC, into formal one has taken place. After carrying out the experiment and analyzing obtained results, we can formulate the following recommendations. It is advisable to recommend courses for students to a lecturer who understands before the start of the course whether theoretical topics are included and whether practical tasks are proposed that are close to the planned ones. The lecturer clearly formulates the criteria for evaluating studying results. The lecturer, studying the courses, can expand his own competences, find more appropriate teaching methods, more successful examples. The lecturer must be ready for conflict situations. For example, a student completed a certain course without consulting the lecturer. Such a course has a related title, but the knowledge that the student gained there is lower in level. At the beginning of his discipline, the lecturer should present the syllabus to the students, where the requirements for the level of knowledge and the criteria for their evaluation are clearly described. The student must confirm the acquired knowledge and skills in the interview or by another evaluation method (testing, performing a practical task, etc.)

IV. CONCLUSION

Non-formal education is in demand. The reasons for the demand for mass open online courses are the lack of consistency of the proposed content of education in educational institutions with modern requirements for the competencies of graduate students [7]. Therefore, it gradually occupies its niche. The results of the research presented above confirm that many young people have a desire for self-development. At the same time, it does not

matter if they live in a big city or not. Access to online courses makes it possible for young people to develop in the direction they choose.

An online course cannot replace a comprehensive education. Education consists not only in the acquisition of skills and knowledge, but also in socialization, exchange of experience, formation of soft skills. Studying should be a social experience, not the interaction between a person and devices [8].

When integrating elements of non-formal education into the education process at the university, it is necessary for the teacher to control the process. He can determine the quality of material presentation and its relevance to the discipline.

The most appropriate element of non-formal education that can be implemented in institutions of higher education is massive open online courses. And this is explained by the fact that the conditions for each student may be different. In conditions of a pandemic, war or other unavoidable circumstances, access to institutions is limited for a certain time. And in order not to lose time and opportunities, it is important for a young person to continue to develop, to form his own competencies and experience.

REFERENCES

[1] SJ. Blackmon, and CH. Major, "Wherefore Art "Thou MOOC?: Defining Massive Open Online Courses,". Online Learning. vol. 21 (4) , pp.195-221, Dec 2017.

[2] U .Nazir, H. Davis, and L. Harris, "Input On Mooc Forums Is Dominated By Completers," Edulearn15: 7th International Conference On Education And New Learning Technologies. 2015. pp.4003-4009.

[3] A. Yurchenko, M. Drushlyak, S. Sapozhnykov, S. Teplytska, L. Koroliova, and O. Semenikhina, "Using online IT-industry courses in the computer sciences specialists' training," International Journal of Computer Science and Network Security. Vol. 21 No. 11. 2021. pp. 97-104. DOI: 10.22937/IJCSNS.2021.21.11.13

[4] A. Yurchenko, V. Shamonia, O.Udovychenko, R. Momot, and O. Semenikhina, "Improvement of teacher qualification in the field of computer animation: Training or master class?" 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021, pp. 631-635. DOI:10.23919/MIPRO52101.2021.9596946

[5] Gamage, D; Perera, I and Fernando, S. What do star rates for MOOCs tell you? An analysis of pedagogy and review rates to identify effective pedagogical model.Think Global Act Local, 2016 24th international conference on computers in education (icce 2016) , pp.300-302.

[6] Lazarus, FC; Suryasen, R. The quality of higher education through MOOC penetration and the role of academic libraries. Insights-The UKSJ Journal. May, 2022. P. 35.

[7] Dehtiarova,N., Hontar,O.& Vernydub,H.(2021). Attitude towards mass open online courses as a form of non-formal education. Physical andMathematical Education,32(6), 18-22.https://doi.org/10.31110/2413-1571-2021-032-6-003.

[8] Warren, J; Rixner, S; Greiner, J; Wong, S. Facilitating Human Interaction in an Online Programming Course. 45th ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE). 2014. pp.665-670.

Comparison of Pandemic and Post-Pandemic Education at Higher Education Computer Science Course

D. Paľová *, J. Štofa* and M. Vejačka*
* Department of Applied Mathematics and Business Informatics
Faculty of Economics, Technical University of Košice, Slovakia
dana.palova@tuke.sk, jan.stofa@tuke.sk, martin.vejacka@tuke.sk

***Abstract* - Many people see the period of COVID-19 pandemic as a challenge. This period causes many challenges in the teaching process. The post-COVID-19 period is also full of challenges. E-learning is an integrated part of classical education. It opened a new space for students and teachers. We provide the subject Informatics I at our faculty. We assumed that this subject is easy to pass, but the opposite is true. The proof is the repetition of the subject by a high number of students. Our goal is to investigate the reasons of failure on the form of teaching during and after the pandemic. Therefore, we approached students who participated in both forms of the educational process of this subject. We investigated the obstacles in education that they had to overcome during individual periods. We were interested in what they consider positive and negative in both forms of learning. These results will better help us design the structure and organization of the subject. This contribution is aimed at identifying problematic parts of the educational process.**

Keywords – comparison, education, pandemic, COVID-19

I. INTRODUCTION

As Albert Einstein once said: *"It is the supreme art of the teacher to awaken joy in creative expression and knowledge."* We, as university educators, are always trying to achieve this in our teaching process.

The education in the field of basics of informatics is provided at our faculty to students at the first year of study. If students are not successful in finishing the course, they might repeat it in next year of study. The rate of the repetition of this course is relatively very high. It is important to identify reasons why some of the students are not successful to pass the subject in the first semester of their university study. Whether it lies in students' attitude to the study (their insufficient preparation, problems to adapt to demanding education process etc.) or it is caused by inappropriate teaching method.

Also the COVID-19 pandemic didn´t help us with answers. Social distancing during the pandemic could also be the cause of students' failure and low level of knowledge to pass the exams. But when the face-to-face learning came back, and results were not any different. Multiple students took part in the course in both forms of education due their repetition of the course. So, students´ experience with both types of learning can be compared, as it might be interesting issue to investigate the reasons of failure in the course completion on the form of education during and after the pandemic. Therefore, we approached students who participated in both forms of the educational process of this subject.

II. COVID-19 PANDEMIC EDUCATION CHALLENGES AND CHANGES

The shift to distance learning has had a significant impact on students' lives, access to resources and technology, digital literacy, and mental health. Many world universities, as well as ours, closed on-campus residence halls, forcing many students to return home to their families and continue their education via distance learning. When learning from home, many students faced challenges, including family conflicts and obligations and lack of adequate resources. Without the right environment, students found it difficult to concentrate while learning. This affected student´s mental health [1], [2], [3]. Up to 42% of university students reported feelings of stress, depression, and anxiety while learning from home [4].

However, the essence is the realization of one fact. A successful transition to higher education requires that the learner learn how to adapt to educational requirements, i.e., how to become an effective learner. Effective learning occurs when the learning leads to the desired outcome. It as long been known that learner´s success, as evidenced by continued study, is influenced by student attributes and experiences combined with institutional factors. These attributes include prior educational achievement, family history, and personal ability, while institutional factors focus on performance during study and interactions with faculty [4], [5].

Because of the switch to distance learning, faculty and students have had to quickly learn how to use the new software for their courses. Also, the need for digital education platforms, internet access, and technological devices increased. Because computer science is also taught during normal teaching in the digital world, teachers were more confident and positive about the transition to distance learning than teachers in other disciplines [4], [6], [7].

Radić et al. also found that despite the change, students' grades were in the average range and students were very satisfied with the quality of instruction [8], [9]. In addition, flexibility and access to course materials during distance learning have been improved. While studies show that Computer Science Education has been influenced in some ways, it is unclear how these students' study habits have changed [1], [10].

Cooperative learning structures and processes increase academic achievement, enhance lifelong learning competences and develop the personal and social competences of every single learner in a more effective and fairer way when compared to traditional structures of learning in schools. In a traditional classroom structure there is usually one type of interaction, with the teacher the principal participant. The teacher makes a presentation, asks and answers questions, shows materials from the Internet, checks and discusses individual learning behaviors. These interactions are mainly conducted in the presence of passive listeners. How can a teacher deconstruct these traditional structures of interaction in a classroom practice to provide personal and individual access to learning interaction as many participants as possible during classes? That is the reason why we completed this principle with „personal involvement", which means that a cooperative teacher should provide interactions in which every single class member can participate with his/her full personality [11].

III. THE COURSE OF INFORMATICS I

The course of Informatics I is a compulsory subject at our faculty. Our students have majored in Economics but it is indisputable that, the basics of informatics are necessary for their future careers. The course Informatics I reflects the demands of the labor market and the situation on the market with information technologies and therefore it is always updated (similarly as the tools used in the course).

A. Course education goals and content

The aim of the subject is to provide students with basic knowledge, skills and competences based on knowledge of informatics and information technologies. The subject deals with the basics of information theory, the basics of ICT, the basics of algorithmizing and programming in a programming language R, the basics of networks and network communication, the security of data protection and the principles of network communication security, and finally, the possibilities of cloud solutions and their use in the business environment.

After completing the course, the student should achieve:

- *Knowledge* - basic terms in the field of information theory and ICT, terms in the field of computer hardware and software, algorithmizing and programming languages, cloud technology and cloud services, terms in the field of computer security and communication security through an open Internet environment.
- *Skills* – passing the course students achieve ability:
    - to safely protect personal data and face attacks from the Internet,
    - to create simple algorithms and programs in the currently taught programming language and algorithmically solve given problems,
    - to process data using selected ICT tools and programming languages,
    - to create a simple HTML page.
- *Competences* – the subject significantly develops logical and critical thinking, provides the ability to navigate in the environment of various ICT tools used and then use them to handle simple problem tasks. Through the defense of an independently prepared semester project, the student develops the ability of presentation, communication and argumentation. The subject is followed by almost all subjects taught in higher years of study.

The mixture of classical and alternative teaching methods is used during the course realization. During the semester, students work out problem assignments in the software currently being taken over, the emphasis is on self-study and then consultations on the problems that have arisen. During the semester, they develop a complex semester project on a given topic and, after handing it over, defend the given project to the teacher through a presentation. As exposure and fixation methods, we use: classic (frontal) and interactive lecture, explanation, repetition, consolidation and practice of the subject matter, as well as discussion during the exercises.

The nature of the subject requires methods of independent student work, work with information sources, self-study of relevant literature.

B. Changes done during the COVID-19 pandemic

COVID-19 attacked world in enormously fast pace. Healthcare professionals, politicians and also us teachers had to adapt in a short time and prepare for functioning in this period. Considering the difficulty of the situation we searched for options how to make it easier and upgrade the realized course to be more interesting for students while they miss social contact and motivation to study regularly, because of the shift to online education. Nobody of us was experienced with full distance learning in our courses, but we try to the best implementing our personal experiences from already attended online courses, webinars and our experience with part-time students who were educated this way before the COVID-19 pandemic. Besides the supportive study materials that were already published on course LMS Moodle page, we came up with necessary news within the course:

- *use of the videoconferencing system MS Teams* to replace missing face-to-face lessons that were held in PC labs. This system was used also to support fast message exchange between students and teachers, so provided us more flexible and supportive approach to students who identify any

kind of problem while they are studying particular topics.

- *different kinds of supportive materials* like video tutorials, e-books, and online activities that are aimed to enhance students learning regularly. One of example were extra activities that are optional, but they were concerned especially on further development of achieved skills and implementation them while they are solving the problem connected to real life. Materials and activities were published to our students on the Moodle course page. The developed video tutorials seemed to be the best option for help for those students that had some problems during our live online session. It provides the to study at their own pace at home. The videos consisted of commenting on the task development process. We also gave students option to participate in teaching by making those videos themselves. After creating the video, they received points for activity. That scoring system helped with better motivation of our students, and they also practiced and consolidated their knowledge.

These materials and activities were updated every year based on the expectation and problem identified by education process. Nowadays they are used as materials for the students, despite the fact that we already teach the subject in an original pre-covid regime, i.e. face-to-face in PC labs.

IV. RESEARCH ON INFORMATICS I COURSE

As was mentioned above, during COVID-19 we were forced to change our education style and also students need to change their learning style. We designed and developed a mass of study materials and different supportive assignments, exercises, etc. to help students successfully pass the course Informatics I. However, we have had students who have not been successful in their studies of this subject. These students attended thanks to the pandemic situation the course in two different modes – fully online and face-to-face learning. We are interested in their opinion on the education process of the course Informatics I. They were asked to compare both modes and define benefits and shortcomings that helped or didn't help in their study of the subject Informatics I.

After passing the course this academic year we asked students to answer the feedback questionnaire that contained the questions concerning the course content, developed study materials, e-books and video-tutorials. They were asked to identify the main issues that problems due to which they could not successfully finish the subject during the pandemic.

To answer the feedback were asked just students who were not successful last year and attended the course again this academic year. Their opinion is important because they can compare both modes of realizing the course – online and face-to-face and achieved outcomes help us to identify the problems and develop materials and implement changes that could be helpful independently in the pandemic situation. We collected data from 28 students who participated in both forms of the educational process of computer science. Using descriptive statistics, we analyzed the problems and came to conclusions that we will try to implement as much as possible in the creation of the course.

A. Data and methods

This is an online study based on a survey of 60 students who participated in both forms (Pandemic and post-pandemic education) of the educational process of this subject. The survey was optional, so unfortunately we received feedback just from 28 students, which represents almost half of the respondents.

To collect the information, an online survey was conducted from December 17 2022 to January 27, 2023. A link to a structural questionnaire (https://tinyurl.com/32zz4k23) created using Microsoft Forms was sent by email to students. It contained 56 questions (closed, opened) about the different components of the education process. The questions were compiled after studying scientific articles [1], [3] with a similar issue and then after a departmental discussion and an open discussion with students.

The simple percentage distribution was estimated to assess learning status, type of learning, opinion on educational decisions, and problems related to studying due to closure.

B. Results and discussion

Because during COVID-19 pandemic the lectures and seminars were shifted to online space, we were interested in the different aspects of online learning: students' need to be interactive during the lessons, the application that was used for the lecture or seminar, the way how the lectures and seminars were realized and content presented, etc. Because the lectures and seminars are realized by different methods, we decided to evaluate every part differently (i.e. online – lectures, seminars and face-to-face – lectures and seminars). Out of 60 survey questions we decided to present the most interesting and unexpected ones in next parts.

1) Online lectures

Of the students surveyed who took online courses, 64.3% took online courses daily, while 35.6% of them took online courses mostly. The lectures were mostly understandable for 71.4% of respondents. Even though 60.7% of students said that the content of the lectures corresponded to the tasks solved at the seminars, for 67.8% of them lectures did not help them solve the problems with the tasks at the seminars. Here the question arises, if the lectures are created in an interesting way, including the practical part of this subject, or is it just a theory that the students aren´t able to connect it with practice. We can deduce the answers from the questions aimed at the online seminars.

2) Online seminars

All of the students were attending seminars on weekly basis. The main problem that answers our question is that only 50% of the students regularly prepared for the seminars and 42.9% of them actually read the materials and instructions. Second problem is active involvement of

students during online seminars. Namely 78.5% of them were passive during the lesson (Fig. 1). We can´t expect high results with low preparation and a pinch of effort. Up to 10.7% of students for whom the content of seminar was not clear, did not try to solve it in any way, left it as it was. Fig. 1 illustrates these results.

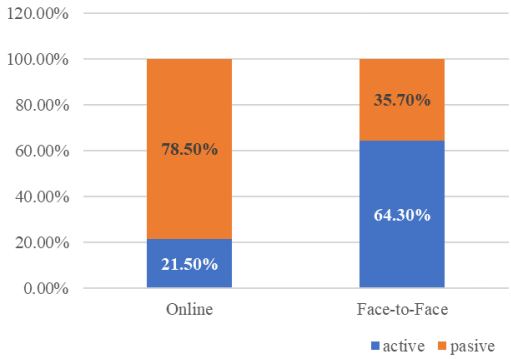


Figure 1. Students' activity during seminars

Communication is the basis of functioning. During online seminars, 64.3% of students felt a deficit of social interaction with classmates, which leads to stress and depressed emotions. In lectures, up to 57.2% (Fig. 2). In compared to West Bengal [4], where up to 42% of student felt stressed followed by the absence of a favorable environment to study at home. Up to 32,4% of their students had problems with internet connection. Even though our students have better conditions at their homes, as for the internet connection, they still felt more stressed and depressed. Still the main problem, due to our conditions, that can reflect the quality of seminars is the fact that 17.8% of our respondents addressed the frequent dropout of the internet connection. The use of cameras during online learning is another interesting capture. People are used to a kind of anonymity in the internet environment. They have freedom of speech, they can write where they want, what they want. This was reflected in online learning.

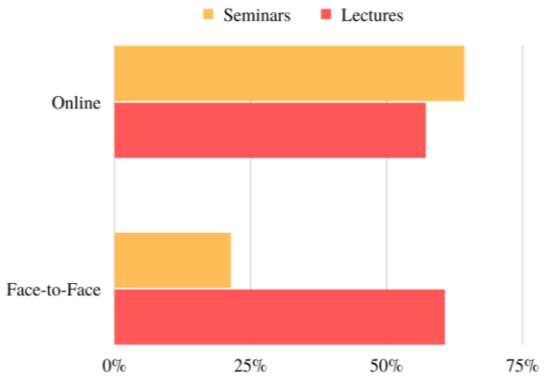


Figure 2. Deficit of social interanction during seminars and lessons

As our survey says, up to 67.9% of students had the camera off during online classes. It was due to discomfort feelings and shyness in front of classmates (Fig. 3).

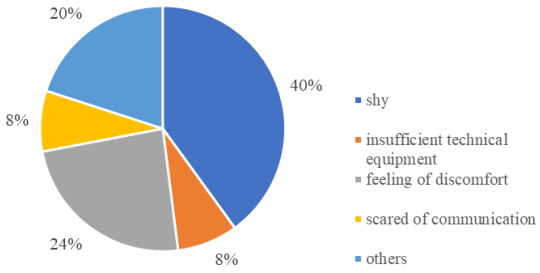


Figure 3. Reasons of not using camera during online lessons

3) Face-To-Face lectures

With the advent of face-to-face lectures, we can see a drop in attendance by 21.4%. The lectures were mostly understandable for only 42.8% respondents and 53.6% of them were confused. Even though 42.9% of students said that the content of the lectures corresponded to the tasks solved at the seminars, we can see a small improvement in lectures. From 67.8% online problems, now it is only 50% of them that lectures did not help them solve the problems with the tasks at the seminars. The question again arises. Why is that so? Did the lectures really improve or was there a problem in personal communication and motivation?

4) Face-To-Face seminars

Mostly all of students (96.5%) were attending seminars on the weekly basis. Yes, it is a bit lower number, but we can say that all of them were directly focused on learning. But still the main problem, which is regular preparation for the seminars, didn´t improve. We are on 50% of students that are preparing regularly and only 42.9% really read the materials and instructions. But a high shift for the better occurred in activity of the students. Up to 64.3% of students were active during seminars (Fig. 1). If something wasn´t clear, 82.2% of them asked teacher for help and 78.5% ask their classmate.

These results came in response to improvements in social interactions at seminars. Through the lectures only 39.3% of students were positively influence by social interaction with classmates and during seminars it was up to 78.6% of them (Fig. 2).

The question is, why less students were positively influence by social interaction at face-to-face lectures than online. Maybe we should focus more on interactive lectures using recommendations from students.

V. CHALLENGES AND RECOMMENDATIONS

A. For students

The most common recommendation was regular preparation for classes. Through our survey we find out that just a small percentage of students were really preparing for classes. Therefore, we are pleased to see students´ awareness of this problem. Also, in Slovakia we tend to say, *"Fortune favors the bold"*. Unfortunately, our students understood this only after completing the course. Because another common recommendation was for

students not being ashamed or scared of asking questions that they are not understand.

Very helpful for them was when they took notes in words, not only functions from the seminar. And the last very important is study hour. Which is alfa and omega at university and a lot of new students are not used to this style of teaching.

B. For teachers

Students perceived the high speed of learning and asked for video recordings of seminars. The video assignment solutions were therefore provided to students. Then teachers could focus on the problematic parts of the assignment and thus slow down the pace.

For better student motivation, they would like to have points from assignments added to the credit not to the exam. In most cases, successful achieving of credit tests is the main problem for students within the Informatics I course. This could be also good motivation in regularly preparing and working out assignments and it could help with better understanding the problematic of subject.

VI. CONCLUSION

As the online world is saturated with information, young people have difficulty distinguishing the importance of information, selecting the unnecessary and consolidating the necessary information. It is necessary to connect this Informatics I course more with practice, to show them the concrete use of the programs with which they work during their study within the course. A significant problem, that was captured by the questionnaire, was the impact of online teaching on the psychological health of students. The lack of direct interaction with classmates and teachers, feelings of loneliness and senselessness of what they do, what they learn, led to depressive states and stress which we analyze in part IV. All these aspects have been underlined by the COVID-19 situation. Our feedback outcomes correspond with the findings of other studies (e.g. [12], [13]).

However, the onset of the post-COVID-19 period did not bring the expected positive changes. The post-COVID-19 period was and still is marked by a tense situation in the world in relation to the war between Ukraine and Russia.

From the results of the responses to the open-ended questions, it can be concluded that in the post-COVID-19 period, with the start of face-to-face classes, there have been slight positive changes in terms of the success of the course Informatics I, despite the negative impact of the political-military world situation on the students' psyche. An important finding is that the face-to-face form of learning is more beneficial and satisfying to students than the online form. From the students' point of view, the problem with both forms of teaching was also the fast pace. It seems that another chapter of failure in passing this course, in both forms of instruction, is the indifference of the students and their lack of regular preparation. This course is taken by first year students. Therefore, failure can also be connected with unpreparedness for university studies, the basis of which are self-study, but also to the student's own inner motivation.

We evaluate this survey as useful for us teachers. It has helped us to think better about an unflattering trend that has accompanied us and the students in recent years: the repeated completion of the course Informatics I. The results of the questionnaire not only gave us a high information value about the Informatics I. teaching itself, but, what is very important in today's difficult times, it captured the issue of the influence of the online world and teaching in it and its impact on students' psychological health. These forces us to reevaluate the content (prioritizing quality over quantity) as well as the form of face-to-face teaching of Informatics I., in the hope that the future new approach to this subject will bring about positive changes in student outcomes.

REFERENCES

[1] I. YeckehZaare, G. Grot, I. Dimovski, K. Pollock and E. Fox. “Another Victim of COVID-19: Computer Science Education.” In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education. Volume 1 (SIGCSE 2022), Vol. 1. Association for Computing Machinery, New York, NY, USA, pp. 913–919. 2022. https://doi.org/10.1145/3478431.3499313.

[2] F. Ahmadon, H. I. M. Ghazalli and H. Rusli , “Studying during Pandemic: A Review of Issues from Online Learning in the Middle of COVID-19,” In 6th Conference on Interactive Digital Media (ICIDM). IEEE, pp. 1–4, 2020. https://doi.org/10.1109/ICIDM51048.2020.9339644.

[3] R. E. Baticulon, et al. “Barriers to online learning in the time of COVID-19: A national survey of medical students in the Philippines.” Medical science educator, 2021, pp. 1–12. https://doi.org/10.1007/s40670-021-01231-z.

[4] N. Kapasia, et al. “Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal,” Children and Youth Services Review, Volume 116, 105194, 2020. https://doi.org/10.1016/j.childyouth.2020.105194.

[5] V. Tinto, “Dropout from Higher Education: A Theoretical Synthesis of Recent Research,” Educational Research 45, 1 , pp. 89–125. https: //doi.org/10.3102%2F00346 543045001089.

[6] J. L. Gaudiot, and H. Kasahara, “Computer Education in the Age of COVID-19,” In Computer 53, 10, 2020, 114–118. https://doi.org/10.1109/MC.2020.3011277.

[7] T. Crick, T. Prickett and J. Bradnum, “Exploring Learner Resilience and Performance of First-Year Computer Science Undergraduate Students during the COVID-19 Pandemic.” Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1, 2022. https://doi.org/10.1145/ 3502718.3524764

[8] G. Radić, B. S. Ristić, S. Anđelić, V. Kuleto and M. Ilić, “E-Learning Experiences of Higher Education Institutions in The Republic of Serbia During Covid-19, Content Analysis and Case Study ITS Belgrade,” Proceedings of the 36th International Business Information Management Association (IBIMA), 4-5 November 2020, Granada, Spain. 2020, https://doi.org/10.3390/su132312935.

[9] E. Riese and V. Kann, “Computer Science Majors’ Experiences of Their Distance Education Caused by the COVID-19 Pandemic, “ In 2021 IEEE Global Engineering Education Conference (EDUCON), 2021., IEEE, pp. 393–397. https://doi.org/10.1109/EDUCON46332.2021.9454105.

[10] W. Simão de Deus, “Emergency Remote Computer Science Education in Brazil during the COVID-19 pandemic: Impacts and Strategies,“ Revista Brasileira de Informática na Educação 28, 2020, pp. 1032–1059. http://dx.doi.org/10.5753/rbie.2020.28.0.1032.

[11] F. Arató. “Towards a Complex Model of Cooperative Learning”. in Da Investigaçãoàs Práticas, Vol.3, No.1, 2013, pp. 57-79.

Available at: https://www.researchgate.net/publication/322683280_Towards_a_Complex_Model_of_Cooperative_Learning [Accessed at 24-Jan-2023]

[12] S. Hammerstein, Ch. Konig, T. Dreisorner, A.Frey, "Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review" Frontiers in Psychology, https://doi.org/10.3389/fpsyg.2021.746289

[13] A. A. Christ, V. Capon-Sieber, U. Grob, A.K. Praetorius, "Learning processes and their mediating role between teaching quality and student achievement: A systematic review", Studies in Educational Evaluation, Volume 75, 2022,101209, ISSN 0191-491X, https://doi.org/10.1016/j.stueduc.2022.101209.

AI Comics as Art: Scientific Analysis of the Multimedia Content of AI Comics in Education

PhD Krunoslav Bedi
School of Civil Engineering, Art and Craft Čakovec
Športska 1, HR-40000 Čakovec, Croatia
bedi.krunoslav@gmail.com

Abstract - This paper examines the advantages and disadvantages of using AI comics as a tool for more effective artistic expression among students. The paper includes a detailed description of the production of AI comics. In addition to the visual, the study also focuses on drawing skills and the presentation of AI comics as an active agent in media creation - not only with content, but also with a careful application of form features. This type of process complements the working methodology that enhances expertise, skills, and knowledge and promotes critical and creative thinking in students. The collaborative processes often used in the educational process have been partially replaced by individualized work with AI comics. Of course, this does not relegate the learning outcomes to the background, but rather incorporates them into the design. In this way, students can develop a wide range of creative activities specific to the medium of AI comics: narrative construction (bridging the virtual world and reality), temporality, space, synergy of text and drawing, organizational skills, publishing, digitization, etc. In addition to the actual creation of the AI comic, learning about storytelling and project work brought students closer to understanding the art of the AI comic.

Keywords: AI, AI comics, multimedia, education, art, design, creativity, artificial intelligence

I. INTRODUCTION

It is human nature to strive for innovation and effective adaptation. For as long as humans have existed, we have created new tools from our environment to better meet our needs. This has been a consistent trend in human development over the centuries, and now it is more than evident as we achieve better and more interdependent connectivity with the advent of computer technology. The whole world is interconnected, concepts and ideas are shared instantly, and perspectives are broader than ever - in an increasingly global society. This has enabled a renaissance of ideas and innovation in various fields, with creative people at the forefront of advanced technologies aimed at improving humanity and its relationship with technology.

Among the significant concepts in modern society are artificial intelligence (AI) and its applicability to almost all areas of life. The effectiveness of artificial intelligence is increasing at an alarming rate. It can provide a simpler solution to a wide range of tasks, from handling everyday routine tasks to producing creative texts, art, and music. Moreover, the use of artificial intelligence and machine learning is currently the best method for collecting and analyzing data and increases the efficiency of certain activities that require massive data processing, such as weather forecasting, business analysis, face recognition, and others.

The aim of this paper is to provide an informative discourse on current trends, technologies, and innovations in education that can utilize AI technology in their field, and to show how it can promote progress in the arts in various fields in combination with multimedia technologies.

II. AI TEACHING IN CLASSROOM

New AI technologies and artificial intelligence management languages are increasingly shaping the field of contemporary art, raising a number of questions about the present and future of artistic creation. This new collaboration between humans and machines is not only democratizing artistic production, but also questioning its very essence: authorship, originality, and creativity in education.

The implication that artificial intelligence will reduce the number of jobs for artists working today in fields such as illustration, animation, and graphic design is overshadowed by the counter question - what is the value of this activity, and can teaching respond to the needs of the labor market, given the historical lens of the post-industrial revolution and the needs of the knowledge market.

Modern educators have a variety of pedagogical strategies to foster students' AI literacy through inquiry-based learning. The use of AI is an effective pedagogical approach based on the exploration and improvement of various language skills and technological achievements. By reflecting on their experiences with AI creations, students understand the importance of its application, but also experience the immediate impact of AI in the workplace and the ethics of its use through various learning activities.

Previous research [1], [2], [3], [4] has shown that students can create a meaningful scenario and implement appropriate new and modern solutions. Applying this

knowledge encourages students to solve real-world problems and not just know concepts, but connect them.

AI was often described as an anthropomorphic technology that possesses cognitive qualities equivalent to those of humans—a conception that notably resembles how AI is portrayed in the media. As a pedagogical implication, findings suggest that it would be valuable to "demystify" AI by exploring its technical principles (i.e., the role of data) of the "human-like" AI solutions students encounter in their everyday lives. Artificial intelligence (AI) competence is recognized as a new competence using comics as an example [5].

The growing number of research initiatives and creative applications that emerge in the intersection of AI and art, motivates to examine and discuss the creative and explorative potentials of AI technologies in the context of art [6].

III. AI LITERACY

Nowadays, the concept of digital literacy in the field of artificial intelligence, which includes skills, competencies, and abilities for its use, is increasingly differentiated.

The first use of the term "Artificial Intelligence Literacy" was in the work of Burgsteiner et al. [7] and Kandlhofer et al. [8] refers to the ability to understand the fundamental knowledge and concepts behind AI-driven technologies. Along with knowledge and ethical use of AI, AI literacy serves as a set of competencies that enable individuals to critically evaluate AI technologies and effectively communicate and collaborate with artificial intelligence [9].

Artificial intelligence (AI) is similar to traditional literacy, such as digital literacy, and should become a new literacy skill in response to this new era of intelligence [10].

On first sight, coming with a definition for "AI arts" does not sound hard. AI (an abbreviation for the term Artificial Intelligence) refers to computers being able to perform many human-like cognitive tasks, such as playing games of chess and Go, recognizing content in images, translating between languages, selecting best candidates in a job search based on their CVs, and so on. This is how AI has been traditionally understood, and we can extend this concept to the arts. Following this logic, "AI arts" would refer to humans programing computers to create with a significant degree of autonomy new artifacts or experiences that professional members of the art world recognize as belonging to "contemporary art" [11].

AI is becoming an essential technological skill, and teachers should equip students with modern skills to live and work with, and the use of AI should definitely begin at the secondary level. Therefore, it is appropriate to think about working with AI and applying such pedagogy in secondary education. As AI has become one of the most important technological skills of the 21st century, teachers must inevitably incorporate AI into students' education and foster their thinking to train them for the world in which they will live, learn, and work. In this context, it is important to point out that a work of art is not created simply by putting pen to paper.

The potential awareness that behind everything there may be potential ethical problems in the use of artificial intelligence will certainly disappear with time and the definition of "standards" for its use.

IV. COMIC

Comics as a media product are familiar to almost all students. Using comics, children transfer an idea onto paper (or another medium) and elaborate a variety of picture elements in the most appropriate way.

Comics are examined as a teaching and learning technique by combining media activities and art.

Comics use sequential images to convey the action and speech windows (speech bubbles) for character dialog, as well as special captions to clarify and reinforce the action. Comics are a multimodal medium; for example, a close-up of a tear can illustrate a character's emotions, or graphic elements can indicate danger, such as an explosion.

In doing so, students form qualitative judgments when selecting a topic and an appropriate method for conveying information.

Comics are an entertaining and familiar medium for storytellers and readers. Today, however, this type of visualization medium is underutilized as a medium of expression.

Comics can be seen as a central medium in the accelerated 'convergence culture" that sits between traditional literature and film. Adaptations of novels, poems, and even songs have become an important part of the so-called 'graphic novel' or 'graphic literature' [12].

V. CREATION OF COMICS

In addition to the visual, it also deals with drawing skills and the presentation of comics as an active process of media creation - not only with content, but also with careful application of formal features. Often, collaborative processes in the educational process are partially replaced by individualized work on comics. This, of course, does not relegate learning outcomes to the background. On the contrary, this type of process complements the work methodology that enhances subject knowledge, skills and knowledge, and critical and creative thinking. Design is undoubtedly integrated into the learning outcomes.

Students developed a wide range of creative activities specific to the comic book medium: narrative construction, temporality, space, text/drawing/synergy, organizational skills, publishing, digitization, etc.

The constant use of existing knowledge (experiences, history, facts) and the creation of their own narratives made it possible to build a bridge between the virtual world and reality.

Besides the actual creation, learning storytelling and project work brought students closer to the art of comics.

VI. OBJECTIVES

The goal of developing comics in digital and physical media is to broaden children's horizons to familiarize them with different aspects of pictorial elements and to achieve further generalization and generalization of concepts and ideas that can be transferred to "paper."

After receiving the task, the process was initiated in the following way:

1. elaboration of the idea and content
2. focus on the way of implementation
3. conceptual sketch of the media content
4. drawing on the computer and design of the cover
5. printing (if necessary)
6. analysis of the content

This process specified the way in which comics create meaning. The evolution of the creation of the comic was presented in the following steps: Moment, Frame, Image, Word, and Flow.

VII. WORKING METHODS

Pupils/students used sequential images to represent actions and processes, speech bubbles (in our example there were none, only text) to represent dialog between characters, and/or text headings to provide additional detail and explanation. By drawing comics, they had a more detailed elaboration of the illustration and activities than is the case when processing only one image. Sequential-temporal procedures were used in the description of attitudes and emotions. The AI comic was created on the website https://www.craiyon.com/ in the form of the former DALL-E Mini.

Taking into account prior knowledge about comics, we can distinguish four basic steps of recognizing activities:

- drawing activities,
- examination of technical procedures,
- layout of the drawing and
- emotions and attitudes.

In this way, qualitative overviews were made by category and an acceptable method of transferring information was found using other programs for digital processing (supplementing the image generated by the AI with text).

VIII. PRESENTATION OF THE COMIC

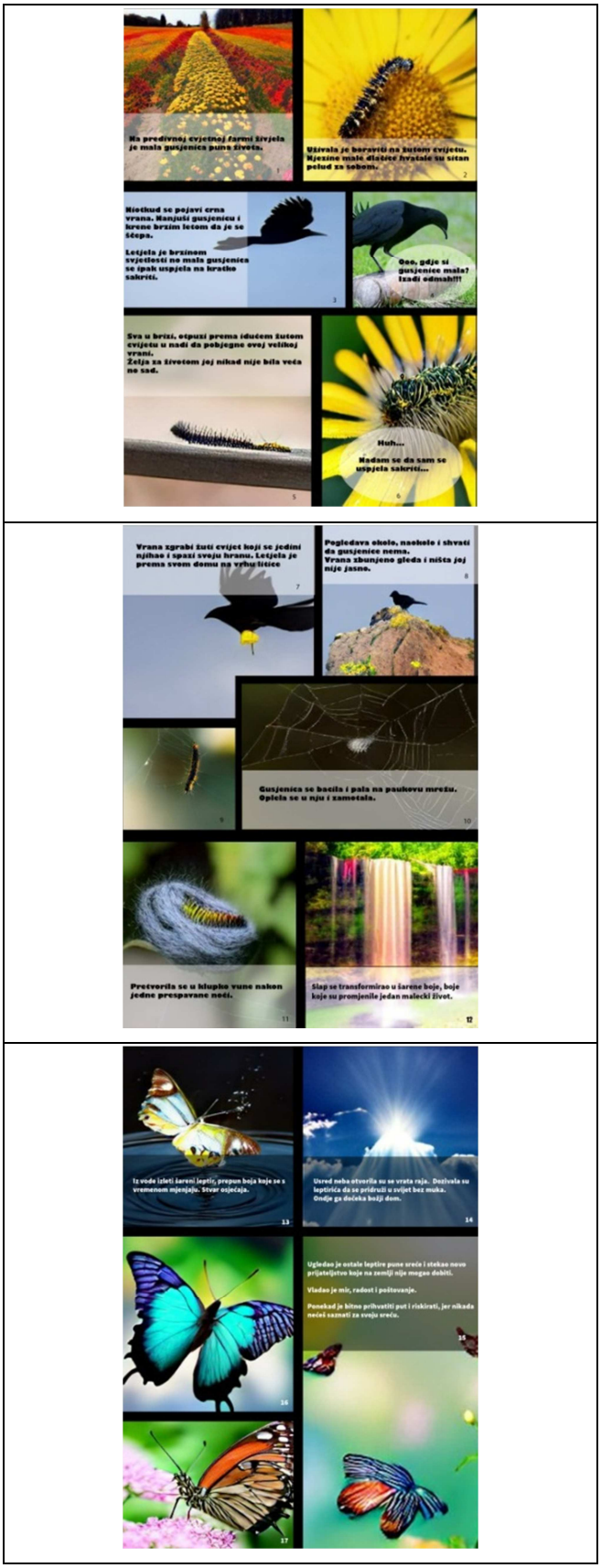


Figure 1. AI comics

IX. CONCLUSION

Artificial intelligence image creation technology is currently gaining traction in a wide variety of scientific, professional, and artistic circles. AI enables artists to inspire and increase the creative lane, i.e., it combines personal self-expression and collective exploration. AI encourages artists to expand and enrich expressive possibilities through innovative digital art.

If we are to take AI image creation seriously, we must create a standard for citing our credentials. Much like writing a scientific paper, a cited page/author is necessary to legitimize the work.

The answer is not to stop technological progress. Instead, we need to improve the way we integrate it into our world now.

REFERENCE

[1] Wallner, L., & Barajas, K. E., Using comics and graphic novels in K-9 education: An integrative research review. *Studies in Comics*, *11*(1), 37-54., 2020.

[2] Akcanca, N., An alternative teaching tool in science education: Educational comics. *International Online Journal of Education and Teaching*, *7*(4), 1550-1570., 2020.

[3] Sentürk, M., & Simsek, U., Educational comics and educational cartoons as teaching material in the social studies course. *African Educational Research Journal*, *9*(2), 515-525., 2021.

[4] Matuk, C., Hurwich, T., Spiegel, A., & Diamond, J., How do teachers use comics to promote engagement, equity, and diversity in science classrooms?. *Research in Science Education*, *51*, 685-732., 2021.

[5] Mertala, P., Fagerlund, J., & Calderon, O., Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence (AI) and their implications for AI literacy education. *Computers and Education: Artificial Intelligence*, *3*, 100095., 2022.

[6] Cetinic, E., & She, J., Understanding and creating art with AI: review and outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *18*(2), 1-22., 2022.

[7] Burgsteiner, H., Kandlhofer, M. and Steinbauer, G, "Irobot: Teaching the basics of artificial intelligence in high schools". In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 30, No. 1), 2016.

[8] Kandlhofer, M., Steinbauer, G., Hirschmugl-Gaisch, S. and Huber, P., "Artificial intelligence and computer science in education: From kindergarten to university". In 2016 IEEE Frontiers in Education Conference (FIE) (pp. 1-9). IEEE, 2016.

[9] Long, D. and Magerko, B., "What is AI literacy? Competencies and design considerations". In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-16), 2020.

[10] Ng, D. T. K., Luo, W., Chan, H. M. Y. and Chu, S. K. W., "Using digital story writing as a pedagogy to develop AI literacy among primary students", *Computers and Education: Artificial Intelligence*, *3*, 100054, 2022.

[11] Manovich, L., Defining AI arts: Three proposals. *AI and dialog of cultures" exhibition catalog. Saint-Petersburg: Hermitage Museum*., 2019.

[12] Peña, L. V. Á., "Adaptación de textos literarios a literatura dibujada". El historietista como traductor. *Estudios de Traducción*, *11*, 33-42., 2021.

Teenagers' Gaming Practices and Their Performance in STEM Courses

M. Homen* and V. Juričić**
* Faculty of Education, Zagreb, Croatia
** Faculty of Humanities and Social Sciences, Zagreb, Croatia
maja.homen@ufzg.hr, vjuricic@ffzg.hr

Abstract - **Millions of teenagers worldwide choose digital games as their primary entertainment. Except amusement, gaming environments enhance the capacity of learning, as it has been proved by many studies. Gaming is the only activity that shows a constant positive coefficient between PISA test scores and intensity of use. This paper evaluates gaming habits in teenagers and analyzes the connection between students' performance in Science, Technology, Engineering and Math (STEM) courses and playing digital games. For the purpose of this paper, we surveyed 6th, 7th and 8th grade Croatian primary school students and evaluated their gaming culture. We analyzed their play time, genre of games they play, complexity of games and used platforms. The aim was also to gather data regarding their self-perceived gamer identity, i.e. do they feel like gamers, do they play alone or with others and their general attitude towards games. We compared the data gathered with students' grades in STEM related courses and analyzed the relationship of each variable separately to identify those who have the impact on their performance. Our results confirm that there is a partial positive relationship between increased gamer identity and success in STEM related courses.**

Keywords – gaming; gamer identity; STEM; students' grades

I. INTRODUCTION

Digital games are a source of primary entertainment for millions of children and teenagers worldwide. The amount of time spent playing is rapidly increasing, from 5.1h/week in 2011 over 6.5h/week in 2017 [1] to 12h/week in 2021 [2]. It has been shown that 89% of children (3-11 years) in Italy were video gamers, of which 19% are categorized as problematic. Although preferences differ by sex, children mostly play first-person shooters (31%), sandbox (14%) and management (11%) games [2].

Except for the obvious primary amusement factor, video gaming also has many secondary consequences. Numerous studies have proved the benefits of gaming, particularly improved spatial cognitive benefits, strategic problem solving, enhanced short-term memory, pattern recognition, focus and multitasking [4]. In 2009 Reeves, Brown, and Laurier [4] has shown that even occasional and amateur first-person-shooter players show improved reaction times, synchronizing movements and actions, sense of opponents' movements and team plays. Consequently, professional and competitive gamers have even greater skills. Because of their capabilities, they can even be included in solving problems that could not be solved by traditional or scientific methods. For example, gamers [5] discovered the correct molecular structure essential for designing drugs by playing "protein playing game", where other attempts had failed.

Numerous prejudices and stereotypes were and still are associated with gaming and gamers. For years gamers have been labeled as antisocial, childish and unsuccessful, and gaming was perceived as the notable cause of children's physical and psychical health problems. It is not so simple and straightforward, as children are influenced by many other factors at the same time. However, it is proven that excessive time in playing video games carries certain risk for the development of behavioral deviations. Internet Gaming Disorder and Gaming Disorder are recognized disorders and included in official psychological books and classifications [6][7].

Beside the previously mentioned effects, gaming can also be viewed from cultural, social and educational perspectives, which is a reason that video games have lately attracted serious academic attention. Recent advances in ICT have provided schools and educational institutions in general a possibility for applying new means of knowledge transfers, reviews and tests. Many studies analyze the effect of using games in the teaching process, particularly their influence on learners' mood, motivation, knowledge and grades or success. The whole new game category appeared - serious games are designed for an educational, rather than an entertaining purpose [8]. Gaming shows a constant positive correlation with PISA test scores, as proven by Biagi & Loi [9], Gubbels, Swart & Groen [10], Kong, Seo & Zhai [11] and many others.

Digital games expand target audience, providing a platform to reach nontraditional learners [12]. They are used to build academic skills in a positive way, because they (can) provide personalized learning material and instructions, learning alone, in groups or (competitive) teams and work collaboratively with different types of people, in different environments and learning conditions. They can develop critical thinking skills, especially when faced with different real-life concepts and situations, requiring them to play a certain role or interact with objects.

As gaming has become an important part of children's and teenagers' lives, many of them become members of cliques or gaming communities. The term gamer describes

a person who plays video games, but this is a different concept from a gamer identity, i.e., from a person who feels or thinks that he is a gamer. “How people identify as gamers, is a different question from who counts as a gamer” [13]. Griffiths, Moore, and Richardson 2007 [14] explored and quantified gamer identity through a series of interviews. Authors examine the relationship to gamer identity of the tested players, how players interact with game texts (identification and interaction), and their thoughts about media representation. Their research shows that objective metrics such as overall time spent playing, frequency of play etc. are not always correlated with one’s perception as a gamer.

One of the tools used for quantification of gamer motivations is Video game uses and gratifications instrument (VGUGI) [15]. VGUGI is a self-report assessment of different motivations underlying video game play. Gamer Identity Strength (GIS) has been introduced by Neys, Jansz, and Tan 2014 [16], which is based on social identity theory three groups of gamers were defined: hardcore, heavy and casual. Authors confirmed the self-reported labels by gamers through investigation of time and frequency played, but also other reasons given for a particular GIS label through classification using systematic content analysis. The central aim of research by De Grove, Courtois, and Van Looy 2015 [17] was to identify relevant indicators of why players attribute the category of gamer to themselves or to others. They presented a model for measuring gamer identity comprising factors such as frequency of play, degree to which gamer identity is attributed to friends, genre impact, frequency of game-talk etc.

A recent study by Howe, Livingston, and Lee 2019 [18] examined the hours spent in the game, the genre of games and the medium on which the players played, and it was concluded that there is a statistically significant difference between men and women, where men feel like gamers more often and spend more time playing games. Also, gamers who spent more time in the week playing games are also more likely to feel like gamers. As for the game genres themselves, the most gamers were found to be among people who play first-person shooters, role-playing games and sports games, which is a shift from previous research in which people who felt like gamers most often played MMORPGs.

Study by Dindar [19] examined the gender-based disparities in the video gaming habits of 479 high school students as well as the connections between video gaming, academic success, and complex problem-solving abilities (CPS). The characteristics of playing video games alone versus with a team, game experience, game time, game frequency, perceived gaming skills, and game type were all measured. CPS was assessed using data from the 2012 Creative Problem Solving section of the Programme for International Student Assessment (PISA). The results indicated that the sampled males spent more time playing video games than their female counterparts and had better expertise and skill in the field. Any correlation between the video game factors examined in the study and CPS or academic success was not found to be of practical importance. The resulting findings show that the transfer of gaming skills to real-world problem solving may not be as straightforward as is claimed in the literature and add to the sparse empirical evidence on the association between video gaming and CPS.

Smith & Mcgregor [20] define collaborative learning as „an umbrella term for a variety of educational approaches involving joint intellectual effort by students, or students and teachers together”. It usually involves students working in groups of two or more, allowing them to pool their skills and knowledge. Cooperative learning represents the most carefully structured end of the collaborative learning continuum. Collaborative learning techniques represent a significant shift from the traditional teacher or lecture centered contexts in the classroom.

Baek and Achraf [21] described an empirical study that examined gender differences in cooperative and collaborative social gaming in connection to accomplishments and attitudes. Additionally, it compared the learning outcomes achieved by students in cooperative versus collaborative digital game-based learning environments. In this study, 164 sixth graders from five separate classrooms at an elementary school in South Korea actively participated. Findings shows that male students exhibit more positive game attitudes in collaborative settings, whereas female students exhibit more positive game attitudes in cooperative settings, based on interaction effects. Furthermore, data indicated that regardless of the circumstances surrounding grouping, male students express higher favorable attitudes toward group work than female students. In terms of academic and gaming success, female students shown greater academic success under cooperative conditions, whilst male students demonstrated greater academic success under collaborative situations. According to the study's findings, gender-balanced groups do much better in collaborative gaming than cooperative gaming. According to research by Gupta, Manju, and Pooja [22], students who were taught using the cooperative learning strategies TAI and STAD outperformed the control group on the post-test, clearly demonstrating the superiority of cooperative learning over traditional methods of instruction. As a result, when compared to traditional teaching methods, cooperative learning was proven to be a more successful instructional paradigm for mathematics.

II. METHODOLOGY

The aim of this paper was to evaluate gaming habits in teenagers; their play time, genre of games they play, complexity of games [23] and used platforms. In addition, information about the participants' self-perceived gamer identities, including whether they play alone or with others and how they feel about games in general, was sought. To determine the factors that affect students' performance, we compared the data collected with their grades in STEM-related courses and examined the relationships between each variable separately. Three hypotheses were derived:

1. There is no statistically significant difference between students who feel like gamers and those who do not feel like gamers and their grades in STEM related subjects.

2. There is no statistically significant difference in students' preference for more complex games (role-playing games, strategy games, action games and first-person shooter games) and their grades in STEM related subjects.
3. There is no statistically significant difference between female and male respondents and their self-perceived gamer identity.

The questionnaire was a paper-and-pencil instrument administrated to participants at schools after obtaining the written consent of participants' parents. The participants were given all the instructions necessary for filling in the questionnaires. Participation in the survey was voluntary and anonymous.

The questionnaire was distributed between 286 13-15 years old students in 6st, 7th and 8th grade of primary education schools in Croatia in Zagreb, Čakovec and Petrinja. The first part of the questionnaire collected their demographic data as well as their grades in STEM related subjects; mathematics, chemistry, biology, physics and informatics. The participants also stated how many hours per week they play, which platforms they use as well as and which game genres they prefer.

The second part of the questionnaire consisted of 22 statements regarding the participant's self-perceived gamer identity. Participants specified their level of agreement with each statement on a five-level Likert scale: 1 – Strongly disagree; 2 – Disagree; 3 – Neither agree nor disagree, 4 – Agree, and 5 – Strongly agree.

III. RESULTS

Participants' demographic characteristics with regard to gender and age was following: Female participants: 50,/%, male participants: 49,3%; 6th grade: 40,9%, 7th grade: 33,6% and 8th grade: 25,5%.

Table 1. shows the average participants' marks in STEM subjects. It can be seen that there is a high average mark in individual subjects as well as the average success of students in all STEM related subjects. Because of that, it would be hard or unsuccessful to perform the analysis and correlations between students' success and other observed variables. So, the participants' success was recoded; they were given a lower success rate for each subject if they have a grade below the sample average, and a higher success rate if they have a grade above the sample average. Therefore, regarding success in STEM subjects, the respondents were divided into two groups, with lower and higher success.

TABLE 1. – MEAN VALUES OF STEM SUBJECTS

Independent variable	Mean
Informatics	4,86
Science (6th grade only)	4,28
Biology	4,17
Mathematics	3,95
Chemistry	3,82
Physics	3,79
Overall achievement in STEM subjects	4,2544

Self-perceived gamer identity was measured with a series of statements such as *I feel like a gamer*, *My friends see me as a gamer*, *Playing games is cool*, *I could not do without playing games*, *I believe that the knowledge I learned in games was useful in class*, *When I play with friends in a group, it is important for me to cooperate in order to win*, *I post my gaming endeavors on social networks (Facebook, YouTube, etc.)*, *I follow famous gamers on Youtube and other platforms* and similar where the participants can show how important gaming is in their lives.

The results show that more than 30% of participants play between 5 and 10 hours a week (Chart 1.), but also that about 25% of them play less than 1 hour a week and possible do not play at all, which is quite unexpected..

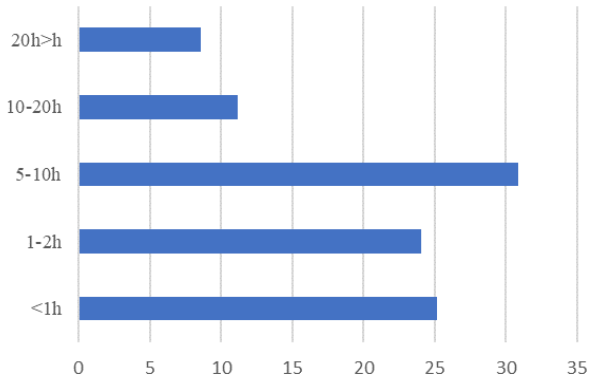


Chart 1. Hours spent playing games per week

Regarding devices they use for playing games the participants could choose one or more devices and the results showed that the majority of students use Smartphones followed by PC, while Wii and Xbox did not prove to be popular devices for playing videogames (Chart 2.)

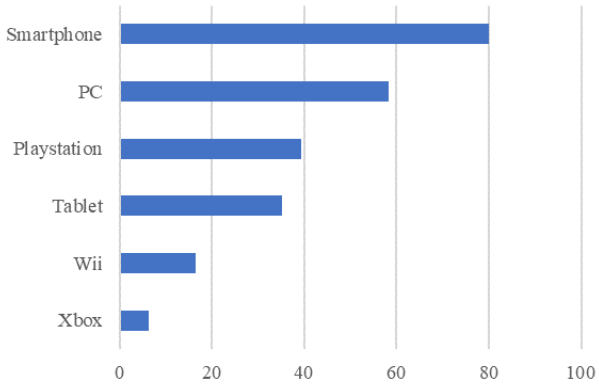


Chart 2. Devices used for playing games

Chart 3. shows the popularity of game genres among participants. The most popular genres seem to be Educational games (like Učilica, Sebran, Glaskalica and other ICT-AAC games) and Role playing games (like Witcher and World of Warcraft). Logical games (like Putt Putt and World of Goo), Simple online games (like Farmville) and Adventure games (like Life is Strange, Sherlock Holmes and Broken Sword) are approximately equally represented. Strategy games (like League of

Legends, Minecraft and Starcraft) and First-person shooter games (like Counter strike and Call of Duty) are played rarely.

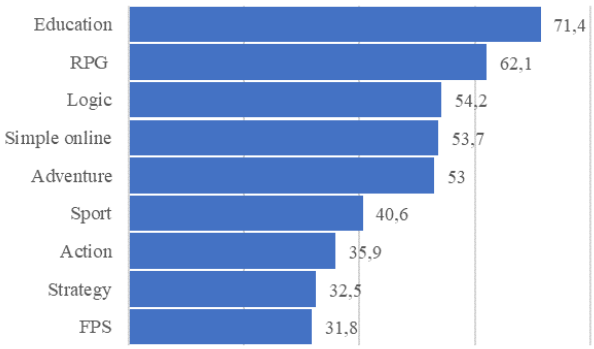


Chart 3. Percentage of game genres played by participants.

Regarding the **first hypothesis:** "There is no statistically significant difference between students who feel like gamers and those who do not feel like gamers and their grades in STEM related subjects" the results of the Mann-Whitney U (MWU) test of differences in the perception of oneself as a gamer showed that respondents with a lower level of success in STEM subjects attach more importance to cooperation with friends in order to win and that they more often put their gamer endeavors on social networks (Table 2.)

TABLE 2. DIFFERENCES IN SELF-PERCETION AS A GAMER WITH REGARD TO SUCCESS IN STEM SUBJECTS

When I play with friends in a group, it is important for me to cooperate in order to win.	Lower level	155,36	MWU	8451,5
			Wilcoxon W	23329,5
	Higher level	135,64	Z	-2,184
			Asymp. Sig. (2-tailed)	0,029
I post my gaming endeavors on social networks (Facebook, YouTube, etc.).	Lower level	151,18	MWU	8435,5
			Wilcoxon W	22970,5
	Higher level	135,12	Z	-2,015
			Asymp. Sig. (2-tailed)	0,044

The results also show that students with a higher level of success in the subject of physics statistically significantly more believe that the knowledge they learned in games was useful in class (Table 3.).

TABLE 3. DIFFERENCES IN SELF-PERCETION AS A GAMER WITH REGARD TO SUCCESS IN PHYSICS

The knowledge learned in games was useful to me in class.	Lower level	52,16	MWU	1323,500
			Wilcoxon W	2451,500
	Higher level	67,35	Z	-2,362
			Asymp. Sig. (2-tailed)	,018

Additionally, differences in gamer identity with respect to success in STEM subjects were examined separately for male and female respondents. The results showed that male respondents with lower success in STEM subjects give statistically significantly more importance to cooperation with friends in the group in order to win compared to male respondents with higher success. The results also showed that male respondents with higher success in STEM subjects statistically significantly more believe that playing games stimulates their imagination and that the knowledge learned in games was useful to them in class. Also, male respondents with higher success in STEM subjects feel statistically significantly more like gamers compared to male respondents with a lower level of success (Table 4.).

TABLE 4. DIFFERENCES IN SELF-PERCETION AS A GAMER WITH REGARD TO SUCCESS IN STEM SUBJECTS - MALE

When I play with friends in a group, it is important for me to cooperate in order to win.	Lower level	78,58	MWU	1979,500
			Wilcoxon W	5060,500
	Higher level	64,88	Z	-2,303
			Asymp. Sig. (2-tailed)	,021
I find that playing games stimulates my imagination.	Lower level	63,14	MWU	1962,000
			Wilcoxon W	3978,000
	Higher level	77,35	Z	-2,181
			Asymp. Sig. (2-tailed)	,029
The knowledge learned in games was useful to me in class.	Lower level	60,73	MWU	1812,500
			Wilcoxon W	3765,500
	Higher level	78,26	Z	-2,603
			Asymp. Sig. (2-tailed)	,009
I feel like a gamer.	Lower level	62,06	MWU	1895,000
			Wilcoxon W	3848,000
	Higher level	76,39	Z	-2,142
			Asymp. Sig. (2-tailed)	,032

Results also showed that when compared to female respondents with higher success in STEM subjects, female respondents with lesser success in STEM subjects were statistically significantly ($p<0,021$) more likely to think that playing games is just for boys.

From this we can conclude that the first hypothesis is partially refuted because for some STEM subjects there are statistically significant differences in the statements used to determine the gamer identity and male respondents with higher success in STEM subjects feel statistically significantly more like gamers compared to male respondents with a lower level of success.

Second hypothesis: "There is no statistically significant difference in students' preference for more complex games (role-playing games, strategy games, action games and first-person shooter games) and their grades in STEM related subjects." was partially refuted. In the entire sample, no statistically significant differences

were found in students' preference for more complex games (role-playing games, strategy games, action games and FPS games) and their grades in STEM related subjects. However, for male respondents with higher success in STEM subjects, there was a statistically significantly difference in the following (complex) genres they play: role playing games (p<0,004), strategy games (p<0,023) and adventure games (p<0,007) than male respondents with lower success in STEM subjects. For male respondents with higher success in mathematics, it was determined that they play role-playing games statistically significant (p<0,028) more often compared to male respondents with lower success in mathematics. By measuring the correlations between the perception of oneself as a gamer and certain types of games, positive correlations were found between the statement *I feel like a gamer* and the shaded statements about the game genres, which indicates that respondents who feel more like gamers choose the shaded (complex) games more (Table 5.).

TABLE 5. CORRELATIONS BETWEEN THE PERCEPTION OF ONESELF AS A GAMER AND GAME GENRES

Spearman's rho		I feel like a gamer.
Role playing games	Correlation Coefficient	,341**
	Sig. (2-tailed)	,000
	N	278
Strategy	Correlation Coefficient	,478**
	Sig. (2-tailed)	,000
	N	279
Action	Correlation Coefficient	,506**
	Sig. (2-tailed)	,000
	N	283
FPS	Correlation Coefficient	,513**
	Sig. (2-tailed)	,000
	N	282
Sport	Correlation Coefficient	,296**
	Sig. (2-tailed)	,000
	N	282
Logic	Correlation Coefficient	,116
	Sig. (2-tailed)	,053
	N	282
Adventure	Correlation Coefficient	,205**
	Sig. (2-tailed)	,001
	N	281
Education	Correlation Coefficient	-,080
	Sig. (2-tailed)	,180
	N	283
Simple online	Correlation Coefficient	-,042
	Sig. (2-tailed)	,487
	N	281
I like more complex games that require a lot of thinking and strategy.	Correlation Coefficient	,228**
	Sig. (2-tailed)	,000
	N	283
I like simpler games for relaxation and distraction.	Correlation Coefficient	,026
	Sig. (2-tailed)	,662
	N	281

Third hypothesis: "There is no statistically significant difference between female and male respondents and their self-perceived gamer identity" was refuted. The results of MWU test showed statistically significant difference (p<0,01) in the perception of oneself as a gamer regarding gender where male subjects perceive themselves as gamers statistically significantly more than female subjects in the following statements: *I think gaming is cool. I couldn't do without playing games. When I play with friends in a group, it is important for me to cooperate in order to win. When I play with friends in a group, it's important to me to be the best in the group. I feel like a gamer. My friends consider me a gamer. I post my gaming endeavors on social networks (Facebook, YouTube, etc.). Famous gamers are my idols. I follow YouTube channels of famous gamers.*

IV. DISCUSSION

Our results are mainly in accordance with the ones mentioned in the Introduction. While Donati, Alessia Guide et all. [2] report the average of 12h of playing per week, our results show 31% of teenagers play between 5 and 10h per week. The same authors report that more than 50% of teenagers play first-person shooters, sandbox and management games, we found out that teenagers mostly play education, role-playing and logic games. However, our results are in accordance with Howe, Livingston and Lee [16] who determined that people who felt like gamers nowadays most frequently play role-playing games. The same research showed that male participants more often feel like gamers and play more often, which was also indicated by our data. Reeves, Brown, and Laurier [4] showed many studies that proved the benefits of gaming, particularly improved spatial cognitive benefits, strategic problem solving, enhanced short-term memory, pattern recognition, focus and multitasking and Biagi & Loi [9] proved that gaming is the only activity that shows a constant positive coefficient between PISA test scores and intensity of use. We also partially confirmed those results by showing that male respondents with higher success in STEM subjects statistically significantly more believe that playing games stimulates their imagination and that the knowledge learned in games was useful to them in class and also, male respondents with higher success in STEM subjects prefer to play more complex games such as role playing games and strategy games. This conclusion is in contrast with Dindar [19] where no positive relationship between playing games and academic success was found.

When it comes to cooperation with friends respondents with a lower level of success in STEM subjects attach more importance to cooperation with friends and especially male respondents with lower success in STEM subjects give statistically significantly more importance to cooperation with friends. This is an interesting conclusion because according to Baek & Archraf [21] male students express higher favorable attitudes toward group work than female students and Gupta, Manju, and Pooja [22] demonstrated the superiority of cooperative learning over traditional methods of instruction. We draw the conclusion that students with lower success in STEM subjects prefer cooperation because it may help them learn more effectively. A paradigm shift in education is necessary because the advantages of collaborative learning have been established.

Regarding gamer identity Howe, Livingston, and Lee [18] proved that men feel like gamers more often and spend more time playing games which was also confirmed by our study. In light of this, we also came to the conclusion that male respondents who reported greater success in STEM subjects feel statistically much more like gamers than male respondents who reported lower success in STEM subjects.

CONCLUSION

The aim of this paper was to analyze teenagers' habits of playing video games and the connection between their performance in Science, Technology, Engineering and Math related courses. Our research showed that about one third of teenagers play between 5 and 10 hours per week, and mostly play educational, role-playing and logic games.

The three hypotheses we set were partially refuted. We determined that there partially exists a difference between students that feel and do not feel like gamers, and their grades in the STEM courses. Also, there partially exists a difference between preference for more complex games and their grades in STEM courses. The third hypothesis, that there exists no statistically significant difference between female and male respondents and their self-perceived gamer identity, was also refuted. More research is needed to further the link between gaming and success in STEM field, and to relate it with other levels of education.

REFERENCES

[1] Heiden, Juliane M. von der, Beate Braun, Kai W. Müller, and Boris Egloff. 2019. "The Association between Video Gaming and Psychological Functioning." Frontiers in Psychology 10 (1731). https://doi.org/10.3389/fpsyg.2019.01731.

[2] Donati, Maria Anna, Cristiana Alessia Guido, Giuliano De Meo, Alberto Spalice, Francesco Sanson, Carola Beccari, and Caterina Primi. 2021. "Gaming among Children and Adolescents during the COVID-19 Lockdown: The Role of Parents in Time Spent on Video Games and Gaming Disorder Symptoms." International Journal of Environmental Research and Public Health 18 (12): 6642. https://doi.org/10.3390/ijerph18126642.

[3] Faust, Kyle, Joseph Meyer, and Mark D. Griffiths. 2013. "Competitive and Professional Gaming." International Journal of Cyber Behavior, Psychology and Learning 3 (1): 67–77. https://doi.org/10.4018/ijcbpl.2013010106.

[4] Reeves, Stuart, Barry Brown, and Eric Laurier. 2009. "Experts at Play." Games and Culture 4 (3): 205–27. https://doi.org/10.1177/1555412009339730.

[5] Khatib, Firas, Frank DiMaio, Seth Cooper, Maciej Kazmierczyk, Miroslaw Gilski, Szymon Krzywda, Helena Zabranska, et al. 2011. "Crystal Structure of a Monomeric Retroviral Protease Solved by Protein Folding Game Players." Nature Structural & Molecular Biology 18 (10): 1175–77. https://doi.org/10.1038/nsmb.2119.

[6] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders DSM-5, 5th ed.; American Psychiatric Association: Arlington, VA, USA, 2013.

[7] World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 11th ed.; WHO: Geneva, Switzerland, 2020; Available online: https://icd.who.int/en

[8] M.Nazry, N.Nazrina, and Daniela M. Romano. 2017. "Mood and Learning in Navigation-Based Serious Games." Computers in Human Behavior 73 (August): 596–604. https://doi.org/10.1016/j.chb.2017.03.040.

[9] Biagi, Federico, and Massimo Loi. 2013. "Measuring ICT Use and Learning Outcomes: Evidence from Recent Econometric Studies." European Journal of Education 48 (1): 28–42. https://www.jstor.org/stable/pdf/23357044.pdf

[10] Gubbels, Joyce, Nicole M. Swart, and Margriet A. Groen. "Everything in Moderation: ICT and Reading Performance of Dutch 15-Year-Olds." Large-scale Assessments in Education 8, no. 1 (2020). https://doi.org/10.1186/s40536-020-0079-0.

[11] Kong, Yiren, Young Sik Seo, and Ling Zhai. "ICT and Digital Reading Achievement: A Cross-National Comparison Using Pisa 2018 Data." International Journal of Educational Research 111 (2022): 101912. https://doi.org/10.1016/j.ijer.2021.101912.

[12] Turner, Patrick E., Elizabeth Johnston, Mansureh Kebritchi, Sally Evans, and David A. Heflich. 2018. "Influence of Online Computer Games on the Academic Achievement of Nontraditional Undergraduate Students." Edited by Shuyan Wang. Cogent Education 5 (1). https://doi.org/10.1080/2331186x.2018.1437671.

[13] Shaw, Adrienne. 2011. "Do You Identify as a Gamer? Gender, Race, Sexuality, and Gamer Identity." New Media & Society 14 (1): 28–44. https://doi.org/10.1177/1461444811410394

[14] Griffiths, Marie, Karenza Moore, and Helen Richardson. 2007. "CELEBRATING HETEROGENEITY?: A Survey of Female ICT Professionals in England." Information, Communication & Society 10 (3): 338–57. https://doi.org/10.1080/13691180701409945.

[15] Sherry, J. L., Greenberg, B. S., Lucas, K., & Lachlan, K. 2012. "Video game uses and gratifications as predictors of use and game preference". Playing video games: 248-262.

[16] Neys, Joyce L.D., Jeroen Jansz, and Ed S.H. Tan. 2014. "Exploring Persistence in Gaming: The Role of Self-Determination and Social Identity." Computers in Human Behavior 37 (August): 196–209. https://doi.org/10.1016/j.chb.2014.04.047.

[17] De Grove, Frederik, Cédric Courtois, and Jan Van Looy. 2015. "How to Be a Gamer! Exploring Personal and Social Indicators of Gamer Identity." Journal of Computer-Mediated Communication 20 (3): 346–61. https://doi.org/10.1111/jcc4.12114.

[18] Howe, William, Dalaki Livingston, and Sun Kyong Lee. 2019. "Concerning Gamer Identity: An Examination of Individual Factors Associated with Accepting the Label of Gamer." Shareok.org. https://hdl.handle.net/11244/324944.

[19] Dindar, Muhterem. 2018. "An Empirical Study on Gender, Video Game Play, Academic Success and Complex Problem Solving Skills." Computers & Education 125 (October): 39–52. https://doi.org/10.1016/j.compedu.2018.05.018.

[20] Smith, Barbara Leigh, and Jean T. MacGregor. "What is collaborative learning." (1992): 233-267.

[21] Baek, Youngkyun, and Achraf Touati. "Comparing Collaborative and Cooperative Gameplay for Academic and Gaming Achievements." Journal of Educational Computing Research 57, no. 8 (January 24, 2019): 2110–40. https://doi.org/10.1177/0735633118825385.

[22] Gupta, Madhu, Manju Jain, and Pooja Pasrija. 2014. "Gender Related Effects of Co-Operative Learning Strategies (Stad and Tai) on Mathematics Achievement." Issues and Ideas in Education 2 (1): 53–68. https://doi.org/10.15415/iie.2014.21005.

[23] Lee, J. H., Karlova, N., Clarke, R. I., Thornton, K., & Perti, A. 2014) Facet analysis of video game genres. IConference 2014 Proceedings.https://doi.org/10.9776/14057

Education via ICT Due to Corona Crisis – Our View on Pros and Cons

L. Révészová
Faculty of Economics Technical University of Košice
Department of Applied Mathematics and Business Informatics, Košice, Slovakia
libusa.reveszova@tuke.sk

***Abstract* - The article summarizes our experience in teaching via ICT due to Corona crisis in the period March 2020 - January 2023, at the Faculty of Economics, Technical University, Košice. We discuss pros and cons of this form of education from the point of view of university teachers and we present also our students views obtained from questionnaires. Today in the fast-changing global environment, knowledge is one of the most important factors determining success and technologies influence all areas of human lives. That is why in our article we deal with the following problems: what methods should we implement to optimize the effects of education via ICT and what will be the impacts on students' knowledge base. We discus also what will be the overall effects of the corona crisis on education.**

Keywords* - *education; ICT; evaluation; knowledge base; skills;

I. INTRODUCTION

COVID-19 has changed processes all around the world. Due to Corona crisis we are forced to look for new ways how to act, how to do business, how to teach and learned ...

We have shifted into survival mode and we focused on digital world and digital economy or more precisely distance world and distance economy. Nowadays we communicate and work via ICT, via ICT (more or less) works e-business, e-learning, e-health, e-government.

It is well known, that every crisis is a great opportunity for rethinking and for new beginnings.

The current situation - after the Corona crisis - has shown that everyone will probably need to update their skills, not only once, but also many times during their career. The McKinsey Global Institute report [4] revealed four key skills areas that will be needed in the future post Corona crisis:

- ability to operate at pace in a fully digital environment;
- cognitive skills for redesign and innovation;
- social and emotional skills to ensure effective collaboration, management, and self-expression;
- adaptability and resilience to thrive during COVID-19 aftershocks;

Confronting with the educational emergences, factors above involve a modification of the education systems, systems of knowledge and work, and, as a consequence, also of educational politics which must promote a personal development of citizens through the development of the necessary knowledge base, skills, and competences.

II. HOW THE CORONA CRISIS AFFECTED US IN THE FIRST WAVE

During the summer semester 2019/2020, in the third week of teaching, March 9th 2020, we were forced, due to the regulations and rules of the government and the university management, to move to fully online teaching. The situation was not easy, we were not prepared, new procedures were unknown, we did not know if and when the situation would return to normal.

Pros and cons of the situation from the point of view of infrastructure, teachers, students and management were as follows:

Pros:

- we had a well-established LMS system Moodle at the faculty, teachers and students were used to work in the Moodle environment - therefore, we could immediately start using it more intensively;
- a computer for the teacher was available in each classroom and the Internet was introduced, an interactive whiteboard was also available in every computer lab;
- we had “department of IT support” at the faculty - this small department was and still is extremely important in managing online teaching;
- we tried to do our best, we were optimistic and enthusiastic, we believed in the early end of the crisis – as we can see now, “state of mind” is extremely important in overcoming problems;

Cons:

- we did not have adequate equipment (tablets, whiteboards) and platforms (webex, MS Teams, Zoom) available for online lectures and exercises and, of course, teachers' and students' experience of using them were minimal;
- the home technical equipment and the quality of the students' internet connection were very different, especially in the case of foreign students (they mostly come from India and Ukraine);
- regulations adopted by the government and university management were not clear, further procedures, processes and practices could not be estimated, the lockdown was extended,

- conditions were confusing and uncertain practically in all areas of life.

The online education during first wave - meant at least ensuring the publication of offline materials - lectures, solved examples, assignments, procedures mostly via LMS Moodle. Later, the instructional videos and some lectures were published e.g. on YouTube, but it can't be called a systematic approach.

During the semester, we provided access and training for employees and students to the WebEx and MS Teams platforms, and we purchased and replenished technical equipment.

III. HOW THE CORONA CRISIS AFFECTED US IN THE NEXT WAVES

We were ready to start the winter semester 2020/2021 "normally" (full-time study in person), but the situation with COVID-19 has worsened and so, in order to protect health and safety, we have started online education via ICT again, but we have moved to a higher level.

We have approached online education according to a fixed schedule (this was not the case in the previous semester), following the rules of full-time teaching. Teachers had fully equipped classrooms at their disposal, and students joined lectures and exercises from home. The platform, mostly WebEx or MS Teams, was chosen by the teacher according to the specifics of the subject. The "common using" of LMS Moodle has moved to a significantly higher level. Many teachers discovered until that time unused functionalities of LMS Moodle and e.g. included various activities and tests in the teaching process.

Due to the ongoing pandemic, we were forced to make difficult decisions before the start of the next three semesters - whether the teaching will take place online or face-to-face. The most important factors that influenced us were the regulations of the government and Public health office of the Slovak Republic and the regulations of university management. Based on this regulations, the number of students who could participate in group tuition was significantly limited, the meeting of teachers and students was conditioned by proof of infection-free status based on COVID-tests, and accommodation and catering facilities for students were closed. The efforts of both teachers and students to protect their health and health of their families were also significant, which resulted in many cases of reluctance to teach/learned face-to-face. That is why we had full online education for 4 semesters. In summer semester, which starts in February 2022, we have online education and credit tests and exams took place face-to-face.

During the Corona crisis time we tried to implement different models at our university - a combination of online and face-to-face teaching, but the results were not convincing. For example: students were divided into small groups in order to comply with government regulations and attend classes every other week. The curriculum was either repeated - which significantly reduced the content of the curriculum, or the curriculum was taken in the usual way and the students alternated between online and face-to-face form.

Based on feedback, teachers' and students' responses to these models were not very favorable.

IV. WHAT IS THE CORONA CRISIS HAS TAUGHT US?

Nowadays in every environment, also in the school system it is necessary to count with a minimal stability and certainty. The only thing that is certain in today's world is a permanent change. For a successful existence in this turbulent environment it is very important:

- to learn how to cope with crisis and changes;
- to catch information that is useful, know how to keep it and use it for our benefit;
- it is necessary for us as educational institution not only to be „educating" but also to be „learning";
- the ratio of our change has to be higher than the ratio of changes in the environment around us;
- a good position can only be the result of a quick adaptation to current needs and changing requirements;
- the implementation of innovations is considered as the key factor of success;

During our attempt to adapt the education we contemplated human's abilities, processes and technologies in their mutual relations and connections. We use a well-known holistic, model of knowledge management (sometimes consciously, sometimes intuitively). This tool allows us to profit from synergic effect. It is well captured in the Figure 1.

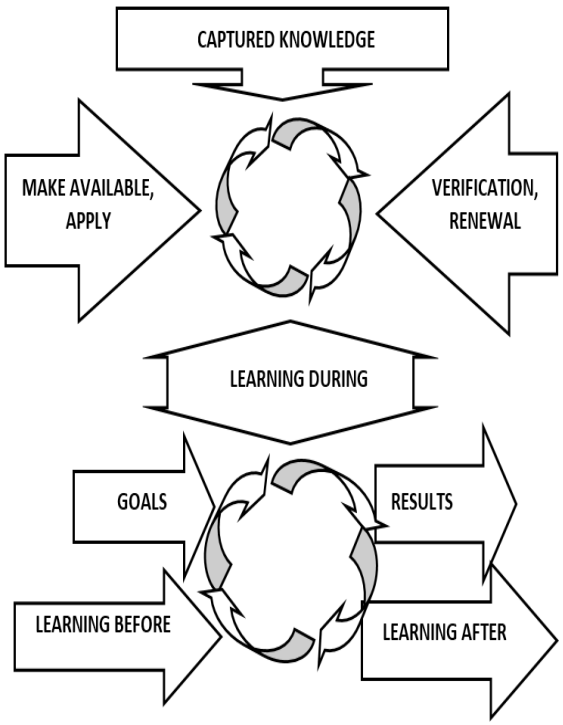


Figure 1. Knowledge management – holistic model according to[6]

V. IMPACTS OF ONLINE EDUCATION ON STUDENTS' KNOWLEDGE BASE

An important question for us, of course, is: What is the impacts of online education on students' knowledge base?

As an example, we present the evaluation of the subjects Informatics II in the Figure 2, Mathematics II in the Figure 3, and Management in the Figure 4, in the year 2018/2019 in "normal" full-time teaching and examination and in the next years 2019/2020, 2020/2021 - online teaching and online examination, and in the 2021/2022 year - online teaching and face to face examination. Grades A, B, C, D, E, and FX represent the success of students according to the commonly used scale: A means 100 - 91%, B 90 - 81%, C 80 - 71%, D 70 – 61 %, E 60 – 51 % and FX is a failing grade 50 % and less.

As we can see in the Figure 2 and Figure 3 there is a significant decrease in subject failure and an increase in grades, especially B and C in the case online examination. Of course we tried to use all available ways/methods to ensure the objectivity of the online assessment of students. For example: collective writing of credit tests, turning on the camera during the testing/exam, mixing test questions, etc. Consequently, in the last year (online teaching but face-to-face assessment) an extreme increase in failure is visible.

Most teachers of subjects based on quantitative methods have similar experiences. We can conclude, that from our point of view, the full online form of teaching is inappropriate for these kind of subjects and it can serve as a support or as a solution to specific situations.

Vice versa, within the full online form of teaching subjects such as Management, were appropriately supplemented and expanded teaching texts and materials with online lectures, demonstrations, case studies, etc. which could have resulted in a higher success rate, as we can see in the Figure 4.

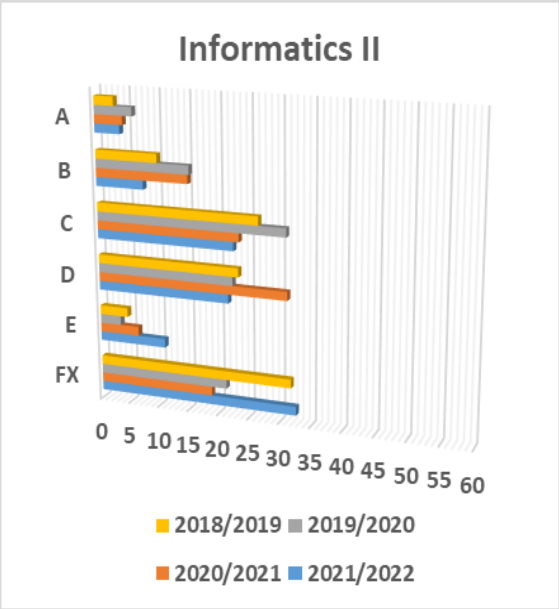


Figure 2. Evaluation of subject Informatics II in percentage, own source

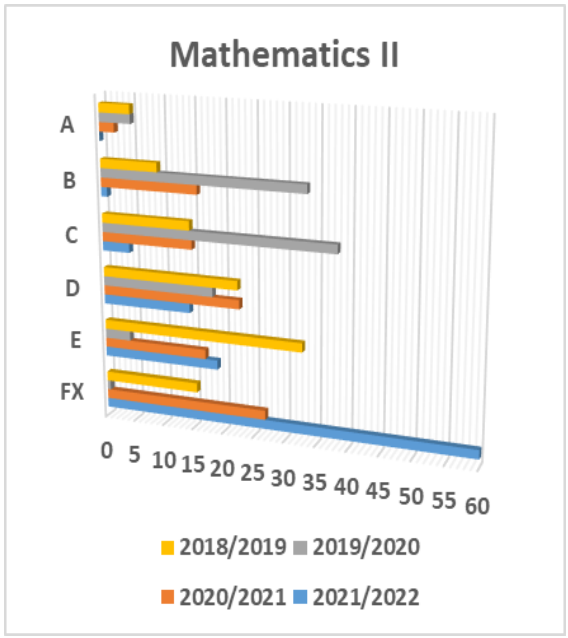


Figure 3. Evaluation of subject Mathematics II in percentage, own source

The negative consequences of online teaching in secondary schools are felt significantly in the first year of study at our faculty. Students' knowledge base in the field of mathematics and informatics is significantly limited compared to years with a standard way of teaching.

VI. WHAT ARE STUDENTS' VIEWS ON ONLINE EDUCATION?

We collect students' opinions on online education from several sources:

- survey initiated by the Student Council of Higher Education Institutions in Slovakia at the end of March 2020;

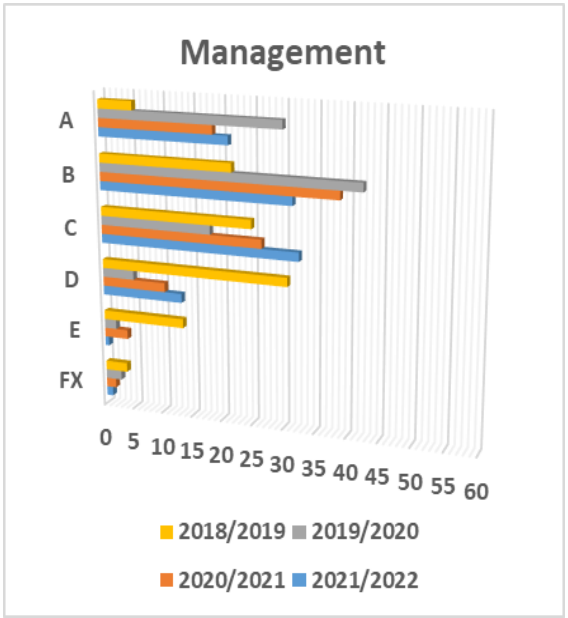


Figure 4. Evaluation of subject Management in percentage, own source

- questionnaire of The student part of the Academic senate TUKE to determine the impact of the Corona Crisis;
- twice a year administered questionnaire. Evaluation of the quality of previous semester teaching at the faculty.

The survey initiated by the Student Council of Higher Education Institutions in Slovakia, with the aim of mapping the situation regarding the distance education at Slovak universities. The result was the participation of 1130 respondents from 22 universities in Slovakia, and the material is available at [7]. 224 full-time students from our faculty participated in the survey.

We can present their opinion and views as follows: The most preferred methods in the first online semester (frequency more than 75%), at the faculty was the recommendation of study literature and scripts to be studied by students (43%), providing presentations created by teachers (38%), distance learning is also realized through writing seminars. works (31%) and e-learning materials (22%). On the contrary, few (less than 25%) used video conferencing, live lectures and archived lectures.

In other questions, students were asked to evaluate the course of distance learning from three perspectives, whether:

- they feel informed by the school to the extent that they can handle distance learning – information sufficiency;
- they are satisfied with the course of distance learning in their field – satisfaction, and
- they consider the distance form of teaching to be a full-fledged substitute for full-time teaching – equivalency.

The data in Figure 5, on a scale (1 - agree; 5 - disagree) indicate that 22% of students feel sufficiently informed by the school, 19% have expressed a neutral attitude and only 10% have expressed that they do not feel well informed by the school. The average scale value of the responses to this item reached 2.58, which is closer to agreeing with sufficient information. 19.5% of students are satisfied with the course of distance learning, almost 14% are dissatisfied. The average scale value of the answers reached the value of 2.87, which means that most respondents do not feel expressed satisfaction or dissatisfaction. Regarding the evaluation of distance teaching as a full-fledged substitute for full-time teaching, the results strongly indicate that more than a third of students (37%) do not consider distance teaching as a full-fledged substitute for full-time teaching, the opposite opinion was taken by only 9% of respondents. The average scale value of the answers in this item was 3.61, which is closer to the disagreement.

Another source of feedback information was a questionnaire survey at the Technical University in Košice (TUKE), organized by the student part of the Academic senate TUKE.

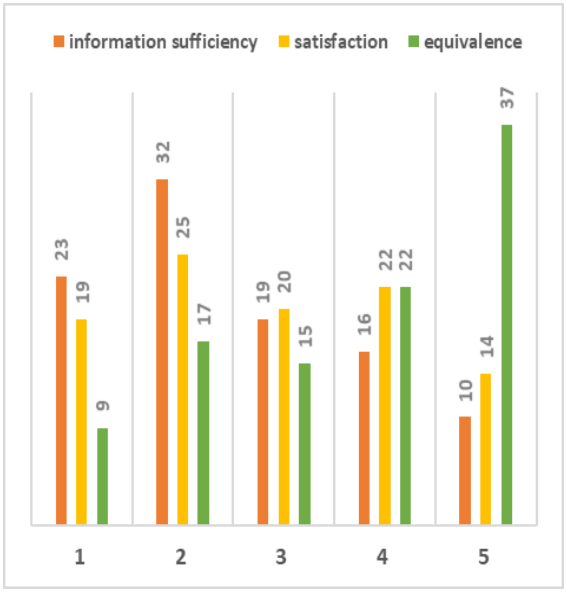


Figure 5. Evaluation of studnets´opinion on online education, source [7]

Basic information about the questionnaire participants:

- number of TUKE students 10059, number of responses 4133 (41.09%);
- number of students of our faculty 1012, number of responses 432 (42.69%).

We chose evaluation of questions related to student satisfaction with the online form of teaching.

The questions: I – I acquired similar knowledge as during face-to-face teaching. II – The distance method was able to fully replace face-to-face lectures. III – The distance method was able to fully replace face-to-face exercises/seminars.

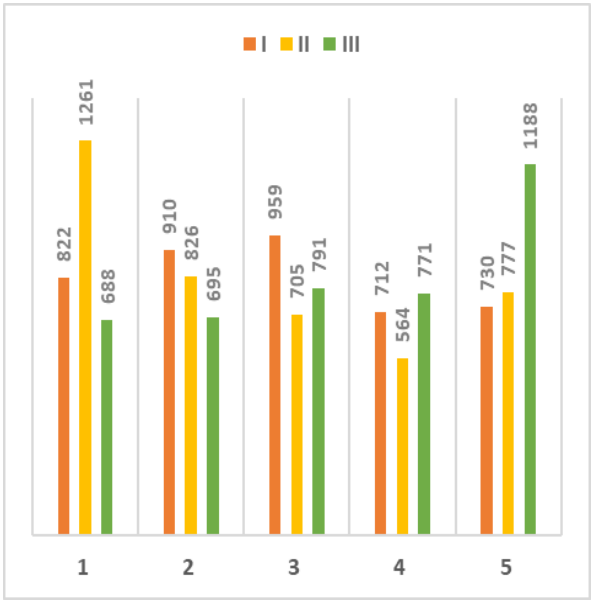


Figure 6. Evaluation of studnets´ online education satisfaction, internal source of TUKE

The scale: 1 - I agree 2 - I rather agree 3 - I do not know 4 - I rather disagree 5 - I do not agree

As we can see in the Figure 6, students perceived online teaching relatively positively only in the case of lectures. In the case of exercises and seminars, online teaching was insufficient. Even in the case of the assessment of acquired knowledge, there are more negative opinions, or the students could not assess that. This assessment corresponds to the experience and opinions of teachers at our faculty.

We also assessed the level of online teaching and its effects through a questionnaire survey at the faculty after the end of each semester. In general, we can say that students' satisfaction with information flows, technical equipment and the course of online teaching gradually grew. However, the percentage of students who demanded a return to face-to-face teaching was always around 50.

VII. Discussion

Nowadays everyone agree, that education is widely accepted to be a fundamental resource, both for individuals and societies, as we can read in [2].

In the current situation we understand education in accordance with [1] as any act or experience that has a formative effect on an individual's mind, character, or physical ability.

In accordance with what we stated above and based on our experience we can see now pros and cons of education via ICT from the point of view of infrastructure, teachers, students and management as follows:

Pros:

- our institution in accordance with [5] (Gartner pandemic response) went through three phases: respond, recover, renew - we were forced to react adequately and we gathered large amount experience;
- now we are ready to act/respond in the case of another crisis better and faster than before the Corona crisis;
- the technical and software equipment has improved significantly;
- the abilities and skills of educators and students in the field of teaching/learning and communicating via ICT have significantly improved;
- both students and teachers have personal experience with the positive and negative sides of e-learning and home office, they are able to assess their strengths and weaknesses;
- crisis empowered educators and student to teach/to learn better through sharing good practice and collaboration;
- now we see more clearly the positive aspects of e-learning – cost effective, saves time, improves performance and productivity, lower environmental impact, in accordance with [3], [5];
- crisis created an opportunity to reset some of our goals and ambitions;

Cons:

- the reactions of the government, the state administration and university management to the rapid changes, challenges and demands of the present are inadequate;
- some problems associated with the ICT using in education are difficult to solve despite technological progress especially when students live in remote areas or because of their socioeconomic situation or special needs;
- it is very time consuming to prepare and to integrate elements such as images, videos, audio and graphics but it is way of keeping learners engaged, compared to traditional learning;
- the crisis confirmed even more that the biggest challenges of education still remain motivation and objective evaluation for both students and educators;

VIII. Conclusion

Even before Corona crisis, revolutions in technologies, workforce preferences, and many new models were affecting the ways of education. The virus has accelerated and intensified this, and is forcing education institutions to shift their organization's crisis operations while fundamentally reimagining strategies, teaching and learning models.

The Corona crisis has created an opportunity to reset our goals and ambitions. But we have to ask: What will our future be? How further technological development and advances will affect education?

We hope that the knowledge and experience gained during the past difficult period will help us face new challenges.

References

[1] M Eurostat, Statistic explained, Glossary, Education and training, 2021 [online] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Education

[2] M. Roser and E. Ortiz-Ospina, "Global Education". Published online at OurWorldInData.org., 2016, [online] Retrieved from:

[3] C. Lawless, "What is eLearning?" 2018, [online] https://www.learnupon.com/blog/what-is-elearning/

[4] A.Feld, A. Reich, K. Störk, and S. Durth, "Thriving after COVID-19: What skills do employees need?", 2020, [online], https://www.mckinsey.com/business-functions/mckinsey-accelerate/our-insights/accelerate-blog/thriving-after-covid-19-what-skills-do-employees-need#

[5] Ch. Howard, "Reset Your Business Strategy in COVID-19 Recovery", 2020, [online], https://www.gartner.com/smarterwithgartner/reset-your-business-strategy-in-covid-19-recovery/

[6] C. Collison, G. Parcel, "Knowledge Management". CPPress, Brno, Czech republic 2005.

[7] ŠRVŠ SR, Sprievodca výučbou počas COVID-19, 29.3. 2020 , [online], https://srvs.eu/covid-19/ in Slovak

Comparative Analysis of Student Assessment during the COVID-19 Crisis

A. Pongrac Pavlina, K. Pavlina
University of Zagreb / Faculty of Humanities and Social Sciences, Zagreb, Croatia
kpavlina@ffzg.hr

***Abstract* - During the coronavirus pandemic (summer of 2019/20) classes at the University were held via videoconferencing tool. The lectures were also available to the students in the form of recordings. For a written online colloquium we designed a specific system. The colloquium was solved while teacher simultaneously held a video conference call for guiding students on time limit of each question. The entire colloquium is divided into smaller tests; first test-the closed-type questions were mixed so that the student received his combination of questions. The open-ended questions were divided into separate tests and the students, via video conference call, received instructions when each question was opened. In this way, all students wrote the same questions at the same time, and a certain amount of time was set for each question. The next generation of students (2020/21) was again the pandemic generation. These students, like the previous generation, also got access to recorded lectures, so they could consult the lectures countless times. This generation wrote colloquia on paper in the classroom. This paper compares the results of these two generations of students in written colloquiums and final success in taking courses that include all the elements that are evaluated during the semester.**

Keywords – student evaluation; coronavirus pandemic; students; Moodle;

I. INTRODUCTION

The education system faced great challenges during the COVID-19 crisis. Due to the lockdown, the teaching work had to be fundamentally changed since contact teaching at the college was no longer possible. For this reason, the teaching relied on e-learning systems [1] through which teachers could organize the teaching work using different tools (video conferences, chat, file storage) [2]. Along with synchronous forms of teaching, experts also strongly recommended the transition to asynchronous forms of teaching [3].

Research [4] that was conducted on 150 teachers and 150 students of 5 universities in Nepal showed the positive aspects of online education: increasing independence during learning (76.9%), independent choice of time and place of learning (69.2%) and great source of knowledge (61.5%). The challenges faced by teachers and students mostly relate to the organization of learning and teaching itself; organization of time and access to the Internet as a basic resource (85.7%), finding one's own learning style (84.6%) and social isolation (71.4%). For successful participation in online education, it was necessary to fulfill certain prerequisites; computer literacy (98.9%), technological preparedness (98.7%) and personal organization of time (96.6%).

Due to COVID-19, the evaluation process itself was either postponed or canceled. Institutions that decided to do the assessment through e-learning systems encountered many challenges using online assessment tools [1].

Research [5] of 13 articles from 9 countries of the world showed that university teachers used Moodle (39.19%) and the videoconferencing tools Skype and Zoom (34.78%) in the highest percentage for evaluation purposes. In order to prevent students from cheating, teachers tried to monitor students (e-proctoring) using cameras, software and detecting plagiarism.

Research [6] conducted over 3 years on 63 students divided into 4 groups in the College of Engineering in France showed that the groups acquired knowledge and skills in different ways (the first two groups had a combination of group work, tutorials and lectures, the third group first had lectures, followed by practical work, while the fourth group worked on projects). They used MS Teams, Zoom and Google Drive. The methods of evaluation were also different. The first and second groups took an online quiz and a final exam that they took in a virtual room with limited time for a written test. They had cameras on and were supposed to be connected to a virtual room via smartphones. In the end, they were given extra time to scan and email the exam. Another way these groups approached knowledge testing was based on time-limited tasks, but the students were not supposed to be connected to a virtual room. The exam was sent to them by e-mail, and at the end they had extra time to scan their answers and send them by e-mail or submit them via drop box. The third group had to write several reports about their practical work. The fourth group submitted a report on the completed project. A comparison of the results of students who in earlier years took the same course face-to-face in the classroom with students who took the course remotely showed that distance education did not affect the mastery of the skill.

Research [7] conducted on 3037 students and 231 teachers from 15 universities in Morocco showed that 70.6% of teachers had no previous teaching experience in distance learning. Despite this, 92.08% of teachers delivered their course via different platforms during COVID-19. Teachers, most often, used several different platforms, and the highest percentage (54.7%) used Moodle, followed by MS Teams (48.8%). Despite the fact that teachers used the platforms for teaching purposes, 64.4% of them believe that it is not feasible to hold exams

remotely. Students are also not happy with the idea of holding exams remotely; 81.45% believe that they are unable to take exams remotely.

II. METHODOLOGY

A. Teaching work in the course Teaching methods in computer education 1

The course Teaching methods in computer education 1 is a mandatory course for students of teacher informatics, which means that all students who want to complete the teaching major must attend and pass the course. The course is held in the summer semester of the first year of graduate studies and is the first methodical course that students enroll to acquire methodical knowledge and develop methodical skills for independently designing and maintaining the teaching process of computer science. Since students need to acquire knowledge about the teaching process of computer science, they listen to two hours of lectures per week, which are held face-to-face in the classroom.

In order to prepare for the independent holding of lessons in the subject of computer science, students attend two hours of exercises per week. Since it is a course on the basis of which students should acquire knowledge and develop the skill of preparing and holding a lesson, a combination of both the theoretical and practical part is necessary.

The teaching work is designed through continuous work during the semester. During the semester, for the purpose of practice, the student writes three individual tasks on the topic of his chosen teaching unit for practice in order to begin the process of designing and preparing the teaching process. At the end of the semester, the student submits a complete teaching preparation for the exercise. During the semester, the student should hold one teaching stage from his teaching preparation for the exercises in order to start his process of practicing the teaching of selected lesson.

During the semester, the students write two colloquiums in order to check knowledge they acquired during the semester. All students write the colloquium at the same time in the classroom. The colloquium consists of a combination of open and closed questions that the student can solve in the order he chooses, but in the exact time given, because each colloquium is combination of a test of depth and speed. At the end of the semester, the student writes another teaching preparation. For all activities, the student receives timely feedback.

Entire teaching work is designed through a combination of lectures and exercises. Since the basic goal of the course is to train the student to independently prepare and hold a lesson, lectures, exercises and colloquiums are held face-to-face in a classroom at the faculty.

B. Teaching work in the course Teaching methods in computer education 1 during COVID-19

In the summer semester of the academic year 2019/20 due to the corona crisis, it was necessary to change the coursework. Lectures, exercises and colloquiums could not be held face-to-face in the classroom. It was necessary to conduct the teaching work using e-learning.

Since the Faculty of Humanities and Social Sciences of the University of Zagreb has had an e-learning system based on Moodle for many years, and researchers from the Department of Information and Communication Sciences are intensively studying this topic, we managed to adapt to the new way of working in a short period of time.

For the purpose of teaching, we used the BigBlueButton videoconferencing tool, on the basis of which we held classes synchronously. Since the Teaching methods in computer education 1 course consists of lectures and exercises, we have successfully modified the teaching methods to adapt them to the new teaching conditions. Lectures were also available to students through recordings posted on YouTube, so that teaching was also possible in asynchronous form. In this way, students could return to any lecture countless times. When listening to a contact lecture in the classroom, this method of operation is not possible, because after the lecturer has delivered the lecture, the student cannot listen to that lecture again. Student practice of teaching stages was successfully held through the videoconferencing tool.

The following summer semester (in the academic year 2020/21), when the course Teaching methods in computer education 1 was held again, it was also the corona semester. The way of working was the same with this generation of students. It is possible to work through the video conference tool BigBlueButton in synchronous form. The lectures were also available for this generation of students in the form of recordings (asynchronous form) and these students could also listen to the recordings of the lectures they wanted countless times. This generation of students also successfully practiced holding teaching stages through the videoconferencing tool.

The biggest challenge the teachers faced was the verification of acquired knowledge and developed skills. The colloquiums, under regular teaching conditions, were held face-to-face in the classroom, but during the summer semester of the academic year 2019/2020 this was not possible. For this reason, all knowledge tests had to be held in oral form or in writing via the e-learning system. The colloquiums, in the course Teaching methods in computer education 1 in the summer semester of the academic year 2019/20, were held in written form through the e-learning system. Students who took the course in the upcoming summer semester (academic year 2020/21) could take the colloquia in the classroom. Thus, during the corona crisis, two generations of students who attended the course Teaching methods in computer education 1 attended the course under changed conditions of teaching work. Both generations attended the course through the e-learning system. The exercises for both generations were successfully held in a synchronous form through a video conference call. The first corona generation (2019/20) took written knowledge tests at the colloquium via the e-learning system, while the second corona generation (2020/21) took them in written form in the classroom.

TABLICA I. TABLE 1. CHARACTERISTICS OF THE STUDENTS INVOLVED IN THE RESEARCH

Group	Number of students	Academic year	Course
1st group	1st colloquium: 24 2nd colloquium: 23 Overall grade: 23	1.year graduate study (2019./20.)	Teaching methods in computer education 1
2nd group	1st colloquium: 20 2nd colloquium: 20 Overall grade: 19	1.year graduate study (2020./21.)	Teaching methods in computer education 1

C. EVALUATING STUDENT ACHIEVEMENT DURING COVID-19

The colloquium, composed of closed and open-ended questions, could not be sent to the e-learning system and allow students to start and finish writing equally. If the written test of knowledge were to be approached in this way, academic integrity could theoretically be compromised. All students would take the colloquium at the same time and, theoretically, through some other communication channels, they could mutually agree on which question should be solved by which student, and towards the end of the designated time, the answers could be exchanged.

Because of this, for the generation of students who took the course in the summer semester, in 2019/20 teachers designed written knowledge assessment system that sought to take advantage of the advantages offered by e-learning. The written test of knowledge began at the agreed time with the initiation of a video conference call initiated by the subject teacher during the course itself. All students who registered to write a colloquium joined the video conference call with audio. The video conference call was designed for the purpose of holding a colloquium. The subject teacher would tell the exact time when a particular question was opened and warn when the time to solve a particular question was about to end. In this way, students would know how much time they still have at their disposal.

The entire written knowledge test was originally divided into two main parts; the first part consisted of closed type questions and the second part consisted of open type questions. The closed-type questions were chained into one written test in such a way that the written test consisted only of closed-type questions through which the student could only move forward, and could not return to the previous question. The written test with closed-ended questions was time-limited, and each student received a random combination of questions and offered answers. After the time expired, that written verification was automatically submitted to the system.

The second part was open-ended questions. That part of the written test was divided into several smaller parts, i.e. each question represented an exam in itself. Students would be instructed again that a particular question is being opened and they would be informed how much time they still have to solve it. After the time was up, the answer to each question would be automatically submitted to the system. In this way, all students wrote every single open-ended question at the same time, but due to the time limit and simultaneous solving of only one question, mutual cooperation in solving the colloquium was impossible.

The next generation of students who took the course in the summer semester of the academic year 2020/21 attended tests of knowledge (first and second colloquium) in the classroom. All students simultaneously wrote written knowledge tests, which, as with the previous generation, consisted of a combination of closed and open questions.

This generation of students received the colloquium in its entirety, i.e. they immediately received all closed and all open questions. Also, like the previous generation, they had a time limit to solve the colloquium, but that time was not divided into each individual question, but the time limit referred to the limit of how much time they had to solve the entire knowledge test. The input parameters for both groups of students were the same, i.e. both groups prepared for colloquiums based on lectures that were also available in the form of recordings and could be listened to countless times.

TABLICA II. TABLE 2. STUDENT TASKS

Group	Tasks	Teaching forms	Colloquiums
1st group (2019./20.)	a) Three individual tasks b) Teaching preparation for the exercise c) Second teaching preparation d) Presentation of the teaching stage e) Two colloquia	Synchronous + asynchronous	Online written colloquiums through the e-learning system
2nd group (2020./21.)	a) Three individual tasks b) Teaching preparation for the exercise c) Second teaching preparation d) Presentation of the teaching stage e) Two colloquia	Synchronous + asynchronous	Contact written colloquiums in the classroom

In order to conduct a comparative analysis, the following hypotheses were set:

H1: there is no statistically significant difference in the results achieved between the first (academic year 2019/20) and second (academic year 2020/21) generation of students at the first colloquium

H2: there is no statistically significant difference in the results achieved between the first (academic year 2019/20) and the second (academic year 2020/21) generation of students at the second colloquium

H3: there is no statistically significant difference in the final grade between the first (academic year 2019/20) and second (academic year 2020/21) generation of students

Statistical analysis was made using MS Excel Data analysis statistical package and Analyze It package. Differences between results achieved on first and second colloquium and between final results of both groups of students where statistically analyzed using t-test which was calculated with 95% level of confidence.

Statistical analysis revealed that there is no statistically significant difference in the results achieved at the first colloquium between the first (academic year 2019/20) and second (academic year 2020/21) generation of students. Therefore, the first hypothesis is accepted.

TABLE 3. STATISTICAL ANALYSIS OF RESULTS ACHIEVED AT THE FIRST COLLOQUIUM (T-TEST)

	2019./20.	**2020./21.**
Mean	0,641383	0,6475
Variance	0,013404	0,025359
Observations	24	20
P(T<=t) one-tail	0,443519	
P(T<=t) two-tail	0,887037	

Statistical analysis revealed that there is no statistically significant difference in the results achieved at the second colloquium between the first (academic year 2019/20) and second (academic year 2020/21) generations of students. Therefore, the second hypothesis is accepted.

TABLE 4. STATISTICAL ANALYSIS OF RESULTS ACHIEVED AT THE SECOND COLLOQUIUM (T-TEST)

	2019./20.	**2020./21.**
Mean	0,731304	0,653
Variance	0,027957	0,032506
Observations	23	20
P(T<=t) one-tail	0,074909	
P(T<=t) two-tail	0,149817	

Statistical analysis revealed that there is no statistically significant difference in the overall grade between the first (academic year 2019/20) and second (academic year 2020/21) generation of students. Therefore, the third hypothesis is accepted.

TABLE 5: STATISTICAL ANALYSIS OF RESULTS ACHIEVED IN THE OVERALL GRADE (T-TEST)

	2019./20.	**2020./21.**
Mean	3,478261	3,315789
Variance	1,16996	0,672515
Observations	23	19
P(T<=t) one-tail	0,291611	
P(T<=t) two-tail	0,583221	

D. CONCLUSION

Due to the COVID-19 crisis higher education had to completely switch to e-learning. Major challenge during the COVID-19 crisis was the issue of student evaluation.

This paper presents results of student evaluation on same course where the teaching process of both groups of students was the same; the e-learning system and the video conference tool BigBlueButton were used, and the students followed the classes through synchronous and asynchronous forms of teaching.

The results of the comparative analysis showed that there is no statistically significant difference between the results achieved between the first and second groups in the first colloquium, the second colloquium and in the final assessment. The results of the research showed that, since there is no statistically significant difference between the groups, that the designed evaluation system has its scientific and practical weight in such specific working conditions.

The results of this research demonstrate how the application of a well-designed evaluation system through electronic learning enables the same quality of evaluation of student learning achievements as the evaluation conducted in the classic way with exams in the classroom.

REFERENCES

[1] P. Tarkar, „Impact of covid-19 pandemic on education system“, International Journal of Advances Sceiece and Technology,“ vol. 29, no. 9s, pp. 3812-3814, 2020.

[2] S. Pokhrel, R. Chhetri, „A literature review on impact of COVID-19 pandemic on teaching and learning,“ Higher Education for the Future, 8(1), pp. 133-141, 2021.

[3] J. Daniel, “Education and the COVID-19 pandemic,” Prospects 49, pp. 91–96, 2020. https://doi.org/10.1007/s11125-020-09464-3

[4] P. Paudel, „Online education: benefits, challenges and strategies during and after COVID-19 in higher education,“ International Journal on Studies in Education (IJonSE), 3(2), pp. 70-85.

[5] M.Montenegro-Rueda, A. Luque-de la Rosa, J.L. Sarasola Sanches-Serrano, J. Fernandez-Cerero, „Assessment in higher education during the COVID-19 pandemic: A systematic review,“ Sustainability, vol. 13, no. 19, 2021, https://doi.org/10.3390/su131910509

[6] S. Jacques, A. Ouahabi, T. Lequeu, „Remote knowledge acquisition and assessment during the COVID-19 pandemic,“ International Journal of Engineering Pedagogy (iJEP), 10, 2020, ⟨10.3991/ijep.v10i6.16205⟩. ⟨hal-03046424⟩

[7] S. Elfirdoussi, M. Lachgar, H. Kabaili, A. Rochdi, D. Goujdami, L. El Firdoussi, „Assessing distance learning in higher education during the COVID-19 pandemic,“ Education Research International, vol. 2020, Article ID 8890633, 13 pages, 2020. https://doi.org/10.1155/2020/8890633

ICT Support for Successful Face-to-Face and Online Teaching Process

Trpimir Alajbeg, Mladen Sokele, Sanja Morić and Frane Brkić
Department of Electrical Engineering
Zagreb University of Applied Sciences, Zagreb, Croatia
trpimir.alajbeg@tvz.hr, mladen.sokele@tvz.hr, sanja.moric@tvz.hr, frane.brkic@tvz.hr

***Abstract* - ICT as a technical support is an important factor for improving the success of the teaching process. Requirements for ICT support are elaborated in detail for Face-to-Face, asynchronous and synchronous Online teaching process as well as for Blended teaching environment used in the post-pandemic period as a selection of the best practices aforementioned teaching methods. The success of the analyzed methods of teaching processes was quantified via key performance indicators for the effectiveness and efficiency of the teaching / learning process in the case of the undergraduate and graduate study of Electrical engineering at the Zagreb University of Applied Sciences.**

Keywords* - *Effectiveness of teaching, Efficiency of teaching, Key performance indicators, ICT support for teaching activities, Biclustering, Jaccard index

I. INTRODUCTION

With the aim of determining the factors that influence the effectiveness of teaching and especially the motivation of both students for learning and teachers for teaching, an internal scientific project was launched at the Zagreb University of Applied Sciences (*Tehničko veleučilište u Zagrebu* - TVZ): *Research on the influence of synchronous and asynchronous forms of teaching on the success of electrical engineering students*. Based on the results of the project research for undergraduate and graduate studies in electrical engineering, factors and guidelines were determined for improving the success of Online teaching as well as for improving the quality of teaching that take place again face-to-face (F2F) [1], [2], [3].

Contemporary teaching activities require intensive use of ICT support, which was particularly evident in the (sudden) transition from F2F teaching to Online teaching due to the Covid-19 pandemic which required the establishment of new and adaptation of existing ICT systems to support teaching [1], [4]:

- Video conference platform (*MS Teams*);
- New version of LMS (learning management system);
- Integration of work authentication software (*Turnitin*) with the LMS system;
- Upgrading the options of the intranet portal intended for teaching (*mojTVZ*) etc.;
- Software for emulating laboratory exercises.

The pre-pandemic teaching process mainly consists of F2F manner with asynchronous Online teaching enhancement in the form of available online materials for students. Some classes were conducted synchronously online via the LMS, but only for laboratory exercises, exams and attendance records.

During pandemic period teaching was organized synchronously online and hybrid (F2F with synchronous online video streaming). In addition, lectures were video recorded thus creating new teaching materials for asynchronous online learning as well as more intensively use of LMS. Post-pandemic teaching process is again mainly F2F with the supplements resulting from the best experiences from Online teaching, presented in Fig. 1.

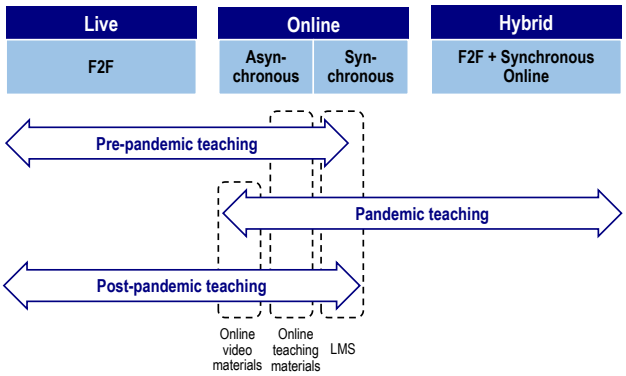


Fig. 1. Teaching process at the study of Electrical engineering at the Zagreb University of Applied Sciences (TVZ)

The requirements for ICT support have been elaborated in detail, depending on the form of teaching (lectures, classroom, laboratory and seminar exercises) and teaching activities (preparation and conducting of lessons, testing of students' acquired knowledge, etc.) in Section II. In Section III. particular examples of ICT systems used for F2F and Online teaching are given. In Section IV. data collection and similarity analysis of the values of key performance indicators (KPI) are presented.

II. ICT SUPPORT FOR TEACHING ACTIVITIES

Based on the analysis of the teaching methods and forms, for the courses held by the authors, Tab. I. was compiled with a list of ICT support items related to each teaching activity depending on the teaching process before, during and after the pandemic period. The importance and / or intensity of use is indicated with the appropriate symbols.

The new ICT support used intensively for Online teaching during the pandemic period was: Video streaming and conferencing, Video material repository, Delivery of documents during exams, Software for emulating laboratory exercises, and on a smaller scale: Intelligent board / graphics

tablet and Instant messaging among students for collaborative tasks.

TABLE I. ICT SUPPORT FOR PRE-PANDEMIC, PANDEMIC AND POST-PANDEMIC PERIOD

ICT support	Teaching / Learning activities: Teacher preparation	Lectures and exercises	Lab. and seminar exercises perform.	Student self-preparation for exams	Performing of exams
Course administering and conducting	○○ ●● ●●		○○○ ●●● ●●●		○○○ ●●● ●●●
Announcement writing	○ ● ●	○ ● ●	○ ● ●		○ ● ●
Classroom reserving		○○ ●●	○○ ●● ●●		○○ ●●
Document creating	○○○ ●●● ●●●				○○○ ●●● ●●●
Projector using		○○○ ●●●	○ ●		
Video streaming and conferencing		●●●	●●		●●●
Intelligent board / graphics tablet using		●● ●●			
Teaching material repository	○○○ ●●● ●●●	○○○ ●●● ●●●	○○○ ●●● ●●●	○○○ ●●● ●●●	⊘ ⊘ ⊘
Video material repository	●●●	●●● ●●●	●● ●●	●●● ●●●	⊘ ⊘
Document delivering			●● ●●● ●●●		●●● ●
Online exams, student work originality checking			○ ● ●		○ ●●● ●●
SW for lab exercises			○○ ● ●●●		
SW for emulating lab exercises			●●● ●	●● ●●	
Instant messaging			●●		⊘ ⊘ ⊘
Legend	○ Pre-pandemic period ● Pandemic period ● Post-pandemic period ⊘⊘ Not allowed *Intensity of use* Three symbols High Two symbols Moderate One symbol Low				

In post-pandemic period ICT support, includes best practices from Online teaching, such as: Video material repository for previously recorded lectures as new teaching material and Software for emulating laboratory exercises, both for self-preparation of students at home. Online exams are more frequent, resulting in more intensive use of the LMS, as well as verification of the originality of student's reports submitted in the text form.

III. ICT SUPPORTING SYSTEMS

Particular examples of ICT systems used to provide ICT support for F2F and Online teaching activities of electrical engineering courses at TVZ are listed in Tab. II.

TABLE II. ICT TEACHING / LEARNING SUPPORTING SYSTEMS

ICT support	ICT systems: MS 365	Intranet portal	LMS	Turnitin	MS Teams	Specific Apps*
Course administering and conducting	●●●	●●	●●●		●●	
Announcement writing		●●●	●			
Classroom reserving		●●●				
Document creating	●●●		●●			●●●
Video streaming and conferencing					●●●	
Intelligent board / graphics tablet using						●●●
Teaching material repository		●●●	●●		●	
Video material repository					●●●	
Document delivering		●●●				
Online exam, student work originality checking		●●	●●●	●●		
SW for lab exercises						●●●
SW for emulating lab exercises						●●●
Instant messaging					●●●	●●●
Legend	*Intensity of use* ●●● High ●● Moderate ● Low ** Note: for PC and/or mobile devices*					

The success of ICT supporting systems is evaluated through the following courses which are in the scope of this research:

The undergraduate study

1. *Personal computers in electrical engineering* (**PCEE**) in winter semester of academic year 2019/2020 held F2F and in winter semester of academic years 2020/2021 and 2021/2022 held Online;

2. *Signals theory and processing* (**STP**) in winter semester of academic year 2019/2020 held F2F and in winter semester of academic years 2020/2021 and 2021/2022 held Online;
3. *Information theory and coding* (**ITC**) in summer semester of academic year 2018/2019 held F2F and in summer semester of academic years 2019/2020 and 2020/2021 held Online, but in summer semester of academic year 2021/2022 held F2F again;
4. *Digital circuits* (**DC**) in summer semester of academic year 2018/2019 held F2F and in summer semester of academic years 2019/2020 and 2020/2021 held Online, but in summer semester of academic year 2021/2022 held F2F again.

The graduate study

5. *Protection of the environment and the quality of life* (**PEQL**) in winter semester of academic year 2019/2020 held F2F and in winter semester of academic years 2020/2021 and 2021/2022 held Online.

The examples of ICT support systems best practices that were primarily used during the transition to Online teaching, also proven useful in the post-pandemic F2F teaching period, are given below:

Intranet portal

The portal is available to all teachers and students. The teaching materials are visible to students and teachers of the entire TVZ, which allows better information exchange. The portal is the result of in-house development as well as its maintenance and enhancements. It enables: news, notifications, course curriculum info, timetable management, student service information, reservations of classrooms, dates and groups. The easy-to-use repository management enables the submission and placement of teaching materials (Teacher => Student) as well as delivery of student works (Student => Teacher).

Learning management system (LMS)

The Moodle system is open-source software for learning management which provides numerous useful options for teaching and learning process. In our case, it is mostly used for: record of attendance, knowledge and skill assessment, record of exam results, notifications, repository of teaching materials, delivery of student works, student work originality check with Turnitin software. During Online teaching, it was used intensively, and teachers adapted LMS to their courses and needs.

Turnitin

It is mandatory for checking the originality of final and diploma theses, and it is proven useful for checking all written works of students, for example: seminar papers, program codes, laboratory reports. In TVZ, it is integrated with LMS [5].

MS Teams

MS Teams is Microsoft's video conferencing platform. It enables the storage of documents in the MS Sharepoint cloud, so it does not consume the TVZ's storage systems. In this way, video recordings of lectures are stored, which requires about 15 GB per course (about 4 MB per minute of video recording). It also supports attendance records, a repository of teaching materials and collaboration in subteams within course teams. Therefore, it is suitable for group solving of seminar and laboratory exercises, when the cooperation of students is required.

Specific software for emulating laboratory exercises

The software suite used to emulate the lab environment at home consists of:

Audacity - free software under General Public License for recording, reproduction, processing, generation and analysis of audio recordings and programming in the Nyquist programming language. It is also possible to develop modules (*plug-ins*) for Signal analysis (multimeter emulation), Signal quantization, Generation of digital transmission signals with additive noise, Signal decimation, Amplitude histogram and aperture, etc.

Logisim - free software under General Public License for simulating digital circuits. Simple implementation of virtual laboratory exercises in form of XML files enables additional interventions such as protection of templates.

Multisim - electronic schematic capture and simulation software tool (Electronic design automation), also available on mobile devices.

GeoGebra - free dynamic math software under General Public License that includes geometry, algebra, tables, graphing, statistics and calculus. It can easily be adapted for tasks and demonstration exercises in electrical engineering.

IV. DATA COLLECTION AND KPI ANALYSIS

The analytical part of the research is based on extensive, multi-year data related to lectures and exams at undergraduate and graduate studies in Electrical engineering. In total, data were collected for 1 554 full-time study students for courses in the scope listed in Chapter III.: in pre-pandemic period 544 students, during pandemic period 888 students and in post-pandemic period 122 students.

One of the goals of paper is to determine variations of key performance indicators for evaluating the effectiveness and efficiency for different forms of teaching (F2F, Online, Blended) [6]. Additionally, the question arises whether these variations can be reduced by using appropriate ICT support.

According to previous research [7], the following Key Performance Indicators (KPIs) for the teaching/learning process were taken into account:

- Average exam grade,
- Exam pass rate,
- Share of top grades,
- Early exam pass rate and
- Success in taking exams.

KPIs definitions are presented in Tab. III. For the courses in research scope, the values for all the listed KPIs were calculated for the pre-pandemic, pandemic and post-pandemic period, and they are presented in Tab. IV.

TABLE III. KEY PERFORMANCE INDICATORS FOR TEACHING / LEARNING PROCESS IN SCOPE

KPI_i	Name	Description)
KPI_1	Average exam grade	Average grades for all students who successfully passed the exam of the course during the academic year. Grades range from 2 (lowest passing grade) to 5 (highest possible grade). This KPI is measure of ***effectiveness***.
KPI_2	Exam pass rate	The number of students who successfully passed the exam of course divided by the number of enrolled students in the course during the academic year. This KPI is measure of ***efficiency***.
KPI_3	Share of top grades	The number of students with top grades: 4 and 5 divided by the number of students who passed the exam of course during the academic year. This KPI is measure of ***effectiveness***
KPI_4	Early exam pass rate	The number of students who successfully passed the early exams of course divided by the number of all students who successfully passed the exam of course during the academic year. Early exam terms include continuous course assessment via midterm exams and exams within 4 weeks after the end of lectures. This KPI is measure of ***efficiency***.
KPI_5	Success in taking exams	The number of students who successfully passed the exam of course divided by the total number of exam attempts of course in academic year (AY). This KPI is measure of ***efficiency***.

TABLE IV. VALUES OF KEY PERFORMANCE INDICATORS FOR COURSES IN THE SCOPE

Course	AY	Period	KPI_1	KPI_2	KPI_3	KPI_4	KPI_5
PCEE	2019/20	Pre-pandemic	2.605	0.783	0.062	0.955	1.000
STP	2019/20		2.609	0.687	0.043	0.761	0.676
ITC	2018/19		2.761	0.826	0.085	0.958	0.910
DC	2018/19		2.653	0.890	0.116	0.917	0.818
PEQL	2019/20		4.552	1.000	0.828	0.828	1.000
PCEE	2020/21	Pandemic	2.594	0.618	0.109	0.930	1.000
PCEE	2021/22		2.699	0.649	0.124	0.903	0.724
STP	2020/21		2.533	0.692	0.089	0.911	0.600
STP	2021/22		2.500	0.556	0.067	0.667	0.588
ITC	2019/20		2.511	0.734	0.106	0.638	0.701
ITC	2020/21		2.615	0.736	0.103	0.974	0.765
DC	2019/20		2.964	0.775	0.273	0.927	0.846
DC	2020/21		2.773	0.838	0.136	0.909	0.815
PEQL	2020/21		4.425	0.909	0.825	1.000	1.000
PEQL	2021/22		4.490	0.961	0.918	0.918	1.000
ITC	2021/22	Post-pandemic	2.714	0.667	0.143	0.929	0.718
DC	2021/22		3.015	0.850	0.279	0.882	0.840

In order to find out if there are deviations in KPI values depending on the teaching period (such as a significant decrease or increase), the data were analyzed using the biclustering technique. Biclustering technique simultaneously clusters both rows and columns of a data (in our case, data on the *y* and x axes), allowing for the identification of subsets of objects and features that exhibit coherent patterns of behavior [8]. Average exam grade (KPI_1) was chosen as the *x*-axis, and each remaining KPI_2 – KPI_5 was analyzed separately, through their value on the *y*-axis. For each KPI separately, an *x*-*y* plot with points showing the KPI values for all five courses in the scope in three teaching environments (pre-pandemic, pandemic and post-pandemic) are shown in Fig. 2 – Fig. 5.

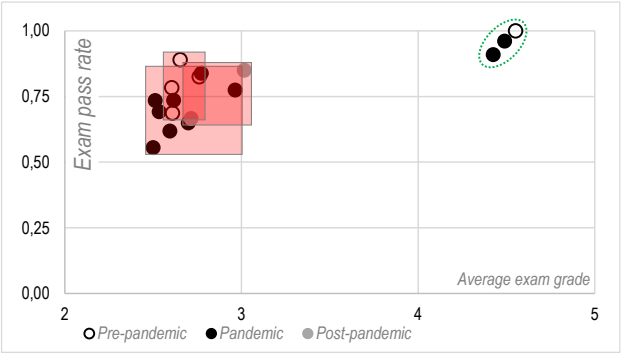


Fig. 2. KPI_2 Exam pass rate (*y* axis) – depending on KPI_1 Average exam grade (*x* axis) for all courses and teaching periods

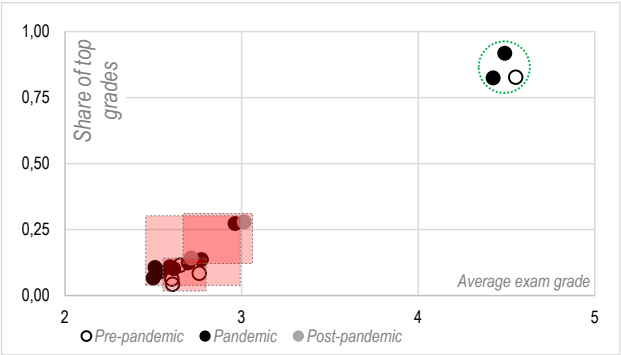


Fig. 3. KPI_3 Share of top grades (*y* axis) – depending on KPI_1 Average exam grade (*x* axis) for all courses and teaching periods

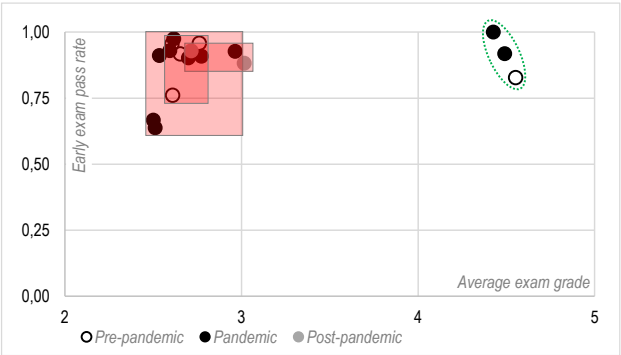


Fig. 4. KPI_4 Early exam pass rate (*y* axis) – depending on KPI_1 Average exam grade (*x* axis) for all courses and teaching periods

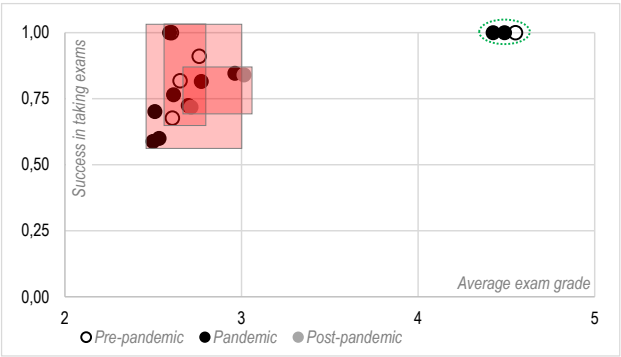


Fig. 5. KPI_5 Success in taking exams (*y* axis) – depending on KPI_1 Average exam grade (*x* axis) for all courses and teaching periods

In addition to points, rectangles are drawn on the *x-y* plots, that cover all the KPI values for a particular teaching period with their surface, which will be described in continuation.

The first feature that can be seen on the *x-y* representation is that the values of the KPIs are grouped into two clusters: the values for the undergraduate courses (covered by red rectangles) and the values for the graduate course (circled by a green line). This is to be expected, since graduate study are attended by the most successful students and therefore the average grades are generally higher. Another feature that can be noticed is that for each KPI the values for all undergraduate courses are mixed, without a disjunctive distribution according to the form of teaching.

In order to determine the closeness of the KPI values according to the form of teaching, a measure of similarity was calculated using the Jaccard index [8]. The Jaccard index measures similarity between finite sets, and it is defined as the size of the intersection divided by the size of the union of the sample sets (1). In our case:

$$J_i(R_A, R_B) = \frac{|R_A \cap R_B|}{|R_A \cup R_B|} \quad (1)$$

Where:

- i is index of observed KPI_i, i = 2, 3, 4 and 5;
- *A* and *B* are different forms of teaching;
- Denominator: the union of the area of the rectangle R_A that covers all the points of the observed KPI_i for the *A* form of teaching and the rectangle R_B that covers all the points for the *B* form of teaching.
- Numerator: the intersection of the sets i.e., the common (overlapping) surface of rectangles R_A and R_B.

TABLE V. SIMILARITY MEASURE BY JACCARD INDEX

i	KPI_i	Pre-pandemic vs. Pandemic	Pre-pandemic vs. Post-pandemic	Pandemic vs. Post-pandemic
1	Exam pass rate	17 %	10 %	30 %
2	Share of top grades	8 %	0 %	31 %
3	Early exam pass rate	20 %	5 %	7 %
4	Success in taking exams	26 %	7 %	15 %
	Average	**18 %**	**5 %**	**21 %**

The results in Tab. 5 show that the KPI values are more similar between the pre-pandemic period and the pandemic period (i.e., when there was a transition from F2F to Online, similarity is 18 % in average), than between the pre-pandemic period and the post-pandemic period (both periods are F2F, similarity is only 5 % in average). The obtained high similarity of KPI values during the transition from F2F to Online teaching proves that the transition was successful and that, in addition to the efforts made by students and teachers, the selected ICT support was effective during the pandemic period.

Additionally, the high similarity of KPI values between the pandemic and post-pandemic periods show that the use of new ICT support, which proved to be effective, continued to be effective in the post-pandemic period.

V. CONCLUSION

Based on experience in Face-to-Face (F2F) and Online teaching, as well as in Blended teaching that resulted from the best practices of the mentioned ones, the paper presents the required ICT support for each teaching activity.

The success of teaching is verified through KPI for efficiency and effectiveness. Efficiency and effectiveness of teaching for each teaching method is analyzed via the following key performance indicators (KPIs):

- Effectiveness: *Average exam grade* and *Share of top grades*;
- Efficiency: *Exam pass rate*, *Early exam pass rate* and *Success in taking exams*.

KPIs were collected for five courses encompassing 1 554 students of undergraduate and graduate studies in Electrical engineering at Zagreb University of Applied Sciences.

By using Jaccard index, it was shown that the efficiency and effectiveness KPIs from pandemic period are more similar to KPIs in pre-pandemic period than KPIs from post-pandemic period to KPIs in pre-pandemic period despite both being F2F. In other words, teaching in the post-pandemic period, even though it is F2F, has adopted the best practices from Online teaching process. The above confirms that, in addition to human efforts in teaching and learning, ICT support for Online and Blended teaching is optimally selected.

The overall conclusion could be that contemporary ICT support is very important in today's education since it increases the availability of learning for everyone, anytime and on any device and improves personalization, collaboration, administration, knowledge assessment and innovation.

Future steps could be activities that will encourage and support teachers for more advanced use of the LMS platform and other ICT related tools: social networks as an attractive way to exchange information; gamification of particular course parts and teacher-student co-creation of learning materials.

REFERENCES

[1] M. Sokele, F. Brkić, T. Alajbeg, S. Morić: Faktori uspješne online nastave, Konferencija o elektroničkim komunikacijskim tehnologijama – KOM 2022, Zagreb, 2022.

[2] Sanja Morić, Trpimir Alajbeg, Mladen Sokele, Frane Brkić: Kvantitativne metode vrednovanja kvalitete nastave na studiju elektrotehnike Tehničkog veleučilišta u Zagrebu, 5. Dani obrazovnih znanosti (DOZ), Knjiga sažetaka pp. 90-91, IDIZ, Zagreb, 2022.

[3] M. Sokele, T. Alajbeg and I. Lujo: The importance of face-to-face teaching at fundamental telecommunication courses, Proc. of the 15th International Conference on Telecommunications ConTEL 2019, pp. 1 – 5.

[4] M. Sokele, T. Alajbeg, F. Brkić: The impact of distance learning on student success for electrical engineering professional courses, 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020, pp. 617-621, doi: 10.23919/MIPRO48935.2020.9245315.

[5] M. Sokele, S. Morić, V. Zuppa Bakša: Development of new complementary metrics for the authenticity of student papers, Proceedings of 44rd International Convention MIPRO 2021 / Computers in education (CE), pp. 730-734, doi: 10.23919/MIPRO 52101.2021.9596857

[6] S. Singh, D. H. Rylander, T. C. Mims: Efficiency of Online vs. Offline Learning: A Comparison of Inputs and Outcomes, International Journal of Business, Humanities and Technology Vol. 2 No. 1; 2012, pp. 93 – 98

[7] S. Morić, M. Sokele, T. Alajbeg, F. Brkić: Evaluating the effectiveness of Online learning for undergraduate and graduate studies of electrical engineering, 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), 2022, pp. 729-734, doi: 10.23919/MIPRO55190.2022.9803763.

[8] A. Kasim, Z. Shkedy, S. Kaiser, et al.: Applied Biclustering Methods for Big and High-Dimensional Data Using R, Chapman & Hall/CRC, London, UK, 2017.

Programmable Questions in Edgar

I. Mekterović *, Lj. Brkić* and V. Krstić**
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
** The Fifth Gymnasium, Zagreb, Croatia
igor.mekterovic@fer.hr

***Abstract* - Automated programming assessment systems (APAS) are a valuable tool that is growing in popularity, particularly in the field of computer science education. They can provide quick and objective assessment and feedback to the programming assignments – those that receive source code as a response. Most APASs treat code as a black box and employ dynamic analysis to assess code. Dynamic analysis is straightforward, easily implemented, explainable, and works well in most situations. Edgar is a comprehensive, state of the art APAS, that has been used daily and has evolved for the past six years. This paper examines the pipeline used by Edgar to assess programming questions and presents our enhancements to the traditional dynamic analysis - programmable templates and scripts. Templates enable customized question texts based on the programmable model, so that each student can receive personalized variation of the question. Personalized questions are a great way to fight potential academic dishonesty. Scripts are plugged into the assessment pipeline after the dynamic analysis and can override the default grade by examining some other aspect of the program. We also offer our thoughts on upcoming plans to include generic static analysis as we move closer to a unified assessment pipeline.**

Keywords - APAS; automated assessment; dynamic analysis; programmable questions, CS education

I. INTRODUCTION

Automated programming assessment systems (APAS) is an information system used in educational environments to (semi)automatically assess students' answers to programming questions. They typically also support other types of questions, such as multiple-choice questions, and provide monitoring and logging of exams, various statistics, and visualizations, etc. Nowadays, they are typically implemented as web applications. A detailed overview of comprehensive APAS features can be found in our previous work [1]. APASs provide fast and objective assessment and feedback but are less capable of producing partial assessments, especially for code that cannot be executed (e.g., does not compile). Manual assessment of code by teachers is still considered the gold standard, but manual assessment is difficult and very time consuming for the teachers. Nevertheless, APASs are being increasingly developed and used (great recent overviews can be found in [2] and [3]), and it is now hard to imagine a larger computer science course that does not use APASs to one degree or another. At the Faculty of Electrical Engineering and Computing, we have been developing and actively using a state-of-the-art APAS called Edgar for six years, and it has become an indispensable part of many computer science courses, relying on it for part or even all of the assessment. For example, in the previous academic year 2021/22, more than 57,000 exams were administered using Edgar, containing approximately 340,000 questions from 19 different courses. In this paper, we present two improvements made to the classic question assessment pipeline. First, we introduced *templates* to enable programmatic question content generation so that each student could obtain a personalized question variation. Second, we enable custom scripts that are appended at the end of the assessment pipeline and can alter the assessment outcome. Although our focus is on programming questions, both concepts can be used on any type of question. In the following chapters, we briefly outline assessment types in APASs and in Edgar, and then present templates and scripts in dedicated chapters, followed by brief discussion and conclusion.

This work was supported by the European Regional Development Fund under Grant KK.01.1.1.01.0009 (DATACROSS).

II. CODE ASSESSMENT IN AUTOMATED PROGRAMMING ASSESSMENT SYSTEMS

Techniques for evaluating program code can be roughly divided into **dynamic** and **static** analyses. Dynamic analysis is much more prevalent, and most APAS bases their assessment on dynamic analysis.

A. Dynamic analysis

In dynamic analysis, the program is treated as a *black box* and tested by running it with different input data and analyzing the outputs. For one program, it is possible to define N test-cases (which can affect the sore with different weights), so it is also possible to achieve partial evaluation. However, the disadvantage of dynamic analysis is that the program must be able to run, so programs that cannot be compiled, for example, will not be able to run and will be rated with the worst rating. On the other hand, this approach is widely applicable because the same principle is applied for any programming language, it is only necessary to establish the process of compilation and execution for a language and to ensure that the execution of the program does not have negative side-effects, i.e., it is necessary to execute the program in a *sandbox* - protected and limited environment. Edgar uses a separate code execution system named Judge0 [4] for secure and scalable code execution.

Dynamic analysis mainly tests the functional aspects of the program, but it is also possible to analyze some non-functional attributes, such as efficiency (CPU, time, memory, etc…). However, latter is much more difficult to implement because it is necessary to ensure that all programs have exactly the same resources at their disposal (CPU, disk, ...), which is almost impossible in virtualized

environments, so it is advisable to repeat the analysis several times and statistically process the results.

B. Static analysis

In static analysis, the program is analyzed without executing the program. It is used to check style, syntax errors, various metrics (number of lines of code, cyclomatic complexity, etc.), program design and structure, special properties (e.g., use of an expression), and even plagiarism detection [5]. It should be noted that the manual review of program code written on paper is actually a static analysis performed by a teacher through visual inspection. In regard to assessment and grading, the typical workflow is as follows: source code is parsed to construct an abstract syntax tree, which is then transformed into a graph representation and compared to a set of reference graphs using graph similarity measures. The best fit is used to find the potential differences, and an overall assessment is given. Static analysis is very dependent on the programming language, and it is difficult to implement it uniformly. Typically, the literature contains works focused on a specific programming language. In addition, it is more difficult to design tasks that are checked in this way (e.g., to provide the set of all correct solutions). The field of application of static analysis is smaller than that of dynamic analysis.

Finally, it should be noted that static and dynamic analysis can be combined into a joint - hybrid evaluation method.

III. ASSESSMENT IN EDGAR

Dynamic analysis has been used in Edgar since the beginning. It was initially developed only for the SQL programming language, where the resulting data sets from two SQL commands were compared, and then dynamic analysis was implemented for standalone languages. In both approaches, the program code is separated into three parts: prefix, main part, and suffix. Only the main part represents the solution that is expected from the student, and the optional prefix and suffix allow the teachers greater options in formulating and testing the questions. Namely, before executing the code, those three parts are merged into one, which means that the teacher can inject their own code before and/or after the student's code. Therefore, for example, it is possible to ask a student to write only a function, and the teacher injects his own code that calls and tests that function. Similarly, it is possible to ask a student to submit just one Java class, which can then be instantiated and/or explored using reflection, and thus even non-functional code properties can be evaluated, e.g., whether the class conforms to naming conventions, etc.

Dynamic evaluation of a question is carried out with an arbitrary number of test-cases. Additionally, test-cases can have an arbitrary weight, that is, they can carry a different number of points. The weight is expressed as a percentage that is subtracted from the initial 100% and is called the penalty percentage (PP). Therefore, the correctness is calculated as follows:

$$correctness := \max(0\%, 100\% - \sum_{fail} PP(fail))$$

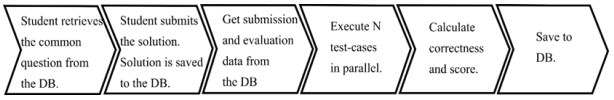


Figure 1. Edgar's initial assessment pipeline

The calculation of the grade itself is carried out in two stages: first, the correctness is evaluated on a scale of 0% - 100% as described, and then the number of points is assigned from the correctness based on the assigned grading model. The grading model is defined as a triplet:

$$gm := (c, i, e)$$

where *c*, *i*, and *e* are values for correct, incorrect and empty answers respectively. As a rule, the incorrect score is set to a negative value for multiple-choice questions to discourage students from guessing. For programming questions, it is always set to zero. Given the grading model, the final score is calculated according to:

$$score = \begin{cases} \boldsymbol{gm.e} & answer\ is\ empty \\ \boldsymbol{correctness} * (\boldsymbol{gm.c} - \boldsymbol{gm.i}) + \boldsymbol{gm.i} & answer\ is\ non\ empty \end{cases}$$

Further details and examples can be found in our previous work [1].

Figure 1. shows the simplified Edgar's dynamic-evaluation pipeline, as used in most APAS-es. Technical details are omitted, and, for the sake of simplicity, it addresses a single question. In reality, an exam consists of an arbitrary number of questions. This model works well, but after several years of use, ideas emerged for some more demanding evaluation scenarios. With this type of organization, all students get the same common question with the same content. If one wants to reduce the possibility of plagiarism, it is necessary to make several similar versions of the same question and then assign them to students by random selection. Although this is possible, such methodology is tiresome for the teacher, error-prone and requires significant additional effort. This problem is addressed by the introduction of **templates** - the possibility to programmatically generate or adjust the text of the assignment for each student. The second improvement that is presented here refers to the possibility to programmatically manipulate the outcome of the assessment performed by Edgar, that is, to override the default assessment at the end of the evaluation pipeline.

Both concepts are described in the following chapters.

IV. QUESTION TEMPLATES

Templates enable the generation of customized question content based on the programming model. They are inspired by the MVC pattern where content is generated in a view based on a model, while the whole process is orchestrated by the controller. MVC frameworks typically generate HTML using programming constructs interleaved with HTML code. The method of interleaving, that is, the syntax that can be used in the view definition depends on the framework and the so-called *view/templating engine* that is used to produce the final HTML. For example, in the .Net MVC framework Razor syntax is used [6], the Node.JS express development framework allows various template engines [7](EJS, handlebars, etc.), etc..

To introduce such a concept into the existing model in Edgar, it is necessary to:

- generate each student's **data object** only **once** when the exam is **first** started (question retrieved)
  - the data object is defined in design time by the teacher using the JavaScript programming language.
- generate the **custom question content**, and
- **store both** the data model and question content in the database

In other words, when the student requests the question, the database is checked for their existing custom question, and if there is no question, the above-mentioned steps are taken. Ultimately, in both cases the question is retrieved from the database. The concept is illustrated with a simple example: Figure 2. and Figure 3. show the definition of a templated multiple-choice question where the assignment is to add two numbers. The model (data object) here consists of two random numbers x and y and three answers: a1, a2 and a3, with a1 being the correct answer:



Figure 2. Template data object definition and a random instance

The data object is constructed by invoking the `init()` method, which produces some random values, such as the one shown in Figure 2. These variables can then be used both in question text and answers, as shown in Figure 3. :

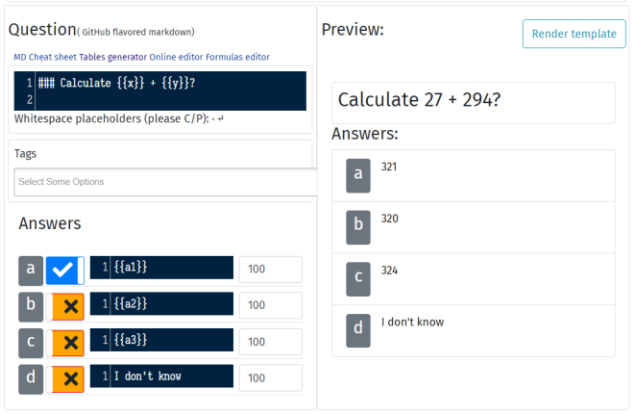


Figure 3. Template definition (left) and preview (right). Tempalate references variables x and y from the data object.

Note the "moustache/handlebar" syntax in the question definition (upper left part of Figure 3.) - string interpolation is performed with double curly braces. Edgar uses GitHub markdown for rich text content and Handlebars.js as a templating engine. Handlebars (and consequently Edgar) support more than just inserting values; it supports expressions, loops, etc. [8]. The template is rendered in the context of the given data object. The right part of Figure 3. shows the rendered template using the ad hoc instantiated data object every time the "Render template" button is clicked. The data object is constructed using a three-level hierarchy:

- **global** data object (same for all courses, cannot be changed by a teacher)
- **course** data object (can be changed by the teachers in the course)
- **question** data object (defined when the question is defined)

Global and course data objects are meant to store utility methods and variables to simplify the question's data object source code (such as the `randomInt()` method which is not part of JavaScript). For instance, if global, course and question data objects are:

Global:

```
{
  randomInt: function(minInt, maxInt) {
    return Math.floor(Math.random() * (maxInt -
minInt + 1)) + minInt;
  }
}
```

Course:

```
{
  randomBoolean: function() {
    return !!this.randomInt(0, 1);
  },
  currAcYear: "2022/2023",
  someOtherConstant: 42
}
```

Question:

```
{
  init() {
    this.x = this.randomInt(13, 37);
    this.y = this.randomInt(133, 337);
    this.a1 = this.x + this.y;
    this.a2 = this.a1 - 1;
    this.a3 = this.a1 + this.randomInt(1, 5);
  }
}
```

Then, the final data object is assembled as follows:

```
{
  randomInt: function(minInt, maxInt) {
    return Math.floor(Math.random() * (maxInt -
minInt + 1)) + minInt;
  }
  randomBoolean: function() {
    return !!this.randomInt(0, 1);
  },
  currAcYear: "2022/2023",
  someOtherConstant: 42
  init() {
    this.x = this.randomInt(13, 37);
    this.y = this.randomInt(133, 337);
    this.a1 = this.x + this.y;
    this.a2 = this.a1 - 1;
    this.a3 = this.a1 + this.randomInt(1, 5);
  }
}
```

Which when init() is called, will produce e.g.:

```
{
  "currAcYear": "2022/2023",
  "someOtherConstant": 42
  "x": 31,
  "y": 203,
  "a1": 234,
  "a2": 233,
```

```
  "a3": 239
}
```


V. OVERRIDE SCRIPTS

Scripted questions enable teachers to append custom scoring codes/scripts at the end of the assessment pipeline. Scripts were inspired by *checker* programs in competitive programming: "In programming competition environment, a checker is a program written for the purpose to check the output of the contestant's program for a task that has many solutions. Usually, a checker is written manually as needed."[9]. The purpose of checkers, or *override scripts* as they are called in Edgar, is to enable custom grading in advanced scenarios. Consider the following example assignment:

Assignment:
Print "hello world" in C programming language without using the `printf` function.

The output produced by the student's solution – "hello world" can be checked with dynamic testing, but the *non-functional* program property "without using the printf" cannot. However, a script that would have access to submitted code could simply search for "printf" in the code and override the score if "printf" is found. If not found, the score should be inherited from the dynamic analysis. Obviously, the script must have access at least to code and dynamic analysis score.

Accordingly, the scripts are embedded at the end of the pipeline, as shown in Figure 4.

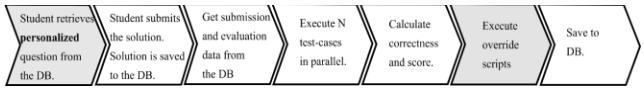


Figure 4. Assessment pipeline in Edgar with newly added features (gray): templates (at the beginning) and override scripts (at the end).

If scripted questions are used, the assessment takes place as follows:

- Edgar performs the built-in default assessment and scoring and forwards all the relevant data further down the pipeline.
- A custom evaluation script written by the teacher is executed:
  - The script has all the data from previous steps at its disposal (question definition, student's answer, evaluation score from dynamic analysis, etc.)
  - The script acquires the data object context. If templates were not used, then the data object is constructed from the course and question data object, thus having utility methods at disposal.

The script must implement the `getScore()` method which returns **the score object**. Thus, the teacher can alter the default score object that was received as an argument.

Figure 5. shows the script for this example and Figure 6. shows the resulting score object. Note that scripts can be used not only to modify the score object, but also to provide helpful hints to students. Actually, having all the data at their disposal, the teacher can program anything.

The script can be written in arbitrary programming language (supported by Edgar, e.g. C, Java, Python…) or in JavaScript, with the difference being:

- JavaScript (recommended) script is evaluated and executed *in-process*, at the web-server machine serving the request. All relevant data is embedded in the script data object and is available simply as `this.something`, no parsing or evaluating necessary.
- Other programming languages (Java, C, …) - scripts are being sent to the Judge0 code execution engine – the very same used to evaluate students' code. This means that Edgar invokes an HTTP request to Judge0, where the code is compiled and executed in the sandbox. Relevant data are serialized as JSON and sent to the script as the single stdin argument. The script/program must deserialize the input to do anything useful. The program must return a new score, again - serialized as JSON, by writing that string to stdout. Edgar subsequently deserializes the stdout and proceeds.

Custom scoring script

```
{
    getScore() {
        let answer = this.getCodeAnswer(); // helper method from the global data object
        answer = this.removeComments(answer); // also, helper method from the global data object
        if (answer.indexOf("printf") >= 0) {
            return this.getIncorrectScore("You cannot use printf"); // helper method to construct the score
        } else {                                                    // according to the grading model
            this.score.hint += " I was also here at " + (new Date()).toISOString();  // just for fun
            return this.score;
        }
    }
}
```

Figure 5. Override script checking for usage of printf function. If found, an incorrect score is returned. Otherwise, the score is unchanged, and the timestamp is appended to the hint.


```
Score is:
{
  "is_correct": true,
  "is_incorrect": false,
  "is_unanswered": false,
  "is_partial": false,
  "score": 1,
  "score_perc": 1,
  "hint": "Correct. Well done! I was also here at 2023-01-
31T09:29:02.893Z",
  "c_outcome": [
    {
      "input": "1",
      "percentage": "100.00",
      "hint": "Correct. Well done!",
      "output": "hello\n",
      "expected": "hello",
      "mode": "check elements order : false, case sensitive : false, ignore
whitespace : true",
      "isCorrect": true,
      "comment": null,
      "stderr": null,
      "is_public": false
    }
  ]
}
```


▶ Click to see the entire object...

Figure 6. Score object produced by the script in Figure 5. for correct submission.

For all these reasons, remote procedure calls being dominant, JS scripts are much faster and easier to write. On the other hand, in certain advanced scenarios, where override scripts need complex libraries to check the program (e.g., NumPy, OpenCV, etc.) the communication overhead is well worth the cost.

Both templates and override scripts can be used on any type of question, not just programming questions. Our final example shows a combination of templates and scripts on a **free-text** question. The assignment is a classic first-year programming task – provide a hexadecimal representation of a number according to the IEEE 754 standard. First, to provide each student with their custom number, we define the data object as follows:

```
{
    toHex(double) {   // convert number to IEEE754 hex
      const BYTES = 4;
      const buffer = new ArrayBuffer(BYTES);
      const float32Arr = new Float32Array(buffer);
      const uint32Array = new Uint32Array(buffer);
      float32Arr[0] = double;
      const integerValue = uint32Array[0];
      const integerBitsHex = integerValue.toString(16);
      return integerBitsHex;
    },
    init() {
        this.x = this.randomInt(100, 200);
        this.x += this.randomBoolean() ? 0.5: 0;
        this.x += this.randomBoolean() ? 0.25: 0;
        this.x += this.randomBoolean() ? 0.125: 0;
        this.x += this.randomBoolean() ? 0.0625: 0;
        this.hex = this.toHex(this.x);
        this.correct = this.hex.toUpperCase();
    }
}
```

```
{
  "x": 143.8125,
  "hex": "430fd000",
  "correct": "430FD000"
}
```

Figure 7. Question data object definition and one random instance. The object provides the random decimal number and corresponding hexadecimal IEEE754 representation which will be required of the student.

The `init()` method chooses a random integer in the [100, 200] range and then randomly adds a few negative powers of two so that the decimal number converts nicely to binary (unlike. e.g., 0.3 which has an infinite number of decimals). The program also immediately calculates the correct answer and stores it in the data object as the "`correct`" variable. The template is defined simply as:

> Provide a hexadecimal representation of the number {{x}} according to the IEEE 754 standard.

which will render different numbers for different students. Finally, the override script must check the student's answer and compare it with the prepared correct answer in the data object (Figure 8.). Figure 9. shows a simplified UML sequence diagram of the assessment pipeline for questions using both templates and override scripts.

```
{
    getScore() {
        this.answer = this.getFreeTextAnswer().toUpperCase();
        this.answer = this.removeWhitespace(this.answer);
        if (this.answer === this.dataObject.correct) {
          return this.score;  // free texts are always correct by default
        } else {
          return this.getIncorrectScore("Failed: " + this.answer
                + "<>" + this.dataObject.correct);
        }
    }
}
```

Figure 8. Student's answer is trimmed and converted to uppercase and compared to the correct answer.

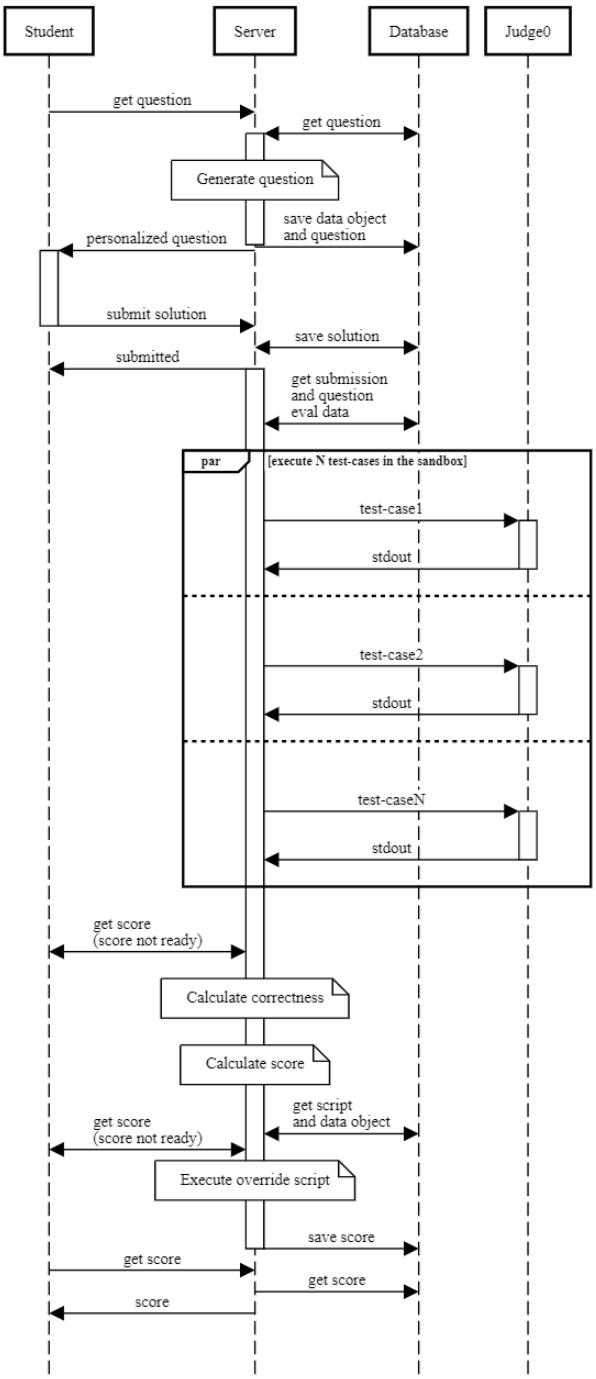


Figure 9. Simplified UML sequence diagram of the assessment pipeline for questions using both templates and override scripts

Override scripts enable teachers to parse free text answers and grade them. In this way, the teacher can prescribe the allowed nomenclature and then programmatically interpret the submitted solutions and evaluate them. Since textual representations can be prescribed for many data structures (graphs, tuples, tables, lists, etc.), this approach opens up great possibilities for (semi)automatic testing. If necessary, the grades assigned in this way can be manually reviewed and changed later, which is also supported in Edgar. We successfully use this type of assessment, where automatic assessment is followed by teacher control, in several courses at FER. Of course, automatic evaluation is of great help to the teacher, especially for correct solutions. In fact, most of the time, it is necessary to review only the answers that were declared incorrect by the automatic evaluation.

It should be noted that override scripts could be used to perform additional static analysis on the submitted code. In this way, it is even possible to create a hybrid model - static and dynamic analysis together. However, we believe that static analysis via override scripts would not be a good user experience for a teacher who would have to do too much programming. In terms of future development, our plan is to:

- Develop a separate standalone component that can perform static code analysis and return results in a standard format (e.g., SARIF [10])
- Allow question authors to include static analysis in the evaluation pipeline, and then evaluate the solution using a hybrid model or static analysis only.

In other words, our goal is to provide a unified assessment pipeline where teachers can opt-in for various assessment features and combine them in configurable ways to provide the assessment. As a side note, although we use term "pipeline", technically some steps are performed in parallel for performance reasons. Of course, these facilities can be used in other settings, not only during grading. For example, in e-learning environments, the student would be helped and guided how to fix the code and come up with a correct solution.

VI. CONCLUSION

In this article, we introduced two enhancements to the standard procedure for creating and evaluating questions in automated code evaluation systems: programmable templates and override scripts. Templates allow teachers to programmatically generate question content. In this way, a single question definition can yield multiple variations and students can receive personalized questions. Scripts allow the teacher to build upon the default grade and programmatically assign a final grade, or just a helpful comment. In addition to programming questions, both mechanisms can be used for other types of questions, which gives teachers much more freedom and opportunities in composing questions.

REFERENCES

[1] I. Mekterovic, L. Brkic, B. Milasinovic, and M. Baranovic, "Building a Comprehensive Automated Programming Assessment System," IEEE access, vol. 8, pp. 81154–81172, 2020, doi: 10.1109/ACCESS.2020.2990980.

[2] J. C. Paiva, J. P. Leal, and Á. Figueira, "Automated Assessment in Computer Science Education: A State-of-the-Art Review," ACM Trans. Comput. Educ., vol. 22, no. 3, pp. 1–40, Jun. 2022, doi: 10.1145/3513140.

[3] S. Combéfis, "Automated Code Assessment for Education: Review, Classification and Perspectives on Techniques and Tools," Software, vol. 1, no. 1, pp. 3–30, 2022, doi: 10.3390/software1010002.

[4] H. Z. Dosilovic and I. Mekterovic, "Robust and scalable online code execution system," 2020 43rd Int. Conv. Information, Commun. Electron. Technol. MIPRO 2020 - Proc., pp. 1627–1632, Sep. 2020, doi: 10.23919/MIPRO48935.2020.9245310.

[5] K. M. Ala-Mutka, "A Survey of Automated Assessment Approaches for Programming Assignments," http://dx.doi.org/10.1080/08993400500150747, vol. 15, no. 2, pp. 83–102, 2007, doi: 10.1080/08993400500150747.

[6] "Razor syntax reference for ASP.NET Core | Microsoft Learn." https://learn.microsoft.com/en-us/aspnet/core/mvc/views/razor?view=aspnetcore-7.0 (accessed Jan. 31, 2023).

[7] "Template Engines." https://expressjs.com/en/resources/template-engines.html (accessed Jan. 31, 2023).

[8] "Handlebars." https://handlebarsjs.com/ (accessed Jan. 31, 2023).

[9] R. I. Hadiwijaya and M. M. Inggriani Liem, "A domain-specific language for automatic generation of checkers," Proc. 2015 Int. Conf. Data Softw. Eng. ICODSE 2015, pp. 7–12, Mar. 2016, doi: 10.1109/ICODSE.2015.7436963.

[10] "SARIF Home." https://sarifweb.azurewebsites.net/ (accessed Jan. 19, 2023).

Inclusion of Digital Competence in Educational Physics Study in Croatia and Slovenia

F.M. Grgurin*, R. Repnik**
* University of Split, Faculty of Science Split, Split, Croatia
fgrgurin@pmfst.hr
** University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor, Slovenia
robert.repnik@um.si

***Abstract* - Technology is all around us in our daily lives, in mobile phones, televisions, computers, and smart devices. Teachers need to incorporate digital tools into their lessons and perform other activities to develop digital competence. There are three key competences that have been identified by the European Union (2019), and we believe there is an opportunity to address them in physics teacher education programs. These are mathematical competence, competence in science, technology and engineering, and digital competence. One of the important issues is experimental work in physics education, especially when we look at specific subgroups of students. For example, an ongoing challenge for teachers is to combine experimental activities with the development of different competences. We will address the questions, what are the essential knowledge, skills and attitudes related to digital competence, and do physics teachers acquire them during their studies? We focused on the inclusion of digital competence in physics teacher studies, and we examined the description of subjects in the physics curriculum at several faculties in Croatia and in Slovenia.**

Keywords – digital competence, educational physics study, study subjects

I. INTRODUCTION

The education of physics teachers should provide necessary knowledge and skills in physics, as well as the knowledge and skills for imparting knowledge to students. We are living in a time of rapid change, and after completing their education, teachers should have a basis for their work to build on through lifelong learning. The biggest changes are the possibilities and application of digital technology. It is necessary for physics teachers to acquire skills to promote the development of digital competences in students.

According to Tibor Navracsics (European Commissioner for Education, Culture, Youth and Sport, 2014-2019), there are necessary skills for coping with daily life, such as critical thinking, media literacy and communication skills, while the fundamental skills for lifelong education are digital skills, literacy, mathematics and science. The Council of the European Union issued a recommendation on eight key competences that should be encouraged and developed in the education system and continue to be developed through lifelong learning. Each competence consists of knowledge, skills and attitudes.

Digital competence includes the confident and responsible use of digital technologies for learning, work, and participation in society, as well as information and data literacy, communication, media literacy, digital content creation, safety, intellectual property, problem solving, and critical thinking. [1]

Following these recommendations, teachers should improve their level of digital competence on a daily basis. Furthermore, acquired knowledge and skills in these subjects, and their application in teaching, can help in assisting students with special needs.

In Chapter 2 we present methodology. In chapter 3, available study program documents from the University of Osijek, Rijeka, Split, Zagreb and Maribor will be analysed. Chapter 4 will describe the possible application of certain digital competencies in teaching physics at school and will give an example. The conclusion of the work will be given in the fifth chapter.

II. METHODOLOGY

The analysis in this paper relies on publicly available documents. Many documents in the public domain have been written by skilled professionals and contain valuable information [2]. According to LeCompte and Preissle, important considerations for document analysis include: the position of the researcher, choice of informant/respondents, social situations and conditions, analytic constructs and premises used, and methods of data collection and analysis [2].

According [1] digital knowledge: Digital literacy requires individuals to have an understanding of how digital technologies can facilitate communication, creativity, and innovation, while also being conscious of their potential limitations, effects, and risks. This involves having knowledge of the general principles, mechanisms, and logic underlying evolving digital technologies, as well as the basic functions and uses of various devices, software, and networks. Individuals should also possess critical thinking skills to evaluate the validity, reliability, and the impact of information and data that is accessible via digital means, and be mindful of the legal and ethical principles related to engaging with digital technologies. Digital skills enable individuals to use technology to support their social and professional goals, including active citizenship, social inclusion, and collaboration with

others. These skills include the ability to use, access, filter, evaluate, create, program and share digital content, as well as manage and protect information, data, and digital identities. Individuals should also be able to recognize and engage effectively with software, devices, artificial intelligence, and robots. Attitudes towards digital technologies should be reflective, critical, open-minded, and forward-looking, and include an ethical, safe, and responsible approach.

Official documents on the description of subjects taught in physics study programs are not standardized between universities. Certain information is not publicly available, so an analysis of the available documents was performed. The parts of the documents related to the development of digital competences are general. The assumption why they are written in such a way is that digital technology is developing very quickly, while the process of changes in course descriptions is much slower.

III. Analysis of study programs in Croatia and Slovenia

We analyzed the available documents of study programs for physics teachers in five universities in Croatia and in one in Slovenia: the University of Osijek, the University of Rijeka, the University of Split, the University of Zagreb and the University of Maribor. Students at these institutions can choose elective courses in which they acquire additional digital competences, but they will not be covered in this paper, since all students are not required to take those courses.

A. University of Osijek

At the University of Osijek, students can enroll in an undergraduate course in physics and after three years, a graduate course, to become a teacher of physics and computer science. Course descriptions for undergraduate studies can be found at [3] and for graduate studies at [4]. The graduate study of physics and information science has more subjects in the area of information science, and students are provided instruction in digital competences.

In the first semester of undergraduate study, students take Basics of Information Science, and Text Processing and Spreadsheets, in which they acquire knowledge about programming tools that will help them to continue their studies and professional development. In the second semester, they take Basics of Programming 1, in which they learn basic concepts of programming and software development. In the third semester, they take Programming Basics 2, a continuation of the course from the previous semester, and Practicum in Physics Basics 1, in which, in addition to acquiring experimental skills, they process the results using a computer. In the fourth semester, they take Practicum in the Basics of Physics 2, a continuation of the course from the previous semester, except that the experiments are in a different area of physics. In the fifth semester, the descriptions of compulsory courses do not explicitly state the development of digital competences. In the sixth semester, they take the Computer Physics course, with the goal of developing skills solving physical problems numerically in Python.

In the first semester of graduate study, students take the Practicum in Experimental Teaching of Physics 1, in which they acquire skills in conducting experiments adapted for teaching in primary and secondary schools using different devices and computers. In Statistical Data Processing Using Computers, they expand their knowledge of statistics. In the second semester they take the Practical Course in Experimental Teaching of Physics II, a continuation of the course from the previous semester. In the third semester, they take Computer Science Teaching Methodology, in which they acquire digital competences for computer science teaching, and Project Management, in which they acquire and apply digital competences for the creation and implementation of projects. In the fourth semester, there is a Practicum in Computer Science Teaching Methodology, in which the acquisition of digital competences is focused on computer science teaching, but it can also be applied to physics teaching with modification.

B. University of Rijeka

At the University of Rijeka, students can enroll in an undergraduate course in physics, and after completing three years, they can take a graduate course to become a teacher of physics and computer science or a course to become a teacher of physics and mathematics. Course descriptions for undergraduate studies can be found at [5] and for graduate studies at [6] and [7]. In our analysis of graduate studies, we will concentrate on joint courses. The graduate study of physics and information science provides students with formal knowledge in the field of digital competences.

In the first semester of undergraduate studies, students take Programming. In the second semester they take Experimental Data Processing in Physics, in which they acquire and apply digital competences. In the third semester, they take Physical Practicum I, in which they acquire experimental skills, and apply digital competences. In the fourth semester they take Physical Practicum II, a continuation of the course from the previous semester. In the fifth semester they take Physical Practicum III, a continuation of the course from the previous semesters. In the sixth semester, they take Physical Practicum IV.

In the first semester of graduate study, students take the courses Methodical Practical Demonstration Experiments in Physics, Didactics 2, and General Pedagogy, in which they acquire advanced digital competences. In the second semester, they take Methodological Practicum of Laboratory Experiments in Physics, in which they learn to use and analyze educational films in physics classes; the titles of the exercises are written, but there is no information on how they are carried out. In the third semester they take Physics Teaching Methodology II, in which they learn information and communication technology (ICT) in physics teaching. In the fourth semester, Methodical Practice in Physics provides competences for connecting professional knowledge in physics with pedagogical knowledge and modern teaching technologies (ICT).

C. University of Split

At the University of Split, prospective students in the Faculty of Science and Mathematics can enroll in the undergraduate study of physics and undergraduate study of physics and mathematics (teaching). After completing the first cycle, they can enroll in graduate studies in physics (teaching), physics and information sciences (teaching), and physics and mathematics (teaching). In our analysis, we concentrated on the physics major because there was not enough publicly available information on the physics and mathematics majors (for teachers). At the graduate level in physics and information science (for teachers), students acquire more digital competences through formal education. We focused on the common subjects that all majors share. Unlike other universities, the subject descriptions at the University of Split explicitly state what percentage of e-learning is represented in which subject. This analysis does not pay much attention to it because there is no explanation of what it entails. E-learning can be merely a different form of work that does not necessarily result in the development of digital competences.

In the first semester, both majors offer textual and graphic programs for physicists, in which the main goal is to enable students to use Gnuplot and LaTeX in order to successfully continue their educations [8]. Physics students attend the course Programming in the Profession, in which the goal is to learn the basics of programming that they will need in their further education and work [9]. In the undergraduate teacher training program in physics and mathematics, this subject is offered in the fifth semester. In the second semester, students take the course Programming Tools in Physics, with the goal of learning different programming tools that can be applied to concrete examples [10]. In the third semester, the compulsory courses do not explicitly mention the development of digital skills [11]. In the fourth semester, students apply digital skills in the course Practicum in General Physics II, which is a continuation of the course from the previous semester [12]. In the fifth semester, they take a continuation of the course Practicum in General Physics III [13] and apply the acquired digital competences in the course Mathematical Methods of Physics III [14]. In the sixth semester, they also have a continuation of the General Physics Practical IV course [15].

In the first semester of graduate studies in all majors, none of the courses explicitly states that they develop digital skills. There are courses in the physics and information science teaching program where digital skills are developed, but these are related to the narrower field of information science. In the second semester, all students attend the Physics Teaching Methodology course in which they acquire digital competences [16]. In the third semester, certain subjects are compulsory in some programs, while in others they are optional. These include E-learning Systems [17] and Physics Teaching Methodology II [18]. In the fourth semester, the courses are partly conducted in schools and, depending on the mentors in the classrooms, students acquire and develop digital skills. They are not described in detail in the subject descriptions because many factors affect the amount and type of digital competences [19] [20].

D. University of Zagreb

At the University of Zagreb, Faculty of Science and Mathematics, students can enroll in four combinations of majors for physics teachers. These are Teaching Physics, Teaching Physics and Information Science, Teaching Physics and Chemistry, and Teaching Physics and Mathematics. Each major lasts five years. The description of the course for Teaching Physics and Mathematics is not publicly available, so we omitted it from this article. The Physics and Information Science Teaching Major has more information science subjects, which is self-explanatory, and thus students acquire and develop more digital competences through their formal education. In the first three years of study in all majors, there is a greater emphasis on general subjects in physics, while in the last two years, knowledge and skills are acquired in pedagogy, psychology, didactics, and physics methodology. We will concentrate on subjects that all majors have in common to gain insight into what competences a physics teacher acquires. Certain subjects require writing a seminar paper or a similar work in which general critical thinking and the use of digital technology for creation and presentation are encouraged. Some courses mention e-learning, which is a form of teaching that is carried out and does not necessarily have to be for the development of competences, which is why it is not emphasized in our study.

In the first semester of the Physics major and the Physics and Information Science major, students take the course Word Processing and Spread Sheets, in which they learn the basics of LaTeX, MS Word and Excel in order to successfully continue their studies and future work at school [21]. The Physics and Chemistry teaching major does not have such a course. In the second semester, students in the programs of Pedagogy of Physics and Chemistry take the subject Computers and Operating Systems, in which they theoretically and practically learn about computers and programs [22]. The Physics and Information Science Teaching major offers a more specialized course called Fundamentals of Programming. All three majors offer the subject Fundamentals of Physics Measurements, in which they use the computer programs Excel, Statistica and Qplot to apply acquired theoretical knowledge to processing experimental data, as well as the results of student work [23].

In the third semester of the teacher's course in Physics and Information Science, students take professional subjects in the field of information science in which digital competences are developed. The Physics major offers the subject Computing Lab 1, in which students improve practical competences, connecting physics with information technology [24]. The teaching department of physics and the teaching department of Physics and Chemistry offer Physics Lab 1, in which students develop experimental skills in conducting experiments in physics, as well as digital competences. Students use programs for statistical analysis and to display numerical and graphical results. In the fourth semester, there are no explicitly stated outcomes of the development of digital

competences in the physics and chemistry teachers' course, but the Physics and Information Science teachers' program offers professional courses in the field of information science. The Department of Physics offers Computing Lab 2, [25] and Physics Lab 2, which continue the subjects from the previous semester, including exercises from other areas of physics [26].

In the fifth semester, students majoring in Physics and Information Science have professional courses in information science, as well as Physics Lab 1, which develops their digital competences. Physics Lab 1 is the same subject taken by students of the teachers' course in Physics in the third semester. For the teaching major of Physics and Chemistry, the outcomes that would develop digital skills in the fifth semester are not explicitly stated. In the fifth semester, physics students take Physics Lab 3 in which they apply their acquired digital knowledge and skills [27]. In the sixth semester, students majoring in Physics take Physics Lab 4, but unfortunately there is no publicly available course description. The Fundamentals of Electronics course mentions the outcome related to the use of computers and different sensors in physics experiments [28]. Physics and chemistry teaching students and physics and information science teaching students take the subject Physics Lab 2, which was previously described.

In the seventh semester, all teaching majors have Laboratory in Physics Education 1, in which they learn to perform experiments properly and use digital technology [29]. In the eighth semester, all teaching majors take Laboratory in Physics Education 2, which is a continuation of the course from the previous semester, except that there are exercises from other areas of physics [30]. The Physics and Information Science teaching major offers Fundamentals of Electronics, which was previously described. It also offers Computers in Education, which develops digital competences according to the stated outcomes [31]. In the ninth semester, all teaching majors take the subject Teaching Practice in Physics 1, in which students go to practice at schools; the description of the subject is very general, so digital competences cannot be explicitly read [32]. Physics teachers take another subject, Laboratory in Physics Education 3, in which they acquire additional digital competences [33]. In the tenth semester, all majors take the Course Teaching Practice in Physics 2, a continuation of the course from the previous semester [34].

E. University of Maribor

At the University of Maribor, prospective students can choose different combinations of teaching majors. In this paper, we will deal with the subjects that students of the teaching major in physics must pass. The entire study program lasts five years. In almost every study subject, teaching methods include the use of information technology, so this will not be emphasized in the text because it is not explained for each subject how it is used.

In the first semester, students attend Computational Physics, about the construction and operation of computers, and learn the knowledge and skills they need for further education and work [35]. Also, students take Mechanics, in which they have laboratory exercises where they learn and acquire digital competences [36]. In the second semester, students attend Physics Experiments 1, where they analyze data using a computer and use computer simulations in class [37]. In the third semester, they attend the continuation of the course from the previous semester, Physics Experiments 2, in which they continue to develop their digital competences [38]. In the fourth semester, Physics Experiments 3 continues, in which they continue to develop their digital competences [39]. In Thermodynamics, students acquire the skills to use modern computer software for quantitative calculations and plotting dependencies among variables with varying parameters [40].

In the fifth semester, students attend Physics Experiments 4, in which they continue to develop digital competences [41]. In Didactics of Physics 1 with Practicum, they acquire digital competences with an emphasis on quality in teaching [42]. In the sixth semester, students in Didactics of Astronomy connect astronomy and digital technology, thereby developing their digital competences [43]. In the seventh semester, they take Didactics of Physics 2 with Practicum, a continuation of the course from the fifth semester, and learn and apply digital technology in other areas [44]. In Measurements in Physics, they process data and use digital technology, further developing digital competences [45]. In the eighth semester, they attend Astronomical Observations, in which they use astronomical computer programs and digitally process data, improving their digital skills [46]. In the ninth semester, they take subjects in which they use digital technology, but it is not explicitly written in what way. In the tenth semester, students write a graduation thesis and take electives that they choose according to their own preferences.

In the table 1. we present the results of the study program analysis, where the number of shaded squares represents the number of courses in that semester with explicitly written digital competences in their course descriptions. Elective courses were not taken into consideration.

TABLE I. PRESENTATION OF THE NUMBER OF COURSES THAT USE DIGITAL TECHNOLOGIES

Semester / University	*Osijek*			*Rijeka*			*Split*			*Zagreb*			*Maribor*		
1	■	■		■			■			■			■	■	
2	■			■			■			■	■		■		
3	■	■		■						■			■		
4	■			■			■			■	■		■	■	
5				■			■	■		■			■	■	
6	■			■			■			■			■		
7	■	■		■	■	■				■			■	■	
8	■			■			■			■			■		
9	■	■		■						■	■				
10	■			■						■					

IV. THE FUTURE OF THE APPLICATION OF DIGITAL COMPETENCES IN PHYSICS TEACHING

By properly including and using digital technologies in physics classes, students' digital competences can be

improved. The example chosen for this paper is the introduction of total resistance when resistors are connected in series. [47]

First, the teacher connects two resistors in series and measures the voltage across those two resistors and the current, then connects three resistors and measures the current and voltage (Figure 1.). If the school equipment allows, it would be even better if the students connect the circuits themselves. When the students have seen a certain regularity, the computer can be included in the experiment. The two previously described circuits are reconnected, but now a computer is used (Figure 2). Phet's simulation can be used to further improve students' digital competences by repeating the entire process (Figure 3.) [48]. Also, additional combinations can be made in the simulation.

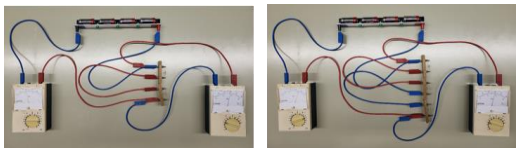


Figure 1. The image on the left shows two resistors connected in series. The picture on the right shows three resistors connected in series. The voltage and current in the circuit are measured using an ammeter and a voltmeter.

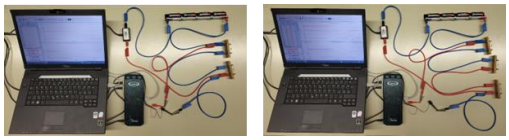


Figure 2. The image on the left shows two resistors connected in series. The picture on the right shows three resistors connected in series. This experimental setup includes sensors that are connected to a computer.

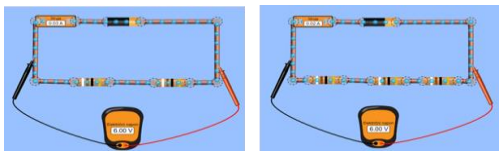


Figure 3. The picture shows the circuit simulated in Phet's simulation [48]

Physics teaching should rely on experiments carried out by the teacher in cooperation with the students. The experiment as a starting point for the study of any teaching material can be modernized by using computers, and sometimes simulations. Research [49] shows that half of Croatian secondary school teachers do not have adequate digital competences and that the attitudes and educational values, computer anxiety, self-efficacy etc. are related to the perceived Croatian secondary school teachers' intention to use e-learning technologies in their teaching after the COVID-19 pandemic.

V. Conclusion

During their formal education, students of the physical sciences acquire digital competences that they can apply in their future work as teachers. Because of recent rapid changes in the digital sciences, it is very important that teachers continue with further education in order to better attain and transfer digital competences to students. Further research is needed in order to gain insight into the quality of the acquired digital competences needed in today's schools. Future research should try to incorporate study programs from many nations in order to acquire a more thorough grasp of digital skills across various educational systems. Even if the present examination of the descriptions of the available physics study programs yielded useful insights, including interviews with the course instructors and polling the students could provide other viewpoints.

References

[1] European Commission. (2018). Key Competences for Lifelong Learning. doi:10.2766/569540

[2] L. Cohen, L. Manion, and K. Morrison, Research Methods in Education, 6th ed. Routledge, 2018.

[3] Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za fiziku. (2022). Izvedbeni plan preddiplomskog studija Fizika akademske godine 2022/2023 [PDF]. Available: https://www.fizika.unios.hr/wp-content/uploads/2022/09/Izvedbeni-plan_preddiplomski_2022_2023_prosireni.pdf [Accessed: Jan. 29, 2023].

[4] Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za fiziku. (2022). Izvedbeni plan diplomskog studija Fizika akademske godine 2022/2023 [PDF]. Available: https://www.fizika.unios.hr/wp-content/uploads/2022/09/Izvedbeni-plan_diplomski_2022_2023_prosireni.pdf [Accessed: Jan. 29, 2023].

[5] University of Rijeka, Department of Physics, 2022. IZMJENE I DOPUNE PREDDIPLOMSKOG STUDIJA FIZIKA Available: https://www.phy.uniri.hr/files/nastava/preddiplomski_studij/planovi_i_opisi_2022/Preddipl%20Fizika_Izmjene%20i%20dopune_svibanj%202022_prihvaceno%20na%20Vijecu.pdf [Accessed: Feb. 2, 2023].

[6] University of Rijeka, Department of Physics, "Izmjene i dopune Studijskog programa Diplomski studij Fizika - smjer Fizika i informatika" [Online]. Available: https://www.phy.uniri.hr/files/nastava/diplomski_studiji/fizika_informatika/Izmjene_i_dopune/Dipl%20FI_izmjene%20i%20dopune_2018.pdf . [Accessed: Feb. 4, 2023].

[7] University of Rijeka, Department of Physics, "Izmjene i dopune Studijskog programa Diplomski studij Fizika - smjer Fizika i matematike" [Online]. Available: https://www.phy.uniri.hr/files/nastava/diplomski_studiji/fizika_matematika/Izmjene_i_dopune/Dipl_FM_obrazac_izmjene_i_dopune_2018.pdf [Accessed: Feb. 10, 2023].

[8] University of Split, Faculty of Science, Department of Physics, Tekstualni i grafički programi za fizičare Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP071.pdf [Accessed: Jan. 30, 2023].

[9] University of Split, Faculty of Science, Department of Physics, Programiranje u struci Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP073.pdf [Accessed: Jan. 30, 2023].

[10] University of Split, Faculty of Science, Department of Physics, Programski alati u fizici Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP074.pdf [Accessed: Jan. 30, 2023].

[11] University of Split, Faculty of Science, Department of Physics, Praktikum iz opće fizike I Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP011.pdf [Accessed: Jan. 30, 2023].

[12] University of Split, Faculty of Science, Department of Physics, Praktikum iz opće fizike II Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP012.pdf [Accessed: Jan. 30, 2023].

[13] University of Split, Faculty of Science, Department of Physics, Praktikum iz opće fizike III Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP013.pdf [Accessed: Jan. 30, 2023].

[14] University of Split, Faculty of Science, Department of Physics, Matematičke metode fizike III Available: https://www.pmfst.

unist.hr/wp-content/uploads/predmeti/PMP102.pdf [Accessed: Jan. 30, 2023].

[15] University of Split, Faculty of Science, Department of Physics, Praktikum iz opće fizike IV Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP014.pdf [Accessed: Jan. 30, 2023].

[16] University of Split, Faculty of Science, Department of Physics, Metodika nastave fizike I Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP050.pdf [Accessed: Jan. 30, 2023].

[17] University of Split, Faculty of Science, Department of Physics, Metodika nastave fizike II Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP150.pdf [Accessed: Jan. 30, 2023].

[18] University of Split, Faculty of Science, Department of Physics, Sustavi E-učenja Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMIK10.pdf [Accessed: Jan. 30, 2023].

[19] University of Split, Faculty of Science, Department of Physics, Seminar iz metodike fizike s nastavnom praksom Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMP152.pdf [Accessed: Jan. 30, 2023].

[20] University of Split, Faculty of Science, Department of Physics, Metodički informatički seminar s nastavnom praksom II Available: https://www.pmfst.unist.hr/wp-content/uploads/predmeti/PMIK61.pdf [Accessed: Jan. 30, 2023].

[21] University of Zagreb, Faculty of Science. WORD PROCESSING AND SPREAD SHEETS Available: https://www.pmf.unizg.hr/phy/en/course/otpt_b [Accessed: Jan. 14, 2023].

[22] University of Zagreb, Faculty of Science. COMPUTERS AND OPERATING SYSTEMS Available: https://www.pmf.unizg.hr/phy/en/course/ros_b [Accessed: Jan. 14, 2023].

[23] University of Zagreb, Faculty of Science. FUNDAMENTALS OF PHYSICS MEASUREMENTS Available: https://www.pmf.unizg.hr/phy/en/course/sofm [Accessed: Jan. 14, 2023].

[24] University of Zagreb, Faculty of Science. COMPUTING LAB 1 Available: https://www.pmf.unizg.hr/phy/en/course/racpra1 [Accessed: Jan. 14, 2023].

[25] University of Zagreb, Faculty of Science. COMPUTING LAB 2 Available: https://www.pmf.unizg.hr/phy/en/course/racpra2 [Accessed: Jan. 14, 2023].

[26] University of Zagreb, Faculty of Science. PHYSICS LAB 2 Available: https://www.pmf.unizg.hr/phy/en/course/fizpra2 [Accessed: Jan. 14, 2023].

[27] University of Zagreb, Faculty of Science. PHYSICS LAB 3 Available: https://www.pmf.unizg.hr/phy/en/course/fizpra3 [Accessed: Jan. 14, 2023].

[28] University of Zagreb, Faculty of Science. FUNDAMENTALS OF ELECTRONICS Available: https://www.pmf.unizg.hr/phy/en/course/osnele_a [Accessed: Jan. 14, 2023].

[29] University of Zagreb, Faculty of Science. LABORATORY IN PHYSICS EDUCATION 1 Available: https://www.pmf.unizg.hr/phy/en/course/pienf1 [Accessed: Jan. 14, 2023].

[30] University of Zagreb, Faculty of Science. LABORATORY IN PHYSICS EDUCATION 2 Available: https://www.pmf.unizg.hr/phy/en/course/pienf2_b [Accessed: Jan. 14, 2023].

[31] University of Zagreb, Faculty of Science. COMPUTERS IN EDUCATION Available: https://www.pmf.unizg.hr/phy/en/course/urun_b [Accessed: Jan. 14, 2023].

[32] University of Zagreb, Faculty of Science. TEACHING PRACTICE IN PHYSICS 1 Available: https://www.pmf.unizg.hr/phy/en/course/mpif1_b [Accessed: Jan. 14, 2023].

[33] University of Zagreb, Faculty of Science. LABORATORY IN PHYSICS EDUCATION 3 Available: https://www.pmf.unizg.hr/phy/en/course/pienf3 [Accessed: Jan. 14, 2023].

[34] University of Zagreb, Faculty of Science. TEACHING PRACTICE IN PHYSICS 2 Available: https://www.pmf.unizg.hr/phy/en/course/mpif2_a [Accessed: Jan. 14, 2023].

[35] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Computational physics Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Racunska_fizika.pdf [Accessed: Feb. 12, 2023].

[36] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Mechanics Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZOB_FIZ_PU_mehanika-maj2022_popravljeno.pdf [Accessed: Feb. 12, 2023].

[37] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, hysics experiments 1 Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZOB_FIZ_Fizikalni-Eksperimenti-1-EMAG-nova-verzija-AD.pdf [Accessed: Feb. 12, 2023].

[38] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Physics experiments 2 Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Fizikalni_eksperimenti_2.pdf [Accessed: Feb. 12, 2023].

[39] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Physics experiments 3 Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Fizikalni_eksperimenti_3.pdf [Accessed: Feb. 12, 2023].

[40] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Thermodynamics Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Termodinamika.pdf [Accessed: Feb. 12, 2023].

[41] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Physics experiments 4 Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Fizikalni_eksperimenti_4.pdf [Accessed: Feb. 12, 2023].

[42] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Didactics of physics 1 with practicum Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZOB_FIZ_Didaktika_fizike_1_s_praktikumom-nov.pdf [Accessed: Feb. 12, 2023].

[43] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Didactics of Astronomy Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZOB_FIZ_Didaktika_astronomije-EMAG.pdf [Accessed: Feb. 12, 2023].

[44] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Didactics of physics 2 with practicum Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZOB_FIZ_Didaktika_fizike_2_s_praktikumom-nov.pdf [Accessed: Feb. 12, 2023].

[45] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Measurements in Physics Available: https://www.fnm.um.si/wpcontent/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Fizikalna_merjenja.pdf [Accessed: Feb. 12, 2023].

[46] Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Astronomical observations Available: https://www.fnm.um.si/wp-content/uploads/predmetniki/1_stopnja/predmetni-ucitelj/izofiz/2022-2023/EMAG_IZO-FIZ_Astronomska_opazovanja.pdf [Accessed: Feb. 12, 2023].

[47] R. Repnik, "Using physics simulation environment for better students' performance," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2018, pp. 0819-0824, doi: 10.23919/MIPRO.2018.8400151.

[48] PhET Interactive Simulations, Circuit Construction Kit DC Simulation Available: https://phet.colorado.edu/hr/simulations/circuit-construction-kit-dc [Accessed: 1-Feb-2023].

[49] S. Babić, S. Križan Sučić, and G. Sinković, "Understanding the factors that influence secondary school teachers' intention to use e-learning technologies for teaching after the COVID-19 pandemic," in Proceedings of the 43rd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2020, pp. 938-943, doi:10.23919/MIPRO49410.2020.92

The Influence of the COVID 19 to the Future of Learning

Marjan Krašna* and Tomaž Bratina**
* University of Maribor, Faculty of Arts, Maribor, Slovenia
** University of Maribor, Faculty of Education, Maribor, Slovenia
marjan.krasna@um.si, tomaz.bratina@um.si

Abstract - In the years of COVID 19 pandemic we have changed the education. Though we wanted to change only the way we teach, study and learn we have changed ourselves. May studies have been done about the drawbacks of the distance education, videoconferencing and social interaction. Much less were focused to the benefits of the changed learning environment. In the short time we have learnt how to use tools for remote work. We have become masters of videoconferencing and in the education, teachers produce a lot of learning materials. But what good comes to the students. Do they miss something from COVID 19 pandemic? The research of the educational study programs' students has shown differences between students' groups. Students from the first year of study much more like to communicate with the teachers in person than the those from later years. Though the communication method seems to be different between students of Faculty of Education and Faculty of Arts, statistical analysis has not confirm this. Interestingly we have discovered that students like e-mail communication the most but they are really patient for the reply (up to 4 days). The very good finding is that students do not expect technically related problems in the case of another lockdown.

Keyword – Education, ICT, Covid-19, Benefits

I. INTRODUCTION

Many studies have been performed during and after the COVID 19 lockdown. In general, we can see that the society managed to function well during such events but there are certain drawbacks of the remote work. Many psychologists warned us of long-term effects on younger population which were unable to develop social skills for two years and this would follow them through the education and life. On the other hand, lock down rapidly increased the ICT competences of students and teachers alike. It would be wasteful to forget these acquired competences. Who can assure us that the next lockdown is not just around the corner?

Was it all bad, is the question that resonate in the heads of many previously highly ICT competent teachers? We cannot support these thoughts. One thing is sure, digital communication competences increased and the education system becomes more resilient not just in our country but also in the other EU countries [1] [2] [3]

It is evident that we, as a society, need to change. This change still needs to find the optimal balance between quality of education and sustainability, particularly in the area of green transition [4] [5]. The two years of COVID-19 lock down may not be enough to really master the ICT in education. We generally did not start as proficient users and the first year was the introduction and learning how to use ICT in education. For those who had used ICT previously this was not a problem and they could start the search for the effective strategies to employ ICT for the best performance. Never the less most of the users (teachers and students) alike were the beginners and they need a lot of time just to master ICT. The second year was the one where these users could start exploit the benefits of ICT. These may be the reasons that in general the lockdown was perceived as the best of the worst options. The teachers who have managed to master ICT started to see the full potential of its use in education [6] and they prepare the list of the needs for the efficient use of ICT in education [7]

Most of the problems was detected in the area of primary education. But we have signs that even in less developed countries the benefits of extensive use of ICT in education is shown [8]. Rapid development of the tools that support the online learning and videoconferencing was evident during COVID-19 lockdown. In our country teachers could use Zoom or MS Teams. Most of teachers opted for Zoom because it did not demand a lot of learning. Those who were more skilled opted for MS Teams which later give them complete solution [9]. In the higher levels of education, the MS Teams was adopted and exploited in full [10]. On the other hand, we could not exploit the hybrid education due to the lack of suitable equipment in our classrooms and insufficient skills to overcome the emerging onsite technical glitches. [11]. The significant changes between education institutions were detected in the different use of learning materials [12] and the students did differentiate according to their learning preferences (visual, reading and auditory kinaesthetic learning styles; have statistically significant influence on grades on some study subjects in high school students) [13]

II. THE RESEARCH

After the COVID-19 our colleagues have agree that lockdown is not a big deal despite they like to be in the classroom much more. We have seen the changes in the student population. They demand more from teachers and they often express they would like to have recording of the lecture. This may be trivial if lectures are performed in the controlled environment (cabinet or home cabinet) but much more demanding in the classroom. Hybrid education just does not work yet. There are many reasons for that but in general only two can emerge as paramount: lack of equipment and personal (mainly negative) attitude toward

two distinctive different group of students that require rapid switching from strategies that work in the classroom and the different for the remote students.

The teachers that have recording of their lectures from the lockdown times find themself in the awkward position. They can give access to the students of these recordings but this means less students in the classroom, therefore the reluctance. Some teachers do not want the recordings because there are things that needs to be said but do not want them to be recorded. Therefore, there are many reasons in favour and against hybrid work.

Described reasons were the foundation of this research. What is the real implication of COVID-19 lockdown on our students from a longer perspective? We have educational study programs students and they will need ICT competences in their work. We could focus to the extensive research but we decide to focus only on three most obvious topics: communication, remote work; and personal opinion. As previously noted, the communication skills increased. Students are much more competent in the communication and they provide better content in the communication. In the remote work we wanted to find out which type of learning materials would they prefer and what was the benefits of the remote work. Many companies have decided that remote work is their future; hence students need to sustain these competences. The last topic is the focused to detect the reluctance to the distance work and what would be the most problematic issue for them.

The questionnaire was sent to the students of Faculty of Arts and Faculty of Education and as expected we have limited success in the acquisition of the data. Nevertheless, we acquire enough data to get sufficiently valid statistical results. From the classification data we could made statistical analysis from groups that were large enough to conduct suitable statistical tests.

The most important learning materials for our students is still the textual learning materials. Video learning materials are little below average and interactive learning materials are in the last place. These feedbacks are not entirely expected but could be guessed. The text can be read fast, skipped and searched; video learning materials on the other hand, are more time consuming. If not properly edited into small segments and indexes, they just take too much time for our students. Video clips that are viral are few seconds long and this is all what students are willing to sacrifice of their precious time.

III. RESEARCH OBJECTIVES

The main research objective was to gain insight into how the students who went through the pandemic developed attitudes toward electronic communication and online study. Further, the aim was to test the differences in the primary objective between students in the 1st year of study and older students from the Faculty of Education and Faculty of Arts. The focus of the research was set on:

- The attitude toward email communication
- Behaviour during the email communication
- Preferred types of communication
- The opinion about eventual return to “online” study

A. Methodology & data collection

The online survey was used to collect data and comprehend questions about the attitude toward email communication and behaviour during the communication and popularity of various types of electronic communication. The survey participants were students in the academic year 2022/23 from the Faculty of Education and Faculty of Arts in Maribor, Slovenia. The balanced sample includes 110 students, about 45 from each institution. Only valid responses were analysed due to some incompletely answered or skipped questions. Data were processed using SPSS statistical software. The statistical methods of descriptive statistics, chi-square test, and chosen non-parametric tests were used to analyse the rating scales.

IV. RESULTS

B. Preferred forms of communication with the teachers by the year of study

TABLE I: FORM OF COMMUNICATION PREFERENCES

Which form of communication with the teachers do you prefer?	Year of study	
	1st	2nd to 5th
personal	21 (48,8%)	12 (28,6%)
electronic	1 (2,3%)	9 (21,4%)
personal and electronic equally	21 (48,8%)	21 (50,0%)
$\chi^2 = 8{,}844$, df = 4, $p < 0{,}05$. Numbers in parentheses indicate column percentages		

Personal communication is preferred by first-year students more than by older and more experienced students (TABLE I). The difference in opinions between both groups of students is statistically significant. Concerning the many years of our teaching experience, we assume that the reason regarding younger students is inexperience or insecurity while they electronically communicate with yet fewer known teachers. Or even the fear of making mistakes and possible consequences. Compared to the personal communication type, both the younger and more experienced students prefer the combination of personal and electronic communication equally. The results are somewhat unexpected because the general opinion about millennials is that they master and prefer electronic communication more than personal communication. On the other hand, some researchers are showing that the most preferred way of communication is still close personal contact (Bratina, 2019). Therefore, we expect that electronic communication will never become primary but will coexist with the personal form.

C. Preferred forms of communication with the teachers by the faculty

TABLE II: FORMS OF COMMUNICATION WITH TEACHER

Which form of communication with the teachers do you prefer?	Year of study	
	Faculty of Education	Faculty of Arts
personal	19 (47,5%)	11 (33,3%)
electronic	3 (7,5%)	2 (6,1%)
personal and electronic equally	18 (45,0%)	20 (60,6%)
$\chi^2 = 1{,}794$, df =4, $p > 0{,}05$. Numbers in parentheses indicate column percentages		

The Faculty of Education and the Faculty of Arts differ in the content of their study programs. The first educates future teachers, while the latter educates experts in various social science disciplines. From this point of view, we expected a different perception of each form of communication. The results show no statistically significant differences between students of both faculties (TABLE II). Despite this result, slightly less than half of the students, future teachers, prefer personal communication, against about a third of the students from the Faculty of Arts. We believe future teachers understand the importance of personal communication better than others. The indication of the mentioned fact is more significant share (60,6 %) of students from the Faculty of Arts prefer a combination of personal and electronic communication. The result is interesting, as we expected a much more similar attitude from all students.

D. The expected time for the recipient to reply to the message

TABLE III: TIME TO REPLY

How fast do you expect the reply to the message	**f**	**f%**
Immediately or in a few hours	9	10,5
One day	27	31,4
Two to three days	21	24,4
Four days	27	31,4
One week	2	2,3
Total *	86	100
*valid cases only		

Students are willing to wait between one and four days for a response to the message (TABLE III). About 10 % students are very impatient, which we think depends on the importance of the matter. There is a very small percentage of students who are willing to wait a week for a response. In today's very dynamic time, a message is outdated after a week or more. Additional analyse show no differences between younger and older students in expected time for reply. But there is a statically significant ($p < 0,05$) difference between students in the two faculties in their willingness to wait for the reply to the message. Almost half of the students (45 %) of the Faculty of Education are willing to wait for one day, while 18,2 % of the students of the Faculty of Arts have the same opinion. About 27,5 % of Faculty of Education students and 18,2 % of Faculty of Arts students are willing to wait between two to three days. However, only a quarter of the Faculty of Education students and about 40 % of the others would wait four days for the reply. We believe that the expected time is shorter for students in the Faculty of Education because future teachers know how important quick feedback is for pupils. Moreover, timely responses to questions about the learning content and other phenomena or complications are crucial in the learning process.

E. Types of communication close to the students

Among the listed types of communication presented to the students, email is the most often chosen type (61,9 %). Followed by Chat (19 %) and phone calls (11,9 %). Even though widely used apps like WhatsApp and Viber allow quality video calls, it was selected by only 3,6 % of students. The result is unexpected and hard to explain. Our first thought is that the reason may be a fear of the expensive data volume consumed, which is inaccurate due to the technical processing of compression. It is likely the kind of behaviour preferring textual messaging instead of visual [14]

TABLE IV: COMMUNICATION PREFERENCE

Which type of communication is closer to you?	Year of study		Faculty	
	1st	2nd to 5th	Faculty of Education	Faculty of Arts
Phone call	7 (16,3%)	3 (7,3%)	6 (15,0%)	4 (12,1%)
Email	24 (55,8%)	28 (68,3%)	24 (60,0%)	20 (60,6%)
Chat	9 (20,9%)	7 (17,1%)	7 (17,5%)	6 (18,2%)
Video call	2 (4,7%)	1 (2,4%)	3 (7,5%)	0 (0,0%)
$\chi^2 = 2,836$, df = 4, $p > 0,05$.			$\chi^2 = 8,489$, df = 4, $p > 0,05$.	
Numbers in parentheses indicate column percentages				

Email, as the type of communication closest to students, predominates regardless of the year of study or faculty (TABLE IV). The percentage ranges from 55 % to 61 % in both criteria groups. Chat is the second closest type of communication to the students, but it does not get past a share of just over 20 %. Video call appears to be the least used type of communication, though it is slightly more popular among younger students and future teachers.

F. Going back to "online" study

We asked students to rate their agreement with the statement that eventual returning to online study would disappoint them. They rated on a 5-point scale from "*Does not apply to me at all (1)*" to "*Applies to me completely (5)*" Only valid cases are included in the analysis.

TABLE V: BACK TO ONLINE DISAPPOINTMENT

Going back to "online" study would disappoint me	Institution	N	$\bar{R}$	U
	Faculty of Education	39	36,59	523,000
	Faculty of Arts	30	32,93	
*p < 0.05. Valid cases only				
Going back to "online" study would disappoint me	Year of study	N	$\bar{R}$	U
	1st	41	43,66	629,000
	2nd to 5th	38	36,05	

*p < 0.05. Valid cases only

The majority of students are undecided (neither/nor; M = 3.15) regarding a possible return to online study. The differences between students of the two faculties nor between younger and older students are statistically not significant (TABLE V). About 39.3 % of students completely agree with the statement, while about 35.5 % completely disagree. Based on our experience working with students, we think the results are surprising. Students often tell teachers that they would rather work online than attend lectures.

G. Expecting technical difficulties if back to "online" study

We also asked students to rate their agreement with the statement that they will deal with the technical difficulties in case of returning to the "online study". They rated on a 5-point scale from "*Does not apply to me at all (1)*" to "*Applies to me completely (5)*" Only valid cases are included in the analysis.

TABLE VI: BACK TO ONLINE TECHNICAL PROBLEMS

Going back to "online" study would disappoint me	Institution	N	$\bar{R}$	U
	Faculty of Education	39	38,27	475,500
	Faculty of Arts	30	30,75	
*p < 0.05. Valid cases only				
Going back to "online" study would disappoint me	Year of study	N	$\bar{R}$	U
	1st	41	45,70	545,500*
	2nd to 5th	38	33,86	

*p < 0.05. Valid cases only

Students generally believe they will not have technical problems when they eventually return to online study (M = 2,15) (TABLE VI). The results show that younger students are more likely to anticipate problems than older students. We suspect that younger students were better supported by their parents during the pandemic. Therefore, it would be more difficult for them to deal with eventual technical problems now. However, we find no differences in the assessment of expected problems between students from both faculties.

V. CONCLUSION

The research has shown that we could transit to the possible future lockdown with the ease. Students use different types of the communication; even different for interpersonal and communication with teachers. The later – formal communication is mostly done with the e-mails. But in reality, they mostly chat by typing (on the phone). The feedback from the younger students have shown that we have missed one type of the communication in our research – voice messages. This just shown the generation gap is all present. Video calls are almost non-existent option in the communication with the teacher and among students.

The interesting discovery is that students are very patient regarding the response from the teachers (up to four (4) days). But there is statistically significant difference between students. Students of Faculty of Education is little more impatient than students from the Faculty of Arts. The feedback in the education is of paramount importance and in the lower levels of the education needs to be faster. This could explain the perceived differences between two students' groups.

The good thing is that student acquired enough technical skills to overcome the possible technical difficulties in the case of new lockdown or remote work. On the other hand, students are equally divided between being disappointed or not being disappointed in case of lockdown or online work.

To our disappointment we could not extract any relevant information what the students would like to retain from the COVID-19 lockdown. As expected, the most valued thing was the time allocation, followed by home environment, and learning material in text format. Others are less important. Even video lectures are below average important. It is interesting that the good old text is still "the king" in the tertiary education. From the comments about what should be used, that have been learned in the lockdown period, in the contemporary education, we could not see any dominant feature. In a conclusion we could just state that education system is resilient and highly adaptable if needed.

Nevertheless, it would be very interesting to expand this research to more study programs and educational institutions worldwide.

REFERENCES

[1] A. Appolloni, N. Colasanti, C. Fantauzzi, G. Fiorani in R. Frondizi, „Distance Learning as a Resilience Strategy during Covid-19: An Analysis of the Italian Context," *Sustainability*, Izv. 13, št. 3, pp. 1-12, 2021.

[2] M. Serfozo, S. Golyan, F. Z. Lassu, B. Svraka in A. V. Pirka, „Digitalization, distance learning in teacher training - students' feedback on distance education," *CIVIL SZEMLE* , 2020.

[3] S. P. Goncalves, M. J. Sousa in F. S. Pereira, „Distance Learning Perceptions from Higher Education Students-The Case of Portugal," *Education Sciences*, Izv. 10, št. 12, pp. 1-15, 2020.

[4] S.-K. Podder, M. Karuppiah, B. Thomas in D. Samanta, „Research Initiative on Sustainable Education System: Model of Balancing Green Computing and ICT in Quality Education," v *2022 Interdisciplinary Research in Technology and Management (IRTM)*, Kolkata, India, 2022.

[5] M. Zajc, „Information Communication Technologies as Enabler of Green Transition: Education and Research Perspectives," v *30th Telecommunications Forum (TELFOR)*, Belgrade, Serbia, 2022.

[6] A. Špernjak, „Using ICT to Teach Effectively at COVID-19," v *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2021.

[7] N. Špur, E. Kranjec, M. Puhek, K. Breznik, L. Klasinc in S. Frumen, „University teachers' needs for the efficient use of ICT in the pedagogical process," v *3rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2020.

[8] C. Zuniga-Velásquez in K. Peralta-Valenzuela, „Work in Progress: Information and Communications Technology (ICT) in Learning in Students of Elementary Education," v *2020 IEEE World Conference on Engineering Education (EDUNINE)*, Bogota, Colombia, 2020.

[9] B. Marčinković, B. Aberšek in I. Pesek, „The Satisfaction of Primary School Teachers with the Introduction and Use of MS Teams in Distance Education," v *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2021.

[10] M. Krašna in I. Pesek, „nfluence of Moodle and MS Teams on teaching-learning-studying (TLS) processes," v *43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2020.

[11] M. Krašna in T. Bratina, „The reasons for limited success of hybrid lectures," v *45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2022.

[12] I. Pesek, M. Krašna in T. Bratina, „Students' Transition from High School to University in COVID-19 Constrained Education," v *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2021.

[13] M. Lapi in M. Krašna, „The Distance Learning Implications on High School Students," v *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, 2021.

[14] T. Rebane, "Why Millennials Prefer Texting [2023 Edition]," 2023. [Online].

[15] T. Bratina, "Mobile phones and social behavior among millennials - future teachers," *Journal of Elementray Education*, vol. 12 vol. 4, pp. 315-330, 2019.

How to Measure Co-Creation in the Digital Environment of Higher Education?

Tea Mijač*, Mario Jadrić * and Maja Ćukušić*
* Faculty of Economics, Business and Tourism, University of Split, Croatia
tea.mijac@efst.hr, mario.jadric@efst.hr, maja.cukusic@efst.hr

Abstract - **Co-creation, as a form of collaborative innovation, has been a topic of growing interest in Higher Education (HE). Regardless of the industry, engaging end-users in the design of products and services has been assessed as important for the success of the initiatives and is often considered a means to increase competitive advantage. As all markets face digital transformation impetus that permanently alters product and service delivery, the physical aspect of the delivery becomes secondary in many cases. Namely, most services nowadays are delivered exclusively online, and the area of (higher) education is no exception. New generations of learners belonging to Gen Z are connected 24/7, their digital experiences and expectations have evolved, and are prepared to engage with the product or service providers. In that regard, previously developed research instruments focusing on general services (as opposed to digital ones) comprising of two dimensions (customer participation and citizenship behaviour), need validation in the new context. This study aimed to test it in the digital service environment of HE with more than 550 students (as end-users) from different HE institutions in Croatia. To that end, confirmatory analysis using structural equation modelling confirmed that not all subdimensions are relevant in the digital context.**

Keywords - co-creation, higher education, digital services, digital tranformation

I. INTRODUCTION

Digital transformation has experienced a rapid increase in today's society, resulting in the widespread adoption of digital technologies as the new normal. This digital transformation is significantly impacting the field of higher education (HE), which is undergoing a fundamental change in its operational processes and business models and the integration of technology platforms being increasingly employed [1], [2], [3]. As a result, all aspects of HE are undergoing transformation, with significant changes in how institutions carry out their daily processes [4].

In today's society, institutions of HE are facing a multitude of challenges, including the globalization of the market and the subsequent internationalization of education. This trend has led to increased competition among higher education institutions (HEI) in attracting prospective students, as well as a rise in the demands and expectations of future students, particularly with respect to the level of user experience provided by HEIs and the expected high standards of quality [5], [6]. To remain competitive, institutions must be able to provide a superior student experience that meets or exceeds the expectations of their students. This involves optimizing the interaction between students and the institution [7]. Institutions are also developing strategies to enhance efficiency, quality, and student experience, as well as to foster loyalty through co-creation, which involves creating collaborative, customer-specific value that is closely aligned with the concept of value-in-use [8], [6]. Value co-creation is a concept that has been defined as the process of creating value through collaborative activities between users and organizations during the consumption and use of a product or service [9]. In other words, it occurs when interactions between the organization and the customer take place, leading to the co-creation of value.

Accompanied by mentioned digital transformation, the web environment became increasingly important for offering digital services that support interactivity, thereby enhancing the student experience [10]. However, due to evident relevancy, empirical research on the operationalization of co-creation in higher education is scarce [11], and there is a need for further research on the topic, especially in the context of technology support for co-creation [12].

In line with this, the study presented herein aimed to *examine how to measure co-creation in this (new) digital environment of HE.*

The paper is structured as follows. Section 2 presents the role of co-creation in higher education and digital transformation. Section 3 describes the research methodology, while section 4 shows the study's empirical results. Highlights and the main contributions of the study are presented in section 5, together with future research directions.

II. LITERATURE REVIEW

A. Co-creation in Higher Education (HE)

Over the past 15 years, the service-centered approach known as service-dominant (S-D) logic has evolved into a meta-theoretical framework that fosters a systemic understanding of value co-creation [13]. Co-creation can be defined as joint innovation, that is, joint development and improvement of a product or service with all stakeholders [14]. Even though co-creation originally comes from marketing it has become a popular topic in HE research [15]. Higher education institutions (HEI) tend to increase loyalty through cooperation and collaboration with students [6].

Co-creation represents an alternative with the help of which it is possible to improve the study experience [6]. In the context of HEIs, co-creation refers to the integration of institutional resources with student opinions, responses, personalities, and academic abilities in such a way as to realize shared value for all stakeholders [16], [17]. Co-creation is based on voluntary participation, which represents an additional challenge for institutions considering that users can withdraw at any time [18]. Authors [11] define co-creation as the creation of value that is common to all stakeholders (organization and users) and enables everyone to build experiences, define problems and solve problems together [11]. Authors [18] and [19] investigated the co-creation construct on the example of HE teaching processes and confirmed that co-creation could be used as a tool for bridging cultural differences.

The authors [20] operationalized the construct of co-creation in the corporate environment and confirmed its multidimensionality. Precisely, they confirmed two dimensions: (1) customer participation behavior (CPB) and (2) customer citizenship behavior (CCB). The authors [12] adopted the instrument [20] for the HE context and confirmed that student satisfaction positively affects co-creation. In addition, authors [21] also examined how to enhance the university's brand image and reputation through customer value co-creation behavior. They confirmed that the university's website positively impacts co-creation behavior in general. Authors who examined co-creation in HE are listed in Table I.

TABLE I. CO-CREATION IN HE

Dimensions	***Subdimension***	***Authors***
Customer participation behavior	Information seeking	[6], [12], [22]
	Information sharing	[12], [22]
	Responsible behaviour	[6], [12], [22]
	Personal interaction	[12], [21], [23]
Customer citizenship behavior	Feedback	[12], [21], [22]
	Advocacy	[12], [21]
	Helping	[21]
	Tolerance	[21]

Eight subdimensions commonly used in previous research regarding co-creation construct are:

- **Information seeking**: Customers seek information for several reasons. Firstly, information helps consumers to comprehend better and manage their co-creation settings. Secondly, acquiring knowledge allows customers to become integrated into the value co-creation process and understand their role as value co-creators. There are various ways in which customers can request information from businesses [20]. Additionally, apart from seeking out alternative sources of information, such as other customers or informal and personal means of information, customers also prefer these sources over official ones [12].
- **Information sharing**: To ensure successful value co-creation, it is necessary for customers to contribute resources, including information, to be utilized in the co-creation process, as noted by Lengnick-Hall (1996). If customers are unable to provide precise and relevant information, the quality of value co-creation may be negatively impacted. Therefore, effective information sharing is essential to promote high-quality value co-creation [20]. For instance, it is expected that students will give university staff the information they need and respond to their inquiries about customer service to enable staff to provide the best possible service [12].
- **Responsible behavior**: It is often required for consumers to adhere to employee instructions and be physically present for value co-creation to be successful. In the service interaction, little value is co-created without customers' responsible behaviour [20]. The consumer (student) must cooperate, follow the rules and policies, behave politely, and take instructions from service provider staff [12].
- **Personal interaction**: The process of creating value in a service environment typically takes place in a social setting. Customers are more likely to take part in value co-creation if the social setting is fun, welcoming, and upbeat [20]. Owing to the social context, interpersonal skills, including politeness, friendliness, and respect, are crucial [12].
- **Feedback**: The behavior of organization affects the experience of students, who are essentially the customers, and thus organization can gain significant benefits by considering suggestions from these customers for improving their services [20]. However, although feedback is valuable, it should be noted that it is not a mandatory but voluntary act [20] where students (customers) engage in making constructive suggestions for improvements [12]. For example, feedback can be obtained via higher education websites [21].
- **Advocacy:** When customers recommend a company to others in a positive manner, it is a sign of their loyalty to that company. This kind of advocacy is highly beneficial for the company in many ways, including improving its reputation, promoting its products and services, receiving higher evaluations for service quality, and growing the customer base [20]. This behavior involves endorsing, supporting, and defending the organization by recommending it to people outside the organization, such as friends and family, and spreading goodwill about it [21],[12].
- **Helping**: When customers (students) are involved in co-creating a service, they tend to offer help and support to other customers (students) [20]. Customers remember their own challenging experiences and feel a sense of obligation to assist others [20]. Helping may include students' willingness to help other students with difficulties

[12], or it could be just considered as helping refers to students' behavior aimed at assisting other students [21].

- **Tolerance**: Tolerance is the degree to which a user is willing to be patient when the service they receive falls short of their expectations for acceptable service [20]. In the HE offline context, it can refer to accepting inconvenient study conditions, accepting alternative class meeting times, etc. [12], or if any HE service is not delivered as expected [21].

B. Digital transformation

Information and communication technologies (ICT) have changed the way in which services are defined and realized in HEIs [10], [24]. Co-creation presupposes connections and interactions with all stakeholders [25], and it is almost impossible to achieve it without using technology (which has become a key element in the functioning of HEIs). Not only is technology necessary for the execution of all processes, but it also provides the basis for joint work and cooperation among all stakeholders [26].

Technology supports and helps students co-create [27]. For most organizations, even websites alone represent a key tool that serves them to strengthen relationships with customers/users [28]. The results of the author's research [21] indicated that the university's website positively impacts co-creation behavior in general. Thus, even the website, which represents the lowest level of maturity of digital services in HE because it is an asynchronous way of communication [29], has a positive impact on co-creation. Precisely, the author [15] listed several examples in which the use of technology influenced co-creation, and it is about students who, with the help of digital resources, stated their ideas about improvement.

Several authors confirmed that social networks could also be used to encourage co-creation thanks to functionalities such as enabling active dialogue/interaction between (future) students and employees of the institution [18], [30]. Different system functionalities encourage the interactivity and creativity of end users and, thus, co-creation [10]. For example, online discussions are also considered a form of co-creation realization [23].

Online platforms have proven to be a good tool for encouraging and conducting co-creation, but the results indicate that the level of student engagement in practice is still low [18].

III. METHODOLOGY

To answer the proposed research question, quantitative research was conducted on students who have experience in using digital service for supporting activities (such as applying for exams, gaining proofs, etc.).

Items were adopted and prepared based on the literature review. The instrument consisted of 29 items and Liker scale was used from 1 (totally disagree) to 5 (totally agree).

This study was conducted with students from ten HEIs in Croatia who use digital services for supporting activities. Total number of fully-completed survey questionnaires was N=551. As for year of studying, majority had 3 years' experience of studying (27%), 22.1% had two-year experience, 18.5% were freshman's, and 17.8% had 4-year experience, while the least had 5 years (14.5%). Also, 70.8% were female which is comparable publicly available gender structure data at Croatian HEIs.

After data had been collected, multivariate statistic methods were applied [31]. Precisely, to examine the dimensionality, confirmatory factor analysis (CFA) was used. Software tools used for data analysis were IBM SPSS Statistics and IBM SPSS Amos for structural equation modelling.

IV. RESULTS

As variables' normality is required for conducting a CFA, skewness and kurtosis have been calculated. Items measuring *personal interaction* subdimension have been excluded from further analysis since values have not met the referent value of +/-3 [32]. Also, three out of four items from the subdimension *responsible behavior* have been removed due to not meeting the referent value for normality.

After excluding the two subdimensions (*personal interactions* and *responsible behaviour*), six subdimensions were retained for further analysis (*information seeking*, *information sharing*, *feedback*, *advocacy*, *helping* and *tolerance*). CFA was conducted in AMOS, and the results are shown below in Figure 1.

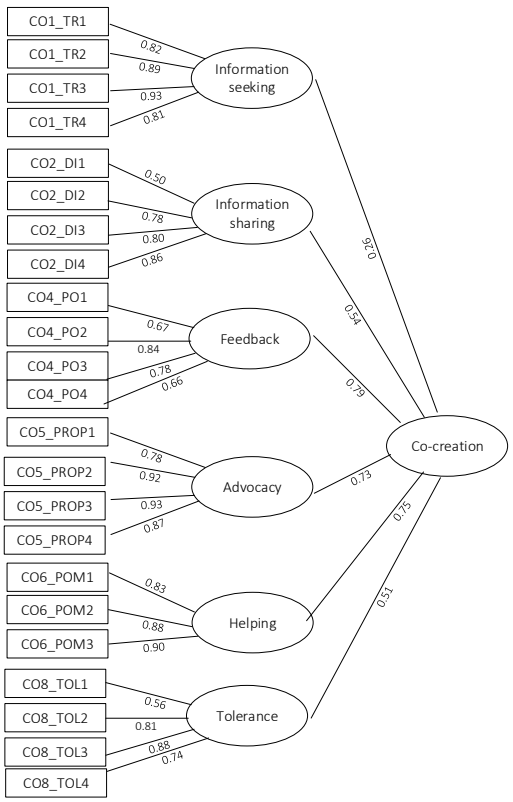


Figure 1. CFA results

TABLE II. VALIDITY AND RELIABILITY RESULTS

	FL	*CR*	*AVE*
Information seeking (I_SE)			
I have asked others about this digital service.	0.82	0.92	0.74
I have asked others how to access this digital service.	0.89		
I have asked others what is contained within this digital service.	0.93		
I have paid attention to how others use this digital service.	0.81		
Information sharing (I_SH)			
I clearly explain to employees/IT department what I want if I need something.	0.50	0.83	0.56
I share necessary information via this digital service.	0.78		
I share true information via this digital service.	0.80		
I share all the needed information so digital service can be realized.	0.86		
Feedback (FE)			
If I have idea for improving this digital service, I would be happy to share it.	0.67	0.83	0.55
If I am satisfied with this digital service, I will share it with others.	0.84		
I will comment with others on my experience with this digital service.	0.78		
If I experience problem with this digital service, I will be happy to report it.	0.66		
Advocacy (AD)			
I share my positive experience about institution with others.	0.78	0.93	0.77
I recommend the institution to other potential students.	0.92		
I encourage other students to apply for the institution.	0.94		
I encourage other students to attend the institution.	0.87		
Helping (HE)			
I assist other students if they need my help with using this digital service.	0.84	0.90	0.76
I teach other students how to use this digital service if needed.	0.88		
I happy give advice on how to use this digital service.	0.90		
Tolerance (TO)			
If problem with this digital service occurs, I will not stop using it immediately.	0.56	0.84	0.58
If problem with this digital service occurs, I am ready to wait and not to use alternative way (ex., F2F visit to institution).	0.82		
If it takes longer than usually while using the digital services, I am ready to wait.	0.88		
If this digital service does not meet my expectations, I will tolerate it.	0.74		

To evaluate reliability and validity, composite reliability (CR) and average variance extracted (AVE) have been calculated. The results show that the scale possesses good reliability, as the CR for each construct is greater than 0.70, and the AVE is greater than 0.50 [33] (see Table II). For accessing the reliability of items, factor loading (FL) has also been calculated and presented in the Table II. The smallest FL value can be 0.4, but only if the condition that the AVE is greater than or equal to 0.5 is met [33]. All results suggest that no further deletion of items is necessary.

Results for calculating reliability and validity are shown in Table II.

To assess discriminant validity, we used Fornell & Larcker criteria [34], and the results are shown in Table III.

The Fornell & Larcker criterion assumes that the second root of the AVE value of each observed construct should be greater than all correlations to other constructs - when this criterion is met, discriminant validity is achieved. Results point out that criteria has been met.

TABLE III. DISCRIMINANT VALIDITY RESULTS

	I_SE	I_SH	TO	FE	AD	HE
I_SE	**0.86**					
I_SH	0.28	**0.75**				
TO	0.07	0.29	**0.76**			
FE	0.28	0.40	0.46	**0.74**		
AD	0.07	0.32	0.32	0.62	**0.88**	
HE	0.22	0.47	0.38	0.54	0.59	**0.87**

To estimate model fit, several measures have been calculated. The value of the chi-square ratio and degrees of freedom (CMIN/df) must not be less than 2 and greater than 5 according to [35] and the calculated CMIN/df is 3.427. The comparative fit index (CFI) has been calculated and equals 0.94. The root mean square error of approximation (RMSEA) should not be above 0.08, according to [31], and it scores 0.66. Therefore, all calculated values are within the acceptable range.

The results confirm the dimensionality of the 29-item, six-dimension scale. However, since two subdimensions (*responsible behavior* and *personal interaction*), which belong to the dimension *Customer participation behavior*, were removed, the third level construct is not present in this study. However, all subdimensions of *Customer citizenship behavior* were retained.

Descriptive statistics have been conducted to calculate the average value for each of the six validated co-creation subdimensions. The results are graphically presented in Figure 2 in continuation.

The subdimension with the highest average value was *helping*, which had an average score of 4.06. This suggests that participants in the study placed a high value on their ability to help other students. The next two subdimensions with the highest average scores were *tolerance* and *advocacy*, both with an average score of 3.97.

The fourth subdimension, *information sharing*, had an average score of 3.91. This is followed by the fifth subdimension, *feedback*, with an average score of 3.71.

The lowest average score was for the subdimension *information seeking*, with an average score of 2.80. This suggests that students were less likely to actively seek information from digital services.

V. Discussion and Conclusion

Co-creation is a marketing concept that assumes

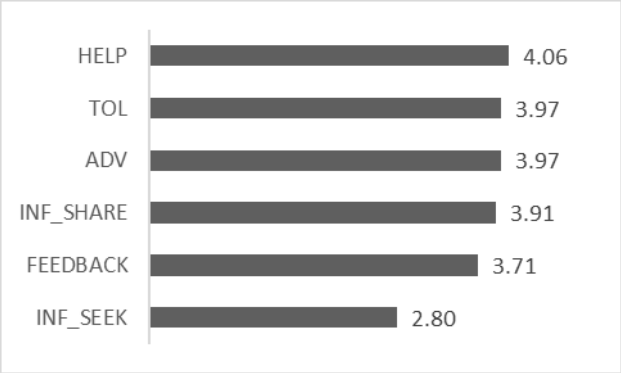


Figure 2. Average values for co-creation dimensions

collaborative value creation, where the customer receives higher value if they participate in the whole process. Co-creation has been examined in HE, but primarily regarding the face-to-face context. For example, authors have confirmed how co-creation can be used to bridge cultural differences [18] or how co-creating the teaching process enriches the overall process and provides higher value for students.

However, due to digital transformation that is changing the whole setting of HE, circumstances have also changed. The majority of services are now being consumed online [29], including the delivery of support services, such as applying for exams, obtaining different proofs, payments, and scheduling. Taking into consideration that the new generation of students has a different mindset [36], it is essential to provide value co-creation in an online setting as well.

The goal of our paper was to examine how to measure co-creation in this new digital environment of HE. To provide an answer to our research question, a quantitative survey was conducted on more than 500 students in Croatia. The results of this study confirmed that measuring co-creation in the digital environment differs from offline co-creation. The multidimensionality of the construct was confirmed; however, some subdimensions were omitted. Precisely, the dimensions *personal interaction* and *responsible behavior* did not confirm as relevant to measure co-creation in the digital environment.

As indicated in the literature review, personal behavior is essential since the service encounter occurs in a social setting [12]. However, if we take into consideration that the majority of services nowadays are being delivered online, it is questionable how this dimension remains important, even in other settings and not just in HE. The results of our study did not confirm the importance of subdimension *responsible behavior*. This can be explained by the fact that from the first day of enrolling in an HEI, it is expected that students follow the rules and procedures. This is something that students take for granted, given the HE context. Due to this, it can be expected that this subdimension may remain relevant in other digital settings as well.

There are a few limitations that need to be addressed in future research. Even though this study included participants from ten different HEIs, it would be interesting to conduct a similar study in another country to check whether these findings can be considered for all Z Generation.

Future research should be oriented toward examining antecedents of co-creation. Also, to use this instrument in the business digital environment, dimensions need further to be validated.

To summarize, the results of this study address a few research gaps and therefore make several scientific contributions. First, literature regarding co-creation in HE is relatively scarce. Thus, the empirical results of this study contribute to enriching the literature. Second, even though there are studies examining co-creation in HE, they haven't been examined in the digital environment. As mentioned, due to rapid digital transformation in HE, it is particularly important to measure and manage co-creation in digital settings. Through empirical research, this study validated the student value co-creation behavior scale.

In conclusion, this study provides valuable insights into the measurement of co-creation in the digital environment of HE. It suggests that while the multidimensionality of the construct is confirmed, some dimensions may not be as relevant in online settings as they are in face-to-face settings. The results of this study can be used to guide the development of co-creation strategies in the digital environment of HE and provide a foundation for future research in this area.

Acknowledgment

The research is supported by the Croatian Science Foundation [grant number IRP-2017-05-7625].

References

[1] D. Biedermann, L. Kalbfell, J. Schneider, and H. Drachsles, "Stakeholders attitudes towards digitalization in higher education institutions," pp. 187–192, 2019, doi: 10.18420/delfi2019.

[2] E. Plotnikova, "Digitalization of education in the leading universities of Saint Petersburg," IOP Conf. Ser. Mater. Sci. Eng., vol. 497, no. 1, 2019, doi: 10.1088/1757-899X/497/1/012047.

[3] A. Rof, A. Bikfalvi, and P. Marquès, "Digital transformation for business model innovation in higher education: Overcoming the tensions," Sustain., vol. 12, no. 12, 2020, doi: 10.3390/su12124980.

[4] T. Obaid, "Digital Transformation in Higher Education UniSZA Case Study," no. February, 2019.

[5] L. S. Rodrigues, "Challenges of digital transformation in higher education institutions: A brief discussion," Proc. 30th Int. Bus. Inf. Manag. Assoc. Conf. IBIMA 2017 - Vis. 2020 Sustain. Econ. Dev. Innov. Manag. Glob. Growth, vol. 2017-Janua, no. November 2017, pp. 4490–4493, 2017.

[6] O. P. Díaz, G. Ribes-Giner, and M. R. Perello-Marin, "The impact of cocreation on the student satisfaction: Analysis through structural equation modeling," Abstr. Appl. Anal., vol. 2016, 2016, doi: 10.1155/2016/3729791.

[7] J. Anttila and K. Jussila, "Universities and smart cities: the challenges to high quality," Total Qual. Manag. Bus. Excell., vol. 29, no. 9–10, pp. 1058–1073, 2018, doi: 10.1080/14783363.2018.1486552.

[8] T. Sutjarittham, H. H. Gharakheili, S. S. Kanhere, and V. Sivaraman, "Realizing a Smart University Campus: Vision, Architecture, and Implementation," Int. Symp. Adv. Networks

Telecommun. Syst. ANTS, vol. 2018-Decem, 2019, doi: 10.1109/ANTS.2018.8710084.

[9] M. Royo-Vela, G. Leszczyński, and M. Velasquez-Serrano, "Sustainable Value Co-Production and Co-Creation in Virtual Reality: An Exploratory Research on Business-to-Business Interactions," Sustain., vol. 14, no. 13, 2022, doi: 10.3390/su14137754.

[10] A. M. Shaltoni, H. Khraim, A. Abuhamad, and M. Amer, "Exploring students' satisfaction with universities' portals in developing countries a cultural perspective," Int. J. Inf. Learn. Technol., vol. 32, no. 2, pp. 82–93, 2015, doi: 10.1108/IJILT-12-2012-0042.

[11] Y. Sutarso, R. E. Halim, T. E. Balqiah, and P. Tjiptoherijanto, "Understanding customer co-creation activities in higher education: Groupings, characteristics and implications," Int. J. Bus. Soc., vol. 20, no. S1, pp. 42–56, 2019.

[12] T. H. Elsharnouby, "Student co-creation behavior in higher education: the role of satisfaction with the university experience," J. Mark. High. Educ., vol. 25, no. 2, pp. 238–262, 2015, doi: 10.1080/08841241.2015.1059919.

[13] S. L. Vargo, K. Koskela-huotari, and J. Vink, Service-Dominant Logic : Foundations and Applications, no. April. 2020.

[14] C. Grönroos, "Conceptualising value co-creation: A journey to the 1970s and back to the future," J. Mark. Manag., vol. 28, no. 13–14, pp. 1520–1534, 2012, doi: 10.1080/0267257X.2012.737357.

[15] M. Dollinger, "Technology for the scalability of co-creation with students," ASCILITE 2018 - Conf. Proc. - 35th Int. Conf. Innov. Pract. Res. use Educ. Technol. Tert. Educ. Open Ocean. Learn. Without Borders, no. November 2018, pp. 346–350, 2018.

[16] M. Dollinger, J. Lodge, and H. Coates, "Co-creation in higher education: towards a conceptual model," J. Mark. High. Educ., vol. 28, no. 2, pp. 210–231, 2018, doi: 10.1080/08841241.2018.1466756.

[17] M. K. Vespestad and K. K. Smørvik, "Co-Creation as a Tool to Overcome Cross-Cultural Differences in Educational Experiences?," J. Hosp. Tour. Educ., vol. 00, no. 00, pp. 1–11, 2019, doi: 10.1080/10963758.2019.1685391.

[18] N. Hasan and A. A. Rahman, "Exploring factors that influence customer engagement in value co-creation in higher education institutions using online platforms," J. Theor. Appl. Inf. Technol., vol. 90, no. 2, pp. 247–260, 2016.

[19] C. Bovill, "Co-creation in learning and teaching: the case for a whole-class approach in higher education," High. Educ., 2019, doi: 10.1007/s10734-019-00453-w.

[20] Y. Yi and T. Gong, "Customer value co-creation behavior: Scale development and validation," J. Bus. Res., vol. 66, no. 9, pp. 1279–1284, 2013, doi: 10.1016/j.jbusres.2012.02.026.

[21] P. Foroudi, Q. Yu, S. Gupta, and M. M. Foroudi, "Enhancing university brand image and reputation through customer value co-creation behaviour," Technol. Forecast. Soc. Change, vol. 138, no. May 2017, pp. 218–227, 2019, doi: 10.1016/j.techfore.2018.09.006.

[22] R. Maxwell-Stuart, B. Taheri, A. S. Paterson, K. O'Gorman, and W. Jackson, "Working together to increase student satisfaction: exploring the effects of mode of study and fee status," Stud. High. Educ., vol. 43, no. 8, pp. 1392–1404, 2018, doi: 10.1080/03075079.2016.1257601.

[23] B. Tari Kasnakoglu, "Antecedents and consequences of co-creation in credence-based service contexts," Serv. Ind. J., vol. 36, no. 1–2, pp. 1–20, 2016, doi: 10.1080/02642069.2016.1138472.

[24] C. L. Mui and J. S. Murphy, "The university of the future: Stiegler after Derrida," Educ. Philos. Theory, vol. 52, no. 4, pp. 455–465, 2020, doi: 10.1080/00131857.2019.1605900.

[25] N. S. Terblanche, "Some theoretical perspectives of co-creation and co-production of value by customers: original research, , 2014.," Acta Commer. no.2, pp. 1–8, 2014.

[26] D. Rico-Bautista, Y. Medina-Cárdenas, and C. D. Guerrero, "Smart University: A Review from the Educational and Technological View of Internet of Things," Adv. Intell. Syst. Comput., vol. 918, pp. 427–440, 2019, doi: 10.1007/978-3-030-11890-7_42.

[27] I. Blau and T. Shamir-Inbal, "Digital technologies for promoting 'student voice' and co-creating learning experience in an academic course," Instr. Sci., vol. 46, no. 2, pp. 315–336, 2018, doi: 10.1007/s11251-017-9436-y.

[28] T. H. Elsharnouby and A. A. Mahrous, "Customer participation in online co-creation experience: the role of e-service quality," J. Res. Interact. Mark., vol. 9, no. 4, pp. 313–336, 2015, doi: 10.1108/JRIM-06-2014-0038.

[29] A. Pasini, E. Estevez, and P. Pesado, "Assessment Model for Digital Services provided by Higher Education Institutions," ACM Int. Conf. Proceeding Ser., pp. 468–477, 2019, doi: 10.1145/3325112.3325268.

[30] A. Fagerstrøm and G. Ghinea, "Co-creation of value in higher education: Using social network marketing in the recruitment of students," J. High. Educ. Policy Manag., vol. 35, no. 1, pp. 45–53, 2013, doi: 10.1080/1360080X.2013.748524.

[31] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, Multivariate Data Analysis. Pearson Prentice Hall, 2010.

[32] T. A. Brown, Confirmatory factor analysis for applied research. New York: The Guilford Press, 2006.

[33] J. F. Hair, G. T. M. Hult, and M. Sarstedt, A promer on partial leat squares structural equation modeling (PLS-SEM). SAGE Publications, 2017.

[34] A. Afthanorhan, P. L. Ghazali, and N. Rashid, "Discriminant Validity: A Comparison of CBSEM and Consistent PLS using Fornell & Larcker and HTMT Approaches," J. Phys. Conf. Ser., vol. 1874, no. 1, p. 012085, 2021, doi: 10.1088/1742-6596/1874/1/012085.

[35] H. W. Marsh and D. Hocevar. (1985). Astrlication of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. Psychological Bulletin. US: American Psychological Association, 97(3), str. 562–582. doi: 10.1037/0033-2909.97.3.562.

[36] M. V Arkhipova, E. E. Belova, Y. A. Gavrikova, T. Pleskanyuk and A.N. Arkhipov (2019). Reaching Generation Z. Attitude Toward Technology Among the Newest Generation of School Students. Perspectives on the Use of New Information and Communication Technology (ICT) in the Modern Economy, str. 1026–1032. doi:10.1007/978-3-319-90835-9_114.

Classification of Text, Image and Audio Messages Used for Cyberbulling on Social Medias

Ermira Idrizi, Mentor Hamiti
South East European University, Faculty of Contemporary Sciences and Technologies, Tetovo, North Macedonia
e.idrizi@seeu.edu.mk, m.hamiti@seeu.edu.mk

***Abstract* - Cyberbullying has become an increasingly significant cultural issue in recent years. A person is affected by cyberbullying in both psychological and emotional ways. Bullying that takes place via electronic devices, such as a computer, a smartphone, or a tablet computer, is known as cyberbullying. Harassment in the digital world can take many forms, including but not limited to transmitting or uploading insulting, harmful, inaccurate, or offensive material about another person via text, message, or program; or on social networking sites, message boards, and online games. Disclosure of personal or sensitive information about another individual might be considered cyberstalking. This paper will analyze several media types (text, images, and videos) posted on social media with the goal of identifying instances of cyberbullying. In this study, a graph convolutional neural network, a pretrained Googlenet, a Mel-scale filter bank speech spectrogram, and a CNN network model are introduced for use in audio post-classification. The study's primary findings indicate that audio post-processing with MFCC's and graph convolutional neural networks generates improved outcomes, including one-dimensional representation, for both text and image properties. To this end, we used a combination of GCN and Melfrequency cepstrum to represent text, image, and video input in this setting, with a resulting 85% accuracy of a bullying class.**

Keywords – Cyberbullying, Image, Text, Audios, Facebook, Instagram

I. INTRODUCTION

As a result of the proliferation of social media platforms and their ever-increasing user bases, internet consumers now maintain an almost constant presence online. Users will post their thoughts, opinions, or feelings on social media, which will then be remarked on or addressed in public or private chats. In recent years, there has been a proliferation of brand new services that make it possible for users to view films, images, and comments pertaining to any kind of event. Many amusing pictures, videos, and comments have come out as a result of this trend. However, harmful videos, pictures, and comments have also been posted on social media. According to [1], some social networks, including Facebook, Twitter, and YouTube, are attempting to erase or prohibit articles that make it easier for people to become victims of cyberbullying. The ways in which individuals communicate with one another have undergone significant transformations over the past few decades, and this trend is expected to continue. Rapid technological progress has been brought about as a consequence of microaggressions that are not restricted to the external surroundings or facial expression settings. According to [2], the definition of conventional bullying is described as repeated and deliberate acts of violence directed toward a powerless victim. This power imbalance may be personal or emotional, and the assault may be verbal (for example, insults), physical (for example, hitting), emotional or contextual, or any combination of the three (e.g., gossiping). The term "harassment" can refer to anything from harmless teasing to overtly violent behavior, including but not limited to threats of self-harm, homicide, or other violent acts. From a different point of view, it's possible that overt and indirect forms of bullying are related. Direct bullying consists of assaults that are clearly visible on the victim, allowing the bully to be identified as the perpetrator. On the other hand, implicit abuse is transmitted to the person by a private entity (such as rumors), with the goal of isolating the victim so that the aggressor cannot be identified. This is done in order to prevent the victim from being able to identify the perpetrator of the abuse. With the advent of modern information and communication technologies (ICT) and the tremendous explosion of social network systems that allow teenagers to communicate online, social connections have gained a new mechanism for communication. This is particularly true in the case of online communication. Social interactions involving offensive online content are common in these online interactions, as this is one of the primary manifestations of violence [3] in cases of cyber harassment, such as cyberbullying. These online interactions take place in the context of social interactions. Abuse that occurs between peers through the use of electronic media can be described in a number of different ways; nevertheless, there is consensus among researchers that this type of behavior must be intentional, harmful, and repeated for it to be considered abuse.

A. Cyberbullying forms

Cyberbullying can take place via chat rooms, e-mail, instant chats, blogs, social networking sites, and texting. It can also take place using personal computers, laptops, mobile phones, and online video games. Verbal and emotional aggression are the most common forms of

cyberbullying. This can take the form of things like spreading false rumors, disclosing personal information that is either genuine or erroneous via texts, or uploading things on social media websites. The use of fire, denigration, masquerading, outing and deceit, exclusion, and cheerful slapping are some of the other sorts of abuse. To prevent and respond to cyberbullying, digital solutions that include automatic cyberbullying identification and messages that encourage behavioral self-reflection can be utilized. Therefore, to reduce instances of cyberbullying, classifiers must be as precise as feasible. However, it is still difficult to detect these events online in near-real time, dependent on the operationalization of cyberbullying given by automated cyberbullying detection systems, among other things [4].

B. Cyberbullying algorithm for detections

Cyberbullying is a major issue in modern society. Unfortunately, perusing the entire Internet to uncover cyberbullying postings is like hunting for a needle in a haystack, while leaving cyberbullying victims exposed to negative words causes them to get depressed, self-mutilate, or commit suicide. Cyberbullying is addressed manually via Internet Patrol. Children and teens use Instagram and Facebook to post photos and videos. Visual (image and video) content accounts for over 70% of internet traffic, and over 80% of adolescents use mobile phones, making them the most ubiquitous technology and a common cyberbullying platform [8]. Deep learning can detect cyberbullying on social media, curation, wikis, tweeting, forums, and bookmarking. These algorithms automatically detect cyberbullying in massive data sets. These self-organizing and self-learning algorithms can help detect online abuse or bullying and automate cyberbullying detection. Computer scientists now use data collection methods to anticipate cyberbullying instead of language and graphics. Data mining is a new field that extracts useful information from massive amounts of data. Data scientists utilized deep learning algorithm to predict cyberbullying messages using picture, text, video, and URL data. Current cyberbullying detection methods either search social media for text messages or integrate text and visual attributes. The current study focuses on developing a deep learning model that predicts the effect of cyberbullying texts on the general sentiment of individuals [5]. Unique to the current technique is the consideration of the impact of sentiment analysis on cyberbullying identification. In addition, the effect of video and URL as attributes and how they can be used to forecast cyberbullying communications. A significant contribution of this study is the use of video and URL analysis in the building of a deep learning model to detect cyberbullying texts.

II. RELATED STUDIES

A. Text-based cyberbullying detection

The first work to detect bullying behaviors that include social media harassment [6], using the CAW 2.0 Twitter corpus, the Naive Bayes(NB) was trained, Sequential Minimal Optimization (SMO), J48, bagging, and digging algorithms using the information benefit algorithm. Utilize the Synthetic Minority Oversampling Technique to construct a sound training dataset (SMOTE). This model included textual and social network components. Twitter's messages are described as "1.5-ego networks." A network consists of nodes and ties. Links represent the number of postings shared between users, while nodes represent the users themselves [7]. This ego network is used to determine which nodes have no contact with their peers and which people are more influential inside their social networks. As a result of the increased number of messages sent between users, bullying messages were described. In the digging classification algorithm, the ROC curve and true positive rate yielded the best results. However, simply the quantity of messages exchanged was utilized to establish whether or not a post was harassing, and demographic factors were not considered in the study. A system was suggested that uses SNA (Social Network Analysis) and text mining to identify key cyberbullies from 650,000 posts collected through web crawling using three sets of features: relationship between users, positive/negative rate of comments, and insulting word rate of comments based on group level and user level with random forest, logistic regression, and SVM classifiers [8]. A random forest classifier gave them a precision of 0.81 on the consumer level and an adapted to specific ratio of 0.74:1 on the group level[9].

B. Cyberbullying Detection: Applying Text and Pictures

Today, image-based cyberbullying is widespread and has far-reaching repercussions on society. They appear to be rapidly spreading through social networks. A VGG-16 and CNN (Convolutional Neural Network) were proposed for picture and text feature extraction [10]. The accuracy, recall, and F1 scores were for Twitter 80%, Facebook 79%, and Instagram 78% data, respectively. A two-dimensional convolution comparison on images and one-dimensional convolution on texts was implemented, as well as TF-IDF vectorization to transform one-dimensional text into a two-dimensional square matrix and integrate image and text into a distinct two-dimensional matrix. Then, a single-layer CNN with a wide filter is compared to a multiple-layer CNN with fewer filters and a 68 percent weighted average F1-score. It was proposed using enough graphic and text weight to distinguish cyberbullying communications in future investigations [22][23][24].

Due to the interdependence of posts like messages, pictures, sounds, and videos, social media cyberbullying

detection is challenging [11][12]. In Table 1 we list some cyberbullying detection methods:

TABLE I. CYBERBULLYING DETECTION METHODS

Obtained Method	Findings	Limitations
Using Synthetic Minority Oversampling, cyberbullying messages can be classified. Huang and colleagues (2014).	Social connections improve precision.	Demographics are avoided during detection.
SVM and KNN algorithms (Hani Nurrahmi et.al 2017)	Provides RBF Kernel and parameter settings for the most accurate recognition of cyberbullying text.	POS tagger may struggle with multiple languages.
Convolutional Neural Network with Shortcuts at the Character Level (Nijia Lu et.al 2019)	The focal loss function and smallest learning unit address class imbalance.	This only works on personality characteristics and is expensive for larger datasets.
Deep Neural Network Architecture(Saima sadiq et.al 2020)	Unigram, Bigram, and TFIDF accurately extract key features.	Single-feature analysis takes time and reduces accuracy.
CNN and pretrained fastText(Sandip Modha et.al 2019)	A browser-based platform detects and visualizes online hostility.	Different hostile and sensitive topic thresholds can affect accuracy.

III. METHODOLOGY

Data was collected from leading online social media platforms including Facebook and Instagram by employing search terms such as fury, impotence, ugly, and scary photographs. By searching Google for image, text, and audio queries, the needed datasets was able to be generated. A graph convolution neural network to classify cyberbullying was used.

A. Cyberbullying Investigation Techniques Using SNA

Social network analysis refers to the statistical and qualitative examination of a networking website (SNA)[13]. SNA monitors and organizes the flow of connections and changes in concepts and relationships with insight. Websites, computers, animals, people, groups, organizations, and nations are examples of simple and complex entities. Consequently, implementing SNA metrics could aid us in researching and understanding many aspects of this network.

Recognition of the *Society A*, where society is a subset of nodes in a graph that are strongly connected to one another but weakly correlated to nodes in the graph's other neighborhoods. A connection with others using sites such has Facebook, Instagram, and Twitter were established. In these social circles, relatives, classmates, coworkers, and others may be included. Sensing societies in a network is one of the most significant jobs of network analysis. There could be millions of nodes and edges in a huge network, such as an online social network. It becomes extremely difficult to identify communities in such networks. Therefore, strategies for detecting societies that can partition the network into distinct neighborhoods are necessary. In the context of the agglomerative approach, we begin with a graph consisting of nodes from the actual line but without edges. The edges are then added one by one to the graph, beginning with the "larger" edges and moving to the "smaller" ones. This edge strength, often referred to as weight, is calculable and represents a subset of the individuals who posted the bullying remark. Analyzeing cyberbullying communications in the network using [14] Girvan-Newman Algorithm for Community Detection.

In this instance, neighborhood identification is used to discover significant and influential themes on social media that involve a big number of persons, and then to investigate the existence of cyberbullying messages behind these significant topics. Figure 1 depicts the community size distribution across the network: The X axis represents the number of communities, while the Y axis represents the size of the communities, which follows a power-law distribution. Power law degree distribution functions determine the likelihood that a node **u** has the degree **du.** Power law distributions and other related variations have the characteristic of being substantially peeled.

Formally, a probability distribution with heavy tails approaches zero for absolute values more slowly than a probability function. This implies that events that are effectively "impossible" in a conventional exponential distribution are assigned a non-trivial probability mass in huge distributions. Due of the heavy tails of degree distributions, the probability of encountering a tiny number of assurance degree nodes is not zero. Figure 1 represents networks large number of nodes with low degrees and a small number of nodes with extremely high degrees.

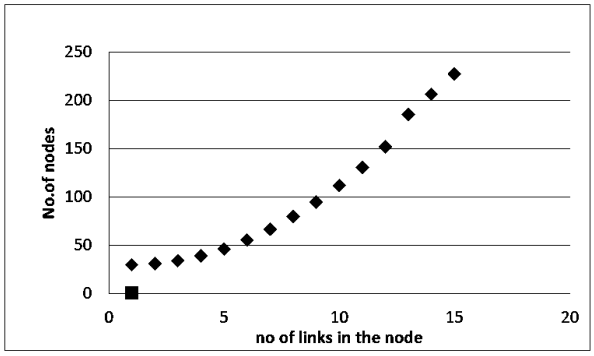


Figure. 1 Distribution of Power-law

B. Text Message Classification for Cyberbullying

It is commonly known that Twitter and Instagram are seeing an increase in cyberbullying towards young users, and that scalable computational solutions must be created to reduce this form of abuse.

Given this, a graph convolutional networks (GCN) was employed to identify cyberbullying [15]. The semi-supervised convolutional neural network-based GCN learns graph-structured data.

Node-level classification for graphs with few labels. This method accepts a feature matrix *ND* and feature description *xi* for each connection point *I*, where ***N*** is the number of vertices and ***D*** is the number of data quality. The amount of output features determines the node level output **Z N*F** from an adjacency matrix **AN*N**. Two dropout **GCN** layers with **L2** regularization were examined. Early Stopping with a 10-minute will be taught patience and 200 training epochs. If validation loss does not decrease for 10 epochs, training will stop. Transudative Learning means we'll train and test utilizing the full graph. Boolean masks divided training, validation, and testing data. These masks feed the sample weight argument. Before training a network, category features must be numeric. Single-pass vectorize categorical values.

Table 2 represents the results of GCC using Text data, with number of nodes 1240, average path length 2 with a F1 score 0.78 which leads to a accuracy of 85%.

TABLE 2. RESULTS OF GCC USING TEXT DATA

Nodes	Edges	Average Path Lenght	Average CC	Average Degree	F1 score	Accuracy
1240	28,975	2.02	0.221	1.876	0.78	85%

C. Classification of cyberbullying images

The initial component of the proposed model is GoogLeNet [16], a pretrained image model, a program for extracting photographic features. Using 3*3 filters, it applies the convolution layers to a 224*224*3 picture. Replace the last three layers of the network to retrain Google Photos to label recent photographs. The layers "loss3-classifier," "prob," and "output" include instructions for converting the network's characteristics into probabilities and symbols. To the layer graph, add three more layers: a fully connected layer, a softmax layer, and a classification output layer. The final, fully connected layer is the same size as the number of classes in the new data set, it increase the learning rate parameters of the entirely linked layer to learn new layers faster than transferred layers. Then "pool5-drop 7x7 s1" connects the newly transmitted network layer to the previously transmitted layer. A graph of the new layer is created and zoomed in on the final layers of the cable network to confirm that the new layers are properly connected. The network is now ready to be retrained with the new images. Optionally the weights are frozen of the network's older layers by setting their learning rates to zero.

During training, the network does not continuously update the variables of the frozen layers upon layers. Freezing the weights of several initial layers can considerably accelerate network training due to the elimination of gradient computations for the frozen layers. Freezing preceding network layers can protect against fitting problems if the new collection of data is minimal. By attempting to extract the layers and linkages from the layer graph, users can select which layers to freeze.

D. Classification of audio cyberbullying posts

Cyberbullying in audio form refers to bullying behavior that involves using audio recordings to harass, threaten, or intimidate someone. To classify audio cyberbullying posts, one could use several approaches:

- Content-based classification: This approach uses the content of the audio recordings to classify them as cyberbullying or not. Features such as speech content, tone, language, and sentiment could be used to train a machine learning model to automatically classify audio posts.
- Context-based classification: This approach considers the context in which the audio recording was made, such as the platform it was posted on, the identity of the speaker and the audience, and the relationship between the speaker and the target.
- Hybrid approach: A combination of content-based and context-based classification could also be used to get more accurate results.

Accurately identifying audio cyberbullying can be challenging, as the context of the audio and the intent of the speaker may not always be clear. Nevertheless, classification systems can be useful for detecting and addressing audio cyberbullying and for improving the overall safety and well-being of those affected by this type of behavior. The classification task for audio cyberbullying consists of two components: feature selection and categorization. A *Mel-scale filter* was applied to bank speech spectrogram as an input feature. The same signal was then shown in a time-frequency domain, allowing to evaluate the signal's time-varying frequency and amplitude. To achieve this result, a Fourier transform on our data was used. Cepstrum is the new name for the output spectrum, whereas Quefrencies refers to the new frequencies. With our new cepstrum and quefrequencies, we can easily discern between our spectral envelope (the vocal tract) and the background noise (the glottal pulse). Moreover, the discrete cosine transformation produces the mel-frequency cepstrum coefficients, or MFCCs, which are the major and most significant coefficients of this cepstrum. Figure 2 shows

the main parts of a MFCC block diagram: frame blocking, windowing, FFT, Mel-frequeny and Cepstrum.

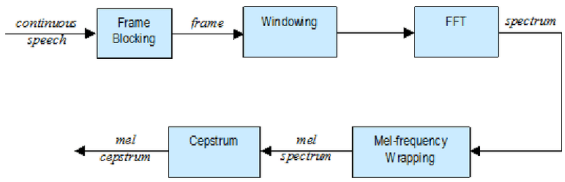


Figure. 2 MFCC Block Diagram[17]

In table 3 results are shown for text, images and voice using different models described apriority. Where all models show a high F1 score with accuracies all above 80%.

TABLE 3. RESULTS OF CLASSIFICATION USING DIFFERENT MODELS

Data	Model	F1 Score	Accuracy
Text	CNN	0.60	88%
	LSTM	0.61	89%
	1DCNN	0.63	91%
	GCNN	0.69	96%
Image	VGG-16	0.70	86%
	VGG-19	0.75	73%
	Resenet	0.78	75%
	GoogLeNet	.079	95.8%
Voice	MFCC	0.81	92%

E. Result analyzes

Using social network analysis, cyberbullying messages were categorized. To represent the data as a network, several SNA techniques were utilized. Using a graph-based deep learning method, the classification task was then performed. All existing methods accomplish the objective, however, using basic machine learning approaches and a limited data set. To assess and differentiate cyberbullying conversations, a novel whale optimization technique was applied on a wide variety of huge datasets, including text, image, and voice.

IV. Discussion and Conclusion

In addition to the benefits of social media, cyberbullying has a variety of well-established negative repercussions. Cyberbullying is one of the most damaging uses of social media, in which digital devices are used to threaten, intimidate, or humiliate an online user. Classification of cyberbullying messages using cutting-edge allometric data was established. Cyberbullying consequences and the degree of distribution and community size follow a power law distribution, and GoogLenet [18] and Mel-frequency cepstrum coefficients are employed to spread information. Each influential almetrics network community contains a number of cyberbullying messages that encourage the transmission of abusive or harsh comments about others as well as the sharing of videos, images, and texts with the intent to injure or dishonor others. Graph convolutional neural networks, GoogleNet, and MFCCs to develop an autonomous classification system for cyberbullying were employed. Our proposed method is capable of accurately categorizing 83% of cyberbullying scenarios. We discovered that GCN with Mel-frequency cepstrum coefficients and a one-dimensional convolution layer with a larger filter size beats multiple convolution layers with fewer filters [19].

For the cyberbullying identification test, only image, text, and voice were considered; however, video and the URL of the post can also be utilized to identify cases of cyberbullying. For the goal of cyberbullying identification, other characteristics of the post, such as network information, URL, and video content, may also be evaluated [20].

Images with multilingual annotations can be used to generate improved futures. To improve the overall effectiveness of cyberbullying comment detection, more optimization strategies for enhanced feature extraction must be investigated.

In conclusion, classifying cyberbullying messages on social media is a crucial task for ensuring the safety and well-being of individuals in online communities. By using NLP and machine learning techniques, it is possible to accurately identify and categorize messages that contain abusive or threatening language. This helps to mitigate the harm caused by cyberbullying and promote a safer and more positive online experience for users. However, it's important to note that this is an ongoing challenge and requires continuous improvement and refinement of these algorithms as cyberbullies adapt and find new ways to harass and harm others online.

References

[1] A. Sánchez-Medina, Galván-Sánchez I, Fernández-Monroy M. Applying artificial intelligence to explore sexual cyberbullying behaviour. Heliyon. 2020 Jan 1;6(1):e03218.

[2] R. Slonje, Smith PK, Frisén A. The nature of cyberbullying, and strategies for prevention. Computers in human behavior. 2013 Jan 1;29(1):26-32.

[3] GW. Giumetti, Kowalski RM. Cyberbullying via social media and well-being. Current Opinion in Psychology. 2022 Feb 19:101314.

[4] A. Bozyiğit , Utku S, Nasibov E. Cyberbullying detection: Utilizing social media features. Expert Systems with Applications. 2021

[5] M. Alotaibi, Alotaibi B, Razaque A. A multichannel deep learning framework for cyberbullying detection on social media. Electronics. 2021 Oct 31;10(21):2664.

[6] A. Perera, Fernando P. Accurate cyberbullying detection and prevention on social media. Procedia Computer Science. 2021 Jan 1;

[7] S.Bharti, Yadav AK, Kumar M, Yadav D. Cyberbullying detection from tweets using deep learning. Kybernetes. 2021 Jul 13;51(9).

[8] J. Zhang, Otomo T, Li L, Nakajima S. Cyberbullying detection on twitter using multiple textual features. In2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST) 2019 Oct 23 (pp. 1-6). IEEE.

[9] V. Balakrishnan, Khan S, Fernandez T, Arabnia HR. Cyberbullying detection on twitter using Big Five and Dark Triad features. Personality and individual differences. 2019.

[10] HC. Chan, Wong DS. Traditional school bullying and cyberbullying in Chinese societies: Prevalence and a review of the

whole-school intervention approach. Aggression and Violent Behavior 2015.

[11] J. Yadav, Kumar D, Chauhan D. Cyberbullying detection using pre-trained bert model. In2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) 2020 Jul 2 (pp. 1096-1100). IEEE.

[12] M. Makram, Ali N, Mohammed A. Machine Learning Approach for Diagnosis of Heart Diseases. In2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) 2022 May 8 (pp. 69-74). IEEE.

[13] L. Cheng, Guo R, Silva Y, Hall D, Liu H. Hierarchical attention networks for cyberbullying detection on the instagram social network. InProceedings of the 2019 SIAM international conference on data mining 2019 May 6 (pp. 235-243). Society for Industrial and Applied Mathematics.

[14] L. Cheng, Mosallanezhad A, Silva YN, Hall DL, Liu H. Mitigating bias in session-based cyberbullying detection: A non-compromising approach. InThe Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP) 2021 Aug (Vol. 1).

[15] F. Elsafoury, Katsigiannis S, Pervez Z, Ramzan N. When the timeline meets the pipeline: A survey on automated cyberbullying detection. IEEE access. 2021 Jul 21.

[16] D. Soni, Singh VK. See no evil, hear no evil: Audio-visual-textual cyberbullying detection. Proceedings of the ACM on Human-Computer Interaction, 2018.

[17] JS. Raj, Anantha Babu S. Smart Cyberbullying detection with Machine Learning. InDisruptive Technologies for Big Data and Cloud Applications: Proceedings of ICBDCC 2021 2022 Aug 2 (pp. 237-248).

[18] GB. Anwar, Anwar MW. Textual Cyberbullying detection using Ensemble of Machine Learning models. In2022 International Conference on IT and Industrial Technologies (ICIT) 2022 Oct 3 (pp. 1-7). IEEE.

[19] J. Qiu, Moh M, Moh TS. Multi-modal detection of cyberbullying on Twitter. InProceedings of the 2022 ACM Southeast Conference 2022 Apr 18 (pp. 9-16)

[20] PK. Roy, Mali FU. Cyberbullying detection using deep transfer learning. Complex & Intelligent Systems. 2022 Dec;8(6):5449-67.

[21] S. Bharadwaj, R., Kuzhalvaimozhi, S., & Vedavathi, N. (2023). A Novel Multimodal Hybrid Classifier Based Cyberbullying Detection for Social Media Platform. In Data Science and Algorithms in Systems: Proceedings of 6th Computational Methods in Systems and Software 2022, Vol. 2 (pp. 689-699). Cham: Springer International Publishing.

[22] W.M Yafooz., Al-Dhaqm, A., & Alsaeedi, A. (2023). Detecting Kids Cyberbullying Using Transfer Learning Approach: Transformer Fine-Tuning Models. In Kids Cybersecurity Using Computational Intelligence Techniques (pp. 255-267). Cham: Springer International Publishing.

[23] Y. Zhao, Chu, X., & Rong, K. (2023). Cyberbullying experience and bystander behavior in cyberbullying incidents: The serial mediating roles of perceived incident severity and empathy. Computers in Human Behavior, 138, 107484.

International Cooperation for Digital Innovations in Primary Schools

Bogusława Denys* and dr Bogusław Klimczuk**
*Zamojskie Towarzystwo Oświatowe, Zamość, Poland
**Zamojskie Towarzystwo Oświatowe, Zamość, Poland
*bdenys22@gmail.com
**klimczuk@wp.pl

The contemporary world is technology-oriented and technology users are younger and younger. Digital literacy has become one of the basic skills that educational institutions focus on delivering for. Together with the technological advancement, the skill of programming has emerged as no longer the need of the hour, but as a lifelong skill. The search for innovative ways of introducing digital content to the youngest users constitutes the main objective for educational institutions, together with an urge to come up with effective educational solutions to provide primary schools with tools and methods to teach programming. Cooperation of varied institutions, ranging from primary schools, universities, cybernetics and robotics societies, to renowned publishing companies, proves to be an effective initiative making exchange of best practices possible and joint efforts realistic. The paper presents the outcomes of the cooperation of European institutions from Poland, Croatia, Portugal, Italy, Finland, Denmark, The Netherlands and Turkey under three international Erasmus+ projects. Jointly created databases, tools and methods to teach about programming and AI include innovative ideas for teaching computational skills with and without technology. International cooperation guarantees wide access to the materials for elementary school educators worldwide and fosters digital innovation in schools.

Key words – ***primary school education; programming; AI; Erasmus+ projects; innovation***

I. INTRODUCTION

Why has technology become so important in our lives? 'It is an essential tool that we cannot avoid, it plays an important part in the majority of our lives. Technology essentially harnesses the tools, technologies and strategies used to help us solve problems and simply make our lives better and easier (…).Technology is inevitable in our everyday lives.' [1] With that said, digital literacy can be defined as an indispensable skill for every individual who needs it to interact with the surrounding world. Digital competencies, being undoubtedly the crucial ones, are no longer to be gained at secondary or high school levels only, the age of the technology user has dropped significantly. When it comes to children the advantages of digital education is far-reaching. Not only is it a powerful tool to expand the students' learning opportunities but it also constitutes an important part of their development and future prospects, both social and professional.

'For many children digital devices and the possibilities they enable are threaded through everyday life from the earliest days, and their early experiences and understandings are patterned by technology use.' [2]. Therefore, educational institutions are in need for innovative solutions towards digital education and progress from the possibly earliest stages of education.

'Fostering digital literacy means going beyond functional skills and the ability to complete basic internet searches and PowerPoint presentations. It means giving students the opportunity to use a wide range of technologies collaboratively, creatively and critically.' [3] Unfortunately, in many cases, primary school teachers do not feel adequately prepared for the challenge of introducing concepts such as coding or Artificial Intelligence to the youngest learners in an attractive and efficient way. What is more, with the digital development progressing so rapidly, working out highly effective educational solutions is not an easy task for a single institution. Therefore, international cooperation of varied institutions serves as an innovative approach towards educational digital innovation so necessary to meet the demands of the contemporary school. It provides an opportunity to equip schools with the innovative tools and solutions to guide the youngest learners to become not only the users, but also the creators of the digital content. This triggered European educational institutions' efforts under Erasmus+ projects to create cutting-edge educational databases for teaching programming and elements of Artificial Education to primary school students. Not only do the databases consist of technology-based curricula, lesson plans and textbooks for teaching programming and technology-related topics, but these are also suggestions for introducing the concept of programming without technology as such, with the innovative MEMA method being the most representative and effective example of this. Without international collaboration of educational institutions the exchange of best practices would not have been possible on such a scale.

As Henry Ford (1863-1947) once said: 'Coming together is a beginning, staying together is progress, and working together is success.'

II. DIGITAL INNOVATIONS IN PRIMARY SCHOOLS

Digital innovations in schools seem inevitable. Why is it so important that schools teach coding? 'In a world where there's a lot of content and games to consume, it's important to have creators of good, engaging, and quality content. By learning how to code students have an

opportunity to be the creator, not just the consumer.' [4] With the digital literacy becoming one of the major skills to be mastered by primary school students, it is vital for the educational institutions to come up with ideas on how to engage young users into becoming active participants in the digital transformation. The ability to passively use the digital contents may not be enough within a decade. Competences related to programming are undoubtedly worth investing the educational efforts into. 'Today, the importance of learning to code rivals even that of reading and writing. It's a core skill that can help a child develop a deeper understanding of how technology works. Given the extent to which technology shapes our lives, learning to code helps develop a better understanding of the world around us.'[5]

A direct answer to the above mentioned expectations is the Erasmus+ CodeInnova project. The initiative created by five European educational institutions: Zamojskie Towarzystwo Oświatowe, Zamość, Poland; I. Osnovna škola, Čakovec, Croatia; Školska knjiga d.d., Zagreb, Croatia; JYU, University of Jyvaskyla, Finland; Agrupamento de Escolas de Fornos de Algodres, Portugal aimed at the creation of uniform school curriculum for teaching programming in primary schools together with the teaching database. The following project results have been created under the project:

- Joint and coherent core curriculum for teaching programming at primary school;
- 50 model lesson plans, coherent with the core curriculum;
- 2 textbooks for teaching programming at primary school: for grades 1-4 and 5-9;
- Internet platform to support teaching programming at primary school, including didactic materials in 5 languages: English, Croatian, Polish, Portuguese and Finnish. (https://www.skolskiportal.hr/projekt/erasmus-2/)

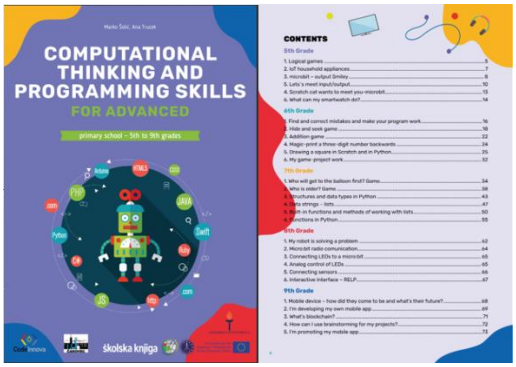


Figure 1: Erasmus+ CodeInnova Textbook Grades 5-9

The idea of introducing the concept of AI into primary schools has been challenged by Erasmus+ project ARTIE: 'Artificial Intelligence in Education - challenges and opportunities of the new era: development of a new curriculum, guide for educators and online course for students". The strategic partnership of: I. osnovna škola Čakovec, Croatia; Agrupamento de Escolas do Barreiro, Portugal; I Społeczna Szkoła Podstawowa im. Unii Europejskiej, Zamość, Poland; Croatian Robotics Association, Croatia; and Nefinia, The Netherlands, aim at the creation of tools and methodologies to guide educators on how to introduce elements of Artificial Intelligence to young technology-oriented learners. The educational final project result 'AI-Edu Kit' to be composed under the project comprises of:

- AI in Education: A new curriculum; (https://www.erasmus-artie.eu/);
- Challenges and opportunities of the new era: A Handbook for Teachers;
- TwinSpace - an open window to other educational realities: Online course for students.



Figure 2.: Erasmus+ ARTIE New Curriculum

The project outcomes constitute ready to use educational materials, leading the educators through the process of teaching primary school students to become active creators of digital contents. Together with the competences of coding, and with the inevitable wave of technological innovation, the understanding of Artificial Intelligence (AI) and introducing its concept to young learners seems to be the next challenge for a modern school. 'The future lies in artificial intelligence and it is best to start teaching children this new world as young as possible.'[6] Nowadays, we are experiencing a wide scope of technologies driven by AI, sometimes even without realizing it. It is important to make sure that young users understand its concept and how it may affect their lives. AI literacy seems to be the next skill to me mastered in order to make the most of what AI has to offer. The question that emerges is what the young users want AI to be like in the future, and how they can contribute to their vision becoming realistic. 'AI Literacy is an organic part of digital literacy for all citizens in an increasingly intelligent society'. [7]

Both European initiatives are the answer to high demand for educational methodological content to be used by teachers and students of primary schools. Materials translated to a few languages: English, Croatian, Finnish, Portuguese, Dutch, Polish, constitute easily accessible database for educators worldwide. Not only do they promote innovation in digital transformation of schools, but they also serve as an example of joint efforts of European institutions to improve the quality of educational systems in general.

III. MEMA METHOD FOR TEACHING PROGRAMMING

With the technological revolution and the necessity to develop digital literacy, technology seems to be the most obvious way to teach coding. But is it the only and most obvious idea? Do we need to computers to teach programming?

'Coding is not just about computers, but about computational and sequential thinking, logical reasoning, creativity, and problem solving skills. This does not just apply to the work of advanced mathematicians and logicians, but to the simplest daily tasks which are familiar even to children.' [8] The truth is that children can learn simple algorithms by playing with LEGO blocks or playing cards. Programming had been invented before the technologically advanced devices were created. In order to provide the educators of primary schools with ideas of how to teach coding without technology, especially to very young learners, the innovative MEMA method has been developed. The initiative undertaken by four European institutions: Cybernetics society Rijeka, Croatia; I. osnovna škola Čakovec, Croatia; Zubeyde Hanim Ortaokulu, Umraniye, İstanbul, Turkey; Zamojskie Towarzystwo Oświatowe, Zamość, Poland under Erasmus+ project 'Cybernetics in Early Childhood Education – MEMA as a Teaching Method'. The project is intended to result in the creation of:

- the Conceptualization of the MEMA method;
- Textbooks for Teachers and Students;
- Lesson Plans corresponding to the Textbooks contents.

What is the MEMA method then? It begins with the simple question: 'When should we start learning programming?' and a very easy answer to follow: 'As soon as possible'. The MEMA method is an innovative method that prof. Ph.D. Marina Čičin Šain designed more than 43 years ago. [9] The MEMA method is an approach to teaching programming to very young learners without any equipment or software. It allows for teachers without any specific knowledge of programming languages to conduct the lessons. It enables children to learn without much effort while doing their regular schoolwork. It involves creativity and teamwork, but without the feeling of being forced to memorize some new contents or having to learn some complex rules and definitions. It involves having fun while revising issues taught in various school subjects. It is about learning in the most natural way, the way we learn a language while being small children, without even realising that. Using the MEMA method is a gradual approach and its biggest advantage is that it can be easily incorporated in most school subjects that children learn at school anyway. Its gradual repetition, learning step by step, revising the school contents the students are familiar with, tasks of different complexity, which stimulate and encourage students to seek for new solutions, all these are the strong points of the method.

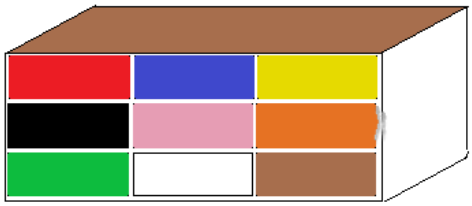


Figure 3.: MEMA toy that helps to understand how the main computer memory works

The modernized MEMA method together with the materials constituting the educational database will be available in English, Polish, Croatian, Turkish to enlarge the scope of their effectiveness and equip primary schools with the innovative idea on how to teach coding without computers in a very easy and efficient way, available to everyone interested.

IV. INTERNATIONAL EFFORTS FOR DIGITAL TRANSFORMATION

'Digital transformation in the education industry helps improve the learning experience for both students and teachers, as well as other people involved in the process. Such changes focus on improving engagement and accessibility through interactive and customizable learning.' [10] Digital transformation involves cooperation, especially when it comes to transformation of educational systems. However, educational innovations are rarely the result of a simple institution's efforts. Cooperation of varied institutions, ranging from primary schools, universities, cybernetics and robotics societies, to renowned publishing companies, proves to be an effective initiative making exchange of best practices possible and joint efforts realistic. The creation of educational tools and methodologies that can be applicable in varied educational systems requires high quality collaboration. International cooperation allows for the exchange of best solutions, ideas, with access to initiatives not previously available to all the parties.

Fast-paced technological development poses new challenges to schools. 'To adapt to changes in the landscape of the jobs market, schools should adjust their curriculum and incorporate coding.'[11] Educational systems need to be adjusted to technology-oriented young students and school curricula equipped in tools providing educators with practical and effective methodological materials to be used in the classroom. Educational institutions need to be open to innovative solutions in order to keep up with the changes that the era of digitalization provides. Cooperation of institutions under international educational projects is much more effective as its impact is much wider. International databases of educational contents , available in many European languages, enhances the scope of the cooperation and constitutes the basis for further improvements worldwide. The international cooperation under Erasmus+ projects constitutes an example of joint efforts to improve educational systems in varied educational environments. Projects ranging from those directly focused on the digital contents, making use of the newest technologies, to projects making use of innovative methods to teach

programming without technology (such as the MEMA method), enable educational institutions experiment with varied and innovative approaches to teaching elements of programming and Artificial Intelligence to primary school students.

The methodological approach of the Erasmus+ CodeInnova, ARTIE and CybeMEMA projects is based on multidisciplinary approach. It allows for the creation of effective tools and learner-oriented methodology that can be used in many school subjects with or without the use of technology. The wide scope of methods, ready-to-use materials such as school curricular, textbooks, guides for teachers, lesson plans not only provides the educators with innovative ideas for digital education, but it also brings educational institutions to a higher level as they can provide service of higher and more effective quality.

The projects bring long-term benefits aiming at one goal - to improve the quality of digital education of primary schools and provide educators with innovative teaching aid databases, also the ones not requiring advanced technological equipment and thus enhancing the possibilities of introducing programming in varied educational environments. Erasmus+ funding made it possible for European institutions to join their efforts in search of such educational innovations.



Figure 4.: Erasmus+ Programme logo

V. CONCLUSION

The transformation of educational processes in the field of digitalization is inavitable and schools need to adjust to the requirements of the technological progress. Digital competences have become the basic ones which all modern school curricula should include and make use of. However, not all educational institutions are in possession of tools and guidelines on how to introduce the concepts of programming and Artificial Intelligence to primary school students, especially the youngest ones. At the same time, the age of the digital contents users is becoming lower and lower, which leads to the rising necessity of educational intitutions to adjust their efforts to the needs of young learners. International cooperation fills the gap of missing educational resources, which primary schools are in high demand for. The collaboration of varied institutions, ranging from primary schools, universities, cybernetics and robotics societies, to renowned publishing companies, makes the exchange of best solutions possible and reliable. This guarantees wide access and high usability of the materials to educators of elementary school in particular, but also to all interested in the idea of introducing teaching programming from the possibly earliest stages.

The creation of school curricula, textbooks, lesson plans, innovative methods to teach programming and AI-related contepts with or without technology, is a huge step forwards and a long-term investment in future initiatives. It contributes to the educational systems' improvement, it ensures effective implementation of innovative educational solutions, and, first of all, it constitutes a crucial milestone towards the digital transformation of schools. All these would not be possible without joint ventures of dedicated international Erasmus+ project partners.

REFERENCES

[1] 'Why Technology Is Important In Our Daily Life?', Technology, Digital Edenz, October 18, 2019, *https://digitaledenz.com/why-technology-is-important-in-our-daily-life/*

[2] C. Burnett, 'The Digital Age And Its Implications For Learning And Teaching In The Primary School', Cambridge Primary Review Trust, July 2016, https://cprtrust.org.uk/wp-content/uploads/2016/07/Burnett-report-20160719.pdf

[3] C. Hague, S. Payton, 'Digital literacy across the curriculum', Curriculum & Leadership Journal, Volume 9, Issue 10, April 8, 2011, http://www.curriculum.edu.au/leader/default.asp?id=33211&issueID=12380

[4] '10 Reasons Why Schools Should Teach Coding', Skill Struck, February 11, 2022, https://www.skillstruck.com/blog/10-reasons-why-schools-should-teach-coding

[5] 'The Benefits of Coding in School and How to Teach It', Nord Anglia Education, August 18, 2020, https://www.nordangliaeducation.com/news/2020/08/18/the-benefits-of-coding-in-school-and-how-to-teach-it

[6] S. Güner, 'Introducing AI in primary school prepares kids for future', Daily Sabah, Istanbul, April 03, 2019, https://www.dailysabah.com/technology/2019/04/03/introducing-ai-in-primary-school-prepares-kids-for-future

[7] W. Yang, 'Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation', Computers in Education: Artificial Intelligence, Volume 3, 2022, https://www.sciencedirect.com/science/article/pii/S2666920X22000169

[8] Can You Teach Coding Without A Computer?', Code Advantage, May 12, 2020, https://www.codeadvantage.org/coding-for-kids-blog/can-you-teach-coding-without-a-computer

[9] S. Babić, D. Bjelanović and i. M. Čičin-Šain, "Programming and Mathematics through Game", 44th International Convention on Information, Communication and Electronic Technology (MIPRO), 2021, pp. 870-874

[10] S. Vaniukov, 'The Impact of Digital Transformation in the Education Sector', Softermii, August 03, 2022, https://www.softermii.com/blog/the-impact-of-digital-transformation-in-education-sector

[11] 'The Benefits of Coding in School and How to Teach It', Nord Anglia Education, August 18, 2020, https://www.nordangliaeducation.com/news/2020/08/18/the-benefits-of-coding-in-school-and-how-to-teach-it

The Role of Informatics Textbooks in Primary School in Croatia

A. Glavaš* and D. Marković**
* Faculty of Humanities and Social Sciences, Osijek, Croatia
** J. J. Strossmayer University of Osijek - Department of Mathematics, Osijek, Croatia
aglavas@ffos.hr, darija@mathos.hr

***Abstract* - Informatics teaching has its own peculiarities, which are reflected in the characteristics of informatics textbooks and their use in the classroom. Although many studies of textbooks from different aspects have been conducted in recent decades, very few of them deal with informatics textbooks. This paper presents a study of the use of informatics textbooks as curriculum resources in Croatian primary schools. The focus of the study is on how and why textbooks are used during the planning phase of the teaching, during the teaching of new content, and during the practice phase in classrooms. The data were collected through a large-scale survey of primary school informatics teachers. The results were obtained by descriptive and inferential statistics (chi-squared test) using the R software. This research will be of interest to the scientific community, creators of educational policy, i.e. textbook standards, as well as authors and publishers of informatics textbooks.**

Keywords - informatics textbooks; teachers' perception; informatics teaching

I. INTRODUCTION

The study of textbooks from various aspects has been the focus of research for several decades. Although research has shown that textbook-based teaching is dominated by the frontal form of work, which is characteristic of traditional teaching, textbooks are an integral and important part of modern teaching as well [1, 2]. This may be due to the textbook being considered to be the most important curricular resource for translating educational policy into pedagogy [3]. A textbook can be defined as a school book that provides an officially approved pedagogical version of human knowledge (Stray, 1994 as cited in [4]), i.e. a didactically designed source of knowledge for teaching that is aligned with the curriculum or syllabus of the subject and as such provides certainty of accuracy of content, methodological validity and appropriateness for students' developmental abilities [5]. Textbooks are mandatory educational material in Croatia for all subjects in primary and secondary schools (excluding art, music, technical, and physical education) and serve as a comprehensive source for achieving all learning outcomes specified in the curriculum [6]. Furthermore, the Regulation on Textbook Standards [7] stipulates that textbooks must be structured in accordance with modern pedagogical, psychological and didactic findings while respecting the subject's didactic requirements and encouraging the use of active learning methods and various teaching strategies that corresponds to the requirements of contemporary teaching (and the constructivist-curricular teaching approach). It is therefore not surprising that they are seen as teaching materials that provide a framework for thinking about what, who, how and when to teach [8], and that researches have highlighted the important role of textbooks in planning and delivering lessons both in other countries and in Croatia, and in various fields such as mathematics [9, 10], science [11], linguistics [12], primary education in general [13] and others. In [2] authors examined textbook studies that focused on the pedagogical use of textbooks in informatics classrooms and identified five main categories: Teaching methods; Learning methods; Motivational aspects; Academic performance; How well the textbook is used as a tool for learning or teaching. However, there is a lack of relevant literature (only 13 studies were found) dealing with the study of informatics textbooks, and most of these studies do not address the issue of their pedagogical (didactic) use [2]. One reason for this could be that informatics has only recently been introduced as a subject, as well as the existing challenges in deciding how to create textbooks (what content, what design and what important components) so that they do not become outdated too quickly due to the ongoing development and constant changes that occur in the field of informatics [2]. One of the peculiarities of informatics teaching is the possibility that students know certain parts of the material better than teachers due to rich extracurricular experiences with technology (Crawford, 1999, as cited in [14]), so informatics teachers frequently face the challenge of motivating students and maintaining their attention, which is also one of the challenges in creating informatics textbooks. The aim of this paper is therefore to examine the role and possibilities of textbook use for the subject of informatics in Croatian primary schools from the perspective of informatics teachers. The term textbook in this paper refers to the textbooks used in the classroom and includes whole textbook sets that teachers and/or students receive from the publisher, i.e. it can include a textbook, a workbook, a methodological manual for teachers and other educational content associated with the textbook, whether digital or physical (e.g. QR codes in the textbook that refer to additional digital educational content in the digital version of the textbook). Currently, three informatics textbooks from different publishers have been accepted for lower classes, while four informatics textbooks from different publishers have been approved for upper classes in primary school in Croatia.

II. REVIEW OF THE LITERATURE

The theoretical framework for developing the questionnaire in this research, taking into account the specifics of informatics teaching and the aim of the paper, is derived from [2], who identified five thematic categories of inquiry that emphasise the pedagogical use of textbooks in the classroom (how teachers use textbooks; how students use the textbook; how the textbook is used for motivation; how well students use the textbook; how well the textbook is used as a teaching tool for learning and teaching) and [9], which systematises research on the pedagogical use of mathematics textbooks into six areas (whether or not textbooks are used in the classroom, the authority of the textbook, who uses the textbook and how, the teacher as a facilitator of the textbook, and the influence of the pedagogical tradition on contemporary teaching).

A. Whether or not textbooks are used in the classroom

Numerous studies in various fields show that textbooks are the most commonly and longest used learning and teaching resource [15]. The use of textbooks by teachers can be categorized into three different levels: textbook direct, textbook indirect and textbook absence [16]. Direct use implies that students work individually or in groups with textbook text, solve textbook tasks, or that the teacher reads the content directly from the textbook. Indirect textbook use occurs when a teacher solves a given example from the textbook on the board, uses an introductory motivation modeled after the textbook, and applies concepts and definitions in the manner prescribed in the textbook, but without explicitly citing the textbook as a source. The absence of textbook use is characterised by situations in which the textbook is not included in the lesson, but the teachers (and students) use other teaching aids and materials, the lesson is introduced methodologically differently than in the textbook, or exercises are assigned that are not taken from or based on the textbook.

B. Textbook authority

There are numerous definitions of textbooks, but each assumes that the content they contain is authorised [17], i.e., that they represent an authoritative pedagogical version of a particular field of knowledge since they convey what should be learned and how it should be learned (Stray, 1994 as cited in [4]). The textbook's authority comes from its definition and description as a book that is officially approved, pedagogically designed, and curriculum aligned in order to teach subject content to pupils (and teachers) [18]. As a result, textbooks closely follow and reflect curriculum requirements. Some authors therefore refer to them as potentially implemented curriculum, which stands between the planned and the implemented curriculum, which is dynamically used in classrooms and thus influences teaching [19].

C. Who uses textbooks

The majority of research on the role of textbooks emphasis the importance of textbooks for teachers in planning and teaching [8, 9]. According to [13], most teachers stick to (and do not reduce) the topics covered in textbooks. Furthermore, [9] notes that textbooks are mostly written for teachers rather than students [11], although they are presented as books for students [4] and their main purpose should be to help students understand the subject matter [20], help them achieve the outcomes and enable them to work independently (at school and at home) [21]. Furthermore, students usually have to purchase them.

D. How and when textbooks are used - teachers as mediators of textbook texts

According to [19] (as cited in [22]), the structure of the textbook is a specific pedagogical model with a plan for a specific set of educational opportunities, which means that the structure (and content) of the textbook influences teaching practices, i.e. pedagogical choices, which means that teachers act as mediators of the textbook text (Luke et al., 1989, as cited in [4]). Due to the specificities of the subjects taught, textbook structures vary. In mathematics textbooks, for example, a structure can be identified in five blocks: introductory tasks, exposition, kernels (basic concepts, definitions, properties, theorems, etc.), solved examples and exercises [10].

III. RESEARCH METHODOLOGY

A. Objectives and research questions

The aim of this paper is to investigate how and when (in what phase of the teaching process) the textbook plays (or does not play) an important role. The following research questions were posed for this purpose:

RQ1: To what extent do teachers use the textbook?
RQ1.1. To what extent do teachers use the textbook to prepare lessons?
RQ1.2. To what extent do teachers use the textbook to introduce new content?
RQ1.3. To what extent do teachers use the textbook for practice activities and summative assessment?
RQ1.4. To what extent do teachers choose teaching approaches based on the textbook?
RQ2. Does the role of the textbook in teaching differ depending on the type of study programme that informatics teachers have completed?

B. Sample and procedure

A total of N=396 people who work as informatics teachers at primary schools participated in the survey. Table 1 provides more specific information about the respondents' socio-demographic characteristics (gender, seniority, type of relevant education, in which grade they teach). A convinient sampling method was used in the study [23], i.e., participants are individuals who voluntarily responded to the invitation to complete the online questionnaire that the authors emailed to all primary schools in the Republic of Croatia. The data were collected during October and November 2022 using Google Forms. The responses from the participants were obtained anonymously.

Because only 5 (1.26%) of the teachers in the study are the authors of one of the textbooks, this variable was neglected in data processing.

TABLE I. THE STRUCTURE OF RESPONDENTS

Participant information	N (%)
Gender	
Female	307 (77,53%)
Male	89 (22,47%)
Teaching experience	
0 – 5 years	173 (43,69%)
6 – 15 years	124 (31,31%)
16 – 29 years	83 (20,96%)
30 or more years	16 (4,04%)
Teaching at	
1.– 4. class (only)	106 (26,77%)
5. – 8. class (only)	97 (24,49%)
1. – 8. class	193 (48,74%)
Teachers education[A]	
M. Ed. in informatics (E1)	123 (31,06%)
M. Sc. in informatics (E2)	70 (17,68%)
M. Primary Ed. with enhanced informatics (E3)	163 (41,16%)
B. Ed. or B. Sc. (E4)	15 (3,79%)
Specialist graduate (E5)	12 (3,03%)
Other (E6)	13 (3,28%)

C. Instrument and data analysis

For the purposes of this study, the authors developed an instrument based on an intuitive-theoretical approach (Fraser, 1986, as cited in [24]). Based on a review of the relevant literature, the authors highlighted key variables for which they independently developed items: A=lesson preparation and planning (6), B=selection of teaching content (7), C=practice and assessment (5), D=selection of teaching methods (4). Variable A indicates the extent to which informatics teachers anticipate the order of lesson topics based on the order in the textbook, follow the structure of individual lessons based on the textbook, use textbooks, didactic manuals and additional content from the publisher when preparing lessons. Variable B investigates the extent to which teachers use definitions and technical terms (e.g., spyware vs. its Croatian translation) as they appear in the textbook, the extent to which they draw attention to parts of the textbook that are highlighted or illustrated and to the elaboration of educational outcomes, and, more broadly, the extent to which they incorporate the content provided in the textbook and relate it to specific outcomes. Variable C refers to whether or not teachers assign exercises and homework from textbooks or workbooks (if they do) and whether or not they rely on these, or even on pre-made exam templates, for written summative assessment of knowledge. Variable D investigates the extent to which teachers use the textbook's didactic procedures. This refers to the order of content, definitions, solved examples, motivating content, etc.

Each item had to be evaluated on a five-point Likert scale, with 1 = never, 3 = sometimes and 5 = always. The results are expressed as the relative frequencies of the response. To check the reliability of the instrument, Cronbach's alpha was calculated, and it was found to be α=0.872 for the whole questionnaire (27 items). Very high reliability was obtained for each variable (α=0.835; α=0.847; α=0.666; α=0.699). The given analysis included descriptive analysis, chi-squared test and Fisher's exact test. The analysis was carried out using the R software.

IV. RESULTS AND DISCUSSION

In this part, we present the results on the use of informatics textbooks. Table 2 shows the relative frequencies of the respondents' answers according to the different variables.

TABLE II. THE RELATIVE FREQUENCIES

N (%)	A	B	C	D
Never	11,8%	2,5%	17,7%	7,7%
Rare	18,7%	8,2%	12,4%	22,2%
Sometimes	26,3%	24,9%	28,6%	41,9%
Often	31,5%	41,1%	29,8%	24,4%
Always	11,7%	23,4%	11,6%	3,8%

We must acknowledge that the answers of respondents who stated that they do not conduct written knowledge tests or assign homework were excluded while analyzing variable C (150 responses, or 37.88% of total number). Table 2 shows that textbooks are used by informatics teachers mainly (always+often) in selecting teaching content (64.5%), as well as in preparing lessons (43.2%) and in practicing and assessing knowledge (41.4%), while the textbook plays the least role in selecting teaching methods (28.2%). These results, particularly related to the variable B, indicate great textbook authority, which is consistent with previous research (e.g., [18]).

It is also interesting to note that teachers use textbooks considerably more for planning and preparing lessons and selecting teaching content than for practicing with students, giving homework, assignments, etc. (variable C). However, these findings are not surprising given that some authors suggest that they are rather teacher books than student books [9, 11], as noted previously.

Approximately 30% of all participants stated they never or rarely use a textbook for planning, practice and assessment activities and teaching methods, whereas only about 10% indicate they never or rarely use a textbook for selecting teaching content (B). Based on these results, it is not possible to reliably determine to which category informatics teachers in Croatia belong according to Johannson's classification of textbook use [16]. Although previous research (e.g. [19]) has shown that the structure of the textbook influences teaching practices in

[A] According to [NN] Narodne novine. Pravilnik o odgovarajućoj vrsti obrazovanja učitelja i stručnih suradnika u osnovnoj školi, NN, 6/2019 [18. siječnja 2019.] Accessed at: https://narodne-novine.nn.hr/clanci/sluzbeni/2019_01_6_137.html

terms of method selection, this study found that the textbook had the least influence on method selection (only 28.2%).

The results of this study are partly consistent with research on textbook use in other areas of education. Textbooks, for example, are most commonly used in mathematics classes for practice and the introduction of new topics [9], in science classrooms for learning new concepts and reading activities [11, 25], and in biology classes for review lessons and independent student work [26]. This research also confirmed that teachers rarely deviate from textbook content or "textbook language" [8, 13] and that textbooks are not usually used to provide solved examples [9]. Furthermore, [25] in their study examined the differences between teachers' and students' perceptions of textbook use (in science classes) and found that their perceptions differed significantly. In fact, students believed to a greater extent that teachers relied on textbooks to deliver lessons than the teachers themselves thought. Thus, [25] wonder if it is possible that teachers report what they do differently than what they actually do. The above reflections in [25] actually warn of the methodological limitations of research whose findings were drawn exclusively from teachers' own perceptions. To obtain a more comprehensive picture and to confirm the results, the authors recommend conducting qualitative research, such as classroom observation, focus groups, and so on.

Although informatics is not a new subject in Croatian schools, it was recently introduced as an elective subject in lower classes in primary school and became a mandatory subject for 5th and 6th grade students with the implementation of the curricular reform in 2019. Consequently, demand for informatics teachers in primary schools was increased, and accordingly the Rulebook on the appropriate type of education for teachers and professional associates in primary schools [7] was adopted. This Rulebook regulates the appropriate type of education of informatics teachers in primary schools (these are persons belonging to the E1-E5 category). It allows the teaching of informatics to persons who have completed one of the 48 prescribed study programmes, from a wide range. It is precisely because of these (potential) differences in teachers' competences (subject knowledge, pedagogical-didactical knowledge, professional skills) that we find important to investigate whether the role of the textbook depends on the type of study programme that informatics teachers have completed.

The chi-squared test was performed to determine whether there is a significant difference in the use of textbooks (in a single variable) based on the respondents' corresponding type of education. Table 3 shows the relative frequencies of respondents' answers in relation to the variable lesson preparation and planning (A), indicating that there is a difference in answers in relation to teachers' education. The chi-squared test showed that in general there is a significant difference (χ-squared=45.318, df=20, $p<0.001$). To determine between which variables there was a significant difference, a post-hoc chi-squared test was conducted. This revealed a

TABLE III. VARIABLE A: LESSON PREPARATION AND PLANNING

	1	2	3	4	5
E1	0,14	0,20	0,27	0,27	0,12
E2	0,10	0,20	0,27	0,31	0,12
E3	0,10	0,18	0,27	0,35	0,10
E4	0,22	0,12	0,17	0,31	0,18
E5	0,10	0,17	0,29	0,24	0,21
E6	0,05	0,13	0,22	0,44	0,17

significant difference in responses between teachers who had completed M. Ed. in informatics (E1) and M. Primary Ed. with enhanced informatics (E3) (χ-squared=13.283, df=4, $p<0.001$), with teachers who had completed a Bachelor degree (E4) (χ-squared=12.597, df=4, p=0.01342) and with teachers in the other category (E6) (χ-squared=14.845, df=4, p=0.005034). From the frequency table (Table 3), it can be seen that teachers who have completed M. Ed. in informatics are less likely to report that they often or always use textbooks in planning and preparing lessons compared to the others. Similarly, there is a significant difference between teachers who have completed M. Sc. in informatics (E2) and teachers who have completed a Bachelor degree (E4) (χ-squared=15.327, df=4, p=0.004068). The latter are less likely to report that they sometimes or often use textbooks when planning lessons. On the same question, there is a significant difference between M. Primary Ed. with enhanced informatics (E3) and those who have a Bachelor degree (E4) (χ-squared=20.063, df=4, $p<0.001$) or a specialist degree (E5) (χ-squared=10.191, df=4, p=0.03733), and between teachers with a Bachelor degree and those from the Other category (E6) (χ-squared=10.929, df=4, p=0.02737). The direction in which the differences go can be seen in Table 3.

Table 4 shows the relative frequencies of the respondents' answers in relation to the selection of the teaching content variable (B). When the chi-squared test is conducted, a statistically significant difference is observed in the responses of the respondents with respect to the type of education and the content selection variable (χ-squared= 96.202, df=20, $p<0.001$). Taking into account the characteristics of the chi-square test and since categories 5 and 6 do not have enough measured values, Fisher's exact test was used for further analysis (the same will be applied in all such future cases).

Again, there is a significant difference between teachers with M. Ed. in informatics (E1) and teachers with M. Primary Ed. with enhanced informatics (E3) (χ-squared=38.11, df=4, $p<0.001$). The former cumulatively state more frequently that they never or rarely use a textbook when selecting teaching content (definitions, technical terms, given content, outcomes, illustrations) and to a lesser extent that they always or often use a textbook when selecting teaching content. Simultaneously, a significant difference can be observed in the answers provided by those with a completed Bachelor degree (E4) (p=0.01007), but in the opposite direction, i.e. the latter adhere to textbooks in the selection of content to a lesser extent than teachers with

M. Primary Ed. with enhanced informatics. In addition, informatics teachers with M. Sc. in informatics (E2) indicate to a significant extent (χ-squared=16.877, df=4, p-value=0.002042) that they always or often use

TABLE IV. VARIABLE B: SELECTION OF TEACHING CONTENT

	1	2	3	4	5
E1	0.03	0.10	0.26	0.39	0.22
E2	0.02	0.09	0.27	0.38	0.23
E3	0.01	0.05	0.23	0.45	0.25
E4	0.11	0.13	0.25	0.33	0.17
E5	0.04	0.17	0.24	0.31	0.25
E6	0.00	0.04	0.26	0.45	0.24

textbooks when selecting teaching content, in contrast to M. Primary Ed. with enhanced informatics (E3), while a significant difference (p<0.001) can also be observed among teachers with completed Bachelor degree (E4), but in the opposite direction. Indeed, the latter state more frequently than M. Sc. in informatics that they do not use textbooks when selecting teaching content. Considering that teachers with M. Primary Ed. with enhanced informatics differ significantly in this respect from teachers in the category E5 (p<0.001), it can be said that teachers with M. Primary Ed. with enhanced informatics to the greatest extent use textbooks when selecting teaching content (always or often 70%). The responses of informatics teachers with M. Primary Ed. with enhanced informatics differ significantly from those of teachers who have completed a Bachelor degree (p<0.001) and Specialist graduate (p=0.01466), who deviate less often from the teaching content given in the textbook (never or rarely only 4%).

Table 5 shows the relative frequencies of respondents'

TABLE V. VARIABLE C: PRACTICE AND ASSESSMENT

	1	2	3	4	5
E1	0,19	0,15	0,27	0,28	0,11
E2	0,14	0,14	0,34	0,28	0,11
E3	0,16	0,09	0,28	0,34	0,13
E4	0,32	0,24	0,16	0,16	0,12
E5	0,28	0,10	0,40	0,18	0,05
E6	0,12	0,12	0,24	0,34	0,18

answers in relation to the type of education and the variable (C) practice and summative assessment. Different behavior can be observed, as confirmed by the chi-squared test (χ-squared=45.318, df=20, p<0.001). By performing a post hoc test, one can also see where this difference is.

Similar to previous analyses, a difference was found between teachers with a Bachelor degree (E4) and teachers with M. Ed. in informatics (χ-squared = 10.526, df=4, p=0.03245), teachers with M. Sc. in informatics (χ-squared=16.587, df=4, p=0.002325), teachers with M. Primary Ed. with enhanced informatics (χ-squared=23,713, df=4, p<0.001) and teachers from the other category (E6) (χ-squared=11.185, df=4, p=0.02456).

In all cases, the difference is in the direction that more than half of the respondents (56%) never or rarely use textbooks when practicing, giving homework, and preparing written summative assignments, i.e., they do not give assignments modelled on textbooks or ready-made test materials from publishers' textbooks. Teachers with M. Primary Ed. with enhanced informatics are significantly more likely to report that they always or often (47%) use a textbook when practicing and preparing written summative assessments (χ-squared=11.059, df=4, p=0.02591) than teachers with M. Ed. in informatics (39%) and a Specialist graduate teachers (p=0.04622; 23%). Furthermore, teachers from the other category (E6) use the textbook to a significantly higher extent (p=0.03988; 52%) during the exercises and written summative assessments, in contrast to Specialist graduate teachers.

Regarding the role of textbooks in the selection of teaching methods, the responses of informatics teachers differed significantly (χ-squared=47.028, df=20, p<0.001). The relative frequencies of respondents' answers are shown in Table 6. Post hoc analysis showed that the responses of teachers with M. Ed. in informatics

TABLE VI. VARIABLE D: SELECTION OF TEACHING METHODS

	1	2	3	4	5
E1	0,10	0,22	0,45	0,20	0,03
E2	0,08	0,26	0,36	0,23	0,07
E3	0,06	0,21	0,43	0,28	0,03
E4	0,12	0,25	0,40	0,17	0,07
E5	0,04	0,25	0,46	0,21	0,04
E6	0,02	0,21	0,31	0,38	0,08

were significantly different from those who had completed M. Sc. in informatics (χ-squared=12.26, df=4, p=0.01552), M. Primary Ed. with enhanced informatics (χ-squared= 15.396, df=4, p=0.003947) or the other category (E6) (p=0.002622). The first mentioned answered significantly more often that they never or rarely use a textbook when selecting methods (methodological approach, motivating examples, lesson structure, etc.) and teaching method (32%), or to a lesser extent they answered they always or often (23%) use textbook recommendations, sequence, motivating content to introduce the topic, etc., in contrast to those who have completed M. Sc. in informatics (34%; 30%), M. Primary Ed. with enhanced informatics (27%; 31%), or the other category (E6) (23%; 46%). On this point, teachers with M. Primary Ed. with enhanced informatics still differ significantly from those with M. Sc. in informatics (χ-squared=16.977, df=4, p-value=0.001953) and a Bachelor degree (p=0.04226). Table 6 shows that the former (E3) relied on the textbook to a greater extent (31% in responses always and often) than E4 (24%) and than E2 (27% in responses never and rarely). The difference in responses is also significant between those with a Bachelor degree and those from the other category (E6) (p=0.04814).

The findings of this part of the research are partly consistent with the prior discussion. For example, in this study, informatics teachers in the E2-E6 category use textbooks more frequently for lesson preparation and when selecting teaching methods than teachers with an

M. Ed. in informatics. Yet, when it comes to using textbooks for teaching content selection or tasks for practicing and homework, teachers with M. Ed. in informatics follow the textbooks more strictly than B. Ed. or B. Sc. teachers.

V. CONCLUSION

Considering the lack of literature related to informatics textbooks from a pedagogical point of view, the aim of this research was to investigate the role of informatics textbooks regarding different phases of teaching as perceived by informatics teachers in Croatian primary schools. Furthermore, the authors intended to determine is there siginificant difference in textbook use depending on the type of formal education of the teacher. According to the findings, informatics teachers use textbooks the most frequently when selecting teaching content and planning lessons, and the least frequently while assigning homework, selecting practice problems, or conducting summative assessments. The findings also indicate that textbook use varies greatly based on the level and type of formal education teachers completed. Due to space constraints, only the most important findings are reported in this paper. More in-depth study is required due to the previously noted methodological limitations of this type of study. Qualitative research methods, such as classroom observation and focus groups, should be used. Other factors should be taken into account, which was not possible to achieve in this paper.

REFERENCES

[1] Gottfried, S. & Kyle, W. (1992) Textbook use and the biology education desired state. Journal of Research in Science Teaching, 29(1), 35–49. https://doi.org/10.1002/tea.3660290105

[2] Hajdin, G. & Divjak, B. (2016) Use of Informatics Textbooks in School Classroom. *Journal of Information and Organizational Sciences*, 40(1), 21–44. https://doi.org/10.31341/jios.40.1.2

[3] Pepin, B., Gueudet, G., & Trouche, L. (2013). Investigating textbooks as crucial interfaces between culture, policy and teacher curricular practice: two contrasted case studies in France and Norway. ZDM, 45(5), 685–698.

[4] Glasnović Gracin, D. & Jukić Matić, Lj. (2016) The Role of Mathematics Textbooks in Lower Secondary Education in Croatia: An Empirical Study. The Mathematics Educator, 16(2), 31-58.

[5] Poljak, V. (1980) Didaktičko oblikovanje udžbenika i priručnika. Zagreb: Školska knjiga.

[6] [NN] Narodne novine. Odluku o donošenju kurikuluma za nastavni predmet informatike za osnovne škole i gimnazije u Republici Hrvatskoj, NN 22/2018 [6. ožujak 2018.] Accessed at: https://narodne-novine.nn.hr/clanci/sluzbeni/2018_03_22_436.html

[7] [NN] Narodne novine. Pravilnik o udžbeničkom standardu te članovima stručnih povjerenstava za procjenu udžbenika i drugih obrazovnih materijala, NN 116/2018, 9/2019 [25. siječnja 2019.] Accessed at: https://narodne-novine.nn.hr/clanci/sluzbeni/2019_01_9_196.html

[8] Nicol, CC, & Crespo, SM (2006) Learning to teach with mathematics textbooks: How pre-service teachers interpret and use curriculum materials. Educational Studies in Mathematics, 62(3), 331–355.

[9] Pepin, B. & Haggarty, L. (2001) Mathematics textbooks and their use in English, French and German classrooms: a way to understand teaching and learning cultures. ZDM – Zentralblatt für Didaktik der Mathematik – The International Journal on Mathematics Education, 33(5), 158-175.

[10] Rezat, S. (2006) The Structures of German Mathematics Textbooks. ZDM, 38(6), 482-487.

[11] Driscoll, M., Moallem, M., Dick, W. & Kirby, E. (1994) How does the textbook contribute to learning in a middle school science class?. Contemporary Educational Psychology, 19, 79–100.

[12] Andraka, M. & Jurković, A. (2015) Potiču li udžbenici za rano učenje stranog jezika razvoj strategija učenja vokabulara? Strani jezici: časopis za primijenjenu lingvistiku, 44(1), 29-53.

[13] Shymansky, JA, Yore, LD & Good, R. (1991) Elementary school teachers' beliefs about and perceptions of elementary school science, science reading, science textbooks, and supportive instructional factors. Journal of Research in Science Teaching, 28(5), 437–454.

[14] Hajdin, G. (2014) Pripremljenost srednjoškolskih učenika za nastavu informatike korištenjem udžbenika. Doktorska disertacija, Sveučilište u Zagrebu. Varaždin: Fakultet organizacije i informatike.Accessed at: https://urn.nsk.hr/urn:nbn:hr:211:609579

[15] Numangolu, G. & Bayir, S. (2009) Evaluation of information and communication technology textbooks according to principles of visual design. Procedia - Social and Behavioral Sciences, 1(1), 2140-2144.

[16] Johansson, M. (2006). Teaching mathematics with textbooks. A classroom and curricular perspective. Doctoral dissertation, Luleå University of Technology. Lulea: Lulea University of Technology.

[17] De Castell, S., Luke, A., & Luke, C. (1989) Language, authority and criticism. Readings of the school textbook. London: Falmer Press.

[18] Glasnović Gracin, D. (2013) Mathematics Textbook as an Object of Research. Croatian Journal of Education, 16(Sp.Ed.3), 211-237.

[19] Valverde, GA, Bianchi, LJ, Wolfe, RG, Schmidt, WH & Houang, RT (2002) According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks. Dordrecht: Kluwer.

[20] Matijević, M., Rajić, V. & Topolovčan, T. (2013) Učenička percepcija srednjoškolskih udžbenika. Život i škola, 59(29), 64-78.

[21] Hutchinson, T. & Torres, E. (1994) The textbook as agent of change. ELT J., 48(4), 315-328.

[22] Jukić Matić, Lj. & Glasnović Gracin, D. (2016) The use of the textbook as an artifact in the classroom A case study in the light of a socio-didactical tetrahedron. Math Didakt, 37, 349–374.

[23] Cohen, Luis; Manion, Lawrence; Morrison, Keith (2018), Research Methods in Education (8th edition), New York: Routledge.

[24] Waldrip, BG, Fisher, DL, Dorman, JP (2008) Students' Perceptions of Assessment Process: Questionnaire Development and Validation. In: D. Fisher, R. Koul & S. Wandpen (Eds.) Proceedings of the Fifth International Conference on Science, Mathematics and Technology Education Udon Thani, Thailand. Curtin University of Technology: Perth, Australia.

[25] Tulip, D. & Cook, A. (1993) Teacher and student usage of science textbooks. Research in Science Education, 23(1), 302–307.

[26] DiGisi, L. L., & Willett, J. B. (1995). What high school biology teachers say about their textbook use: A descriptive study. Journal of Research in Science Teaching, 32(2), 123–142. https://doi.org/10.1002/tea.3660320204

The Use of Smartphones and Tablets for Teaching Biology at a Distance: A Comparison between Rural and Urban Schools

V. Lang*, and A. Šorgo*, **
* Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor, Slovenia
** Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia
vida.lang1@um.si

Abstract – Smartphones and tablets (SPT) have become an integral part of daily life, and their added value for biology education was seriously tested during the COVID -19 pandemic in distance education. We collected data from 355 elementary-aged children attending 6th to 9th grade; 30% of the participants were from rural schools and 70% from urban schools. Results show that most children, 75.7%, used SPT very frequently to search for information on the World Wide Web and 65.9% used SPT frequently to participate in distance learning (videoconferencing). Very rarely, if ever, was SPT used to conduct experiments, virtual field trips, or hands-on work in virtual laboratories. PCA analysis revealed 2 components. The first component explained the non-use of SPT for hands-on work (33.6%, Cronbach's alpha = 0.85) and the second component explained the use of SPT for communication and collaboration (28.8%, Cronbach's alpha = 0.77). We were also interested in whether the environment of the school (rural or urban) influenced the use of SPT. There is no statistically significant difference in access to wearable smart technology between urban and rural students.

Keywords – biology; distance learning; lower secondary; rural school; school students; smartphones and tablets; urban school.

I. INTRODUCTION

The outbreak and spread of the 2019 coronavirus pandemic (COVID -19) have had a profound impact on education systems around the world. Schools have been partially or completely closed in more than 190 countries, affecting 1.6 billion school-aged children [1]. Many education systems have begun using distance learning.

The power of mobile phones and tablets was felt more than ever during the COVID -19 pandemic, when the entire world came to a standstill and billions of people relied on their mobile devices to explore the world, stay connected with their families, friends, teachers, students, services, and for current events. Time spent on mobile devices has increased in 2020 [2]. Many educational institutions have offered distance education to their students. This has been facilitated by mobile technologies such as smartphones, tablets, laptops, e-readers, and other practical devices. It is the development of technologies that has reduced the barriers to distance education. The use of mobile technologies has enabled teachers and students to interact and create a classroom anytime, anywhere [3]. Mobile technologies have improved the connectivity, interactivity, and accessibility of educational materials on the Internet [3]. Smartphones and tablets with built-in cameras have become an integral part of the daily lives of students and teachers in distance education. According to a study [2], educational app downloads doubled during the pandemic COVID -19. Another advantage of smartphones is that they provide access to a wide variety of educational apps and resources. The importance and impact of smartphones on the methods, forms, and outcomes of education can be recognized in at least three aspects: Mobility of technology, mobility of learning, and mobility of students [4]. Smartphones can play multiple roles in biology education. Access to educational resources: smartphones allow students to access a wide range of biology resources, such as e-books and online lectures. Smartphones have high-quality cameras and screens that can be used to capture and display images and videos of biological concepts such as coagulation [5] or colorimetry [6] or visual aids for viewing microscopic organisms [7]. Smartphones can be used to collect and analyze data, such as environmental data or biological samples, for laboratory experiments or field studies. Overall, smartphones can be a valuable tool in the biology classroom, enhancing students' learning experiences and expanding their access to information and resources.

Smartphones have become an integral part of our lives, and their presence in the classroom is no exception. According to recent Slovenian studies [8,9], the majority of students in lower secondary, upper secondary, and tertiary education own smartphones, and the use of tablets has also been reported [5,6,10,11,12], so the availability of the technology cannot be considered a limiting factor for its adoption in daily classroom practice. This article examines how smartphones are used in urban and rural schools in COVID -19 distance education.

The aim of the study was to:

- find out the extent to which students use SPT for distance learning during school closures.
- to find out if there is a statistically significant difference in the use of SPT between urban and rural students.

II. MATERIALS AND METHODS

A. Sample and sampling

The population of interest was elementary school students from 6th to 9th grade of Slovenian 9-year-old compulsory school. The survey instrument, in the form of an online questionnaire based on the 1ka platform (www.1ka.si), was made available to students through various channels, online social media, contacts with schools and individual teachers. The sample began in May 2022 and ended in July 2022. We collected data from 355 school-aged children (44.6% boys and 55.4% girls) attending 6th (5.7%), 7th (6.5%), 8th (54.9%), and 9th (33.0%) grades of elementary school. 30% of participants were from rural Slovenian schools from different regions and 70% of participants were from urban Slovenian schools from different regions. They reported that 96% of them owned a smartphone.

B. Instrument

A structured questionnaire was used as the primary instrument for data collection. The questionnaire is part of a larger questionnaire on the added value of smartphones in biology education [13]. In this paper, we present part of the questionnaire in which we asked students about the use of SPT during the COVID -19 pandemic and the resulting distance learning. For each item, participants indicated their agreement on a 6-point Likert scale ranging from never (1) to very often (6).

In the second question, we asked students whether they owned smart mobile devices and whether they shared them with their siblings or family members.

In the third part of the questionnaire, we were interested in demographic data: Gender, class, and the neighborhood where the school they attend is located (rural or urban).

C. Statistical analyses

Statistical analyses were performed using the open-source statistical programme Jamovi, 2.3.21. Each research variable was analysed for mean, median (Me), Mode (Mo), and standard deviation (SD).

Cronbach's alpha was calculated to assess the reliability of the constructs. Principal component analysis (PCA) with direct oblimin rotation was performed to assess the unidimensionality of the constructs. The reliability of the components resulting from the PCA analysis was assessed using the Cronbach's alpha coefficient. Values greater than 0.7 indicate reliability of the components.

The nonparametric Mann-Whitney U test was used to analyse differences between rural and urban schools. The Mann-Whitney U test, also known as the Wilcoxon rank sum test, tests for differences between two groups on a single ordinal variable with no specific distribution [14,15]. Results with a significance coefficient of less than 0.05 ($p< 0.05$) were flagged as statistically significant differences. When the effect size was less than 0.02 ($r<0.02$), statistically significant differences were also considered nonsignificant.

III. RESULTS

A. Use of SPT for distance learning biology classes during school closures

TABLE I. Measures of central tendencies on the question of the use of a SPT for distance learning biology classes during school closures (N = 355, Cronbach's alpha = 0.85). Results are presented as percentages and sorted by decreasing Mode (Mo).

Use of SPT for distance learning during school closures	***Mean***	***Me***	***Mo***	***SD***
Participation in distance learning (video conferencing).	4.80	5	6	1.50
To find information on the Internet.	5.11	6	6	1.23
For the assessment and verification of knowledge.	4.61	5	6	1.70
Communicating with teachers via email.	4.46	5	5	1.42
To view and analyze video experiments.	4.01	4	4	1.55
To take a virtual video tour.	2.77	2	1	1.76
For working in the virtual lab.	2.55	2	1	1.69
For photo search (collect photos, identify, and add information).	3.63	4	1	1.83
To perform experiments using the SPT built-in measurement devices.	2.97	3	1	1.76

Note: Likert scale ranging from never (1) to very often (6), Me- median, Mo – Mode, SD – standard deviation.

In the statistical analysis, the mode shows us the value that was chosen most often. Table I shows that students used SPT very often (Mode > 5) for participation in distance learning (video conferencing), to find information on the Internet, for the assessment and verification of knowledge, for communicating with teachers via email. The most common response of " never" (Mode = 1) was to use SPT for a virtual tour, to work in a virtual lab, to search for photos, or to conduct experiments.

The PCA analysis, the results of which are shown in Table II, revealed two components that explained 62.4% of the variance.

The first component (PCA1) explained 33.6% of the variance and consisted of statements about the use of smartphones for practical teaching purposes (participation in distance learning, video conferencing. To find information on the Internet. For assessing and reviewing knowledge. To communicate with teachers via email. To view and analyze video experiments.). The PCA1 statements describing when smartphones were not used during distance learning (Mode = 1) were distributed here. The reliability of the first component is indicated by the high value of Cronbach's alpha of 0.85.

The second component (PCA2) had 28.8% of variance and consisted of the statements that smartphones were used for information search, learning, and communication with teachers. The value of Cronbach's alpha was 0.77, indicating high reliability of the component.

An analysis of differences between rural and urban areas was conducted. It was found that there are no statistically significant differences ($p > 0.05$ and $r< 0.02$) in the use of SPT for distance education. There is also no

statistically significant difference between genders (results not reported).

TABLE II. Measure of Principal Component loading on the question about using SPT for distance learning during school closures and comparison between groups (N = 355, Cronbach's alpha = 0.85).

Use of SPT for distance learning during school closures	**PCA1**	**PCA2**	***U***	***p***	***r***
To take a virtual video tour.	0.89		12967	0.789	0.018
For working in the virtual lab.	0.88		12732	0.583	0.035
To perform experiments using the SPT built-in measurement devices.	0.81		12025	0.176	0.089
For photo search (collect photos, identify, and add information).	0.68		12871	0.708	0.0258
To view and analyze video experiments.	0.45	0.49	12522	0.437	0.051
Participation in distance learning (video conferencing).		0.84	12292	0.275	0.069
To find information on the Internet.		0.81	11624	0.052	0.120
Communicating with teachers via email.		0.79	12798	0.642	0.030
For the assessment and verification of knowledge.		0.44	5282	0.506	0.052
Explained variance (%)	33.6	28.8			
Eigenvalue	4.18	1.44			
Cronbach's alpha	0.85	0.77			

B. Do students own their smart mobile devices?

The question about owning a smartphone or tablet was answered by 354 students. The data obtained are summarized in Table III. From the results in the table III, 96% own a smartphone and 79% own a computer. Tablet computers are less common among 6th-9th grade students. (42.1%). The results also show that 35.9% of the participating students share a computer with their family. Only 7.6% of the participants share a smartphone with their family.

TABLE III. Frequency (%) analysis of the ownership of smart mobile devices and ICT (N = 354) and difference between rural and urban groups.

Item	**YES [%]**	**NO [%]**	***U***	***p***	***r***
I own a smartphone.	96.0	4.0	12824	0.283	0.024
I own a computer.	79.1	20.9	12939	0.742	0.016
I own a tablet.	42.1	57.9	12722	0.576	0.032
I must share my computer with my family.	35.9	64.1	12264	0.230	0.067
I must share my tablet with my family.	28.8	71.2	13603	0.907	0.006
I must share my smartphone with my family.	7.6	92.4	12805	0.404	0.026

An analysis of the differences between rural and urban areas was conducted. It was found that there were no statistically significant differences at the $p < 0.05$ and $r < 0.02$ levels in the ownership of smart mobile devices. There is also no statistically significant difference between genders.

IV. CONLUSION

Using smartphones and tablets for distance learning in biology can be an effective way to enhance students' learning experiences, especially in areas where traditional classrooms are not available. A comparison between rural and urban schools can provide insight into the potential benefits and challenges of using these technologies in different contexts.

Based on the results presented in the previous chapter and the research questions posed, we can draw a conclusion about the use of smart mobile phones and tablets in distance education (COVID -19).

Our first research objective was to find out how SPT is used in distance education. The results show that SPT was mainly used during distance education for participating in distance education (video conferencing), searching for information, checking and evaluating knowledge with online quizzes, communicating, and watching video experiments. However, smart mobile devices were never or very rarely used for hands-on work, such as taking a virtual video tour, working in a virtual lab, conducting experiments, and using already built-in sensors. We believe that the use of SPT for schoolwork has increased, but only for communication and solving current problems. SPT offers much more. In our work with students [5,10,11], we have found that smartphones can be used in a variety of ways, but they cannot replace professionally developed equipment (e.g., a microscope). However, they can be a useful, practical, and engaging didactic tool that we always have in our pockets or at hand. If we want the use of smartphones to be introduced even more into the school system, it is certainly first necessary to inform teachers and school management about the advantages of smart mobile devices, to prepare workshops on the smart use of smartphones or pedagogical publications with examples of use.

The second research objective was to investigate whether the use of smart mobile phones differs between students from urban and rural areas. The results show that 96% of students own a smartphone and 79% own a computer. Tablet computers are less common among 6th-9th grade students. (42.1%). They very rarely share their smartphone with their siblings or family members. These data are not surprising, since according to the reports of the Statistical Office of the Republic of Slovenia [16] for the year 2020, 97% of the inhabitants of Slovenia use mobile phones, 81% of them use smart mobile phones. From the research of Macuh [17] we can conclude that primary school students are also big users of smart mobile devices and tablets. Therefore, we assume that it should not be a problem if the BYOD (bring your own device) work method were introduced in classes, i.e., using and bringing one's smartphone or tablet to school for the purpose of schoolwork.

In summary, smartphones have become an integral part of the educational experience in both urban and rural schools. There is no statistically significant difference in access to wearable smart technologies between urban and rural students. Overall, the use of smartphones and tablets for distance education in biology can be a valuable tool to

enhance the student learning experience, especially in areas where traditional classrooms are limited. However, it is important to carefully consider the potential benefits and challenges associated with these technologies in different contexts and provide adequate support and resources to ensure their effective use.

REFERENCES

[1] UNESCO. (2020). UNESCO COVID - 19 education response: How many students are at risk of not returning to school?. https://unesdoc.unesco.org/ark:/48223/pf0000373992 Accessed Feb. 5, 2022.

[2] W. Robert (2020). Coronavirus boots mobile usage in China, Italy. Published Online by Mobile Marketer. Retrieved from https://www.marketingdive.com/news/coronavirus-boosts-mobile-usage-in-china-italy/574353/ Accessed Feb. 5, 2022.

[3] E. M. Onyema, A. U. Eucharia, F. S. Gbenga, A. O. Roselyn, O. Daniel, & N. U. Kingsley (2020). Pedagogical use of mobile technologies during Coronavirus school closure. *Journal of Computer Science and Its Application*, 27(2). https://doi.org/10.4314/jcsia.v27i2.9.

[4] R. Abd Karim, & A. G. Abu (2018). Using mobile-assisted mind mapping technique (mammat) to improve writing skills of esl students. *Journal of Social Science and Humanities*, 1(2), 1-6. https://doi.org/10.26666/rmp.jssh.2018.2.1.

[5] V. Lang, J. Senekovič, š. Majcen, & A. Šorgo (2022). Smartphone as a coagulation sensor. *Dianoia*, 6, 37-50.

[6] C. T. Gee, E. Kehoe, W. C. Pomerantz, & R. L. Penn, (2017). Quantifying protein concentrations using smartphone colorimetry: a new method for an established test. *Journal of chemical education*, 94(7), 941-945. https://doi.org/10.1021/acs.jchemed.6b00676

[7] M. A. Schaefer, H. N. Nelson, J. L. Butrum, J. R. Gronseth, & J. H. Hines (2023). A low-cost smartphone fluorescence microscope for research, life science education, and STEM outreach. *Scientific Reports*, 13(1), 2722. https://doi.org/10.1038/s41598-023-29182-y

[8] K. Dolenc, & A. Šorgo (2020). Information literacy capabilities of lower secondary school students in Slovenia. *The Journal of Educational Research*, 113(5), 335-342. https://doi.org/10.1080/00220671.2020.1825209

[9] A. Šorgo, T. Bartol, D. Dolničar, & B. Boh Podgornik (2017). Attributes of digital natives as predictors of information literacy in higher education. *British Journal of Educational Technology*, 48(3), 749-767. https://doi.org/10.1111/bjet.12451

[10] V. Lang, & A. Šorgo (2022). Elementary School Students and their Satisfaction with Smartphone use In Biology classes. In *INTED2022 Proceedings* (pp. 7516-7516). IATED. https://doi.org/10.21125/inted.2022.1904

[11] Lang, V., & Šorgo, A. (2022b). Added value of the Pl@ ntNet smartphone application for the motivation and performance of lower secondary school students in species identification. In *ICERI2022 Proceedings* (pp. 4534-4540). IATED.

[12] M. Volk, M. Cotič, M. Zajc, & A. I. Starcic (2017). Tablet-based cross-curricular maths vs. traditional maths classroom practice for higher-order learning outcomes. *Computers & Education*, 114, 1-23. https://doi.org/10.1016/j.compedu.2017.06.004

[13] V. Lang, & A. Šorgo (2023). Recognition of the Perceived Benefits of Smartphones and Tablets and Their Influence on the Quality of Learning Outcomes by Students in Lower Secondary Biology Classes. *Applied Sciences*, 13(6), 3379. https://doi.org/10.3390/app13063379

[14] H. B. Mann, & D. R. Whitney (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60. https://www.jstor.org/stable/2236101

[15] F. Wilcoxon (1945). Some uses of statistics in plant pathology. *Biometrics Bulletin*, 1(4), 41-45. https://doi.org/10.2307/3002011

[16] SURS, 2020. https://www.stat.si/statweb. Accessed Feb. 5, 2022.

[17] B. Macuh, A. Raspor, M. Sraka, & A. Kovačič (2018). Media exposure and education of first to six grade children from Slovenia-parent opinions. *International Journal of Cognitive Research in Science, Engineering and Education*, 6(3), 49. https://doi.org/10.5937/ijcrsee1803049M

Using VIDI-X Microcomputers in High School and College Education

Branko Balon* , Aleksandar Skendžić* and Milenko Simić**
*College for Information Technologies, Zagreb, Croatia
**The First Technical School Tesla, Zagreb, Croatia
branko.balon@vsite.hr, aleksandar.skendzic@vsite.hr, milenko.simic@skole.hr

Abstract **- This paper describes the application of the VIDI-X development microcomputer for educational purposes. VIDI-X is a platform developed entirely in Croatia and another great way to popularize STEM and learn new skills. VIDI-X is based on a dual-core ESP 32 processor and comes with a temperature sensor, a microphone, and a 2.8" touch screen. The possibilities of connection and application are enormous considering the diverse communication possibilities (Wi-Fi, Bluetooth, IR). VIDI-X is a significantly more powerful platform than Arduino, and unlike the Raspberry Pi computer, it does not need an operating system, and it can be programmed on a Windows, Linux, or Mac OS computer. It is possible to work in Python, C, C++, JavaScript, and other popular programming languages. Our students first installed the Arduino IDE on their laptops, and then connected the laptops to VIDI-X, and a series of exercises followed that combine the use of software and hardware. Our experience with this platform in the undergraduate computing curricula, at both the high school and college levels is presented here.**

Keywords - VIDI-X, ESP 32, STEM, Education,

I. INTRODUCTION

VIDI Project X was presented on December 5, 2019 in Croatia at conference for innovative projects awards powered by VIDI magazine. The whole project was developed by the VIDI media company in cooperation with a number of Croatian technology companies that financed the project.

VIDI-X is a microcomputer board that in its vision combines the "dream" and "realize" phases of the project, which we believe also requires innovation. The definition could be many, depending on the purpose you use it for. On VIDI-X microcomputer it is possible to run the original Doom or Duke Nukem in the emulator. You can connect additional sensors and manage the smart home system, you can use its Wi-Fi and Bluetooth communication capabilities to connect to an external system.

VIDI-X makes good use of the capabilities of the powerful dual-core ESP32 processor, so it can serve as a multifunctional microcomputer or as a development board, as an efficient HUB for a cluster of sensors or for some completely new purpose that is expected from the community that will gather around it to design.

Why did we choose VIDI-X microcomputer for education purpose at both the high school and college level? We wanted more powerfull platform, totaly new and completelly developed in our country - Croatia. We allready have experience with other widely used microcomputers boards like Arduino and Raspberry Pi microcomputers in education [4][6].

VIDI-X is a significantly more powerful platform than Arduino, and unlike a Raspberry Pi computer, it does not need an operating system. To run, a power source is required via a 3xAA battery or via USB cable, and it can be programmed on a computer with Windows, Linux or Mac OS system. Since its introduction, VIDI-X soon became very popular in Croatia also in the education, especially in STEM field [1] [7].

In II. chapter VIDI-X hardware and programming environment is described in detail. The examples of VIDI-X projects, that were introduced at laboratory exercise course in High school Tesla in Zagreb are shown in III. Chapter. The VIDI-X micro server project, that was used for demonstration in the *Computer Network course* at the College for Information Technologies - VSITE in Zagreb is described in last chapter of this paper.

II. VIDI-X HARDWARE AND PROGRAMMING

The VIDI-X development microcomputer is based on an ESP32 processor, a 2.8" integrated screen, a light sensor, an IR sensor, a temperature sensor, integrated speaker and the possibility of Wi-Fi, IrDA and Bluetooth connectivity.

Here is a quick overview of main parts od VIDI-X microcomputer The parts are numbered from 1 to 21 and are indentified on the front and the back side of VIDI-X as shown in Figure 1. and Figure 2.

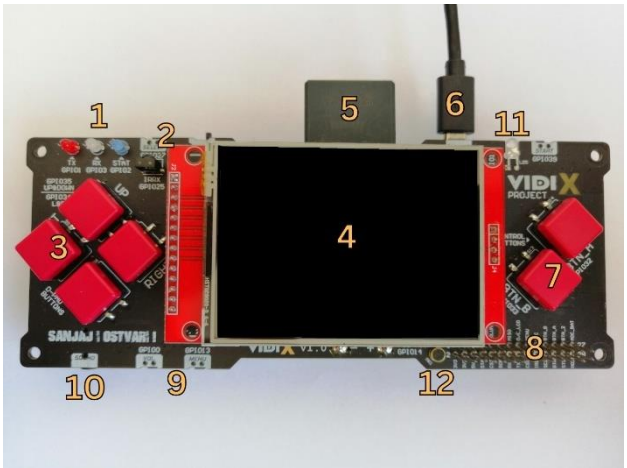


Figure 1. VIDI-X Front side

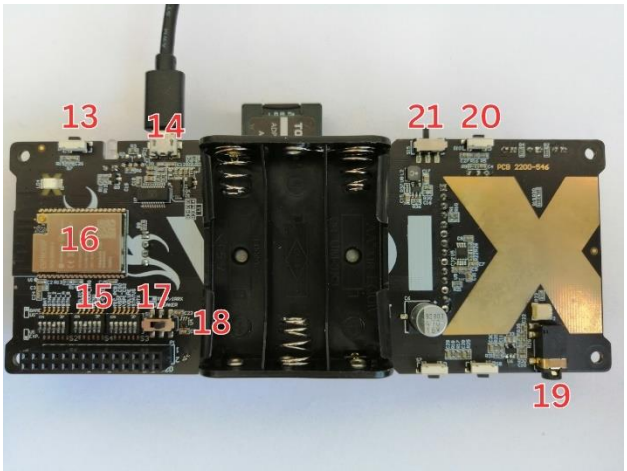


Figure 2. VIDI-X Back side

1. Red, white and blue LEDs - are control lights so we can visually check whether there is activity on the GPIO1, GPIO2 and GPIO3 pins. We mark GPIO2 as STATUS while the other two show serial TX and RX.

2. InfraRed receiver - will enable reception of ESP32-WROVER-B infrared signals from various remote controls of your lines, TV or other VIDI X microcomputer. It is connected to GPIO25.

3. BUTTONS - Two buttons are connected to one pin. Up - down on GPIO35, and left - right on GPIO34

4. Touch LCD screen - The screen has a resolution of 320x240 pixels and is touch sensitive. The screen itself can be easily separated from the rest of the board if necessary. A pen is included for easier use.

5. SD card - will enable more permanent storage of collected data from various sensors

6 . Micro USB - Power from the USB connector and serial communication for programming

7. Buttons labeled A and B - They are connected to GPIO32 and GPIO33 as marked on the board itself

8. Expansion header - it is available with a female version from the back and a male version from the front. On the front, next to the male pins, their markings are written for easier use

9. Volume and Menu - Volume and Menu buttons are connected to GPIO0 and GPIO13

10. Combo audio connector - The four-pole 3.5 mm audio connector for external microphone and headphones

11. InfraRed transmitter - connected to GPIO15 enable the transmission of infrared signals

12. Microphone - The microphone is connected to GPIO14 which serves as an audio input.

13. Start - The start button is connected to GPIO39

14. Micro USB - Power from the USB connector and serial communication for programming

15. Expansion header - In order to use the 28-pin expansion header, it is necessary to switch the microswitches next to it to "USE EXP." position. If they are in the "GAME USE" position, it is possible to use the buttons and sensors on the VIDI-X

16. ESP32-WROVER-B - VIDI-X microcomputer is based on the ESP32 module with a 32-bit dual-core processor that has 600 DMIPS (Dhrystone benchmark). 600 DMIPS means that the processor is capable of processing 600 iterations of the main loop code per second.

17. Switch - With this switch we determine whether we want to use the audio amplifier connected to GPIO25 and GPIO26 to have sound, or in another position the InfraRed receiver and temperature sensor. The temperature sensor is connected to GPIO26 while the InfraRed receiver is connected to GPIO25

18. Temperature sensor - MCP9700AT-E/LT has a measurement range from -40°C to 150°C with an accuracy of ±2°C

19. Speaker connector - This connector connects the speaker

20. Select - The Select button is connected to GPIO27

21. On/Off switch

VIDI-X also has three buses, UART, SPI and I2C, which it uses for direct communication with sensors, a computer or some other micro-controllers. Direct communication with other devices is achieved through the serial UART bus and General Purpose Input/Output (GPIO) connection pins as shown in Figure 3.

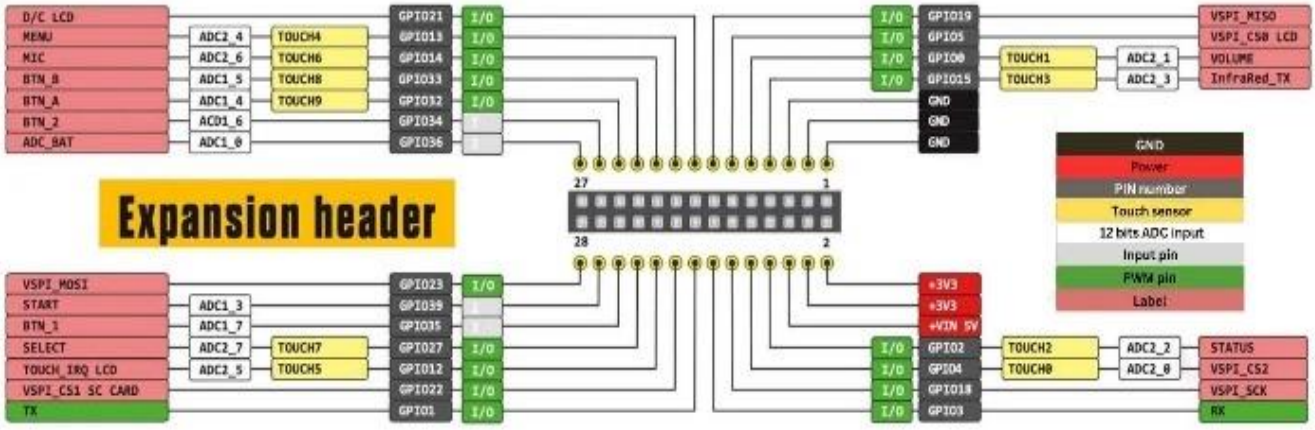


Figure 3. GPIO connection Source: https://VIDI-x.org/

Basic technical specifications of the VIDI-X processor module ESP32 are shown in Table I.

TABLE I. BASIC TECHICAL DATA - ESP32 MODULE
Source: https://VIDI-x.org/

Producer:	Espressif Systems
Wi-Fi:	802.11 b/g/n, BT
IrDA	yes
Frequency:	2.4 GHz to 2.5 GHz
Data transfer:	150 Mb/s
Security:	WPA/WPA2/WPA2-Enter-prise/WPS
CPU:	Dual-core Tensilica Xtensa LX6 microprocessor @ 240 MHz
Memory:	520 KB SRAM, 8 MB SPI flash memory

The VIDI-X microcomputer supports multiple development environments such as Arduino IDE, Visual studio CODE and PictoBlox, and Windows and MacOSX platforms are supported [1]. The development environment used in the projects is the Arduino IDE (Integrated Development Environment), and it was used for writing/uploading to the VIDI-X microcomputer. The Arduino IDE development environment is also used for compatible Arduino microcontrollers [2] [5] . The Arduino IDE environment supports the C and C++ programming language with code structuring rules and the necessary libraries used to compile the ESP32 processor code.

III. VIDI-X PROJECTS IN HIGH SCHOOL TESLA

VIDI–X development microcomputers are used at laboratory exercise course at the First Technical School and computer technicians program. Laboratory exercises are organized in small groups (maximum 10 students). Exercises are divided into two main parts. The first part refers to getting started with VIDI–X development microcomputer and to become familiar with its design. In the second part of this laboratory exercises students do some practical electronic ojects. They combine various electronic components and write code with Arduino IDE (Integrated Development Environment) to control them.[3] [4] [5].

A. LED control on VIDI-X

VIDI-X has 3 control LEDs so that we can visually check whether there is activity on the GPIO01 (red), GPIO02 (blue) and GPIO03 (white) pins. The red and white LEDs induce serial communication and are not programmable. GPIO02 (blue) LED is status and we can program it. The designation GPIO represents a general purpose input/output. Here is the example of the code:

```
int LEDPin1 = 2;
void setup(){
 pinMode(LEDPin1, OUTPUT);
 digitalWrite(LEDPin1, LOW);
}
void loop(){
 digitalWrite(LEDPin1, HIGH);
 delay(1000);
 digitalWrite(LEDPin1, LOW);
 delay(1000);
}
```

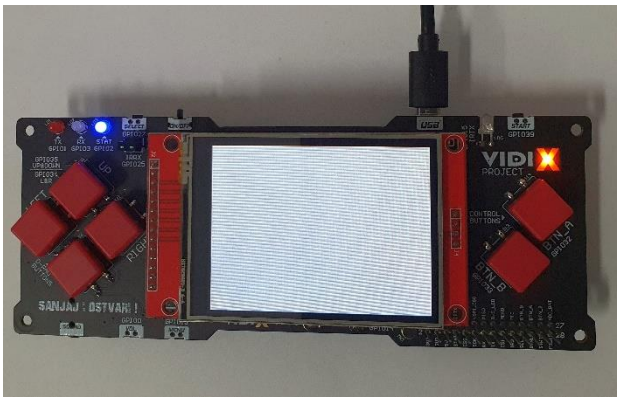


Figure 4. GPIO02 blue LED is programmable status light

B. *Temperature sensor management on VIDI-X*

As mentioned before on the VIDI-X board there is a temperature sensor MCP9700AT. It is located on the back of the board and is connected to GPIO26. In order to be able to use this sensor, the switch must be set to the TEMP/IRRX position. The temperature is read in 12-bit resolution, which means that it is possible to read 2^{12} = 4096 different values. The temperatutre readings are sent to serial port as shown in Figure 5.

Example of the code:

```
int PinTemp = 26;
int temp;
float tempV;
float tempC;
void setup() {
 pinMode(PinTemp, INPUT);
 Serial.begin(9600);
}
void loop() {
 temp = analogRead(PinTemp);
 tempV = map(temp, 0, 4095, 0, 3300);
 tempC = tempV / 10 - 40;
 Serial.print((String)"Temp in C = ");
 Serial.println(tempC);
 delay(1000);
}
```

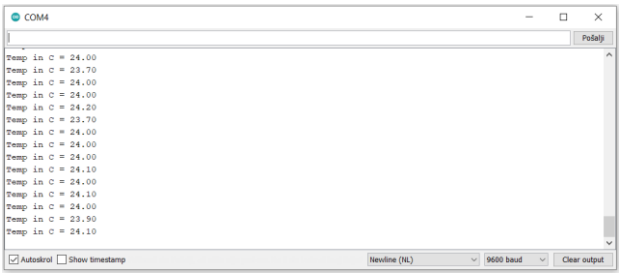


Figure 5. Temperature readings

C. *Servomotor control on VIDI-X*

Servomotors are DC motors that enable precise positioning of the motor shaft. Most hobby servomotors can rotate their axis by 180°, but there are also those that can rotate up to 270° or even 360°. In this project , the SG-90 motor is used, which can make 180°. In addition to the angle the shaft can make, servomotors differ physically in the size and construction of the gears (plastic or metal). They can be controlled analog via PWM signals or digitally via I/O signals. Most hobby servos are powered by 5V-6V, so the servo can be directly connected to the VIDI-X without an additional battery. However, if you want to connect more servomotors, you need to add an external power supply. In order to be able to manage the servo motors on the VIDI-X, it is necessary to download the ESP32Servo command library from the Internet by selecting Tools, then Manage Libraries, within the Arduino IDE interface. Here is the example of the code:

```
#include <ESP32Servo.h>

// create servo object to control a
   servo

Servo myservo;

int pos = 0;    // variable to store the
                   servo position

void setup() {
  myservo.attach(13);  // attaches the
     servo on pin 13 to the servo object
}

void loop() {
  for (pos = 0; pos <= 180; pos += 1) {
// goes from 0 degrees to 180 degrees
    // in steps of 1 degree
    myservo.write(pos);
    delay(15); // waits 15ms for the
        //servo to reach the position
  }
  for (pos = 180; pos >= 0; pos -= 1){
 // goes from 180 degrees to 0 degrees
    myservo.write(pos);
    delay(15); // waits 15ms for the
        //servo to reach the position
  }
}
```

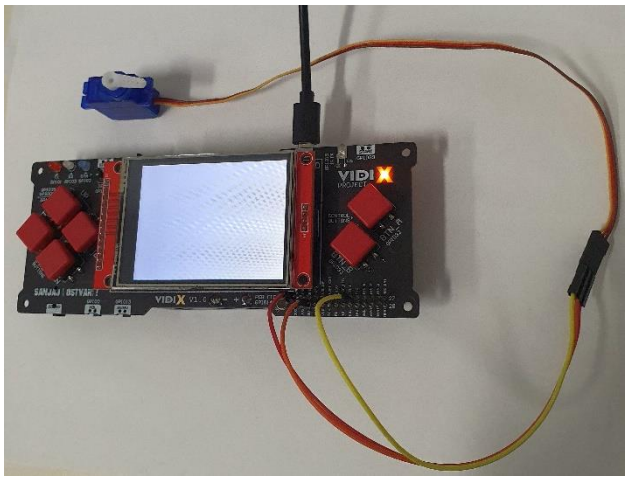


Figure 6. Servomotor controlled by VIDI-X

D. Controlling the TFT LCD on VIDI-X

The screen that comes with VIDI-X is a 2.8" ILI9341 SPI TFT LCD screen with a resolution of 320x240 pixels and is touch sensitive. The screen itself has many pins, some of which are intended for power supply, some for management, some for sending data that will be displayed on the screen, and some for receiving information if you write on the screen.

The ILI9341 driver is used to control the display, which must be additionally downloaded from the Internet and installed. To download and install it is necessary to select Manage Libraries in the Tools menu within the Arduino IDE. Here is the Example of the code:

```
#include <Adafruit_ILI9341.h>
#include <Adafruit_GFX.h>

// ILI9341 TFT LCD pin declaration
#define TFT_CS 5
#define TFT_DC 21
// create a screen object that will be
called tft
Adafruit_ILI9341 tft
    = Adafruit_ILI9341(TFT_CS, TFT_DC);

void setup() {
 tft.begin(); // screen initialization

 tft.setRotation(3); // set orientation
        // screen color
 tft.fillScreen(ILI9341_BLACK);
 tft.setTextColor(ILI9341_WHITE);
 tft.setCursor(45,105); position
 tft.setTextSize(5); // text size
 tft.print("VIDI - X"); // print text

 tft.setTextColor(ILI9341_RED);
 tft.setCursor(0,0); // position
 tft.setTextSize(5); // text size
 tft.print("TESLA"); // print text
 tft.setTextColor(ILI9341_BLUE);
 tft.setCursor(170,200); // position
 tft.setTextSize(5); // text size
 tft.print("VSITE"); // print text
}
void loop() {
}
```

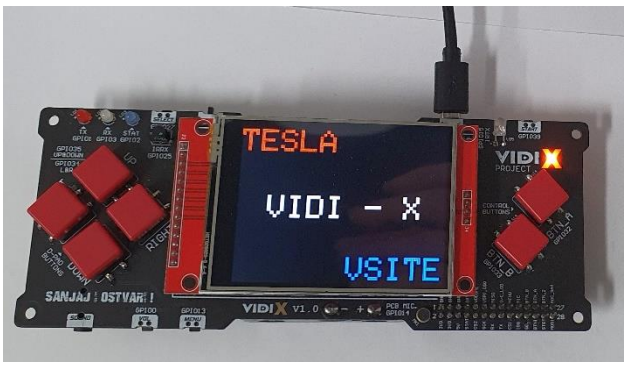


Figure 7. Displaying on VIDI-X TFT LCD

IV. VIDI-X VSITE COLLEGE PROJECTS

The VIDI-X micro server project was used for demonstration purposes in the *Computer Network course* at the College for Information Technologies VSITE in Zagreb. The basic idea of the project is the creation of a micro server with the possibility of network connection (soft Access Point) on which a simple HTML application is installed, through which the signal LED on the VIDI-X microcomputer is turned on/off. The network connection of the VIDI-X micro server is achieved via a wireless connection (SoftAP function), since there is no ethernet network connection on board. The goal of the project was to demonstrate the capabilities of the VIDI-X microcomputer in the role of a simplified micro server in a network environment. The project, in addition to the on-screen display of the network connection status, has been further extended with the functionality of an application that executes HTML code that turns on/off the integrated signal LED.

A. Development environment and used program libraries

The development environment used for the VIDI web server project is the Arduino IDE in version 1.8.10. [6]. The necessary program libraries were also used:

- WiFi.h - library for working with integrated Wi-Fi
- WebServer.h - server functionality library
- SPI.h - serial communication library
- gfxfont.h – library for exploiting the possibilities of TFT screens and printing on the screen
- AdaFruit_ILI9341.h - library for exploiting the possibilities of TFT screens and printing on the screen

B. *Basic network settings*

Network settings are specified for demonstration purposes. By calling the *IPAddress* function, the static IP address of the SoftAP server (192.168.1.1), the default Gateway and the name of the Wi-Fi network (Service Set Identifier - SSID) are defined. The defined network settings are used in the function *WiFi.softAPConfig*. The server listens to traffic on port 80 for http requests, which is accomplished with the *WebServer server(80)* function. Part of the program code of the used libraries and network settings is shown:

```
/* libraries needed */
#include <WiFi.h>
#include <WebServer.h>
#include <SPI.h>
#include <gfxfont.h>
#include <Adafruit_ILI9341.h>

/* SSID and password for wifi */
const char* ssid = "VIDI-X-PROJECT";
const char* password = "12345678";

/* Server IP address */
IPAddress local_ip(192,168,1,1);
IPAddress gateway(192,168,1,1);
IPAddress subnet(255,255,255,0);

/* Server listens on port 80 */
WebServer server(80);
```

C. *Wireless network settings – Soft Access Point*

Since the server has the possibility of wireless connection, the wireless network settings that enable this have also been defined. A wireless access point has been established, which enables the connection and reception of clients who authenticate. For this purpose, the following functions were used [7]:

- `server.begin();` – function that starts the server;
- `WiFi.softAP();` –function that defines VIDI-X as an access point
- `WiFi.softAPConfig();`- configuration parameter of the softAP function, i.e. micro server access point
- `WiFi.softAPgetStationNum();` - a function that obtains the number of clients currently connected to the micro server - SoftAP;
- `server.handleClient();` - listens to incoming HTTP requests;
- `server.onNotFound(handleNotFound);` - when the client requests an unknown URL (different from "/", for example), the "handleNotFound" function is called;
- `server.handleClient();` - a function that listens for HTTP requests from clients.

Definition of the softAP function for wireless connection to the micro server:

```
WiFi.softAP(ssid, password);
WiFi.softAPConfig(local_ip,gateway,sub
net);
delay(100);
```

Part of the program code that defines the number of wireless clients connected to the access point (server):

```
 server.begin();
 Serial.println("HTTP server started");
}
void loop(){
 tft.setTextWrap(false);
 Serial.println("No. of connected
clients");
 Serial.println
         (WiFi.softAPgetStationNum());
 tft.setCursor(10, 130);
 tft.println
         (WiFi.softAPgetStationNum());
 delay(100);
 server.handleClient();
 if (LED1status)
 {digitalWrite (LED1pin, HIGH); }
 else
 {digitalWrite (LED1pin, LOW); }
}
```

D. *Defiinition of screen display*

The *gfxfont.h* and *AdaFruit_ILI9341.h* libraries were used to take advantage of the built-in TFT screen and screen printing. The network configuration parameters and the number of active connections are displayed on the screen.

For this purpose, the TFT_DC and TFT_CS pins are defined, to which the TFT screen of the VIDI-X microcomputer is connected. The font colors used in the server status screen are also defined. Font colors are determined by hexadecimal notation (e.g. WHITE – 0xFFFF). The added functionality of the micro server is to turn on/off the LED connected to PIN 2 controlled by a simple HTML application placed on the server. To fill the screen with black, the command tft.fillScreen(BLACK) was used, which resulted in a better contrast, i.e. white letters on a black background. The command *tft.setRotation* rotates the print on the screen to the appropriate position, depending on the user's needs. The command *tft.setCursor(10, 0)* defines the initial position of text printing on the screen along the X or Y axis, while the commands *tft.setTextColor* and *tft.setTextSize* define the

color of the text to be printed on the screen as well as the size of the text. [7].

The following is the program code for printing the settings of the VIDI-X micro server, and the printout on the screen is shown in Figure 8.

```
/* definition when LED is ON */
void handle_led1on(){
 LED1status = HIGH;
 Serial.println("GPIO2 Status: ON");
 server.send(200, "text/html",
            SendHTML(true));
 tft.fillScreen(BLACK);
 tft.setCursor(10, 0);
 tft.setTextColor(WHITE);
 tft.setTextSize(1);
 tft.print("Server IP address: ");
 tft.println(local_ip);
 tft.setCursor(10, 10);
 tft.print("Gateway: ");
 tft.println(gateway);
 tft.setCursor(10, 20);
 tft.print("Subnet mask: ");
 tft.println(subnet);
 tft.setCursor(10, 100);
 tft.setTextColor(GREEN);
 tft.setTextSize(2);
 tft.println("VIDI Project X Web
              server");
 tft.setTextSize(1);
 tft.setTextColor(WHITE);
```

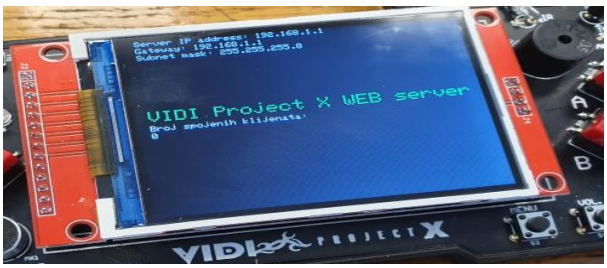


Figure 8. VIDI-X micro server. Display of server network settings and the number of connected wireless clients. Source: vidilab.com [7]

E. HTML Application and controll of LED lights

The VIDI-X micro server project has been expanded with additional functionality, which refers to the possibility of turning on/off the signal LED via an HTML application that is executed on the micro server. For this purpose, two functions are defined in the programming part of the code:

- `void handle_led1on()`
- `void handle_led1off()`

Functions *handle_ledon()* and *handle_ledoff()* **status code 200** (HTTP status) is sent, which corresponds to a confirmation response (OK). The content to be sent is "text / html", while at the end the *SendHTML()* function is called, which creates a dynamic HTML page that turns the LED light on and off. The *tft.printl* command prints the content on the screen. Figure 9. shows the interface of the HTML application running on the VIDI-X micro server.



Figure 9. Mobile application that turns on/off the VIDI-X LED
Source: vidilab.com [7]

V. CONCLUSION

The goal of the projects described in this paper was to demonstrate the capabilities of the VIDI-X microcomputer. Our experience with VIDI-X as an educational tool is exceptionally good. It is easy and fun for students to work with VIDI-X at both, the high school and college levels. All in all, it is another great way to popularize STEM and learning new skills. Anyone who is interested in more information will find them at the VIDI-x.org link, which presents hardware in more detail, the options that are significantly wider than these mentioned in this paper, explained examples of projects that can be made. Startup companies, hobbyists, researchers, as well as students receive a platform by which they can manage other devices, develop their own projects or use it as part of existing solutions.

REFERENCES

[1] VIDI-X, [Online] Available: https://VIDI-x.org/VIDI-x-razvojno-okruzenje/ [Accessed: 20. February 2023.]

[2] https://docs.arduino.cc/software/ide-v1/tutorials/Windows [Accessed: 18. February 2023.]

[3] El-Abd, Mohammed, " A Review of Embedded Systems Education in the Arduino Age: Lessons Learned and Future Directions" International Journal of Engineering Pedagogy (iJEP). 7. 79-93. 10.3991/ijep.v7i2.6845, 2017.

[4] B. Balon and M. Simić, "Using Raspberry Pi Computers in Education", 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, pp. 671-676, 2019.

[5] Novak, Milan & Kalova, Jana & Pech, Jiři, "Use of the Arduino Platform in Teaching Programming", 1-4. 10.1109/INFORINO.2018.8581788, 2018.

[6] B. Balon and M. Simić, "Arduino Platform as Learning Tool in High School and College Education" , 44th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, pp. 740-745. 2021.

[7] A. Skendžić "Vidi ProjectX #98,", VIDI Computer magazine No. 287, [Online] Available: https://www.vidilab.com/vidi-project-x/4753-vidi-project-x-web-server [Accessed: 20. February 2023.]

How Well Students Perceive Their Understanding of Logic Programming Course Content?

S. Lovrenčić, V.Sekovanić
University of Zagreb Faculty of Organization and Informatics, Varaždin, Croatia
sandra.lovrencic@foi.unizg.hr
vlatka.sekovanic@foi.unizg.hr

Abstract **- Teaching and learning programming languages is challenging in many ways. Students are often more familiar with imperative paradigm, and therefore programming languages that belong to declarative paradigm are even more demanding. Logic programming languages are usually taught when students already have prior knowledge of programming concepts and logic. This should make their learning easier, and also their perception to which level they have mastered course content prior to exam should be quite good. The paper describes analysis and results from five-year research on a course where logic programming is taught in the Prolog programming language, and includes one year of online teaching. Students' feedback on their perception of understanding the course content, especially practical part is analysed and compared to their success in passing the course at the end of the semester. Differences between perception and success are discussed together with results of previous research of learning problems perceived by students, with the aim of investigating the role of opinions and perception in learning analytics and improving the learning design of this type of course.**

Keywords - logic programming course; learning analytics; perceived understanding

I. INTRODUCTION

Logic programming, as part of declarative programming paradigm, is often taught at computer science study programmes as compulsory or elective course. Declarative programming paradigm is considered as contrast to imperative programming paradigm, which has procedural and object-oriented programming as most known representatives. Contrast relates to the different way of encoding information and procedures. For example, logic programming is based on logic and is mostly used for representing declarative knowledge (facts and rules) with reasoning (search) procedures already defined in the programming language [1]. It is descriptive by nature and programmers' task is to describe chosen domain in a knowledge base by entering what is true [2].

Programming languages belonging to imperative programming paradigm, such as C or C++, or multiparadigm programming languages, for example Python, are often taught at introductory programming courses in the first year of undergraduate study. Logic programming languages, whose most prominent representative is Prolog [3,4], language that is based on first-order predicate logic [5], are taught at higher years. Knowledge of those languages continues to be very important - Prolog is today widely used in artificial intelligence field [6,7,8,9,10] and it is well-known that it was used for programming the parts of the the IBM's Watson supercomputer [11]. Therefore it is important to ensure that students understand the theoretical concepts important for understanding and learning logic programming.

Since logic programming is taught when students already have prior knowledge of programming concepts from previous programming courses and basic knowledge of logic from mathematics course, this would imply that they can learn new programming languages easier, but it is also understandable that this may not be the case if they belong to different programming paradigm. However, it would be expected that they can relatively correctly perceive whether they understand the course content well. The purpose of the research presented in this paper is to investigate whether this supposition is true, which can be done by comparing students' perception of understanding the course content with their success in passing the course. Results can also help in better understanding how student perception can be used in combination with learning analytics and help in learning design.

The paper is organized as follows. In Section II is reviewed previous research on the problems in teaching and learning programming in general as well as logic programming and the use of learning analytics in programming education. The research is described in Section III and results of the student survey during five year period are presented in Section IV, focusing on how students perceive their understanding of the course content in general and aditionally practical part, compared to their grades. In Section V are presented concluding remarks.

II. BACKGROUND

Programming knowledge is today one of the most important skills for prospective career. It is also one of the skills that is not easy to teach and learn, because students often have problems with understanding programming concepts, feel fear or anxiety and have lack of motivation [12,13,14]. They have different learning styles (for example, alone or in groups), which combined with the choice of programming language and the fact that programming is actually a set of skills and involves more processes, can make teaching and learning very difficult [15]. Important elements in teaching programming to

beginners are curricula (what to teach), pedagogy (managing teaching and learnign process), choice of programming language and type of tool for chosen language (for example, visualisation tool or programming support tool) [16].

Since logic programming languages demand different approach to programming process than programming languages that students are already familiar with, above mentioned problems in teaching and learning do not become easier with students' prior experience. Most of the research about problems of teaching logic programming is based on experience with Prolog, since it is the most known and used logic programming language, with common conclusion that the different programming paradigm and its concepts are difficult to learn [17, 18, 19, 20], and therefore the same elements in teaching are important as in beginner courses [16], with balancing theoretical abstract and concrete practical approach, taking into consideration different learning styles [19] and structuring the teaching material properly [17].

Learning analytics is a discipline that is gaining moment and growing in potential for use in education for more than ten years, enabling collection, analysis and understanding the data about learners (students) with the goal to improve (optimize) the learning process [21]. Various methods of data collection and analysis as well as tools can be used in different areas of application, such as modelling user knowledge or domain, personalization or trend analysis [21, 22]. The use of learning analytics enables better understanding of students' learning process and can improve learning support and teaching [23], as well as retention of students [24]. Since it emerged, the discipline was also used to understand the learning of programming [25] with research oriented to tracking students' activity and comparing it to tinkering and changes in quality of programs [26] or tracking students' progress in code development and exploring the correlation with course performance and grades [27]. It can also be valuable tool in detecting students that have risk of failing some assignments or the programming course [28], especially because very detailed data can be gathered and analysed, including specific data during programming, for example line-level edits [29].

III. RESEARCH

At the University of Zagreb Faculty of Organization and Informatics (UNIZG FOI), from the academic year 2022/2023, the Introduction to Formal Methods course is offered as a revised Introduction to Knowledge Modeling course. The course redesign was also based on the results of the research that included analysing a five-year data (academic years 2016/2017-2020/2021) obtained from student surveys and their course performance, including final grades. The goal of the whole research was to improve the quality of teaching and learning process. The part of the research included comparison of students' perception of their understanding of the content of the course compared to the actual understanding that is represented with their final grades.

At the end of each execution of the course students completed surveys that collected information on understanding the course content, learning preferences, problems in understanding the Prolog, as well as their comments and recommendations. In Table 1 can be seen that out of 864 enrolled students during 5 years, 675 completed the survey. After online classes during COVID-19 pandemic in academic year 2020/2021 there was a lower response to the survey. Number of responses to questions that show students' perception about understanding course contents in general and of the practical part (Q5, Q6, Q11 and Q14) are also presented in the table. In Question 5 students were asked to asses their understanding of the course content in general and in Question 6 to assess their understanding of the practical part (exercises on computers); for both questions possible answers were low, medium and high. In Question 11 students were asked to asses the ammount of learning material for Prolog as insufficient, optimal or too extensive, and in Question 14 to asses whether team assignment helped them in learning Prolog (yes or no).

TABLE I. ANNUAL NUMBER OF ENROLLED AND SURVEYED STUDENTS

Year	No. of enrolled students	No. of completed surveys	Q 5	Q 6	Q 11	Q 14	% of completed surveys
2016	182	154	154	151	154	*	84.6
2017	205	177	175	177	177	176	86.3
2018	159	126	126	124	124	126	79.2
2019	165	144	144	144	144	142	87.3
2020	153	74	74	74	74	74	48.4
Total	**864**	**675**	**673**	**670**	**673**	**518**	**78.1**

Learning outcomes adoption during each semester was tested with several assesment elements that contributed to final grade in the following percentages: two midterm exams - 60% (each 30%), seminar paper - 30%, activity - 5%, and seminar attendance - 5%. Direct knowledge of Prolog programming was assesed during midterm exams with 45% of the exams points (27% of the grade), knowledge of resolution reasoning procedure that Prolog uses for reasoning with 31.7% of the exams points (19% of the grade) and theoretical knowledge of logic needed for understanding practical part with 23.3% of the exams points (14% of the grade). Therefore, testing the practical knowledge of logic programming (Prolog and resolution) directly contributed to the final grade with 46%. In Table 2 are presented grades that students received after course completion. For comparison with survey questions about perception, grades are represented as low (grade 1), medium (grades 2 and 3) and high (grades 4 and 5) course understanding. Each point is 1% of the grade.

TABLE II. ANNUAL NUMBER OF ENROLLED AND PASSED STUDENTS

Year	Number of enrolled students	Percentage of grades (1; 0-49 points)	Percentage of grades (2 or 3; 50-75 points)	Percentage of grades (4 or 5; 76-100 points)
2016	182	31	67	2
2017	205	47	52	1
2018	159	41	54	5
2019	165	42	54	4
2020	153	53	47	0
Total	**864**	**43**	**55**	**2**

IV. RESULTS

A. *Student Perception of the Understanding of the Entire Course and Exercises*

Results were obtained using simple descriptive analytics of student surveys and grades. Questions 5 and 6 focus on student perception how well they understand the course content in general and additionally practical part – exercises on computers. Fig. 2 shows the results of the perception of understanding the entire course content (the theoretical knowledge of logic, knowledge of reasoning procedure and knowledge of logic programming). It is evident that the students to the greatest extent (73%-81%) perceive an average understanding of the entire course content. The remaining students perceive their understanding of the entire course as low (9%-18%) and high (7%-11%) throughout the observed years. The overall perception of the entire course content presented in Fig. 3 shows that on average 77% of students perceive their understanding of the entire course as a medium, 14% as low, and 9% as high.

Similar results were obtained when perceiving understanding of the exercises on computers (logic programming in Prolog) was observed. According to Fig. 4 and 5, students still perceive the understanding of the content of the exercises as medium (67% in total or 57%-71%, depending on the observed year). The only difference (in terms of the perception of the overall content) is that more students perceive the understanding of the exercises as high (22% overall, i.e. 9%-31% over the years), and slightly fewer students perceive it as low (11% overall, 3% -22% throughout the observed years).

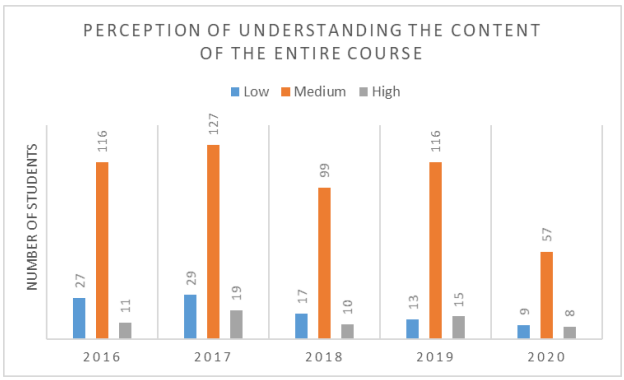


Figure 2. Perception of understanding the content of the entire course (2016-2020; N = 673)

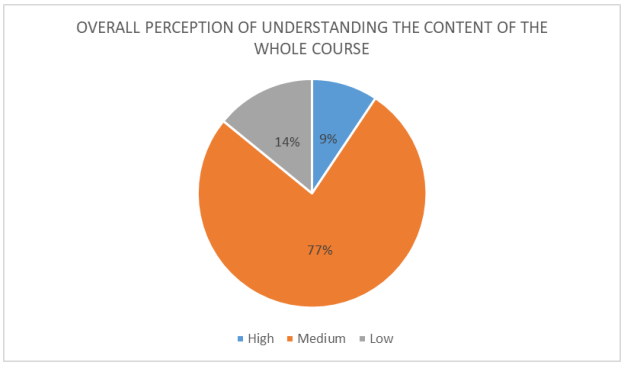


Figure 3. The overall perception of understanding the content of the entire course (2016-2020; N = 673)

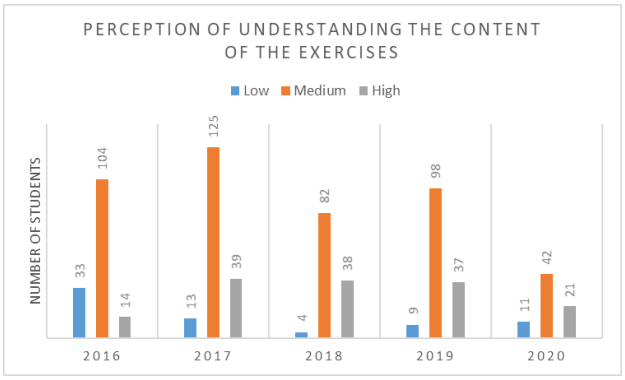


Figure 4. Perception of understanding the content of the exercises (2016-2020; N = 670)

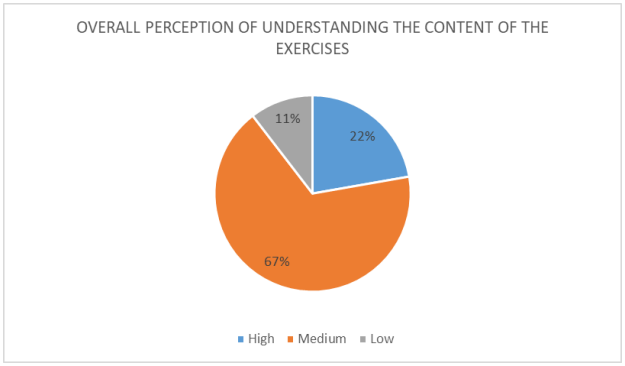


Figure 5. The overall perception of understanding the content of the exercises (2016-2020; N = 670)

Comparison of students' perceptions of understanding the course content (entire or only exercises) and final grades (Fig. 6. and 7.) indicates considerable differences between the perception and the real understanding of the content. According to the analysed data, 43% of all students had a low understanding of the course content (they achieved 0-49 points at assesment elements during semester and did not successfully pass the course). The data show that a majority of all students, 55%, showed a medium understanding of the course content (between 50 and 75 achieved points). The actual percentage is admittedly lower than the perceived one (concerning the content of the entire course and the exercises). However, it is still the closest to perception. Moreover, only 2% of all students (as opposed to 9% and 22% according to self-perception) showed high understanding of the course content (according to 76-100 achieved points).

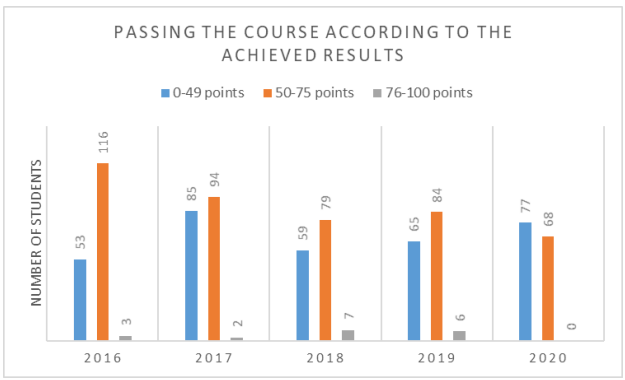


Figure 6. Passing the course according to the achieved results (2016-2020; N = 798)

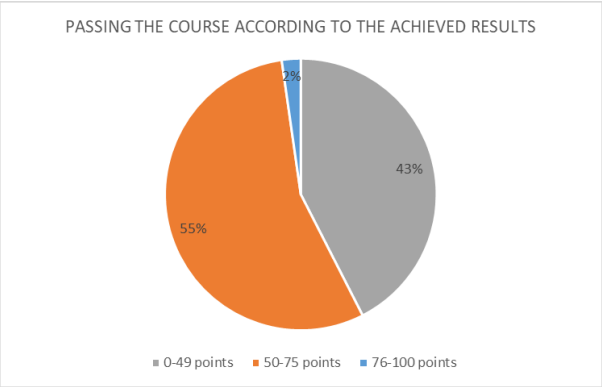


Figure 7. Passing the course according to the overall achieved results (2016-2020; N = 798)

Previous research with other survey information about bigest problems in mastering the Prolog (syntax, logic/semantics, reasoning procedure, declarative paradigm, and examples used in class) identified declarative paradigm as the biggest problem and other elements as less problematic [30], which additionally contributes to student perception that they understand practical part of the course at least at the medium level.

B. Student Perception of the Learning Materials and Team Assignments

Additional information about student perception of their understanding of the course can be obtained with Questions 11 and 14, since practical knowledge of logic programming (Prolog and resolution) directly contributed to the final grade with 46%.

Answers to Question 11 show whether students think that the ammount of learning material available is optimal for learning, thus indirectly showing how they perceive their understanding. Most students over years answered that the ammount was optimal (as shown in Fig. 8). In total, 87% of students believe that the material that includes exercises for learning the basics of Prolog is optimal, 8% that it is too extensive and 5% that it is insufficient. When compared to grades in Table 2, even bigger difference can be noticed than with Questions 5 and 6. Although it must be taken into account that this is indirect indicator with smaller significance, the difference is too big to be dismissed.

Students were given team assignments from academic year 2017/2018, because survey from previous year showed that they think that it would help them to better learn Prolog. This is why results for Question 14 in Fig. 9. that explore how students perceive the role of team assignment in understanding Prolog are not given for academic year 2016/2017. Possible answers were yes and no, so they can not be compared in the same way with grades as previous questions. Additionally, givent that results show that on average little more than half of students - 60% - think that the team assignment helped them in understanding Prolog and that practical knowledge of logic programming directly contributed to the final grade with 46%, results from this question can not give valid information whether students perceive their understanding of the course content correctly.

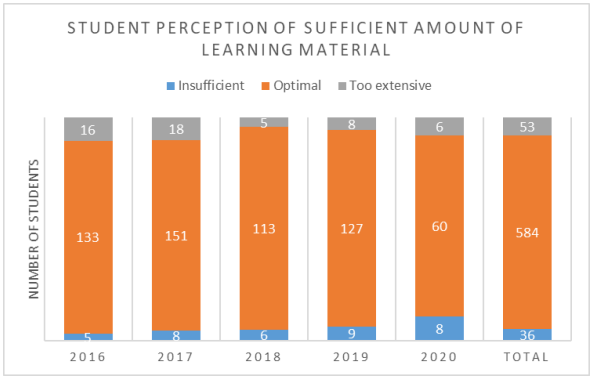


Figure 8. Student perception of sufficient amount of learning material (2016-2020; N = 673)

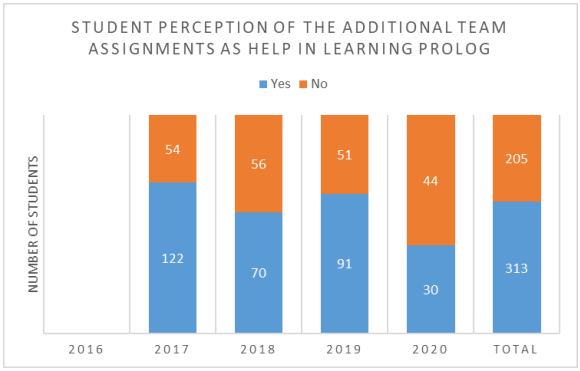


Figure 9. Student perception of the additional team assignments as help in learning Prolog (2016-2020; N = 518)

The paralel with the previous research [30] mentioned in subsection IV.A can be applied also here. Results of this research about problems in learning Prolog [30] show that students think that the language itself does not present a problem, although they consider declarative pardigm as one. Most of the students do not perceive practical application of Prolog, which decreasses motivation and consequently understanding, as well as final grade [30].

V. CONCLUSION

Following the research, the complexity of the details that need to be taken into account and the importance of the students' opinion can be seen. In the previous research, the contemporary importance of Prolog and the paradox of its unpopularity was pointed out [30]. Also, problems in learning its basics through theory and practice were analysed and the effort invested in motivating students, creating exciting materials, and challenges in teaching methodology and achieving results within the online environment forced by COVID-19 pandemic were described [30].

To more understand the problems students face when learning logic programming, in this research student perception of the understanding of the course content was analyzed and then compared with the grades at the end of the semester. Results of the comparison showed that students perceive their understanding of the course content higher than the grades can confirm. Only 14% of students perceived their understanding of the entire

course content as low, and most of them (87%) thought that the ammount of material is optimal, but 43% of them didn't pass the course after assesment during semester. In this case only final grades were taken into consideration, but many other elements of programming can be analysed and the difference in perception of understanding specific elements and real understanding can give valuable information, if observed throughout semester to enable adjustment of teaching pace and other teaching elements.

It is also necessary to refer to the shortcomings of the research. In the first observed year there was only one teaching assistant (later were two), which can possibly explain less negative grades in this year. In the first year there was also no team assignment, which can affect the results to some extent. In the last observed year teaching was conducted only online, resulting in significantly less completed surveys, again affecting the result to some extent. It is interesting that in this last year answers to questions were relatively similar, but it was the only year in which there were more negative than positive grades.

Future research will include analysis of other student engagements throughout semester and comparison with specific assesment results and assesment timing. Analysis of student perception of their understanding will be performed throughout semester and compared to results of assesements with the goal to optimize student learning processs and course performance.

ACKNOWLEDGMENT

This work has been fully supported by the Croatian Science Foundation under the project IP-2020-02-5071.

REFERENCES

[1] V. Lifschitz, "Foundations of logic programming," in Principles of knowledge representation, G. Brewka, Ed. Stanford: CSLI Publications, 1996, pp. 69-127.

[2] M. R. Genesereth, and M. L. Ginsberg, "Logic programming," Communications of the ACM, vol. 28, no. 9, pp. 933-941, September 1985.

[3] V. Aleksić, and M. Ivanović, "Introductory programming subject in European higher education," Informatics in Education, vol. 15, no. 2, pp. 163-182, October 2016.

[4] P. Körner et al. "Fifty years of Prolog and beyond," Theory and Practice of Logic Programming, vol. 22, no. 6, pp. 776-858, November 2022.

[5] A. Colmerauer, and P. Roussel, "The birth of Prolog," in History of programming languages II, T. J. Bergin and R. G. Gibson, Eds. Boston: Addison-Wesley Professional, 1996, pp. 331-367.

[6] P. Brna et al., "Prolog programming techniques," Instructional Science, vol. 20, no. 2-3, pp. 111–133, March 1991.

[7] I. Bratko, "Prolog Programming for Artificial Intelligence," 4th Edition, North York: Pearson Education, 2012.

[8] W. Ertel, "Introduction to artificial intelligence," Cham: Springer, 2018.

[9] F. Fei, J. Xie, and Z. B. Zhong, "Analysis on the application of intelligent robots based on Prolog language in primary and secondary education in the era of digital intelligence," in Proc. of 3rd International Conference on Education, Knowledge and Information Management (ICEKIM), pp. 305-309, January 2022.

[10] R. Buscaroli et al. "A Prolog application for reasoning on maths puzzles with diagrams," Journal of Experimental & Theoretical Artificial Intelligence, in press (published online 19 April 2022)

[11] A. Lally, and P. Fodor, "Natural Language Processing with Prolog in the IBM Watson System," The Association for Logic Programming (ALP) Newsletter, 9, 2011.

[12] D. Radošević, T. Orehovački, and A. Lovrenčić, "New approaches and tools in teaching programming," in Proc. of Central European Conference on Information and Intelligent Systems, pp. 49-57, September 2009.

[13] C. S. Cheah, "Factors contributing to the difficulties in teaching and learning of computer programming: A literature review," Contemporary Educational Technology, vol. 12, no.2, ep272, October 2020.

[14] F. Demir, "The effect of different usage of the educational programming language in programming education on the programming anxiety and achievement," Education and Information Technologies, vol. 27, no. 3, pp. 4171-4194, April 2022.

[15] T. Jenkins, „On the difficulty of learning to program," in Proc. of the 3rd Annual LTSN-ICS Conference, pp. 53-58, 2002.

[16] A. Pears et al., "A survey of literature on the teaching of introductory programming," ACM SIGCSE Bulletin, vol. 39, no. 4, pp. 204-223, December 2007.

[17] D. Callear, "Teaching Programming: Some Lessons from Prolog," in proc. 8th of Annual Conference on Teaching and Learning, 2000.

[18] S. A. Rebelsky et al., "Why I do declare! declarative programming in the undergraduate curriculum," in Proc. of SIGCSE technical symposium on Computer Science Education, pp. 398-399, February 2001.

[19] S. Yang, and M. Joy, "Approaches for Learning Prolog Programming," Innovation in Teaching and Learning in Information and Computer Sciences, vol. 6, no. 4, pp. 88-107, 2007.

[20] A. Stathaki et al., "i-Prolog: a web-based intelligent tutoring system for learning Prolog," in Interactivity, Game Creation, Design, Learning, and Innovation, A. L. Brooks, E. Brooks and N. Vidakis, Eds., Cham: Springer 2017, pp. 337-346.

[21] G. Siemens, "Learning analytics: The emergence of a discipline," American Behavioral Scientist, vol. 57, no. 10, pp. 1380-1400, October 2013.

[22] D. Clow, "An overview of learning analytics," Teaching in Higher Education, vol. 18, no. 6, pp. 683-695, 2013.

[23] O. Viberg, M. Hatakka, O. Bälter and A. Mavroudi. "The current landscape of learning analytics in higher education," Computers in human behavior, vol. 89, pp. 98-110, December 2018.

[24] M. Hernández-de-Menéndez et al., "Learning analytics: state of the art," International Journal on Interactive Design and Manufacturing, vol. 16, no.3 pp. 1209-1230, September 2022.

[25] U. Omer, et al., "Learning analytics in programming courses: Review and implications," Education and Information Technologies, in press (published online 16 February 2023)

[26] M. Berland, T. Martin, T. Benton, C. P. Smith, and D. Davis, "Using learning analytics to understand the learning pathways of novice programmers," Journal of the Learning Sciences, vol. 22, no. 4, pp. 564-599, October-December 2013.

[27] P. Blikstein, M. Worsley, C. Piech, M. Sahami, S. Cooper and D. Koller, "Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming," Journal of the Learning Sciences, vol. 23, no. 4, pp. 561-599, October-December 2014.

[28] D. Azcona, I-H. Hsiao and A. F. Smeaton, "Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints," User Modeling and User-Adapted Interaction, vol. 29, no. 4, pp. 759-788, September 2019.

[29] P. Ihanthola et al., "Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies," in Proc. of 2015 ITiCSE on Working Group Reports, pp. 41-63, 2015.

[30] V. Sekovanić and S. Lovrenčić, "Challenges in teaching logic programming," in Proc. of the 45th International Convention MIPRO 2022, pp. 658-662, May 2022.

Prototyp of an optical - electrical measuring system for the determination of the gravitational constant for teaching purposes

Ch. Ungermanns*, W. Werth *, A. Schöndorfer* Ch. Madritsch*
* Carinthia University of Applied Sciences (CUAS), Europastraße 4, 9524 Villach, Austria
c.ungermanns@fh-kaernten.at
w.werth@fh-kaernten.at
Alexander.Schoendorfer@edu.fh-kaernten.ac.at
c.madritsch@fh-kaernten.at

***Abstract*—Newton's gravitational constant plays an outstanding role in physics and is the most imprecisely determined natural constant from an experimental point of view. To determine this constant, one uses a set-up according to Cavendish, the so-called Cavendish experiment. The aim of the present work is to reconstruct this experiment and to implement a suitable measuring system with automated data acquisition and evaluation. This system is to be used in the physics lecture as a demonstrator and didactic support. Two tasks were addressed and solved:**

1.) mechanical construction of a torsion balance
2.) automated measurement data acquisition and evaluation

This paper deals with the second part. The result of this work is a complete characterization of the mechanical properties of the torsion balance and the construction of a functional, automated optical-electrical measuring and evaluation system for determining the gravitational constant γ.

Keywords—Cavendish experiment, torsion balance, gravitational constant, optical-electrical measuring system, teaching experiment

I. Introduction

Gravity is one of the most important natural phenomena in our universe and is one of the four basic forces in physics. It was first discovered by Sir Isaac Newton (1643 - 1727) in 1687 with a mathematical formula. It is known as Newton's law of gravitation and, together with the axioms of motion, forms the basis of Newtonian mechanics. Newton's law of gravitation describes in mathematical form the force relationship between two masses. Newton found out how the masses of bodies and their distance must be put in relation to each other so that the mutual force effect is correctly represented. The law contains a constant of proportionality which is neither dependent on the masses of the bodies nor on their distance. This constant is called the gravitational constant γ, and its determination is done experimentally. The value for γ is very small because the force between everyday things is present but not perceptible. Henry Cavendish first succeeded in determining the magnitude of this constant. This was realised in the so-called Cavendish experiment with the help of a torsion balance. The currently recognised value for the gravitational constant γ is [1]

$$\gamma = 6.67430(15) \cdot 10^{-11} \frac{\mathrm{m}^3}{\mathrm{kg} \cdot \mathrm{s}^2}.$$

Within the scope of a project a torsion balance was constructed and realised at the Carinthian University of Applied Sciences according to the idea of the Cavendish experiment. Usually, the measurement with such an apparatus (recording and evaluation of the measurement data) is done manually, which means that a person who logs the data must be present during the entire measurement run. In this paper, an automated set-up for the entire experimental procedure is presented.

II. Problem and Task Definition

The measurement of the gravitational constant requires a very careful experimental procedure, as the effect of gravity is extremely weak. Already the smallest disturbances make the measurement unsuitable. The risk of the experiment being disturbed by the presence of people or even by the experimenter is significantly greater than a measurement in an isolated room. The first tests with the torsion balance proofed that the results are strongly distorted due to external influences. Vibrations or air draughts caused by moving persons in the room have a negative effect on the experiment. In high-precision measurements of the gravitational constant, these influences are completely excluded by the experimental setup [2], which is not possible in this work. But these impacts should at least be kept as small as possible. Therefore, the design of an automated measurement system is favourable. Due to the setup, a simultaneous data evaluation can also be carried out. This concept should make it possible to leave the room after the start of the experiment, which results in a reduction of disturbing influences, since no one has to intervene in the measurement operation any more. Furthermore, it is expected that the results will be more reproducible and have a higher accuracy due to the automated data aquisition and evaluation.

III. Basic Setup of Cavendish Experiment

For a better understanding, this section briefly summarises the principle procedure of the Cavendish torsion balance [3]. It consists of two pairs of balls, where the balls of a respective pair attract each other due to gravity. The effective gravitational force is measured with the help of a so-called torsion balance. The basic layout of the system is shown in figure 1.

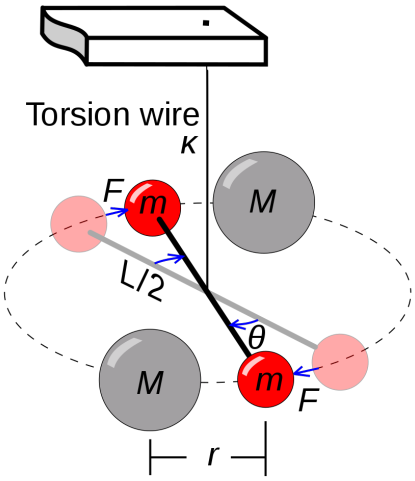


Fig. 1: Schematic diagram of a torsion balance for determining the gravitational constant [4]

At the beginning, the torsion balance is in an equilibrium position, this means that the torsional and the gravitational force mutually cancel out. By deflecting the heavy spherical masses M to a new position, the equilibrium between both force is disrupted and the movable small ball masses m, which are attached to the torsion wire, start to oscillate in a damped manner. After a certain time, the small masses m remain in a new equilibrium position. In this position, the effective gravitational force between the spherical masses and the torque of the torsion wire cancel each other out again. This approach and the period of the oscillations are used for the calculation of the gravitational constant. The task now is to detect the oscillations and the position of the small masses m as accurately as possible. A proper measuring module is needed for this purpose.

IV. Selection of used Components

To develop a suitable measurement system, different measurement principles were first considered and compared in order to decide which measurement method would be appropriate for the job. In addition to suitability, the financial budget also played a role in the selection of components. Furthermore there was a self-imposed requirement to recycle as many existing components from broken devices as possible. Therefore a contact image sensor from a scrapped flatbed scanner was used as the sensor for the measurement system. This device has the ability to detect a laser beam with positional accuracy. MATLAB software was provided to evaluate the measurement data, as it is very useful to process and display large amounts of data with this programme. The STM32f407 microcontroller from STMicroelectronics was applied to control the sensor, process the measurement data and supply power to the measurement circuit. The interface between the microcontroller and the sensor is managed by an SD card module. With this concept, the measurement data can be stored without the aid of a computer.

Figure 2 shows the main components of the measuring system. In particular, one can recognise the sensor removed from the flatbed scanner. It is perfectly suited for the locally resolved detection of the laser beam. In a first step, the data sheet of the sensor had to be found. This is necessary to determine the timings and the pin assignment. The knowledge is also required to control and read out the device with the microcontroller. In addition, the electrical circuit can be taken from the data sheet, which is important to supply the pins of the sensor with the correct voltages and currents. To be on the safe side, the sensor signals were checked using an oscilloscope.

Figure 3 illustrates an example of the output signal after a proper electrical setup. A recorded period of the sensor output data is shown. The first part of the signal (yellow circle in the figure) indicates that the following segment contains the sensor information. In order to better visualize the functionality of the sensor, a box was placed on the sensor surface and covered it partially (red circle in the figure). The position of the object on the sensor surface is detected, which can be clearly seen in the output signal by the reduced amplitude at the location of the placement.

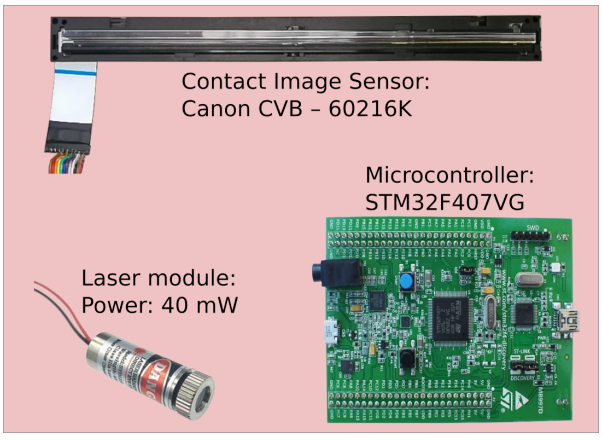


Fig. 2: Main components of the measuring system used

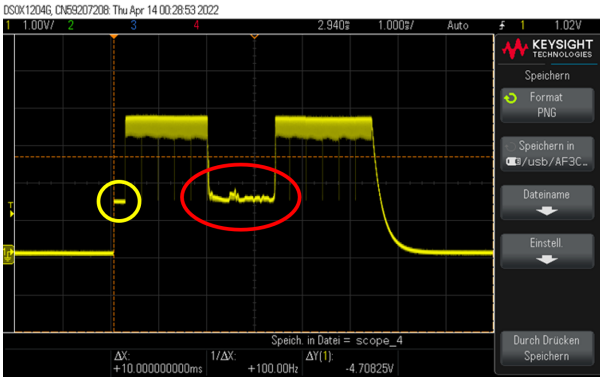


Fig. 3: Measured output signal of the partially covered image sensor demonstrating the functionality of the sensor

V. Concept Implementation

The measurement signal of the Cavendish experiment (deflections of the masses) is very small and must therefore be amplified considerably. The amplification is best achieved by reflecting a laser beam from a mirror attached to the torsion wire and projecting it onto a distant wall. The displacement of the laser beam can be detected with different methods, namely manually, with the help of a camera or directly with a contact image sensor. The latter method provides the best and most accurate measurement results and is also significantly cheaper than a high-resolution camera system. The measurement signal recorded by the sensor is processed further with the microcontroller and evaluated by using a MATLAB script.

VI. Results

In this work, the implementation of the measurement system and the determination of the inertia characteristics and the wire properties of the entire torsion balance have been realised. The final assembly and the main components of the measuring system are shown in figure 4. A detailed description of the construction is documented in bachelor thesis [5], which is published as part of the investigations. The central element is the torsion balance. As one can see, the laser is pointed to the mirror that is attached to the torsion wire and reflected onto the wall of the laboratory or the image sensor, depending on the experimental setup. The displacement of the reflected laser beam is tracked and logged either manually (note the measuring tape on the wall), by a camera system or by the image sensor. As already mentioned, the highest precision is achieved with the image sensor.

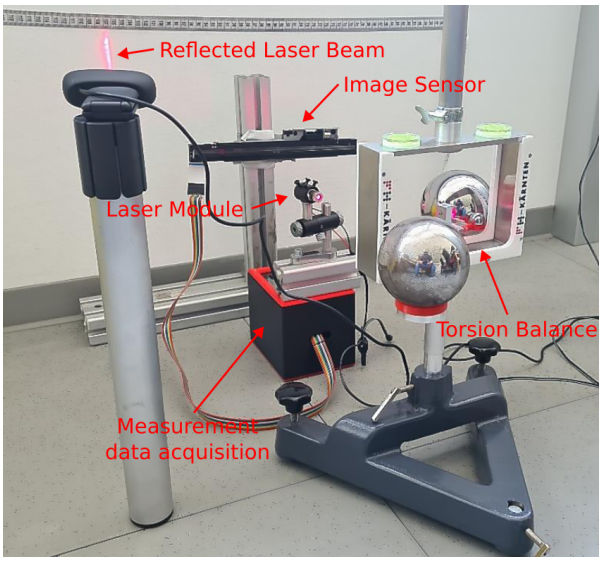


Fig. 4: Torsion balance and automated measurement system

Now the characterization of the torsion balance is carried out. This step is absolutely necessary to ensure an independent determination of the gravitational constant. Figure 5 shows an example of the measured oscillations, which were recorded with the automated measuring system.

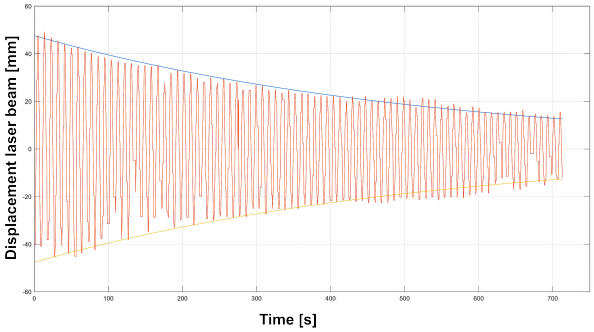


Fig. 5: Measured oscillations - tungsten wire thickness 0.2 mm

The evaluation of these oscillations for different test setups provide the parameters listed in TABLE I and TABLE II respectively.

One can clearly see the exponential decrease of the amplitude, which is typical for a damped oscillation. From the temporal intervals of the maxima, the average period durations were determined. With the help of these parameters, the geometry of the torsion balance and the masses of the moving spheres, the properties of the wire can be calculated as follows (e.g. [6] or [7]):

The moment of inertia of a sphere around its center of mass is given by the formula

$$J_{\text{sphere}} = \frac{2}{5} m_{\text{sphere}} r_{\text{sphere}}^2 \qquad (1)$$

In this setup the mass of one sphere is $m_{\text{sphere}} = 32$ g,

which is a typical size for mobile torsion balances[1]. The balls are positioned on the torsion balance at a distance of d offset from the wire (axis of rotation). Therefore, according to the parallel axis theorem the modified moment of inertia of the sphere can be calculated as

$$J'_{\text{sphere}} = J_{\text{sphere}} + m_{\text{sphere}} d^2 \tag{2}$$

Since the system consists of two spheres, this expression must be multiplied by a factor 2. In this way one obtains J_{w} and J_{s}. The following differential equation holds for a torsional oscillation:

$$J\ddot{\varphi} = -D\varphi \tag{3}$$

From this equation follows the relationship between moment of inertia, period of oscillations and directing moment. It results in

$$J_{\text{xx}} = \frac{1}{4\pi^2} T_{\text{xx}}^2 D, \tag{4}$$

where "xx" has to be chosen according to the experimental setup. With this approach the unknowns D and J_0 can be determined experimentally from the measured quantities T_{ow} and T_{os}. Note that the following relationship holds

$$J_{\text{ow}} = J_0 + J_{\text{w}} \quad \text{and} \tag{5}$$

$$J_{\text{os}} = J_0 + J_{\text{s}} \tag{6}$$

Abbr.	Meaning
J_{w}	Moment of inertia tungsten balls (shifted by 5 cm) without arm
J_{s}	Moment of inertia steel balls (shifted by 5 cm) without arm
J_{ow}	Moment of inertia arm + tungsten balls
J_{os}	Moment of inertia arm + steel balls
T_{ow}	Period duration arm + tungsten balls
T_{os}	Period duration arm + steel balls
J_{o}	Moment of inertia arm without balls
T_{o}	Calculated period arm without balls
D	Directing moment (material constant)
G	Shear modulus

TABLE I: Abbreviations of evaluated parameters

Now the constants J_{o} and D are known. These parameters were used to make a prediction for the period T_{o} in the case that the arm oscillates without balls. This period T_{o} was then used for comparison with the actual measured period $T_{\text{o, measured}}$.

A summary of the calculated results can be seen in the following TABLE II

[1]For the mass of the heavy balls, 1.5 kg has turned out to be a good choice.

Experimental Characterisation of the System
$J_{\text{o}} = 7.275 \cdot 10^{-6}$ kg m^2
$J_{\text{ow}} = 1.689 \cdot 10^{-4}$ kg m^2
$D = 8.779 \cdot 10^{-5}$ kg/m^2
$J_{\text{os}} = 5.005 \cdot 10^{-5}$ kg m^2
$T_o = 1.809$ s
$G = 167.684$ GPa
$T_{\text{o,measured}} = 1.885$ s
$G_{\text{literature}} = 158$ GPa
$\Delta T_o = 4.22\%$
$\Delta G = 6.13\%$

TABLE II: Results of wire and system properties

The percentage difference between the measured period $T_{\text{o,measured}}$ and predicted period T_{o} is 4.22% (1.885 s to 1.809 s). The shear modulus G was also determined [8] using the first setup and compared with the literature value. The deviation of the literature value of the shear modulus G from the calculated value is 6.12% (158 GPa to 167.684 GPa). The small differences between the results can be attributed to uncertainties in the wire measurement (thickness and length), for example.

VII. Summary and Outlook

In this contribution, a torsion balance was successfully equipped with an optical-electrical measurement and evaluation system. Essential for the determination of γ is the sole effect of the gravitational force. Unfortunately, the available spheres have strong magnetic properties (proven with a magnetic field meter) and are therefore unusable for the actual experiment. In the next step, new spheres (made of a non-magnetisable material, e.g. lead) must therefore be installed. Thus, nothing stands in the way of determining γ and using the experiment in teaching.

VIII. Acknowledgement

A special thanks belongs to Mr. Peter Lippitz, who realised the mechanical construction of the torsion balance within the scope of his bachelor thesis.

References

[1] Codata value: Newtonian constant of gravitation. [Online]. Available: https://physics.nist.gov/cgi-bin/cuu/Value?bg

[2] J. H. Gundlach and S. M. Merkowitz, "Measurement of Newton's constant using a torsion balance with angular acceleration feedback," *Phys. Rev. Lett.*, vol. 85, no. 14, pp. 2869–2872, 2000.

[3] M. SHAMOS, *Great Experiments in Physics. Edited by M.H. Shamos, Etc.* New York, 1960. [Online]. Available: https://books.google.at/books?id=uU4OMwEACAAJ

[4] Wikipedia contributors. Cavendish experiment. [Online]. Available: https://en.wikipedia.org/wiki/Cavendish_experiment

[5] P. Lippitz, *Konstruktion einer teilbaren Gravitationswaage mit Messsytem.* Carinthia, Austria: Carinthia University of Applied Sciences, 2022.

[6] D. C. Giancoli, *Physics: Principles with applications*, 7th ed. Upper Saddle River, NJ: Pearson, 2013.

[7] H. J. Paus, *Physik in Experimenten und Beispielen*, 3rd ed. Munich, Germany: Hanser, 2007.

[8] H. J. Eichler, H.-D. Kronfeldt, and J. Sahm, *Das neue Physikalische Grundpraktikum*, 3rd ed. Wiesbaden, Germany: Springer, 2016.

Computers in Education

H. Jaakkola*, J. Henno ** and J. Mäkelä***
* Tampere University, Pori, Finland
** Tallinn University of Technology, Tallinn, Estonia
*** University of Lapland, Rovaniemi, Finland
hannu.jaakkola@iki.fi; jaak.henno@taltech.ee; jumakela10@gmail.com

***Abstract* - Computers, or in wider scope, information and communication technology (ICT), have had revolutionary impact on education. ICT is both merged to the education process and is supporting it. Computers and information technology are typical enabling technologies – not in the focus itself, but when available, these are adopted in use. Performance of computers has grown exponentially: According to Moore's law processing power and RAM capacity is doubled in 1.5 years. The similar growth rate relates to mass memories and data transmission speed. During the era of computing (from the middle 1940s) information technology has gradually transferred to new areas of use. In the beginning, the computers themselves were in focus. In the 1970s microelectronics has been the key to changes: first personal computers, then networking and cloud technologies, further embedded intelligence of devices, distributed computing, complex software, wide access to heterogeneous data sources etc. have brought the computer to the growing number of applications, also in the education sector. The purpose of this paper is to have a look at the concept "computers in education". It has been the title of the conference track in Mipro over decades. It is widely handled in literature, journals, and other publications. These provide forum for experience transfer in the area, as well as encourage researchers to study the topic from the scientific direction. The paper is partially based on the experiences of the authors and aimed to clarify the term from a variety point of view.**

Keywords - computers; education; computing trends; teacher training.

I. INTRODUCTION

The role of *computers in education* has been in focus in last decades. The progress relates both to computers and the spread of information and communication technology (ICT) in the education sector. Computers and related technologies are seen as enablers of changes in this progress. The spread of ICT in education has not been an isolated phenomenon but followed the general changes in the role of ICT in society. In this progress, the education sector has not been the frontier field, more in the role of follower which uses the opportunities than the developer of new innovative uses. Why to be a follower? The main reason might be the lack of resources, especially funding (at all levels), which is mainly coming from limited sources, mainly public but in some amount also private. The education sector is neither in focus of funding for research and development if compared with other sectors of society and economy. Investments in the use of new technologies are targeted to such sectors whose productivity *is seen* as higher. There seems to be also some force of inertia in the education sector, which makes the progress slower than it could be in optimal situation – conservatism in attitudes, lack of required skills to adopt ICT based solutions and resistance to changes.

The matter can also be looked at from the opposite direction: to what extent are computers and ICT needed in education. In some cases, it looks that real needs have not been the driving force, but ICT itself has become such. In recently published studies [1; 2] Mertala et al. analysed the characteristics of 200 highly cited research articles in the field of educational technology. They found the phenomenon called "*positive bias*": the studies emphasized the benefits and positive findings of using ICT in education but ignored the negative aspects. The studies showed that the results of otherwise well-achieving students improved when using the researched technology but ignored the fact that the results of low-achieving students simultaneously deteriorated. These findings indicate a strong desire to find positive effects, even in the case that the results are not so promising. Professor Mertala is distinguished in the field of critical educational technology (Ed-Tech) research. He has provided empirical evidence for the qualities of the phenomenon known as *Ed-Tech speak*, having focus especially in research articles reporting experiences in the use of IPad in education.

In long run, outside the research environment, the adoption of new technologies in practice are based on rationale decision-making and long-term experiences in real environment of use. The desire and the opportunities meet, and the use of ICT finds appropriate solutions. A lot of studies also confirm this finding. Research and tests of new kinds of solutions are needed to get the best ones to find their way to the practice. However, the focus of this paper is not on the discussion about good and bad solutions in the use of ICT in education, but to have look at the *concept "computers in education"* from different perspectives.

The title of the paper comes from the title of the track CE in Mipro conference. The authors have participated in it already since the 1990s and seen presentations having a variety of approaches to the topic area and temporal changes in trends. Quite many of the papers in the conferences report the results of tests and trials, based on statistical analysis, in real teaching / learning situations, some propose new kinds of solutions and the use of new technologies, some represent the wider approach to the

topic and show the flexibility of the conference scope interpretation.

The "*research problem*" of the paper is to find (at least a partial) answer to the question *what means the concept "computers in education*". We approach the solution by looking at the temporal changes in the role of computers and the use of ICT in general (Section II); we see steps in this progress as *enablers*, which provide opportunities for the education sector. How the opportunities were taken in the beneficial use, what means "*computers in education*" in different periods, is discussed in Section III. To get the more objective view, we have a *light-weight look* at the literature in Section IV. The discussion is concluded in Section V.

Our approach is based on the findings in literature and on our (decades long) experience based views, which in this article we organize as *a kind of framework*. Emphasis is given more to the *changes and trends* than in listing the applications. The topic is handled also in some of our earlier papers: papers [3; 9] handle the machine and human learning and intelligence, papers [4; 5] the role of artificial intelligence in education and papers [6; 7; 8] technological changes (the trends in enabling technologies). The topic is handled mainly from the technical point of view, not pedagogical.

II. ABOUT THE ENABLING TECHNOLOGIES

A. *The Embryonic Era and the Centralized Use of Computers*

The era of computers, based on the current concept, started in 1946 when ENIAC was released. It was developed to provide a platform for "complex" scientific calculations. The spread of computer uses to other sectors of society was happening little by little in the 1950s although only the 1960s can be called the decade of the expansion. Even then, computers were rare and mainly used by large companies and organizations. These decades could be called the "*embryonic era of computers*". The use of computers was *centralized*, and users were the computing experts. They provided their services to allow the development and use of the applications. This was called "*closed shop*" use of computers.

B. *Towards Distributed Use f Computers*

The widespread use of computers in society dates to the 1970s. There were two *key enablers*: improvements in computer technology (transistors, mass memories) and new type of operating systems (*time-sharing*). The transition to mass production lowered the prices of and made it possible for smaller (mid-size) organizations to acquire a computer. The time-sharing operating system, in turn, made *direct computer use* via terminals available to the users from different application areas. In a way it was question on a *radical change*, because the first time the *end-users* were able to operate computers themselves. The computer architectures became simpler and did not expect any more special environments of location. The scale of computers became wider – all from large to mini and micro computers were available in the market. Transfer from *the closed shop to the open shop* use had started.

Simultaneously, development in programming languages was active: high-level programming languages FORTRAN, COBOL, Algol and C were published and application development was active. Software supporting the use of computers in different application sectors appeared and the use of computers became more and more common in wide sectors of society and application areas. Even the software development skills became possible for application area experts, because of the easier-to-use programming languages and environments. Even at that time, IT skills were not widespread, but rather belonged to IT professionals and a group of advanced pioneer users from a variety application areas.

C. *Unmanaged Distribution – Introduction of PC*

If the change from closed to open use was called by the term *radical*, then the introduction of PC to the market represents a *revolution*. The first PCs (in wider use) appeared in the *early 1980s*. The key enabler was a new technology – *microprocessor* and in more common *VLSI* technology. Fast growing performance [1], lowering computer prices brought computers in the new sectors of society, including schools, and eventually in homes as well. This progress, in turn, created the conditions for a *promising market for applications*

The fast (unmanaged) spread of PCs resulted in the dramatical change in the use of computing resources. Before *PC era*, computers were centrally managed by the experts. Now, every PC was an independent device managed by the user(s), which mostly had no deep education in computing. The *era of centralized computing* changed to *the era of unmanaged distributed computing*. The term "unmanaged" points out the fact that in organization level computing resources (computers, data) were not manageable and wild out-of-strategy use became possible. This was the beginning to a variety of *security risks*, too.

D. *Centralized Distributed Computing - Networking*

The Networking of computers partially solved the problem of management of computing resources. It provided means for *collaborative and distributed computing*, in which computers were able to share and distribute data between computing units, also the parallel use of separate units became possible. The technical invention of the computer network (Ethernet; competing solutions existed) has roots in the 1970s, but its adoption in practice has started one, almost two decades later *in the 1990s*. Along with networks, layered architecture (two-tier, n-tier, client-server) also became common. These made possible to provide *centrally managed services* to the distributed clients in the network.

Further, progress in networking was seen when wireless (Wi-Fi) networks were introduced in the early 2000s. As in Ethernet, the invention itself has its roots in the 1970s, but transfer to practice took a couple of

[1] First published in "Gordon E. Moore: Cramming more components onto integrated circuits. Electronics, Volume 38, Number 8, April 19, 1965". Simplified interpretation of the law: processing power and memory capacity (in fixed price) of computers doubles every 1.5 years because of the progress in VLSI technology.

decades: Wi-Fi networks started to become *common in the 2000-2010s*.

E. Cloud

Cloud technologies can be seen as an innovative integration of several components of ICT. More than technology, it is a *computing paradigm*, which connects users to their needs in computing ecosystem. Largely, it is question on the access to the services and data resources. Cloud technology (at the conceptual level) has its roots in early network solutions (ARPANET) and in the aim of implementing the concept of *virtualization* (a virtual machine that acts like a real machine from the user point of view). Despite having first references to cloud computing very early in the 1960s, the practical solutions are scheduled in the first decade of 2000 and spread in wide use in the 2010s.

F. VLSI Technology, Embedded Systems, Digital Convergence

Thanks to the rapid changes in ICT infrastructure belongs to *VLSI technology*. First, it has provided the new implementation base for computers. Secondly it has provided an opportunity to *embed* "computer kind of intelligence" in different devices (*Robotics, Internet of Things, IoT*). *Digital convergence* is a phenomenon, which has a huge impact on today's society: (tele)communication, computing and media are integrated and available for use as multi-purpose devices: mobile phone and television are good examples of this progress. The background of this development is also haunted by VLSI technology and the "convergence-based devices" are typical examples of "computer inside" devices.

G. The Role of the Enabling Technologies Summarized

The discussion above has focus in the progress of computing from *technical perspective*. The aim of this paper, to have a look at the concept "*computers in education"*, represents one of the application areas of ICT. In this context it is good to bring up a saying of Edsger Dijkstra[2], one of the pioneers in computing: "*Computer science is no more about computers than astronomy is about telescopes*". This describes the situation: the aim of computer science is to support the progress towards the easy and beneficial use of computers, which has been tightly bounded to the progress steps in related technologies as *enabler*.

ICT has been adopted by the education sector within its resources, needs and the expected skills of users, following the technical progress providing opportunities for the adoption. Because of that the concept "Computers in education" has *different interpretations in different periods*. These aspects are handled in Section III.

[2] Quoted from https://www.brainyquote.com/quotes/edsger_dijkstra_201168

III. HISTORICAL PERSPECTIVE TO COMPUTERS IN EDUCATION

A. Step 0: Training of Professionals

On the early 1950s computers were rare, mainly used for research purposes. Later (the late 1950s, the 1960s) user base extended to government organizations and bigger companies. Education related to computers was (professional) *training oriented* and mainly on the responsibility of manufacturers' organizations. On this period "*computers in education*" was professional training (not education) to use computers. Some research-oriented workshops were also organized by universities and research institutes, mostly in collaboration with manufacturers. A similar cooperative relationship also concerned the development of computers and their software.

B. Step 1: Computer Science Education

Early initiatives to teach "*computer science*" date back to the 1940s and 1950s. The study [10] reports about the first academic credit courses in computing in 1946 in Columbia University, based on the collaboration between the University and IBM. In 1947 Harvard University (Howard Aiken) established a computing laboratory and a degree program in computing there. This was bound to the development of Harvard MARK I computer. In 1953 Computer Laboratory at Cambridge University UK (Maurice Wilkes) also established a computer degree program in 1953. However, these activities are either related to teaching the use of data processing equipment at the time (Columbia) or are part of the history of computer development (Harvard, Cambridge). They therefore do not represent the actual teaching of computer science, the creation of which they have had a great influence, however.

In the 1960s, interest (and need) in the general education of computing began to arise. As a result, university level education started. In the USA both Purdue University and University of North Carolina at Chapel Hill established *computer science department* and related curricula. Very same time the first computer science curriculum started *in Finland by the University of Tampere* in 1965. This was the first initiative in Europe, too. The concept "*computers in education"* would be interpreted as being equal to *computer science education* and could be called the *first step on "computers in education"*. Focus was set to provide skills in the professional use and development of computers and their environments. Manufacturers' training courses to use computers played still an important role, too.

C. Step 2: Spread of Computer Use to Application Areas – Closed Shop Based Services

To enable computer science studies, universities acquired computers, the use of which gradually *spread in other disciplines* as well: science, statistics and solving a variety of research problems. Typical of this era was that all operations needed the support of computer specialists in "*closed shop*" manner and the tasks were executed in the *batch processing* mode without interaction with the end-user. This was the *second (historical) step*.

"Computers in education" was to transfer of computing skills to a variety of application areas. In practice this means *education to consider the opportunities computers (of that time) provide* in these disciplines, mainly in the form of complex mathematical analysis and calculations.

D. Step 3: Open Shop - Distributed Use of (Centralized) Computers

In the 1970s, the interactive use of computers opened the use of computers for the wider group of users. The use of computers spread widely in universities and computers were also acquired to lower-level institutes, however, not yet in schools. Terminals (dummy input / output devices) connected to the central computer allowed the direct *interactive use* (open shop) of computers even for end-users from the application areas. This led to an increased need to include *skills in the use a computer* as part of curricula in different disciplines. It included not only teaching the *use of applications* in one's own research field, but also *teaching programming* in some cases. "*Computers in education"* in that time was related to the ability to *independently utilize the opportunities* offered by the computer in one's own field of research and the ability to *independently develop his / her own computer applications* based on programming skills.

E. Step 4: The PC Revolution

The appearance of *personal computers* (PC) in the early 1980s changed the computer ecosystem dramatically. In the past, computers were special instruments, the use of which was possible only for a few and in limited environments. Now the computer was in principle available for everybody, and computers spread widely everywhere in the society, even *to the schools*. Simultaneously, the development of applications accelerated, and these were available in wide scope of application areas for low price. In the beginning the role of PC in education focused, as traditionally, to provide skills to use it, but with the growing range of applications, they became an integral part of education in *a variety of ways*, and it was also seen as a *pedagogical object*.

The first steps in using computers (PC) in schools were taken in the second half of the 1980s. However, computers in schools were still rare, and we had to wait until the 1990s to see the more *systematic and organized progress* in this area. Nationwide strategies were developed in several countries: Finland [11], Estonia [12]. These "strategies" provided the basement for systematic development to use computers as a part of education. The BSc thesis [13] describes the transfer of ICT to Finnish elementary schools. First steps were based on the volunteer work of teachers in mathematics and sciences by organizing "outside working hours" activities as a study circle. Resistance was noticeable among teachers and their continuing education was not well organized, almost not available. The national syllabus for elementary schools from 1985 had some references to include ICT skills in education as a goal to increase interest in technology, embedded in some topic areas, but not as a study subject itself.

The early era of PCs (the 1980s) blew up the use of computers in education, but only in reasonable limited scope. Partially it was question on lack of (computer) resources and partially on missing skills among teachers. "*Computers in education"* in this era can be described as *seeking opportunities and adapting to a new culture* where the computer plays a growing role.

The improvement of PC technology increased the possibilities of using it in new ways and contexts. Growth of processing power and memory technology (Moore's law) including new mass memories (capacity to store data), transfer from desktop to laptop (portability), graphical user interfaces (usefulness, easy-to-use) diversified the possibilities of using computers in education and *diversified* the interpretation of the concept "*computers in education*".

F. Step 5: Transfer to the Networked World

First steps towards the expansion of computers in education were taken by using computers for a variety and growing number of purposes as an independent unit. Networks (wired from the 1990s, Wi-Fi from the 2000s onwards) provided an enormous number of new opportunities, which were further complemented by the transition to the large-scale use of mobile devices (smart phone, tablet). Additional dimension was provided by the fast-spreading use of cloud-based technologies and services. A large amount of data (knowledge) and services became easily accessible and easy to use in education as well. Today we live in an information society, which is strongly based on the use of information technology. Computers and mobile devices are widely used in homes, workplaces, and schools. Ownership of the computer (or a kind of device) is no more important because students are using their own devices at school (BYOD phenomenon), computers are used to make remote studies possible, as well as virtual classes have become a part of education culture. The use of ICT has become a *part of everyday life*, and so has the use of computers in education. "*Computers in education" has gradually become a part of information society technology and services, more or less equal to ICT in other sectors of society*. This has happened step by step in parallel with other information society developments. In the initial phase, the development focused on the utilization of network solutions, gradually on the utilization of new applications and spread in new application areas, and currently the focus is effective *knowledge creation* using the network capabilities and available data sources (cloud, big data, open data, data pipelines, structured and non-structured sources). Further, the number of applicable technologies and solutions are growing continuously.

G. Historical Perspective Summarized

The aim of this section is to have a look into the role of computers in education in somewhat longer time perspective than just today. The steps are triggered by *technological changes as enablers.*

In addition to the direct use of computers (and computer kind devices) ICT provides wide variety of opportunities for education. ICT based systems in the form of embedded VLSI technology (Lego Education), cheap microcomputer based (Arduino, Rasberry) construction kits, IoT devices, and robotics provide new

opportunities for the use of computers in education; in principle the wide spectrum of current technology is valid this purpose. With the development, the interpretation of the concept "*computers in education*" also changed over time and finally ended up covering the use of information technology in its various forms, adapted to the needs of education.

There seems to be no rational reason to specify exactly the coverage of the concept "computers in education" by listing topics related to it. Nowadays pretty much everything is up to the imagination, both from student and teacher side. From technology and service point of view almost everything is possible.

However, it is worth distinguishing *innovative ad hoc use* from *serious strategic use*. It is always worth making trials to get new experiences, but the skeleton of the use of computers in education must come from the *pre-planned strategic approach*. Both approaches are reported by various publications. Section IV looks at the topic from this point of view.

IV. PUBLICATIONS BASED APPROACH

A. Computers in Education – About the Publications

To find a common view to the topic we conducted a light review of the publications in the area. Not systematic, all coverage, but just to get an overview about the texts available. *Google Scholar* was used because we see it as the source of books and articles having the widest coverage, even outside scientific importance. The search term "*computers in education*" was used. The result covered close to *six Million publications*; even in the limited search covering the 2020s over 0.5 Million publications were listed. last year (2022) over 16,500. We did a rough analysis of over one hundred top-ranked findings. The following topic areas were handled:

- *education technologies and organizing the studies*: teaching and learning platforms, the role of technology, technology-based education methodology, e-learning, mobile learning, game based learning, social media, the use of tablets, online examinations, virtual classroom, automatic feedback, MOOC;
- *learning and teaching*: motivation, investigating the impacts, comparing traditional and computer-based learning, automatic assessment, impact on learning and satisfaction, blended learning, flipped learning;
- *abilities*: the readiness of teachers and students in the use of new technologies;
- *cognition:* cognitive view in the use of computers, modern knowledge, computational thinking;
- *social aspects*: COVID-19, equality;
- *technology*: augmented reality, virtual reality, AI, cloud, chatbots, virtual assistants, IoT, robots, the use of open resources.

The spectrum is wide and covers *both soft and hard topics*. From the point of view of this paper, we found a couple of interesting articles handling the history (long term changes, development) in the area. The papers [14; 15; 16; 17] complement and were also used as a source of ideas in, our history perspective in Section III. An interesting article [18] was found in Scientific American from 1966: it reports about the computer supported instruction trial and learning results in the very early era of computers. Mainframe computer was used with a simple terminal as an interaction tool. Promising learning results results in mathematics teaching were reported. One important topic does not seem to be widely handled: *preparing teachers using ICT in education*. Organized studies are not widely available, much is based on the volunteer work of teachers. Especially preparing *support for the older generation of teachers* is missing although some studies are already now provided as a part of teacher studies.

The shortcomings of continuing education were also noted in the recently published research report [19] handling the impact of digitization on learning in middle schools in Finland. According to it, digital technology is used in the limited scale and the use is rather one-sided. Digital technology use and learning outcomes seem to have negative correlation (as indicated in [1; 2]), but in general it seems not to weaken the learning results. Digitization enables the differentiation of learning and assessment, which increases the motivation and the engagement in learning situations.

B. Computers in Education – The ACM Approach

Association for Computing Machinery (ACM) has published the classification of computing related topics. Along the progress of computing the classification has gone through seven revisions. The first version was published in 1964 and revised versions appeared in 1982, 1983, 1987, 1991, 1998 and the current version is from 2012 [20]. It is implemented as an interactive WWW artefact which represents the classification of different areas of computing as a tree structure.

In the classification the concept "*computers in education*" locates in the branch CCS -> Applied computing -> Education. The topics listed cover:

- Digital libraries and archives,
- Computer-assisted instruction,
- Interactive learning environments,
- Collaborative learning,
- Learning management systems,
- Distance learning,
- E-learning,
- Computer-managed instruction.

The subclass that complements this locates in another branch of the structure: CCS -> Social and Professional topics -> Professional topics -> *Computing education*. This covers the topics of professional education in the area.

The earlier version of the classification from 1998 [21] lists the following topics in the branch Computing Milieux -> Computers and Education -> Computer Uses in Education: Collaborative learning, Computer-assisted instruction (CAI), Computer-managed instruction (CMI) and Distance learning. Difference between two versions is remarkable and points out the *dynamics in the area*. However, comparing these categories with the findings of the literature review above, topic listing is reasonable inadequate and generalized.

C. Computers in Education – The AI Approach

One of the topics under discussion today is the use of artificial intelligence (AI) in problem solving. A lot of publicity is given to the company OpenAI. The company reports having a mission "*to ensure that artificial general intelligence benefits all of the humanity*". To fill the mission it has brought a variety of tools to a public use; one of these is ChatGPT, the question-answer system based on deep learning having capability to multilingual interactive discussion with the user.

We made a trial with the ChatGPT tool for seeing, how the concept "*computers in education*" is seen by it. The discussion below was made in [22] in late January 2023.

What means computers in education: "*Computers in education" refers to the use of computer technology in the classroom or for educational purposes. This can include using computers for teaching and learning, for research and data analysis, for administrative tasks, and for communication and collaboration among students and teachers*".

Would it be someway more wide interpretation of the concept: "*A more comprehensive interpretation of "computers in education" would include not only the use of computer hardware and software in the classroom, but also the integration of technology into the educational curriculum and pedagogy. This could include the use of online resources, educational apps, and digital learning tools to enhance the learning experience and support student engagement and progress. Additionally, "computers in education" can also encompass the use of technology for professional development, such as online teacher training and the use of technology to improve administrative tasks such as attendance tracking and student data management*".

The discussion was continued by asking the **current trends in the area**. The following topics were listed: "*blended learning, personalized learning, gamification, online assessments, virtual and augmented reality, artificial intelligence-powered education, remote/distance learning"*.

The topics listed above are worth comparing with the findings in earlier sections of this paper and the list of the topics of Mipro CE[3]; the match is quite good.

We also asked about **relevant publication forums** in the area. We got the list of *seven journals*. In the question about conferences, also *seven conferences* were listed. To assess the *quality of the reply* the forums were compared with the ranks (scientific level) given by the Federation of Finnish Learned Societies[4]. This classification is used in ranking publication forums for Finnish universities to guide them towards high-impact publishing. The ranking classes cover "not-known", 0 (no value), 1 (basic), 2 (leading), 3 (top). In listed journals *none reached level 3, two were in level 2, 4 in level 1* and *one was not listed*. In the listed conferences, *none was ranked*. The quality of the answer in journals was acceptable but in the conference category the reply was inferior. **About MIPRO conference** ChatGPT says: "*it is a well-established and respected conference in the field of information and communication technology*". As a reference, Mipro belongs to ranking class 1, as most conferences ranked by the system.

[3] http://www.mipro.hr/MIPRO2023.CE/ELink.aspx

[4] https://www.tsv.fi/julkaisufoorumi/haku.php

V. CONCLUSIONS

The aim of this paper is to clarify the concept "*computers in education*". We selected the approach to look for its interpretation in *history perspective*. *Enabling technologies* were selected to explain the changes. The key technical changes were discussed (Section II) and the improvement steps in the use of computers in education (Section III) were mirrored against the technical changes.

TABLE I. STEPS IN THE USE OF COMPUTERS IN EDUCATION

Step	Enabler	Period
0: Training oriented	Early computers	1950s -
1: Computer science education	Computers to Universities	1960s -
2: Computers to application areas, closed shop	Application availability, in Universities.	Late 1960s -
3: Independent utilization in applications, open shop	Operating systems (time-sharing), programming skills. Spread of computers.	1970s -
4: Seeking for new opportunities, widening use	PC	1980s -
5: Merging gradually to Information Society: 5a. Network based and mobile use. 5b. Knowledge based use.	Networking. Cloud. Diversification of devices. Access to wide knowledge resources.	1990s-

Table I summarizes the analysis and the discussion in Sections II-III: Six steps triggered by enabling technologies adjusted to the timeline. Section IV indicates the topics under discussion, based on publications. Mipro CE conference seems to be on the frontier.

The use of *computers in education* follows the general trend in the use of ICT. The focus has shifted from *technical topics* to the ability to find and *utilize accessible knowledge*. The concept "*computer*" is diversified and the use of it is changed in time. The computer has become an extension of thinking, also partially the replacement of it.

In the future (already today) artificial intelligence (AI) based applications will play an important role. The authors have handled this topic in their pearlier apers [3; 4]. In this paper, we provided an example about using one of the AI applications, ChatGPT, in finding answer to our research problem. One of the hot topics in Finnish universities is the acceptability of the use of such tools in studies. The Intranet of Tampere University informs: "*Tampere University welcomes the use of AI-based applications and language models (eg ChatGPT) in research and teaching"*. In the same text challenges are listed, *ethical issues* are handled, and problems related to "*plagiation-kind*" text production are pointed out. In the thesis and reports the use of ChatGPT is approved within the defined rules. Is this the beginning of one *new era* in *computers in education*? In addition to learning, ability to use intelligent tools in beneficial way, having ability to critical review will be in focus.

REFERENCES

[1] P. Mertala, E. Moens, and M. Teräs, "Highly cited educational technology journal articles: a descriptive and critical analysis," Learning, Media and Technology, pp. 1-14, 2022. https://doi.org/10.1080/17439884.2022.2141253.

[2] P. Mertala, "'It is important at this point to make clear that this study is not "anti-iPad"': Ed-Tech speak around iPads in educational technology research," Learning, Media and Technology, vol. 46, no. 2, pp. 230-242, 2021/04/03 2021. https://doi.org/10.1080/17439884.2021.1868501.

[3] J. Henno, H. Jaakkola, and J. Mäkelä, "Machine learning vs human learning," in MIPRO 2021 - 44th International Convention, September 27-October 1, 2021, Opatija, Croatia, Proceedings, K. Skala, Ed. Rijeka, Croatia: Croatian Society for Information, Communication and Electronic Technology – MIPRO and IEEE, 2021, pp. 594-600.

[4] H. Jaakkola, J. Henno, A. Lahti, J.-P. Järvinen, and J. Mäkelä, "Artificial Intelligence and Education," in MIPRO 2020 - 43rd International Convention, September 28, 2020 – October 2, 2020, Opatija, Croatia, Proceedings, K. Skala, Ed. Rijeka, Croatia: Croatian Society for Information, Communication and Electronic Technology – MIPRO and IEEE, 2020, pp. 633-640.

[5] H. Jaakkola, J. Henno, J. Mäkelä, and B. Thalheim, "Artificial intelligence yesterday, today and tomorrow," in 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019 - Proceedings, K. Skala, Ed. Opatia May 2019: IEEE and MIPRO, 2019, pp. 985-992.

[6] H. Jaakkola, J. Henno, J. Mäkelä, and B. Thalheim, "Today is the Future of Yesterday, What is the Future of Today?," in MIPRO 2017 - Proceedings of the 40th Jubilee International Convention. May 22-26, 2017, Opatija, Croatia. , P. Biljanović, Ed. Opatija, Croatia: Mipro and IEEE, 2017, pp. 741-749.

[7] H. Jaakkola, J. Henno, and J. Mäkelä, "Technology and the Reincarnation Cycles of Software," in SQAMIA 2017 - Proceedings of the Sixth Workshop on Software Quality Analysis, Monitoring, Improvement, and Applications. Belgrade, Serbia, September 11-13, 2017., vol. Vol-1938, Z. Budimac, Ed. Belgrade, Serbia: CEUR Workshop Proceedings, 2017, pp. 5:1-10.

[8] K. Yrjönkoski, H. Jaakkola, T. Mikkonen, K. Systa, and J. Henno, "Software business: a short history and trends for the future," in 8th Workshop on Software Quality Analysis, Monitoring, Improvement and Applications, SQAMIA 2019 Proceedings, vol. Vol-2508, Z. Budimac and B. Koteska, Eds.: CEUR Workshop Proceedings, 2019, pp. 18:1-18:18. H. Jaakkola, B. Thalheim, and J. Henno, "About the Essence of Intelligence – Will Artificial Intelligence (Ever) Cover Human Intelligence?," in Information Modelling and Knowledge Bases XXXIII, M. Tropmann-Frick, H. Jaakkola, B. Thalheim, Y. Kiyoki, and Y. Nofumi, Eds. 343 ed. (Frontiers in Artificial Intelligence and Applications, Amsterdam: IOS Press, 2022, pp. 19-43.

[9] H. Jaakkola, B. Thalheim, and J. Henno, "About the Essence of Intelligence – Will Artificial Intelligence (Ever) Cover Human Intelligence?," in Information Modelling and Knowledge Bases XXXIII, M. Tropmann-Frick, H. Jaakkola, B. Thalheim, Y. Kiyoki, and Y. Nofumi, Eds. 343 ed. (Frontiers in Artificial Intelligence and Applications, Amsterdam: IOS Press, 2022, pp. 19-43.

[10] IBM. "The Origins of Computer Science", In B. Wood, H. Aiken, F. B. Brooks Jr, (Contibutors), Icons of Progress, 2022. https://www.ibm.com/ibm/history/ibm100/us/en/icons/compsci/.

[11] M. Sinko & E. Lehtinen, Tieto- ja viestintätekniikka opetuksessa ja oppimisessa Osaamisen haasteet ja tietotekniikan mahdollisuudet [ICT in teaching and learning: Competence challenges and information technology opportunities], https://urn.fi/URN:ISBN:978-952-03-2780-4Tulevaisuusvaliokunnan teknologiajaosto. Eduskunnan kanslian julkaisu 2/1998, Helsinki, 1998.

[12] Wikipedia, Tiigrihüpe [Tiger leap], Retrieved from https://en.wikipedia.org/wiki/Tiigrih%C3%BCpe on February 1st, 2023.

[13] S. Rahikainen, Tietotekniikan murros Suomen peruskouluissa 1984—1986 Opettaja-lehtien tekstien kuvaamana [Information technology revolution in Finnish elementary schools 1984—1986 as described by Opettaja magazine texts], University of Jyväskylä, URN:NBN:fi:jyu-201610164374.pdf, 2016.

[14] A. Molnar, "Computers in Education: A Brief History", The Journal, 1997. Retreived from https://thejournal.com/Articles/1997/06/01/Computers-in-Education-A-Brief-History.aspx on February 2nd 2023.

[15] J. T. Fouts, "Research on Computers and Education: Past, Present and Future", Seattle Pacific University, Seattle, WA, February, 2000. Retrieved from https://ic.unicamp.br/~wainer/cursos/2s2004/impactos2004/TotalReport3.pdf.

[16] J. Lee, "History of Computing in Education," International Federation for Information Processing Digital Library; History of Computing in Education, vol. 145, 04/11 2006. DOI:10.1007/1-4020-8136-7_1.

[17] X. Chen, D. Zou, G. Cheng, and H. Xie, "Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education," Computers & Education, vol. 151, 2020. https://doi.org/10.1016/j.compedu.2020.103855.

[18] P. Suppes, "The Uses of Computers in Education," Scientific American, vol. 215, no. 3, pp. 206-223, 1966.

[19] S.Oinas, M-P. Vainikainen et al.., Digitalisaation vaikutus oppimistilanteisiin, oppimiseen ja oppimistuloksiin yläkouluissa: Kansallisen tutkimushankkeen ensituloksia suosituksineen [The impact of digitization on learning situations, learning and learning outcomes in middle schools], University of Tampere, Faculty of Education, Monographs and series [1199], https://urn.fi/URN:ISBN:978-952-03-2780-4, 2023.

[20] Association for Computing Machinery (ACM), ACM Computing Classification System. Published in https://dl.acm.org/ccs.

[21] Association for Computing Machinery (ACM), ACM Computing Classification System (1998). Available in https://www.acm.org/publications/computing-classification-system/1998/ccs98.

[22] OpenAI, ChatGPT, Available in https://chat.openai.com/chat.

ICT and Changes in Education and Professional Activity of Journalists

Marin Galić*
* Faculty of Political Science, Department of Journalism and Media Production, Zagreb, Croatia
galicmarin@gmail.com

I. INTRODUCTION

The coronavirus disease (COVID-19), caused by the SARS-CoV-2 virus, brought about enormous and far-reaching changes in Croatia and all over the world, As expected, the epidemic also affected the teaching process at universities. Initially, the objective was mere survival, whereas now the world is faced with a situation in which what proved to be a good solution during the pandemic ought to be implemented in the system of education. In this context, the task of journalists is multifaceted – they are supposed to educate themselves and at the same time provide education to others, as well as to report on the situation. The question arises as to how ready we were to face the situation in which we found ourselves as a society and as a profession.

Throughout the entire process of society's development, technology was the driving force and every profession used it with regard to the material possibilities and skills it was developing. In the last twenty years, the forms of teaching have been altering due to the development of technology. New technical possibilities, including online teaching, have been increasingly taken advantage of. It is worth noting that these technical options mainly referred to experimental methods that were gradually introduced into syllabi and classes until the sudden spread of the COVID-19 disease. However, under pandemic conditions, experimental distance learning soon became the only reality and the only possible form of holding classes. Teachers and students had to adapt literally overnight to new methods and a new type of teaching.

On the other hand, apart from the changes in the delivery of classes caused by the pandemic, the implementation of ICT in the study of journalism at the Faculty of Political Science (FPZG) in Zagreb is necessary in terms of its educational content. Like many other professions, journalism has changed and adapted to new technological achievements. It is a socially important profession, and it is precisely for this reason that ICT needs to be implemented throughout the entire process of educating journalists. For the purposes of this paper, the courses at the Department of Journalism and Media Production of the Faculty of Political Science in Zagreb were analysed with a view to obtaining an insight into the stage of the process of implementation of the ICT system within Croatia's oldest higher institution offering study programmes in Journalism.

II. PREVIOUS KNOWLEDGE

In the contemporary era of Web 2.0 and beyond, in which the advent of chatbots in the news media has created a new journalistic narrative, the evolution in technology cannot be neglected in the educational phase. [1]

Rapid technological development has undoubtedly facilitated distance teaching [2]. What most terms (e.g., online learning, open learning, web-based learning, computer-based learning, blended learning, m-learning) have in common is the use of a computer connected to the network, by means of which you can learn anywhere, anytime, at any pace, in any way [3].

A more innovative and flexible way of learning via the Internet is supposed to allow for greater focus on each individual student. Online learning is defined as learning using different devices that have access to the Internet (e.g., mobile phones, laptops, etc.) in a synchronous or asynchronous environment. Students are independent of location when learning and when communicating with professors and other students [4]. The synchronous learning environment is structured in a way that allows students to attend live lectures, interact with teachers and students in real time. It makes it possible for students to receive instant feedback. Synchronous learning means that a group of students engages in learning at the same time. Thus, synchronous online learning methods include videoconferencing, teleconferencing, live chat, and live lectures. In an asynchronous environment, the lecture content is not available live, yet students access it through different learning systems. Instant feedback and immediate response are not possible in such an environment [5]. Asynchronous learning is based on the idea that students learn the same material at different times and in different places. It is also called Location Independent Learning. It differs from synchronous learning, in which students are simultaneously involved in activities such as attending lectures or working in a laboratory. Asynchronous learning takes place according to the student's own schedule. During the studies, the lecturer provides materials for reading and viewing, final assignments and evaluation exams. The student has the ability to access, thus fulfilling the required tasks within a flexible time frame. Asynchronous online learning methods include self-directed learning modules, streaming video content, virtual libraries, published lecture notes, and sharing on discussion forums or social networking

platforms. Synchronous learning, in contrast to asynchronous learning, preserves a certain form of social interaction [2]. Anyhow, it is necessary to clearly distinguish regular online teaching from online learning in a pandemic situation, which is actually a crisis situation in which crisis solutions are applied [6].

Andreas Veglis and Andreas Pomportis highlight the study results demonstrating the lack of journalists' knowledge in certain ICT areas [7]. According to the survey Journalists at Work [8], conducted by the National Council for the Training of Journalists (UK) in 2018, 64% of surveyed journalists (N=73) said they needed new or additional skills to be fully effective [8]. In particular, 46% of respondents mention Media Analytics in Social media, 45% – Video editing, 37% – Photoshop-skills, 36% – Data journalism [8]. Furthermore, they have to be multi-skilled and be able to work with a wide range of media platforms [8].

Andrea Veglis also suggests grouping the skills required for this into five categories [10]. The first covers basic ICT skills (work with a set of office applications and the use of basic Internet services). The second category refers to web publishing skills (basic knowledge of HTML and CSS and the ability to work with content management systems – CMS (WordPress, Drupal, Joomla! etc.)). The third category includes Web 2.0 skills: content creation and distribution through platforms and technologies (social networks, blogs, Wikis, social bookmarking, RSS). The fourth category of ICT skills is webcasting – creating podcasts and streams, while the fifth category combines the skills of data journalism, which are used to process a large body of data and visualize it.

According to J. Pavlik, the technology development is rapidly transforming journalism in four areas at the same time: (1) how journalists do their work; (2) news content; (3) structure or organisation of newsrooms; and (4) relationships between news organisations, journalists and their numerous publications [10].

Cindy Royal [11] substantiates the advisability of introducing courses in the field of ICT for future journalists in order to respond to the challenges of today.

Knowledge of new technologies and the monitoring of their evolution can make journalists actively involved in the ever-changing landscape of modern information distribution. [12]

Type of classes	In-person	Hybrid	Online
Academic year 2019 2020 Semester 1	Yes	No	No
Academic year 2019 2020 Semester 2	No	No	Yes
Academic year 2020 2021 Semester 1	No	No	Yes
Academic year 2020 2021 Semester 2	No	No	Yes
Academic year 2021 2022 Semester 1	No	Yes	No
Academic year 2021 2022 Semester 2	No	Yes	No
Academic year 2022 2023 Semester 1	Yes	No	No

Picture 1. Teaching at the FPZG 2019-2023.

III. METHODOLOGY

The method used in this research is study programme analysis. Job openings in this field were analysed, the requirements expected to be fulfilled by applicants were singled out and a comparison was made in the context of the competencies expected to be developed within individual courses and entire study programmes. Another part of the methodology involves qualitative data analysis. Therefore, a questionnaire was submitted to all members of the teaching staff (holders of either teaching and scientific-teaching academic titles) at the Department of Journalism and Media Production (expert sample). This department was selected because of the fact that it offers the largest number of specialised courses directly related to the journalism profession.

IV. RESEARCH

The research is divided into four parts. In the first part, the introduction of ICT tools into teaching activity at the Department of Journalism and Media Production is analysed. In the second part, the introduction of ICT systems into the process of carrying out administrative duties by students during their studies is explored. The third part examines the use of ICT systems directly in the teaching process of future journalists at the Department of Journalism and Media Production (the study programmes of Journalism), while the fourth part analyses the student media and the possibility of cooperation among students of the Faculty of Political Science.

A. Classes within the study programme of Journalism

At the beginning of the pandemic in 2020, the administration of the Faculty of Political Science immediately decided to switch to an online lecture system. The Microsoft Teams programme was chosen and all employees in teaching professions started using it for distance teaching. In academic year 2021/2022, the online

Courses in radio journalism	Courses in print journalism	Linguistic courses	Multimedia courses	Other courses
Basics of Radio	Short Journalistic Forms	Croatian Language and Journalistic Stylistics	Introduction to Media Journalism	Media Management
Radio Newsroom	Basics of Print Media	Croatian as a Media Language	The Art of the Interview	Journalism Profession: Regulations and Practice
Forms of Radio Expression	Forms of Journalistic Expression		Business Journalism	Media and Violence
Radio Journalism	Newspaper Newsroom		Sports Journalism	Journalistic Research Laboratory
	Writing for Newspapers		Thematic Journalism	Journalistic Ethics
			Journalism in the Intermedia Space	Introduction to Journalism
				Media and Diversity
				News History
				Investigative Journalism

Picture 2. Courses

form of lectures was retained, as the FPZG additionally faced the problem of holding in-person classes due to the fact that the Faculty building was partially damaged in the 2020 earthquake [12]. In the first semester of academic year 2021/2022, a hybrid method of teaching was accepted (classes were partly in-person and partly online). In academic year 2022/2023, all anti-epidemic measures were lifted, and classes almost completely resumed the regular in-person form.

Research among the teaching and scientific-teaching staff of the Department of Journalism and Media Production confirmed that everyone returned to in-person classes, except in some extraordinary and sporadic situations.

Online teaching showed certain shortcomings, primarily when it comes to the activity and participation of students in classes. Students use their own computer equipment and Internet connection, so there is no obligation for students to have their cameras on during class. Consequently, it happens that the majority of students only join the classes via the Internet, without actively participating in them. An additional problem are examinations as it is extremely difficult to control the conditions in which students take exams, that is, it is almost impossible to completely prevent their mutual communication. Furthermore, online lectures with more than 20 students proved to be ineffective because it is impossible to monitor their participation. The experiences of the past two years have shown that this method of teaching is applicable to smaller, interested groups, while it manifests shortcomings in working with larger groups (more than 20 students).

The pandemic brought about a long-lasting change in the communication among the members of the FPZG Council, which remained effective even after the end of the pandemic during the academic year 2022/2023. The Faculty Council sessions are still held in a hybrid model, which means that one can attend the sessions either in person or via the online platform., The majority of the FPZG Council members voted in favour of the hybrid model of holding sessions at the session held in January 2023. The only difference compared to communication with students is that Council sessions are held on the Zoom platform.

Most of the communication between students and teachers takes place via the Microsoft Teams platform, which has put the previous communication via the Faculty's Intranet site on the back burner. As a consequence of a temporary relocation of the Faculty to smaller, temporary premises due to the renovation of the main building, most of the office hours with students (according to the survey of the teaching staff at the Department of Journalism and Media Production) also take place online – via Microsoft Teams, i.e., the Zoom application.

B. Facilitating the administrative aspect of the studies

Every September, long queues of students of both junior and senior years used to be formed in front of the FPZG'S Admissions Office. The administration announced that they would finally take advantage of the technological possibilities, in the sense that registrations would be done electronically. In the academic year 2021/2022 the registration of grades in student record books was abandoned, and all grades have been entered through the teacher's university portal ever since. Topics of graduation theses are registered online, the same as the thesis supervisors' verification procedure. From academic year 2022/2023, students have been using the same platform to submit essays in all years of their studies.

C. Use of ICT tools in teaching

A questionnaire was sent to all colleagues in teaching and scientific-teaching positions (expert sample) of the Department of Journalism and Media Production in order to obtain relevant data on the use of ICT tools in teaching. According to the content analysis of the study programmes of Journalism, 14 members of the Department of Journalism and Media Production teach a total of 30 courses, out of which 11 are compulsory and 19 are optional.

ICT tools	Courses in television journalism	Courses in radio journalism	Courses in print journalism	Linguistic courses	Multimedia courses	Other courses
Microsoft Teams	X	X	X	X	X	X
Zoom	X	X	X	X	X	X
Story Editor			X		X	
Google Forms			X		X	
Google Tables			X		X	X
Google Classroom	X				X	
Adobe Photoshop			X		X	
Adobe Premier	X				X	
Adobe After Effects	X				X	
Adobe Audition	X	X			X	
AVID	X				X	
Kinemaster	X				X	
Kahoot	X				X	
Canva	X				X	
Infogram	X				X	
Weebly	X				X	
Animaker	X				X	
Audacity		X			X	
iMovie					X	
Knight Lab	X				X	
InShot	X				X	
Microsoft office	X	X	X	X	X	X

Picture 3. ICT Tools

Most courses are closely related to journalistic practice and provide the best insight into the use of ICT tools in the process of training of future generations of journalists. The teaching staff surveyed confirm that they use the Microsoft Teams and Zoom applications for online communication. The PowerPoint programme and Windows Media Player are used for screening teaching video materials in class.

In the courses related to radio (e.g., Forms of Radio Expression or Radio Newsroom), students learn to use audio content production applications such as Audacity or Adobe Audition.

Learning to work with an editorial system is part of the syllabi of the courses dealing with print journalism. In courses such as Newspaper Newsroom or Writing for Newspapers, an important tool is the specialist programme Story Editor, as well as aids such as Google Forms or Google Tables. In order to be able to process photos, students learn to use the Adobe Photoshop programme.

The largest number of programmes and applications are used in the courses related to television. Video production is also an important segment of multimedia courses. Thus, for the purposes of the courses, such as Forms of Television Expression or Television Newsroom, students learn to use the Kinemaster programme, as well as the Adobe Premiere and Avid programmes (when editing TV and video clips recorded with their mobile phones). As far as other Adobe programmes are concerned, Photoshop and Audition are taught (e.g., for the purposes of sound processing in podcasts). For the needs of "television" courses, as well as multimedia ones, such as Journalism in a Multimedia Environment, Business Journalism and Sports Journalism, students use various programmes on smartphones for editing and recording, depending on the device model. Students usually use Canva programmes for the purposes of creating graphics and animations, and sometimes use them to create presentations of their works. Moreover, they use Infogram and Animaker tools as well. Knight Lab is used to create interactive journalistic content (maps, timelines, galleries, etc.).

Special emphasis in the courses of the Department of Journalism and Media Production is laid on Google fact-checking tools. Students are taught to use Weebly for the purpose of creating websites.

Certain television and radio courses (e.g., Television Journalism) use the Google Classroom tool for written exams, as well as Kahoot (for occasional in-class testing of students' knowledge).

D. Student media

The FPZG is the only higher education institution in Croatia that offers students the opportunity to work on all existing types of media platforms, which makes it unique. Student Television has been operating at the FPZG since 6 November 2012. It has complete technical capabilities for recording, editing and broadcasting journalistic reports. In this way, students can learn absolutely everything regarding television journalism (reporting, technical support, production, directing). Television Student has its own website www.televizijastudent.com where students gain additional experience in web journalism.

The FPZG's radio channel Radio Student has been operating since 31 October 1996. This radio broadcasts in Zagreb at a frequency of 100.5 MHz. Students have the opportunity to practically master all journalistic and technical skills related to radio broadcasting. Radio student also has its own website www.radiostudent.hr where you can listen to the programme via streaming. This website provides additional information about the programme being broadcast and about various current events, with an emphasis on the music scene.

The FPZG is the publisher of Global, a student newspaper which has been published in a print format since 2014 and in a digital format since 2017. Due to financial reasons, the printed edition is not published regularly, yet occasionally. However, the portal www.globalnovine.eu is regularly updated and offers an excellent opportunity for practical journalistic training on the web platform.

V. CONCLUSION

The research showed that students at the Faculty of Political Science in Zagreb have adequate opportunities to get acquainted with modern ICT tools and systems, which will later facilitate their entry into the world of professional journalism. The FPZG offers its students the possibility of active work in all the media (TV, radio, press, Internet and social networks). In the course of their studies, students get to know most of the ICT tools currently available on the market. The problem is the fact that the Faculty is not able to pay for IT equipment (computers and smartphones) for students, and therefore segregation occurs, that is, students with better payment options have better tools at their disposal. It is precisely in this field that we should try to find a solution, so that IT equipment becomes available to all students under the same conditions.

REFERENCE

[1] A. Veglis.; T. Maniou, "Chatbots on the Rise: A New Narrative in Journalism", *Stud. Media Commun*, 7, 2019.

[2] J. L. McBrien, R. Cheng, and P. Jones, "Virtual spaces: Employing a synchronous online classroom to facilitate student engagement in online learning", *The International Review of Research in Open and Distributed Learning*, 10(3), 1–17, 2009.

[3] V.-M. Cojocariu, I. Lazar, V. Nedeff, and G. Lazar, "SWOT analysis of e-learning educational services from the perspective of their beneficiaries", *Procedia-Social and Behavioral Sciences*, 116, 1999–2003, 2014.

[4] V. Singh, and A. Thurman, "How many ways can we define online learning? A systematic literature review of definitions of online learning (1988–2018)", *American Journal of Distance Education,* 33(4), 289–306, 2019.

[5] J. Littlefield, "The difference between synchronous and asynchronous distance learning", https://www.thoughtco.com/synchronous-distance-learning-asynchronousdistance-learning-1097959, 2018. [18. 12. 2022.]

[6] C. Pace, S. K. Pettit, and K. S. Barker, "Best practices in middle level quaranteaching: Strategies, tips and resources amidst COVID-19", *Journal of the Georgia Association for Middle Level Education*, 31(1), 2, 2020.

[7] A. Veglis, and A. Pomportis, "Journalists in the Age of ICTs: Work Demands and Education Needs", *Journalism & Mass Communication Educator* 69(1), 61–75, 2013. doi:10.1177/1077695813513766

[8] M. Spilsbury, "Journalists at work: their views on training, recruitment and conditions", *National Council for the Training of Journalists*, London, 2018.

[9] A. Veglis, "Education of journalists on ICTs: Issues and opportunities", *Journal of Applied Journalism & Media Studies* 2, 265–279, 2013. doi:10.1386/ajms.2.2.265_1

[10] J. Pavlik, "The Impact of Technology on Journalism", *Journalism Studies* 1(2), 229–237, 2000. doi:10.1080/14616700050028226.

[11] C. Royal, "Coding the curriculum: Journalism education for the digital age", in *Global journalism education: Challenges and Innovations in the 21st Century*, R. Goodman and E. Steyn, Eds. Austin: University of Texas at Austin, 2017, pp. 383–407.

[12] M. Matsiola, P. Spiliopoulos, R. Kotsakis, C. Nicolaou, A. Podara, "Technology-Enhanced Learning in Audiovisual Education: The Case of Radio Journalism Course Design", *Education Sciences,* 9(1), 62, 2019. https://doi.org/10.3390/educsci9010062

[13] M. Gelenčir, "Dva zagrebačka fakulteta odlučila: Online nastava i u ljetnom semestru, no ima izuzetaka", Srednja.hr, 2021. https://www.srednja.hr/faks/jos-dva-zagrebacka-fakulteta-odlucila-nastava-online-i-u-ljetnom-semestru-no-ima-izuzetaka/ [20. 12. 2022.]

Physics-Electronics Practical Work Model Accessible and Manipulable Remotely via the Web

L. Yade, A.D. Gueye, P.L.T. Sow and A. Ndiaye
Alioune Diop University / Bambey, Diourbel, Senegal
dahirou.gueye@uadb.edu.sn

***Abstract* - Today, hands-on laboratory work is the most effective way to develop skills in the field of STEM (Science, Technology, Engineering, and Mathematics). However, this practical work usually takes place in local laboratories, which poses a problem in the context of higher education marked by the massification of students. Researchers have proposed solutions for remote practical work for computer science courses (programming, networking) by using screen sharing, improved functionality of LMS platforms, virtualization and the cloud. However, remote practical work in the field of physics-electronics is very limited and does not take into account the flexibility and simultaneous access to practical work benches. In this paper we propose a model of physics-electronics practical work accessible and manipulable remotely via the web. To illustrate the relevance of our model, we propose a device to detect a laser signal via an optical fiber controlled by a Raspberry pi from our web platform.**

Keywords – E-lab; Physics-electronics; Internet of things; Flask; Remote practical work via the web

I. INTRODUCTION

Senegalese universities, like those in the South, particularly in sub-Saharan Africa, are experiencing increasing difficulties in providing their students with a suitable practical work (PW) environment. The available infrastructures, especially the local laboratories for practical work, are largely insufficient to accommodate the high number of students. These difficulties are linked to the increase in the number of students, the cost of equipment and often to the non-existence of practical work done by teachers, to the point where practical work is no longer done in certain STEM (Sciences Technologies Engineering and Mathematics) disciplines. This situation weakens the field of STEM which requires specialized skills and this requires a systematic revival of practical work sessions in STEM training. Today the question of carrying out practical work at a distance in the field of STEM is raised and becomes a real research problem [1][2]. In the research, authors have proposed solutions for remote practical work in the field of computer science (programming and networking) through the use of distance learning platforms [4][5] and emerging technologies especially in the field of cloud [3], virtualization [15] and the use of telecom network protocols [6][7][8].

In this paper, we are interested in remote practical work in physics-electronics, which is one of the STEM domains for which the proposed remote practical work solutions are very rare. For this purpose, we propose a model of remote practical work that manipulates from a web browser physical-electronic devices connected to IoT objects. This will allow to mutualize the resources of laboratories in order to have a wide range of benches of simulations and experiments. This will also participate in maintaining a large network of collaborators, and consequently guarantee a maintenance and a continuous improvement of the quality of the experiments/simulations.

To bring the proof of our concept, we will set up a PW bench dealing with the detection of the signal intensity of a laser via optical fiber controlled by a Raspberry pi. Thus, we give the possibility to the learners to manipulate via the web the Raspberry which controls and drives the practical benches.

To carry out the work, we present in section 2 related work on remote practical work PW solutions. In section 3, we present our model of remote PW access for practical work benches that need to be connected to connected objects. In section 4, we present a device for detecting the intensity of a laser via optical fiber accessible and manipulable via the web. In the section 5, we present the results of our implementation.

II. RELATED WORK ON DISTANCE LEARNING IN PHYSICS AND ELECTRONICS

Remote laboratories come in various types [9][10] and are nowadays an essential means for the correct performance of remote practical work (PW). They allow solving the problems related to the realization of the practical works that are an obligatory component of the STEM learning and especially in the field of physics-electronics. During this last decade, many solutions resulting from research works [11][12][13][14] have emerged to fill the lack of pedagogical infrastructures in front of the massification of learners which increases from year to year in the Senegalese public universities in particular and in Africa in general. In the same way, we had to propose some solutions as a contribution to the numerous challenges to be taken up in order to have an effective educational system for the realization of practical work at a

distance. In [15] and [16], we have set up a virtual environment based on Cloud Computing that gives access to virtual machines and a web platform integrating a video conferencing solution based on WebRTC technology [16] for synchronous monitoring of the practical work. The web platform through which users authenticate themselves to access the virtual resources provided by the cloud, allows teachers to schedule lab sessions and share materials that will be accessible by students while remaining on the same platform to perform computer science lab sessions. Subsequently, in [3], we proposed a container-based remote lab environment as a replacement for virtual machines to optimize resources, and provide a collaborative and easy-to-access lab interface. This system is mainly based on Docker technology [3].

Other researchers have proposed some solutions to realize practical work based on simulation [17][18] and the Internet of Things [19][20][21] among others. In [18], the author tries to provide solutions to many pedagogical and technical problems often encountered in higher education distance learning courses. These problems have been identified and reported by the teachers in charge of the courses. All the implemented solutions have been thoroughly discussed between developers, teachers and learners; this explains the methodology adopted in the [18]. However, the contribution of the paper is mainly based on the simulation of educational materials.

An Arduino remote lab is proposed in [20] to provide support for e-learning experimental environments in the Internet of Things IoT domain. It provides quality online education programs. The authors in [21] proposed an Internet of Things based remote laboratory that enables experiments in the field of moving image analysis. This paper is carried out in the context of the European project IoT4SMEs and aims at improving distance learning. The paper in [22] presents a technique for motivating students in learning digital electronics and improving practical work in this field. The proposed system gives learners the possibility to perform/retake some experimental work at home, which increases their motivation. This system is mainly based on the Moodle distance learning platform and a prototyping board that includes the FPGA board.

However, most of the proposed solutions are largely based on simulating teaching materials and/or remote access software that teachers and learners have to install on their machines beforehand. This is a good practice because it is better to start with simulation before attacking the physical components which are often very sensitive. However, it would be more interesting if these solutions would allow the learners, after having carried out the PW by simulation, to carry them out on physical material while remaining remotely and on the same platform which centralizes all the process of correct realization of a PW.

It is in this perspective that our contribution will focus on solutions to improve the remote practical work in the field of physics-electronics. The proposed solution is based on the Internet of Things and the Flask web technology.

III. PROPOSAL FOR AN IOT-BASED PHYSICS-ELECTRONICS LAB MODEL ACCESSIBLE VIA THE WEB

A. Proposal for a PW model

The proposed model is structured in three levels: the e-lab (PW material), the web/LMS platform and the users. It is illustrated in figure 1.

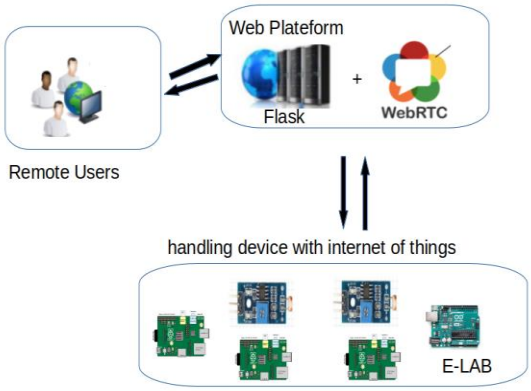


Figure 1. Remote practical work model in physics-electronics

- E-lab:

  This first layer is the basis of the proposed system. It contains the physical and virtual equipment necessary for the realization of remote practical work. These equipments are essentially composed of classical laboratory materials, IoT objects for remote manipulation, a camera for remote visualization of the manipulated objects, a configuration and control system for the IoT objects and a database for storing the data from the IoT objects. The equipment is arranged in the form of practical work benches.

- WEB/LMS Platform

  The second layer is a web platform that serves as an interface for users to access, after authentication, the educational resources located in the remote laboratory. The web platform allows, among other things:

  - to plan practical works and to reserve benches of practical works (teachers);

  - to consult the supports and instructions of the practical works planned by the teachers (learners);

  - to carry out simulations on virtual objects before starting the manipulation on real objects at a distance;

  - to visualize via a simple browser the real objects manipulated remotely from the laboratory.

- Users

  The third layer is the users, which are the teachers, the learners and the administrator.

IV. ADAPTATION OF OUR MODEL FOR A TYPE OF PRACTICAL WORK BASED ON THE DETECTION OF THE INTENSITY OF A LASER VIA AN OPTICAL FIBER

A. Presentation of a type of practical work on the detection of the intensity of a laser via an optical fiber

In this part we will present a typical case of realization of a device of detection of the intensity of a laser via an optical fiber in the laboratory of the department of physics of the University Alioune Diop of Bambey. This practical work is realized by the physicist Dr Papa Lat Tabara Sow co-author of this paper. The architecture of the PW is presented in figure 2.

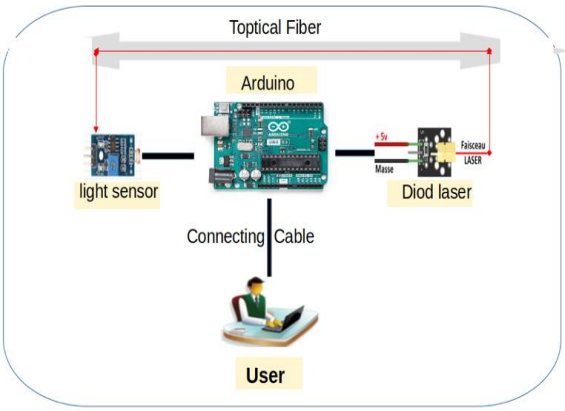


Figure 2. Initial architecture of the practical work with Arduino

Initially, the learners realized the PW using an Arduino board as a control acquisition center, an optical fiber as a waveguide and a laser that plays the role of information transmitter. To carry out this PW, the learner must first install the Arduino technology on the machine in which the Arduino board must remain connected to upload the scripts (figure 3). This obliges the learners to be gathered in the same room of PW where the material is.

This is a project of realization of a device of detection of the intensity of a laser. Its general objective is to show the students the vast field of application of lasers and to put into practice concepts such as pumping, population inversion and laser cavity learned in class. Specifically, it will enable learners to:

- Carry out the transmission of a digital signal by laser via an optical fiber;

- Discover and experiment an application of lasers in the field of telecoms;

- Discover the use of an Arduino board to drive a light source and sensor;

- Recall and experiment with the physical effects involved: total reflection for light guidance.

However, the realization of this type of practical work requires a physical presence in the laboratory and is not flexible for not only other types of practical work but also for a large number of learners wishing to work in collaboration. Therefore, in the following we will propose a more suitable lab architecture.

B. Architecture of the type of practical work proposed

The architecture of the PW model is illustrated in Figure 3.

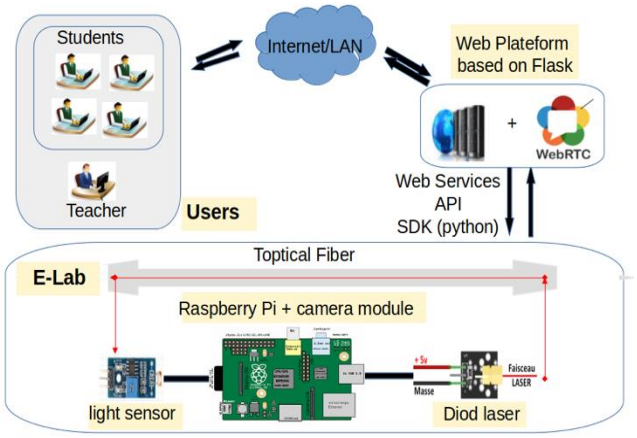


Figure 3. Architecture of the PW of detection of a laser via optical fiber based on the web

The proposed architecture aims to bring ease in the correct realization of the same PW and to achieve the same objectives but with a lot of flexibility, mobility thanks to Raspberry Pi technologies, Web Services and WebRTC for the communication between teachers and learners. For this, we replaced the Arduino board with a Raspberry Pi that allows to control and manipulate the whole device from a web platform. What is interesting here is that the Raspberry itself is remotely controllable from the web platform and can be used for simultaneous access to several learners or other types of practical work. This solves the problem of optimizing educational resources and the massification of learners in public universities. With such a system, the learner will no longer need to install or deploy anything on his machine to perform a practical work. He only needs a simple browser because everything happens on the web platform.

V. RESULTS

To prove the relevance of our model, we have implemented and realized a practical example in electronics about the detection of the intensity of a laser via optical fiber. The laser and the light detector are controlled and manipulated by a Raspberry pi from our web platform as shown in figure 3. To carry out this work, we have deployed the Flask web technology (a python framework) which will serve as our access interface to the implemented IoT device. The following part of the paper shows the different steps of the practical work illustrated by screenshots.

A. *Authentication page*

We have implemented an authentication mechanism to access the web platform; this allows us to control and secure access to the PW devices set up.

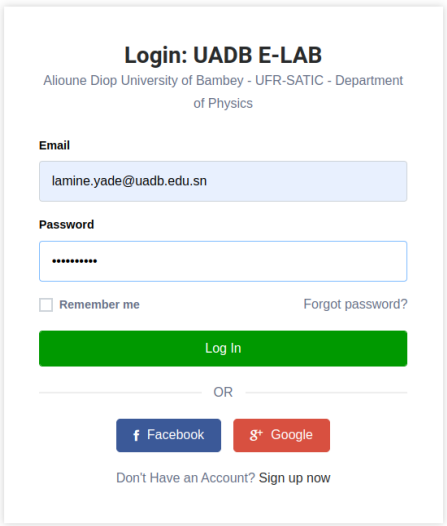


Figure 4. Authentication page

B. *Practical work list*

After successful authentication, the user has access to the list of scheduled practical work. He can select one of them and click on the Start button to start the lab. This scenario is illustrated in figure 5.

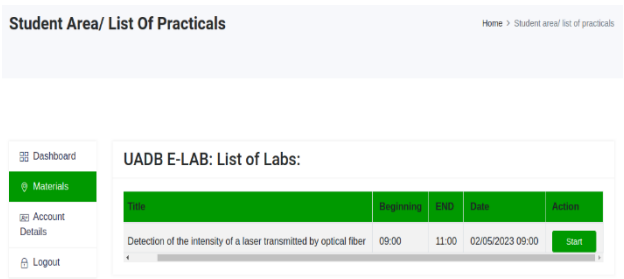


Figure 5. Viewing page for the list of PW

C. *PW realization pages*

There are several applications, protocols or network services to access and manipulate the Raspberry pi remotely. Among them we can mention vnc, spice, ssh, containers, web ssh (Shellinabox), code-server, among others.

In this paper we have chosen Shellinabox and code-server.

- Shellinabox offers a remote access service to the Raspberry terminal from a browser based on the ssh protocol. It is light and easy to deploy. The following picture shows the access to the Raspberry pi from our web platform set up with Flask using the ssh web protocol (Shellinabox).

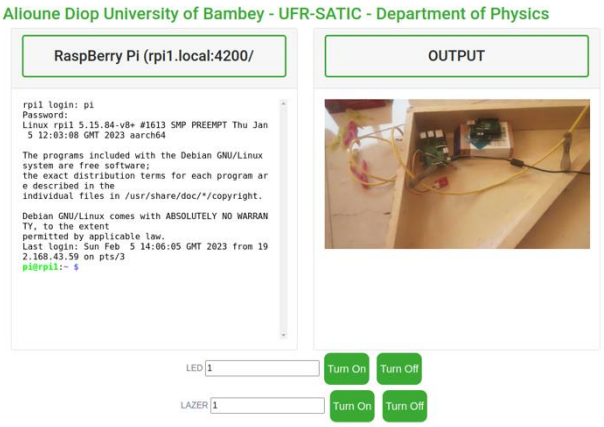


Figure 6. Access to the Raspberry from the web platform with ssh

- Code-server offers a complete remote development environment on Python, C, Java, PHP, JavaScript among others. It integrates a terminal and allows full access to the Raspberry tree. From the same interface, the user can access all the folders/files of the Raspberry, write scripts and execute them on the integrated terminal. Figure 7 shows the page for a practical work with code-server:

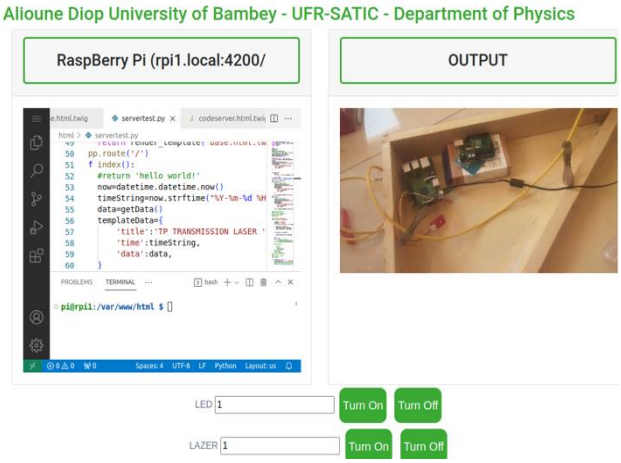


Figure 7. Remote programming on the Raspberry from a web interface

From these two interfaces the user can, on the one hand, install and configure protocols and applications and, on the other hand, write or execute scripts to control the Raspberry GPIOs from the web platform. After starting the python script prepared for this practical work, he can click on the Turn ON/Turn Off buttons to manipulate and control the laser while visualizing the results from the video installed in the lab and displayed in the same page.

VI. CONCLUSION AND PERSPECTIVES

Generally, the practical work in the STEM field more specifically in physics-electronics is done face-to-face in the local laboratory room and does not offer flexibility in the correct realization of remote and collaborative practical

work. In this paper, we have shown the possibility of correct realization of remote practical works in this domain by allowing the learners to have on one hand the hand on IoT objects connected to the laboratory device and on the other hand to manipulate remotely these same devices from a web platform configured and adapted to the need of the PW. From the web platform, the learner acts on buttons to manipulate the remote device and can have in return the result of his manipulations visible on the same web platform.

In perspective, we intend to propose a model of mutualization of resources of practical work on real devices of laboratories of universities and institutes in Africa within the framework of a collaboration or a win-win partnership.

REFERENCES

[1] K. Gaglo, M. Seck, S. Ouya and G. Mendy, "Proposal for an online practical work platform for improving the teaching of STEM," 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, Korea (South), 2018, pp. 1-2, doi: 10.23919/ICACT.2018.8323903.

[2] I. Sanogo, S. Ouya, A. D. Gueye and C. Lishou, "Proposal of cloud-based online laboratory model for practical training in the telecoms and networking fields," 2016 IEEE Global Engineering Education Conference (EDUCON), Abu Dhabi, United Arab Emirates, 2016, pp. 1101-1105, doi: 10.1109/EDUCON.2016.7474691.

[3] L. Yade and A. D. Gueye, "Remote Practical Work Environment based on Containers to replace Virtual Machines," 2022 IEEE Global Engineering Education Conference (EDUCON), Tunis, Tunisia, 2022, pp. 1285-1290, IEEE Xplore, doi: 10.1109/EDUCON52537.2022.9766480.

[4] Degboe, B. M., & Ouya, S. (2019, March). Contribution to the Setting of an Online Platform on Practical Application for the Science, Technology, Engineering and Mathematics (STEM):The Case of Medical Field. In The Challenges of the Digital Transformation in Education: Proceedings of the 21st International Conference on Interactive Collaborative Learning (ICL2018)-Volume 1(Vol. 916, p. 297). Springer.

[5] B. M. Degboe, G. M. S. -J. Kossingou, K. Gaglo, S. Ouya and G. Mendy, "Proposal of a platform for practical bimodal work in STEMs to respond COVID-19 for developing countries," 2021 IEEE World Conference on Engineering Education (EDUNINE), Guatemala City, Guatemala, 2021, pp. 1-5, doi: 10.1109/EDUNINE51952.2021.9429107.

[6] F. G. Loro, A. Macho, E. Sancristobal, M. R. Artacho, G. Díaz; and M. Castro, "Remote laboratories for electronics and new steps in learning process integration," 2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV), 2016, pp. 112-117, doi: 10.1109/REV.2016.7444449.

[7] D. May, B. Reeves, M. Trudgen and A. Alweshah, "The remote laboratory VISIR - Introducing online laboratory equipment in electrical engineering classes," 2020 IEEE Frontiers in Education Conference (FIE), 2020, pp. 1-9, doi: 10.1109/FIE44824.2020.9274121.

[8] G. M. S.-J. Kossingou, B. M. Degboe, S. Ouya, and G. Mendy, "Mutualisation of ICT laboratory resources between West and Central African universities in post-crisis situations: the case of Senegal and the Central African Republic," in *2020 Sixth International Conference on e-Learning (econf)*, Sakheer, Bahrain, Dec. 2020, pp. 1–5. doi: 10.1109/econf51404.2020.9385470.

[9] H. Saliah-Hassane, A. Kourri and I. D. La Teja, "Building a Repository for Online Laboratory Learning Scenarios," Proceedings. Frontiers in Education. 36th Annual Conference, San Diego, CA, USA, 2006, pp. 19-22, doi: 10.1109/FIE.2006.322603.

[10] F. Luthon and B. Larroque, "LaboREM—A Remote Laboratory for Game-Like Training in Electronics," in IEEE Transactions on Learning Technologies, vol. 8, no. 3, pp. 311-321, 1 July-Sept. 2015, doi: 10.1109/TLT.2014.2386337.

[11] O. Bauzha, B. Sus, S. Zagorodnyuk, A. Kozinetz and T. Chaikivskyi, "Applying the Communication Technologies for Remote Laboratory Works and Reseach Exteriments," 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2021, pp. 97-102, doi: 10.1109/UkrMiCo52950.2021.9716597.

[12] I. Ugarte, V. Fernández and P. Sánchez, "Motivation of students in the learning of digital electronics through the double integration: Remote/presential work and theoretical/laboratory classes," 2012 Technologies Applied to Electronics Teaching (TAEE), 2012, pp. 63-67, doi: 10.1109/TAEE.2012.6235408.

[13] C. Onime, J. Uhomoibhi and M. Zennaro, "Demonstration of a low cost implementation of an existing hands-on laboratory experiment in electronic engineering," 2014 11th International Conference on Remote Engineering and Virtual Instrumentation (REV), 2014, pp. 195-197, doi: 10.1109/REV.2014.6784253.

[14] S. Farah, A. Benachenhou, G. Neveux, D. Barataud, G. Andrieu and T. Fredon, "Flexible and real-time remote laboratory architecture based on Node.js server," 2015 3rd Experiment International Conference (exp.at'15), 2015, pp. 155-156, doi: 10.1109/EXPAT.2015.7463251.

[15] Lamine Yade, Amadou Dahirou Gueye, Bounama Gueye, Claude Lishou (2020), "Video Conferencing Solution for Synchronous Follow-Up of Cloud-Based IT Practical Work". Published in: 2020 International Journal of Online and Biomedical Engineering (iJOE) - eISSN: 2626-8493, Vol 16, No 14 (2020), https://doi.org/10.3991/ijoe.v16i14.17045

[16] A. D. Gueye, L. Yade, B. Gueye, O. Kasse and C. Lishou, "Cloud and WebRTC based Laboratory Solution for Practical Work in Computer Science for a Traditional University," 2020 IEEE Global Engineering Education Conference (EDUCON), 2020, pp. 1119-1124, doi: 10.1109/EDUCON45650.2020.9125201 https://doi.org/10.1109/educon45650.2020.9125201

[17] D. May, B. Reeves, M. Trudgen, and A. Alweshah, "The remote laboratory VISIR - Introducing online laboratory equipment in electrical engineering classes," in *2020 IEEE Frontiers in Education Conference (FIE)*, Uppsala, Sweden, Oct. 2020, pp. 1–9. doi: 10.1109/FIE44824.2020.9274121.

[18] A. V. Fidalgo et al., "Using remote labs to serve different teacher's needs A case study with VISIR and RemotElectLab," 2012 9th International Conference on Remote Engineering and Virtual Instrumentation (REV), 2012, pp. 1-6, doi: 10.1109/REV.2012.6293149.

[19] H. Guerra, A. M. Garcia, L. M. Gomes and A. Cardoso, "An IoT remote lab for seismic monitoring in a programming course," 2017 4th Experiment@International Conference (exp.at'17), 2017, pp. 129-130, doi: 10.1109/EXPAT.2017.7984361.

[20] A. Fernández-Pacheco, S. Martin and M. Castro, "Implementation of an Arduino Remote Laboratory with Raspberry Pi," 2019 IEEE Global Engineering Education Conference (EDUCON), 2019, pp. 1415-1418, doi: 10.1109/EDUCON.2019.8725030.

[21] C. Scaffidi and S. Distefano, "A Remotely Configurable Hardware/Software Architecture for a Distance IoT Lab," 2021 IEEE International Conference on Smart Computing (SMARTCOMP), 2021, pp. 341-346, doi: 10.1109/SMARTCOMP52413.2021.00070.

[22] M. Leisenberg and M. Stepponat, "Internet of Things Remote Labs: Experiences with Data Analysis Experiments for Students Education," 2019 IEEE Global Engineering Education Conference (EDUCON), 2019, pp. 22-27, doi: 10.1109/EDUCON.2019.8725070.

How to Use E-Learning App in Education and Presentation of Erasmus Project

M. Sertić*, T. Ćosić*, M. Kerstner**
*Industrijsko-obrtnička škola, Slavonski Brod, Croatia
**OŠ Šestine, Zagreb, Croatia
sertic_79@yahoo.com

***Abstract* - The base of this paper was project *Effective e-Learning System Based on Digital Competences* in the field of strategic partnerships at the level of Erasmus KA2 activities. It will be presented several applications for interactive digital content and tools in education. The idea of this project was to show how we can increase the level of digital competence for students and teachers by using digital tools in teaching classes. The goals of this project were sharing experience with using e-learning and finding new possibilities to improve digital competences of students and teachers.**

Keywords - Erasmus, e-learning, education content, learning and new idea in ICT content, learning app.

I. INTRODUCTION

Digital competences, and related e-learning with distance communication between students and teachers, is one of the more propulsive areas in the education of the European Union in the last decade. Applications that are created exclusively to help in the implementation of digital teaching are defined as "digital tools in education". Such tools - with the development of smartphones, watches and laptops, tablets have become very popular and useful, both for teachers and students, but also for authors and publishers of textbooks and educational materials. These apps and digital tools help teachers who teach geography, foreign languages or biology, as well as for primary and secondary school students. By using digital tools and apps we can better understand James Cook's journey to the South Seas in the 18th century, the syntax of the Latin language or the chemical evolution of the universe and the formation of the first planets thirteen billion years ago.

In addition to programmers and computers experts who create such digital tools with the help of subject experts, and teachers and students, who use them, there are those who practical understand and justify the role of interactive digital tools in the learning and teaching processes of the 21st century.

At this point, we can mention some previous research, but also very recent scientific articles dealing with various forms of e-learning. For example, Awaz Naaman Saleem, Narmin Mohammed Noori and Fezile Ozdamli in the article "Gamification Applications in E-learning: A Literature Review" from 2022 [1] tried to present an overview of the complete literature concerning gamification in digital learning.

The application of some elements of gamification is also present in the tools that we will mention in our project. Gamification is particularly useful in language learning, but also in all other subjects, from an early age. Some other scientists rightly warn of a significant and logical huge increase in e-learning during and after the SARS-CoV-2 virus pandemic. A condition for the quality use of online and digital tools is a solid hardware and software infrastructure, a power connection and an Internet network. Indonesian and Indian children and teachers may have problems with the latter [2, 3], but in Europe such a thing should be flawless (as shown by the work of Virkus et al. in [4]).

Furthermore, during the use of digital tools, it was realized that certain previous teaching models are not applicable for this form of learning. Flora Amiti from Macedonian South East European University tried to explain the difference between synchronous and asynchronous learning in her article "Synchronous and Asynchronous e-Learning" [5]. Namely, despite the importance of the living word and ex cathedra teaching, asynchronous learning in the digital environment is increasingly used, even though Amita presents secondary research that proves the benefits of combining both learning methods. After all, the latest artificial intelligence program ChatGPT and its newer versions show that grading will have to be done live or in a more secure way, and that digital platforms are only a learning aid, not a means of getting better grades.

II. EDUCATION AND E-LEARNING

The project *Effective e-Learning System Based on Digital Competences* connected six EU countries: Croatia, Latvia, Portugal, Romania, Sweden and Greece, and present use of interactive digital apps and educational tools in teaching like *Kahoot*, *Kami*, *Micro:bit*, *Geogebra*, *Python*, *Storyboard*, *Google Classroom*, *Quizizz*, *Genially* and others.

Kahoot is a popular game-based learning platform, and it is a kind of deviation from the usual traditional notion of teaching by using online quizzes [6]. Kami, on the other hand, is an original educational tool for teachers and students whose platform offers a whole range of benefits, from video and voice comments by teachers and

online assessment, through blended learning and learning from home, all the way to group projects and interschool collaboration [7].

Furthermore, the Micro:bit is a very handy little computer tool, actually a pocket computer suitable for teaching computer science (coding), mathematics and mechanical engineering [8, 9]. It can be used by children, students, teachers, as well as another similar tool, Geogebru [10]. Also popular is Python [11], which is used in programming, creating applications, digital content, graphic forms, etc. Furthermore, Storyboard is a useful graphic tool that organizes images and illustrations, as well as moving animations, and is useful in the teaching of applied arts [12]. Google Classroom is as the name suggests, an educational product of the company Google intended primarily for teachers in primary, secondary and higher education. The platform is free, and lecturers use it to monitor and evaluate student assignments [13].

Quizizz is another application for learning with a game, intended for both elementary school students and the elderly who acquire new knowledge more easily with the game [14, 15].
Genially has only been around for a few years (since 2015), but it has already gained a large number of users. It is an online tool for creating visual and interactive content. The tool can be used by individual users (pupils, students), but also by whole classes. Students who are fed up with classic PowerPoint presentations distributed by Microsoft Office can use more creative tools from this manufacturer, as well as other services such as infographics, interactive images, gaming educational tools, etc. [16].
All these possibilities and all other digital educational tools are also mentioned in Tim Slade's 2018 book: The eLearning Designer's Handbook: A Practical Guide to the eLearning Development Process for New eLearning Designers [17].

In general, this paper is presentation and dissemination of project Effective e-Learning System Based on Digital Competences in the field of strategic partnerships at the level of Erasmus KA2 activities and possibilities of use more digital apps in teaching classes. The long- term effects of this project were based on education and creating opportunities for an individual to a higher standard to find a better job, higher satisfaction with life and career, and higher level of inclusion in social development thanks to meaningful use of digital competences and digital tools in education. Project Effective e-Learning System Based on Digital Competences took place from 01.09.2018. to 31.08.2021., with a budget co-financed by the European Union in the amount of 123.193,00 € [18, 19].

The project focused on sharing experience in using of e-learning and finding new possibilities to improve digital competences of students and teachers.

Methods that are used for that was teaching and learning, presentations, workshops on practical use of tools, lessons with round table discussions, creative work, on collaboration and team work.

The key outcome of project will be to increase of the level of digital competence for students and teachers by using digital tools on effective, confident and critical way [18].
Expected results of project:

- New knowledge and experience in the field of e-learning,
- New and high level of digital competences necessary for studies and career improvement,
- Motivation to use new digital skills and e- learning apps effectively and creatively [18].

A. E-learning in project

This project has lots of positive objectives and someof them are easy applicable in the field of ICT and digital tools.

- Developing digital competences of students and teachers focusing on the useful use of ICT for the teaching and learning,
- Sharing experience on creating interactive and motivating study materials and lessonplans and using e-learning platforms and tools for creating e-content,
- Preparing students for better exploitation of ICT potential and useful learning,
- Preparing teachers for effective and motivating use of ICT in the classroom and e-learning platforms,
- Assisting educational institutions to implement innovative practices related to e-learning platform supporting the needs of all learners [18, 19, 20].

B. Activities

All activities will focus on developing digital skills and implementing the digital competences into the schools regular teaching process. In this project attention will be paid to improving digital competences and using them for better study purposes and job opportunities. Students will teach each other skills and make some products using their ICT skills. During Short Term Exchange students will work into international teams to make the best of their ICT skills. During Joint Staff Training teachers will share experience on e-learning platforms and digital tool [18, ,19, 20].

C. Project results

As are results of project it will be made a guidebook for teachers and webpage of project, Padlet materials will be a collection of lesson plans as used interactive materials that is made by students and teachers.

Other results of project will be to improved e-learning system and learning platforms at schools, shearing knowledge and experience gained in the field of e-learning, and increased competence in foreign languages.

All that will apply in digital competences of teachers and students necessary for their studies and career in future [19, 20].

All the participating schools will share results of project in many fields, and some of them will be:

- implemented innovative practices related to e-learning platform supporting the needs of all the learners and effective and motivating study materials adapted to the individual needs of every student,
- better exploitation of ICT potential for teachers and students, increased competence in foreign language,
- shared experience and good practices on using e-learning tools at schools focusing on the best practices that have developed to serve the needs of teachers and students [19,20].

D. E-learning apps

Through the use of certain applications in class, students can improve and increase their level of knowledge. They can do digital crosswords and tasks through applications, and solve tests. Through such applications, teachers can more easily and objectively monitor the student's progress.

Today, there are various digital applications on the Internet that connect teachers and students from all over the world. Applications are ready-made for all areas of education, from computer science, mathematics, logic, philosophy, mechanical engineering, library science [17, 19].

III. EFFECTIVENESS OF E-LEARNING

The conducted study was made about teaching and learning digital tools and educative content in the field of ICT. The research involves teachers and students from Romania, Sweden, Croatia, Latvia, Portugal and Greece. This are some of results of that research [18].

A. Study participants

In study participated six international school from Croatia, Latvia, Romania, Portugal, Sweden and Greece.

A survey with a questionnaire with a random sample of teachers was held online, and embraced mainly ICT knowledge and skills in primary and secondary schools. In the survey participated 201 teachers [18].

The results of the research mainly covered the area related to today's digital education, the application of digital educational content, the Internet in teaching. The research also included an insight into the possible equipment of technical and electronic devices in the school, such as computers, smart boards, projectors, tablets, printers and others.

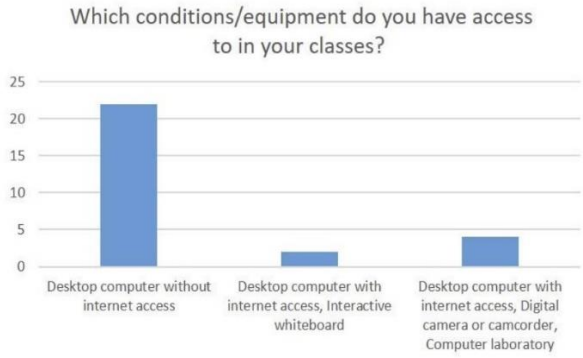


Fig. 1. ICT equipment in schools

From (Figure 1.) we can see that almost all teachers use computers and internet in all six countries for preparing lessons. Only 4% gave negative answer, which is a small number, but we have some teachers who never use computer or internet for preparing lessons [20].

We can see that results are different in all schools. In Sweden, Latvia and Romania teachers usedlaptops or tablet that are provided by school for them. In Croatia, Portugal and Greece the most of the teachers don’tused laptops or tablet that are provided by school for them.

Regarding on the use of digital devices, the possibility of additional education of teachers in the field of ICT were also mentioned. Most teachers regularly attend trainings and actively participate in seminars (Figure 2) [21].

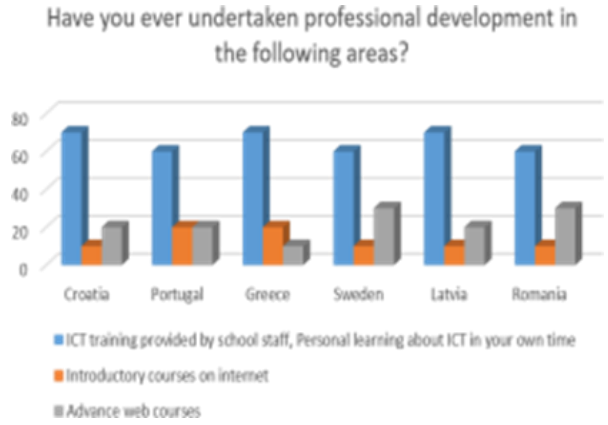


Fig. 2. Attend trainings

B. General questionnaires

The results of the research mainly covered the area related to today's digital education, the application of digital educational content, the Internet in learning like doing homework. The results from the line diagram show that the values how often do students use computers, tablets, and the Internet in class. And does schools provide students with laptops, computer, tablets for their own use we can see from (Figure 3).

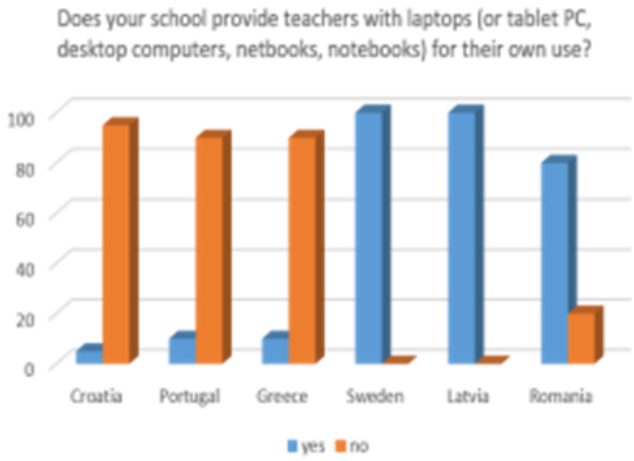


Fig. 3. Students and ICT use

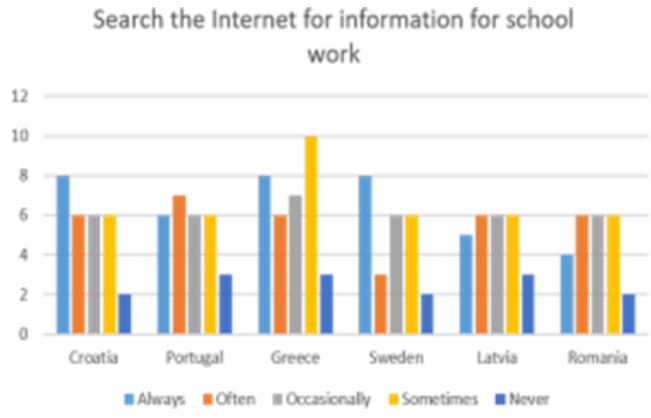


Fig. 4. Internet search for information

That most of students in Croatia, Sweden and Greece participate in online communities or forums related to subjects they study, and that most of students in all of countries occasionally or sometimes participate in online learning by using Instant Messenger, Facebook to contact other students about school work. Most of them use this tool occasionally or sometimes [21]. Most of students are allowed to use in class tablets and notebook for teaching and learning (Figure 4.) [21].

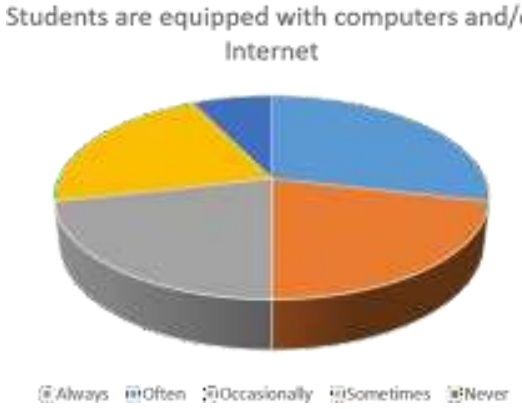


Fig. 5. Computers in teaching

From (Figure 4.) we can also see how often students use devices at school for learning in Greece, Croatia, Latvia, Portugal,Romania and Sweden.

From (Figure 5.) we can see how often students use the Internet to search for information for school work in Greece, Croatia, Latvia, Portugal, Romania and Sweden. And from that we can see how often students visiting online communities or forums related to the subjects they study [21].

IV. CONCLUSION

Most teachers says that they use ICT for "browse/search the internet to collect information to prepare lessons"; "browse/search the internet"; "create your own digital learning materials for students" and "prepare exercises and tasks for students. In most of countries all classrooms are equipped with a PC with internet connection and projector to be used by the teacher. Every teacher and student have an email and internet account so that they can access it. Although we can see that there were some teachers who already use ICT in their practice [19].

Given that online cooperation between teachers and students is essential in interactive teaching. Students should have hardware and software support at home, but also in the school itself. The results of this research showed that in some countries that participated in the project, there is good cooperation between schools and students in terms of IT support (Sweden, Latvia, Romania), while in other countries this needs to be drastically improved (Croatia, Portugal, Greece).

The research showed that students often use tablets in class. Also, the Internet is mostly available to students, as well as educational tools that help them learn. Students use forums to exchange information and knowledge, and receive and send assignments and tasks to their teachers through the before mentioned tools. However, the research shows that part of the students in the studied countries still do not use the Internet or educational tools to a satisfactory extent (Portugal, Greece, but also Latvia). Not all students use social networks, forums, other aids in education; Namely, there will always be a part of students - even in the most developed and IT-equipped countries of the world - who simply have a certain aversion to technological possibilities, and who will always prefer to take a physical book and learn from it, take notes by hand and thus improve their knowledge. But now we have touched on a completely new topic for which there is no more room in the work conceived in this way.

REFERENCES

[1] Saleem, A.N., Noori, N.M. & Ozdamli, F. "Gamification Applications in E-learning: A Literature Review." Tech Know Learn 27(2022), 139-159.

[2] Adenan, Faisal. "How is E-Learning-Based Distance Learning Implemented in Elementary Schools?" Al-Ishlah Journal Pendidikan Abstracted Vol.14, 4(2022), 6527-6538.

[3] Sharma, Deepak, and Amandeep Singh. "E-learning in India during Covid-19: challenges and opportunities." European Journal of Molecular & Clinical Medicine 7.7 (2021), 6199-6206.

[4] Virkus, S., Kirinić, V., Begičević Ređep, N. "The Role of E-Learning and Information Culture in Educational Institutions in Transforming European Education." In: Moos, L., Alfirević, N., Pavičić, J., Koren, A., Čačija, L. (eds) Educational Leadership, Improvement and Change. Palgrave Studies on Leadership and Learning in Teacher Education. Palgrave Pivot, Cham (2020).

[5] Amiti, Flora. "Synchronous and asynchronous E-learning." European Journal of Open Education and E-Learning Studies 5.2 (2020), 60-70.

[6] Wang, Alf Inge, and Rabail Tahir. "The effect of using Kahoot! for learning–A literature review." Computers & Education 149 (2020): 103818.

[7] Kami. Add a little Kami magic to your classroom. Available: Online; Accessed: https://www.kamiapp.com (10.3.2023.)

[8] BBC Micro:bit. Dostupno na: https://microbit.org/hr/ (10.3.2023.)

[9] Cápay, Martin, et al. "Programmable Hardware BBC Micro: Bit as a Tool for Developing Teacher Competencies." 2022 IEEE Global Engineering Education Conference (EDUCON).

[10] GeoGebra. Available: https://www.geogebra.org/?lang=en Online; Accessed: (10.3.2023.)

[11] Medina, Alberto P. Lorandi, et al. "A simple way to bring Python to the classrooms." 2020 IEEE International Conference on Engineering Veracruz (ICEV).

[12] Storyboard. Available: https://www.storyboardthat.com Online; Accessed: (10.3.2023.)

[13] Google Classroom. Available: https://edu.google.com/workspace-for-education/classroom/ Online; Accessed: (10.3.2023.)

[14] Quizizz. Available: https://quizizz.com/admin Online; Accessed: (10.3.2023.)

[15] Zhao, Fang. "Using Quizizz to integrate fun multiplayer activity in the accounting classroom." International Journal of Higher Education 8.1 (2019), 37-43.

[16] Genially. https://genial.ly/we-are/ Online; Accessed: (10.3.2023.)

[17] Slade, Tim. The eLearning Designer's Handbook: A Practical Guide to the eLearning Development Process for New eLearning Designers. Independently Published, 2020.

[18] Effective e-Learning System Based on Digital Competences. Available:https://erasmus-plus.ec.europa.eu/projects/search/details/2018-1-LV01-KA229-046952 Online; Accessed: (15.2.2023.)

[19] Web Tools, Educational Tools Brings the World of ICT. Erasmus+ KA2. Available: https://webskola53.wixsite.com/mysite Online; Accessed: (25.2.2023.)

[20] Hallonqvist, Jessica. EES Our Favourite Digital Tools and Websites. Padlet.
Available: https://padlet.com/jessica_hallonq/ees-our-favourite-digital-tools-and-websites-jv0g10q0zp0e Online; Accessed: (25.2.2023.)

[21] Hallonqvist, Jessica. Lesson Plans and Web Tools. Padlet.
Available: https://padlet.com/jessica_hallonq/9liybnypuxm3 Online; Accessed: (25.2.2023.)

Educational Processor with Single-cycle Instructions SVEU16

S. Ribić*, K. Hodžić*
* University of Sarajevo/Faculty of Electrical Engineering, Sarajevo, Bosnia and Herzegovina
samir.ribic@etf.unsa.ba, kerim.hodzic@etf.unsa.ba

***Abstract*—Courses in principles of digital computers usually begin with elementary logic circuits, proceed through increasingly complex ones, and often end with the design of a central processing unit. Such processors are typically simpler to explain than commercial microprocessors, but also often have very limited capabilities. The educational processor presented in this paper has a good balance between simplicity and capabilities. All its instructions, including multiplication and bit rotation, are executed in one clock cycle. The method of encoding and decoding instructions is quite simplified so that the encoding of instructions can be done even manually without tables, and the decoder unit is a simple forwarding of parts of the instruction word to the control bits of multiplexers. The processor is symmetric around the number 16: it has 16 three-operand instructions, 16 registers, each register is 16 bits wide, as well as the address and data bus. It is simulated at the logic gates level, a Verilog implementation on FPGA, and an emulated computer run by an implementation of a Forth interpreter written partly in its machine language.**

Keywords—processor, education, machine language

I. INTRODUCTION

Most computer science courses include the description of fundamental principles of digital computers. It begins with the basics of Boolean logic, followed by the analysis and synthesis of simple combinational and sequential circuits, progressing towards the memory and CPU (Central Processing Unit) structure. The processor, however, is not a simple circuit, especially if we are talking about modern processors, designed by thousands of people, with huge financial resources and years of development. Full understanding of real microprocessors is impossible, partly because of intellectual property; partly because they have thousands or even billions of logical gates. Therefore, real processors are explained at a higher level of abstraction as sets of larger blocks regarded as black boxes.

An alternative to real processors are educational processors, described in books and articles. They are simplified so that their operation can be understood at lower levels of abstraction, sometimes even down to the logic gate level. They do not have many instructions, although they can be designed as CISC (Complex Instruction Set Computer) or RISC (Reduced Instruction Set Computer). Recently they can be implemented cheaply using programmable chips, but they are mainly used in education as virtual circuits, due to their modest capabilities, often even more modest than commercial microprocessors made decades ago.

One group of educational microprocessors uses microcode withs supplied microcode listing. The Mic-1 [1] is an accumulator-based 16-bit processor, with an address space of 4 K, that decodes instructions quite slowly, bit by bit in microcode. The papers [2] [3] mention simplified eight-bit versions of Mic-1 which are more suitable for simulating in graphical simulators such as Logisim. A newer version of Mic-1, known as IJVM (Integer Java Virtual Machine) [4] is a stack-based 32-bit machine with a subset of the Java virtual machine instruction set.

Processors with a hardwired control unit can be more complex or simpler than microcoded ones, and their instruction sets are more difficult to change. The LC-3 (Little Computer 3) [5] is a well described 16-bit processor, with 16 instructions, 8 registers, and a complex state machine with 63 states. MProz-3 [6] has only three instructions and a 9-state machine, simple enough to be fully drawn using gate-level logic diagram. More advanced educational CPUs usually omit some implementation details. DLX [7] is a MIPS RISC variant focused on pipelining. ANT-32 [8] is a 32-bit CPU with paged virtual memory, 64 instructions, and exception handling.

II. DESIGN OF THE CPU SVEU16

In order to achieve the desired balance between simplicity, capabilities and speed of the processor, it is designed symmetrically. A RISC approach was chosen, using a hardwired implementation of instructions, with just two states determined by the clock signal level. Most of the work is delegated to the ALU (Arithmetical Logical Unit). All instructions share the same format: an opcode with three register operands. The CPU does not have flags, but uses special comparison instructions. Only two instructions have access to memory (load/store architecture). Even without a pipeline, all instructions are executed in one clock cycle.

The processor is completely sixteen-bit: in width of registers, address bus, data bus, and addressable memory cells. The processor has 16 registers, called R0, R1,... R15 available to the programmer. The register R15 is also called a program counter (PC). 16 operation codes are mapped to 16 ALU operations. All above inspired us to name the processor SVEU16, literally meaning "all in 16", local phrase describing operation at full capacity.

A. Instruction set and encoding

The instruction encoding method of SVEU16 processor is one of the simplest. All instructions share the mnemonic form *OP DEST*,*SRC1*,*SRC2*, where *OP* represents the operation performed on registers *SRC1* and *SRC2*, and *DEST* represents the destination register. The operation codes are presented in the table (I). Conversion from the mnemonic to a binary representation is trivial. Each instruction is 16 bits wide, and it is divided into groups of four bits. The highest 4 bits of the instruction represent the

TABLE I: Four bit opcode table

Opcode	Mnemonic	Operation
0000	LOD	*DEST*:=MEM[*SRC2*]
0001	ADD	*DEST*:=*SRC1*+*SRC2*
0010	SUB	*DEST*:=*SRC1*-*SRC2*
0011	AND	*DEST*:=*SRC1* $\wedge$ *SRC2*
0100	ORA	*DEST*:=*SRC1* $\vee$ *SRC2*
0101	XOR	*DEST*:=*SRC1* $\oplus$ *SRC2*
0110	SHR	*DEST*:=*SRC1* shifted *SRC2*
0111	MUL	*DEST*:=*SRC1* $\cdot$ *SRC2*
1000	STO	MEM[*SRC2*]:=*SRC1*; *DEST*:=*SRC1*
1001	MIF	IF *SRC1*<>0 THEN *DEST*:=*SRC2*
1010	GTU	*DEST*:=*SRC1* >*SRC2*
1011	GTS	*DEST*:=*SRC1* >*SRC2* (signed)
1100	LTU	*DEST*:=*SRC1* < *SRC2*
1101	LTS	*DEST*:=*SRC1* < *SRC2* (signed)
1110	EQU	*DEST*:=*SRC1* = *SRC2*
1111	MAJ	*DEST*:=*SRC1* ; R15:=*SRC2*

operation code; the next 4 bits encode the *DEST* register; then 4 bits encode *SRC1*, while the lowest 4 bits represent the *SRC2* register. If *SRC1* or *SRC2* have value 1111 (R15), it actually means R15+1, because it is more useful.

Instruction encoding is especially easy in hexadecimal system. Thus, the instruction `XOR R2,R5,R11` is encoded in hexadecimal as `525B`, while the instruction `EQU R1,R3,R5` is encoded as `E135`. Three digits of the hexadecimal instruction code are therefore immediately recognized from the instruction, while the highest digit is read from the small and easy to remember table.

Arithmetic and logical instructions perform standard operations between registers in *SRC1* and *SRC2*, and store the lower 16 bits of the result into register *DEST*. They are `ADD` (addition), `SUB` (subtraction), `MUL` (multiplication), `AND` (bitwise conjunction), `ORA` (bitwise disjunction) and `XOR` (bitwise exclusive disjunction). Comparison instructions have codes between binary 1010 and 1110, and the result of the comparison can be 0 or 1. Eg. the instruction with the hexadecimal code `E423` will write the value 1 to R4 if the registers R2 and R3 have equal values. Their results may be used in conditional assignment instruction `MIF` (code 1001) which writes the value of the *SRC2* register to the *DEST* register only if *SRC1* is non-zero.

Instructions that access the memory are `LOD` (0000) and `STO` (1000). The instruction `LOD` reads the value from the memory location pointed to by the *SRC2* register into the *DEST* register, while the *SRC1* field has no meaning. The instruction `STO` writes the contents of the register specified in the *SRC1* field to the memory location pointed to by *SRC2* as well as to the register specified in the *DEST* field. Instruction `LOD`, when *SRC2*=R15, exceptionally increments R15 by 2 after loading. This makes it possible to read and skip 16-bit constants encoded in a separate word immediately after the instruction. Technically this word is not a part of the instruction, because RISC processors have fixed instruction length, but the final effect is similar to loading an immediate operand.

The `MAJ` instruction with code 1111 is useful for implementing subroutines, interrupts, jumps and transferring data between registers. It writes register *SRC1* to *DEST* and register *SRC2* to R15. If *SRC1*=R15 (code 1111), then a jump will be made to the location pointed to by *SRC2*, remembering the contents of the R15+1 in *SRC1*. This will allow a return from the subroutine, by jumping to the location stored in the *DEST* register. `MAJ` is not the only way to implement jumps, because the program counter can be the operand and destination of all instructions. This capability is symmetric and spares opcodes, because absolute, indirect, relative indirect, conditional jumps and jumps with link are just special cases of general `LOD`, `ORA`, `ADD`, `MIF` and `MAJ` instructions, but it has two disadvantages. It should be used with care because nonsense operations over program counter like `SHR` are possible, and it is not a best way for learning jumps in assembly language because quite a few other processors (ex. PDP-11 and TI MSP430) use it. The problems may be overcome by changing signals (15) or at assembler level: by issuing warning when R15 is *DEST* and introducing alias mnemonics resembling instructions from other CPUs: for example indirect jump `JMP R5` as alias for `MAJ R0,R0,R5` or `ORA R15,R5,R5`; conditional `JNZ R1,R2` for `MIF R15,R1,R2`, etc.

The `SHR` instruction with code 0110 allows four types of bit shifting and rotating. The register specified in *SRC1* is shifted or rotated using values from *SRC2* and the result is written to the *DEST* register. Bits 0 to 3 (the register *SRC2*) determine the number of repeated shifts/rotates. The bits at position 5 and 4 of the *SRC2* register denote the operation type: 00 – aritmetic shift right (sign bit pushed), 01 – logical shift right (zero pushed), 10 – bit rotation to the right, 11 – logical shift left (zero pushed)

B. Structure of the processor

On the processor scheme, shown in Fig.1, we can see the structure of registers and their connections, program counter circuits, an arithmetic-logic unit, a control unit, a bus connection block, and bus signals.

Multiplexer blocks: Multiplexers have 2^n single-bit inputs, n control bits, and one output bit, where the binary value of the control bits determines which of the inputs is passed to the output, much like railway station switches. They may be packed into blocks, to make it possible to select between several multi-bit values. On SVEU16, each multiplexer block consists of $16n$ input lines denoted by d_{ij}, where the first index denotes one of 16-bit values, while the second index is a bit position in the 16-bit value. The multiplexer block has n control lines that are connected to the same control lines of all multiplexers in the block, and 16 outputs o_i that are connected to the outputs of the multiplexers within the block.

The 16x2:1 multiplexer block is realized by a set of logical expressions (1). SVEU16 uses 4 such blocks.

$$o_i = \overline{c_0} d_{0i} \vee c_0 d_{1i}, \quad 0 \leqslant i \leqslant 15 \tag{1}$$

The 16x4:1 multiplexer block is realized by a set of logical expressions (2). SVEU16 uses 2 such blocks.

$$o_i = \overline{c_1}\,\overline{c_0} d_{0i} \vee \overline{c_1} c_0 d_{1i} \vee c_1 \overline{c_0} d_{2i} \vee c_1 c_0 d_{3i}, 0 \leqslant i \leqslant 15 \tag{2}$$

The 16x16:1 multiplexer block is realized by a set of logical expressions (3). SVEU16 uses 7 such blocks.

$$\begin{aligned} o_i = {} & \overline{c_3}\,\overline{c_2}\,\overline{c_1}\,\overline{c_0} d_{0i} \vee \overline{c_3}\,\overline{c_2}\,\overline{c_1} c_0 d_{1i} \vee \overline{c_3}\,\overline{c_2} c_1 \overline{c_0} d_{2i} \vee \\ & \overline{c_3}\,\overline{c_2} c_1 c_0 d_{3i} \vee \overline{c_3} c_2 \overline{c_1}\,\overline{c_0} d_{4i} \vee \overline{c_3} c_2 \overline{c_1} c_0 d_{5i} \vee \\ & \overline{c_3} c_2 c_1 \overline{c_0} d_{6i} \vee \overline{c_3} c_2 c_1 c_0 d_{7i} \vee c_3 \overline{c_2}\,\overline{c_1}\,\overline{c_0} d_{8i} \vee \\ & c_3 \overline{c_2}\,\overline{c_1} c_0 d_{9i} \vee c_3 \overline{c_2} c_1 \overline{c_0} d_{10,i} \vee c_3 \overline{c_2} c_1 c_0 d_{11,i} \vee \\ & c_3 c_2 \overline{c_1}\,\overline{c_0} d_{12,i} \vee c_3 c_2 \overline{c_1} c_0 d_{13,i} \vee c_3 c_2 c_1 \overline{c_0} d_{14,i} \vee \\ & c_3 c_2 c_1 c_0 d_{15,i}, 0 \leqslant i \leqslant 15 \end{aligned} \tag{3}$$

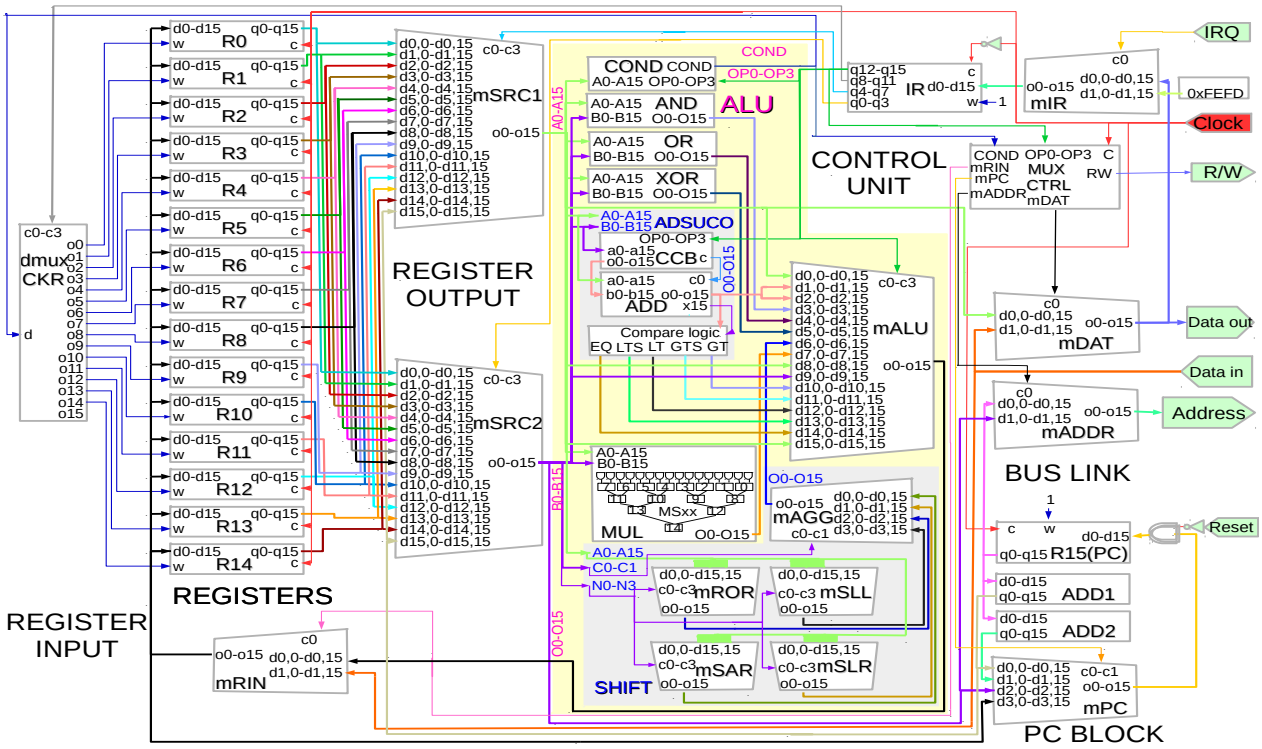


Fig. 1: Schematic diagram of the SVEU16 CPU

Demultiplexers do the reverse operation of multiplexers: depending on the value of n control bits, they pass a single-bit input value to one of 2^n outputs. The 1:16 demultiplexer is realized by (4).

$$\begin{array}{lll} o_0 = \overline{c_3}\,\overline{c_2}\,\overline{c_1}\,\overline{c_0}d & o_1 = \overline{c_3}\,\overline{c_2}\,\overline{c_1}\,c_0 d & o_2 = \overline{c_3}\,\overline{c_2}\,c_1\,\overline{c_0}d \\ o_3 = \overline{c_3}\,\overline{c_2}\,c_1 c_0 d & o_4 = \overline{c_3}\,c_2\,\overline{c_1}\,\overline{c_0}d & o_5 = \overline{c_3}\,c_2\,\overline{c_1}\,c_0 d \\ o_6 = \overline{c_3}\,c_2 c_1\,\overline{c_0}d & o_7 = \overline{c_3}\,c_2 c_1 c_0 d & o_8 = c_3\,\overline{c_2}\,\overline{c_1}\,\overline{c_0}d \\ o_9 = c_3\,\overline{c_2}\,\overline{c_1}\,c_0 d & o_{10} = c_3\,\overline{c_2}\,c_1\,\overline{c_0}d & o_{11} = c_3\,\overline{c_2}\,c_1 c_0 d \\ o_{12} = c_3 c_2\,\overline{c_1}\,\overline{c_0}d & o_{13} = c_3 c_2\,\overline{c_1}\,c_0 d & o_{14} = c_3 c_2 c_1\,\overline{c_0}d \\ o_{15} = c_3 c_2 c_1 c_0 d & & \end{array} \tag{4}$$

C. Arithmetical-logical unit

Boolean logic blocks: Blocks for the implementation of instructions `AND`, `ORA` and `XOR` are quite simple. Each of them has two 16-bit inputs A_i and B_i, outputs O_i, using 16 elementary logic gates that perform AND (5), OR (6) and XOR (7) operations.

$$O_i = A_i B_i, \quad 0 \leqslant i \leqslant 15 \tag{5}$$

$$O_i = A_i \vee B_i, \quad 0 \leqslant i \leqslant 15 \tag{6}$$

$$O_i = A_i \overline{B_i} \vee \overline{A_i} B_i, \quad 0 \leqslant i \leqslant 15 \tag{7}$$

Addition, subtraction and comparison block: ADSUCO (Add Subtract Compare) is a block built around an adder that adds, subtracts or compares two numbers. There are many kinds of multibit adders with different complexity and speed. The simplest, but the slowest one is the Ripple Carry Adder, which consists of cascaded full adders. Each of them has inputs a_i, b_i, c_i and outputs o_i, x_i, which perform the operations given by (8).

$$\begin{array}{l} o_i = \overline{a_i}\,\overline{b_i} c_i \vee \overline{a_i} b_i \overline{c_i} \vee a_i \overline{b_i}\,\overline{c_i} \vee a_i b_i c_i \\ x_i = a_i b_i \vee a_i c_i \vee b_i c_i, \quad 0 \leqslant i \leqslant 15 \end{array} \tag{8}$$

16 full adders are connected in a cascade so that the output x_i is connected to the input c_{i+1}. The output x_{15} is the output carry and the input c_0 represents the input carry. We have also tried much faster but more complex cascaded four-bit Carry Lookahead Adder and Kogge-Stone Adder.

CCB (Conditional Complementer Block) complements the number a_i if the control bit $c = 1$, performing the operations given by (9). Unlike XOR block, which performs similar function, all bits of the second operand have the same value. The outputs o_i of CCB are connected to the inputs b_i of the adder. The control bit c of CCB is connected to the input carry c_0 of the adder. So, if $c = 1$, the adder adds one 16-bit number with a two's complement of the other 16-bit number, basically subtracting them.

$$o_i = a_i \bar{c} \vee \overline{a_i} c, \quad 0 \leqslant i \leqslant 15 \tag{9}$$

The control bit of CCB is zero only for `ADD` instruction (opcode 0001) so it is calculated using the expression (10).

$$c = \overline{\overline{\mathrm{OP}_3}\,\overline{\mathrm{OP}_2}\,\overline{\mathrm{OP}_1}\,\mathrm{OP}_3} \tag{10}$$

The inputs to ADSUCO are the values A_i, B_i ($\leqslant i \leqslant 15$), and the opcode bits OP_i ($0 \leqslant i \leqslant 3$). The values of A_i are connected to the inputs a_i of the adder, and the values of B_i to the inputs a_i of CCB. The output bits are O_i, which represent the result of addition or subtraction, and the bits: EQ (A=B), GT (A>B unsigned), LT (A<B unsigned), GTS (A>B signed) and LTS (A<B signed) which represent the results of the comparison between the numbers A_i and B_i, calculated by (11), but valid only if $c = 1$. Although they look like flag bits, they are actually least significiant digit of different comparison operations, while remaining 15 digits are zeros.

$$\begin{array}{l} \mathrm{EQ} = \overline{o_{15} o_{14}} \cdot \ldots \cdot \overline{o_1 o_0} \quad \mathrm{LT} = \overline{x_{15}} \quad \mathrm{GT} = \overline{\mathrm{EQ} \vee \mathrm{LT}} \\ \mathrm{LTS} = a_{15} \overline{b_{15}} \vee a_{15} \overline{x_{15}} \vee \overline{b_{15}}\,\overline{x_{15}} \quad \mathrm{GTS} = \overline{\mathrm{EQ} \vee LTS} \end{array} \tag{11}$$

Multiplication block: Like the blocks for realizing logical operations, the multiplier has inputs A_i, B_i, and outputs O_i. The output is the result of multiplication modulo 65536, i.e., the lower 16 bits of the product. The multiplier is realized using 15 adders, each 16-bit size, called $\mathrm{MS}_0, \mathrm{MS}_1, ..., \mathrm{MS}_{15}$. All of them have the input carry $c_0 = 0$. Firstly we generate 16 partial products m_{i*} using logical conjunction and shifting(12).

$$m_{ij} = \begin{cases} A_{j-i}B_i & \text{if } i \leqslant j \\ 0 & \text{if } i > j \end{cases}, 0 \leqslant i, j \leqslant 15 \qquad (12)$$

The partial products are sent to 8 first-level adders, i.e., $m_{2i,j}$ is connected to input a_j of MS_i, $m_{2i+1,j}$ to input b_j of $\mathrm{MS}_i, 0 \leqslant i \leqslant 7; 0 \leqslant j \leqslant 15$. The results of these additions are sent to 4 second-level adders: o_j of MS_{2i} is connected to input a_j of MS_{8+i}, o_j of MS_{2i+1} to b_j of MS_{8+i}, $0 \leqslant i \leqslant 3; 0 \leqslant j \leqslant 15$. Second level addition results are sent to 2 third-level adders: o_j of MS_{8+2i} is connected to input a_j of MS_{12+i}, o_j of MS_{9+2i} to b_j of MS_{12+i}, $0 \leqslant i \leqslant 1; 0 \leqslant j \leqslant 15$. The fourth-level adder MS_{14} receives inputs a_j from o_j of MS_{12} and b_j from o_j of MS_{13}. Outputs o_j of MS_{14} are final result O_j.

Block for shifting and rotating: Bit shift and rotate operations are quite common, so it is useful to dedicate a separate block for them. The inputs to this block are the 16-bit number A_i $(0 \leqslant i \leqslant 15)$, the four-bit number N_i $(0 \leqslant i \leqslant 3)$ that determines how many times the operation is repeated; and two control bits C_1 and C_0 that determine the type of operation. The output from the block is a 16-bit number $O_i (0 \leqslant i \leqslant 15)$, which represents the result.

Within this block, four shift and rotate operations are implemented using four 16x16:1 multiplexer blocks called mSAR, mSLR, mROR and mSLL. Their control bits determine how many times the shift/rotate operation is performed. They are further connected to a 16x4:1 multiplexer block called mAGG, whose control bits determine which operation it is. The input control bits N_i are connected to the control bits c_i of all four 16x16:1 multiplexer blocks. The input control bits C_1 and C_0 are connected to the corresponding control bits of mAGG.

The input value bits A_i should be connected to multiplexer blocks inputs. Bits A_{j+i} for $i + j \leqslant 15$, or A_{15} for $i + j > 15$, are connected to mSAR inputs d_{ij}. Bits A_{j+i} for $i + j \leqslant 15$, or 0 for $i + j > 15$, are connected to mSLR inputs d_{ij}. Bits A_{j+i} for $i + j \leqslant 15$, or A_{i-16+j} for $i + j > 15$, are connected to mROR inputs d_{ij}. Bits A_{j-i} for $i \leqslant j$, or 0 for $i > j$, are connected to mSLL inputs d_{ij}. Outputs o_i from mSAR are connected to mAGG inputs d_{0i}, o_i from mSLR to d_{1i}, o_i from mROR to d_{2i} and o_i from mSLL to d_{3i}.

ALU multiplexers: ALU performs simultaneously all the operations described above. Its inputs are the first operand A_i, the second operand B_i and four control bits OP_i that serve to select the operation. The outputs are the operation result O_i and the condition bit COND. Inputs A_i and B_i are connected to the corresponding inputs of all blocks for arithmetic and logical operations. The exception is the shift and rotate block, where only bits B_0 – B_3 of the ALU input are connected to the corresponding inputs N_i, while bits B_4 and B_5 are connected to control bits C_0 and C_1. The result is determined by the 16x16:1 output multiplexer block called mALU, which has outputs from the arithmetic and logic blocks as its inputs.

The inputs of ALU A_j are connected to d_{0j}, d_{8j} and $d_{15,j}$ of mALU. The inputs of ALU B_j are connected to d_{9j} of mALU. The outputs of the ADSUCO block O_j are connected to d_{1j} and d_{2j} to realize addition and subtraction. The single-bit outputs from ADSUCO are connected to mALU inputs as follows: GT to $d_{10,0}$, GTS to $d_{11,0}$, LT to $d_{12,0}$, LTS to $d_{13,0}$, EQ to $d_{14,0}$ while the remaining bits d_{ij} for $14 \leqslant i \leqslant 10$ and $j > 0$ are connected to 0. The outputs O_j from the circuits for AND, OR and XOR are connected to the corresponding inputs d_{3j}, d_{4j} and d_{5j}, respectively. The outputs of the multiplication circuit O_j are connected to d_{7j}. The outputs of the shift and rotate circuit O_j are connected to d_{6j}.

Thus, the operation codes of the instructions are directly mapped to the corresponding operations of the ALU, which is redirected by mALU to give the correct result. Instructions with codes 0000, 1000, 1111 and 1001 are partly implemented outside ALU, so for these codes only the input value is passed.

The output of the conditional assignment COND (13) from the ALU should be 0 only for the instruction with code 1001 if the input A_i has all bits equal to 0.

$$\mathrm{COND} = \overline{\mathrm{OP}_3\overline{\mathrm{OP}_2}\overline{\mathrm{OP}_1}\mathrm{OP}_0\overline{A_{15}}\overline{A_{14}} \cdot ... \cdot \overline{A_1}\overline{A_0}} \qquad (13)$$

D. Registers

There is no need for other sequential subcircuits besides registers. Registers are built from rising edge-triggered D flip-flops with write enable that have input values d (data bit), w (write enable) and c (rectangular clock signal), and the output value q. They are made of 3 SR latches with 2:1 multiplexer and described by (14), where z, k, l, m, and n are auxiliary values.

$$\begin{gathered} z = \overline{w}q \vee wd \quad k = \overline{ml} \quad l = \overline{kc} \quad q = \overline{rl} \\ n = \overline{lcm} \quad r = \overline{qn} \quad m = \overline{zn} \end{gathered} \qquad (14)$$

A register contains 16 edge-triggered flip-flops. The clock and write enable signal are connected to the c and w inputs of all D flip-flops in the register, while the corresponding d inputs and q outputs are independent. Therefore, the registers have a 16-bit input d_i, a clock signal c, a write enable signal w and a 16-bit output q_i, where $0 \leqslant i \leqslant 15$. There are 17 registers: general purpose registers R0 – R14, program counter R15 and instruction register IR.

General purpose registers $R_i (0 \leqslant i \leqslant 14)$ forward their outputs $q_j, (0 \leqslant j \leqslant 15)$ to the inputs d_{ij} of two 16x16:1 multiplexer blocks called mSRC1 and mSRC2. This allows registers to be operands of ALU operations. Inputs $d_{15,j}$ of mSRC1 and mSRC2 are connected to the value of register R15 (PC) increased by 1. The outputs $o_i, (0 \leqslant i \leqslant 15)$ from mSRC1 are connected to the ALU inputs A_i. The outputs o_i from mSRC2 are connected to the ALU inputs B_i, and to the inputs $d_2 i$ of the mPC multiplexer block.

The inputs to all general-purpose registers d_i are fed from the output o_i of a 16x2:1 multiplexer block called mRIN. The inputs to mRIN are outputs from ALU O_i and the DataIn value from memory. Only for the `LOD`

instruction the control bit c_0 of mRIN will be 1 and mRIN passes the memory value; otherwise the ALU value.

The write enable signal inputs w of all general-purpose registers $R_i (0 \leqslant i \leqslant 14)$ are received from the outputs o_i of a 1:16 demultiplexer called CKR. The input d of CKR receives the output COND from the ALU. For the `MIF` conditional assignment instruction, if COND=0, the *DEST* register has $w = 0$ and register value is unchanged.

E. Program counter block

The program counter R15 or PC is handled separately from other registers (although updated simultaneously) with them, because its value comes from five different sources. The clock signal is connected to the c inputs of R15. Write enable signal w is at logical 1. The inputs d_i of R15 are connected to the outputs o_i of the 16x4:1 multiplexer block called mPC in such way that each of them is multiplied with a negated reset signal, i.e. $d_i = o_i \overline{\mathrm{RESET}}$. When the reset signal appears, R15 will get the value 0.

The outputs q_i of R15 are connected to the inputs a_i of two special adders ADD1 and ADD2. The second operand of addition, b_i on ADD1 is connected to the hardwired value 1, and on ADD2 to 2. The outputs o_i from ADD1 are connected to the inputs d_{0i} of mPC, o_i from ADD2 to d_{1i}, o_i from mSRC2 to d_{2i} and o_i from mRIN to the inputs d_{3i} of mPC.

F. Control unit

The control unit generates the control signals of all multiplexers and demultiplexers based on clock signal state and currently executed instruction.

Instruction register: The main part of this unit is the IR (Instruction Register). IR is the only register whose value changes on the falling edge of the main clock signal. That is why the negated value of the clock signal is connected to its c input. IR receives the value from the memory location pointed to by R15 or from the fixed value in the case of an interrupt signal. The value source of IR is determined by a 16x2:1 multiplexer block called mIR, whose outputs o_i are connected to the inputs d_i of IR.

The simplicity of the design of this processor is especially noticeable when looking at the output structure of the IR. Bits $q_{12} - q_{15}$ are directly connected to the control bits $\mathrm{OP}_0 - \mathrm{OP}_3$ in ALU (determining the operation to be performed), $q_8 - q_{11}$ to $c_0 - c_3$ of the CKR demultiplexer (register into which the result is passed, unless it is R15), $q_4 - q_7$ to $c_0 - c_3$ of mSRC1 (first operand of the ALU operation), and $q_0 - q_3$ to $c_0 - c_3$ of mSRC2 (the second operand of the ALU operation or the address register for the `LOD` and `STO` instructions).

Interrupt: The multiplexer block mIR may pass to IR the hardwired value `FEFD` (hex) connected to inputs d_{1j} of mIR, when the signal $\mathrm{IRQ} = 1$. That value actually represents the instruction `MAJ R14,R15,R13` which saves the program counter in R14 and the program continues execution from the memory location specified in register R13. The IRQ signal must be a short pulse because if it remains active until the next clock transition, returning from the interrupt handling routine will not be possible.

Multiplexer control signals: The MUX-CTRL block generates control signals for remaining multiplexer blocks, based on IR outputs, COND signal, and clock phase.

The control signal c_0 of mRIN, which determines whether the new destination register value is taken from the ALU or from the memory, is equal to 1 only for the `LOD` instruction (code 0000), so we send the value $\overline{q_{15}}\overline{q_{14}}\overline{q_{13}}\overline{q_{12}}$ from the register IR to that control bit.

The signals of the mPC multiplexer block c_1 and c_0 denote a two-bit value z. The value $z = 01$ (incrementing R15 by 2) occurs if *OP* $= 0000$ (`LOD`) or *OP* $= 1000$ (`STO`), *SRC2* $= 1111$ (R15), but *DEST* $\neq 1111$.The value $z = 10$ (a new R15 value is taken from mSRC2) occurs if *OP* $= 1111$ (`MAJ`). The value $z = 11$ (a new R15 is taken from mRIN) occurs if COND=1, *DEST* $= 1111$ (R15), but *OP* $\neq 1111$ (not `MAJ`). The value $z = 00$ (incrementing R15 by 1) happens in other cases. Expressions (15) generate the control signals of mPC from the output values q_i of IR and COND from ALU.

$$\begin{aligned} r_1 &= \overline{q_{14}}\overline{q_{13}}\overline{q_{12}}q_4q_3q_2q_0(\overline{q_{11}q_{10}q_9q_8}) \\ r_2 &= q_{15}q_{14}q_{13}q_{12} \\ r_3 &= \mathrm{COND}q_{11}q_{10}q_9q_8(\overline{q_{15}q_{14}q_{13}q_{12}}) \\ r_0 &= \overline{r_1}\overline{r_2}\overline{r_3} \\ c_0 &= r_0 \vee r_2 \quad c_1 = r_1 \vee r_3 \end{aligned} \tag{15}$$

G. Bus connection

The 16x2:1 multiplexer block mADDR determines what will be sent to the address bus. Its outputs o_i are connected to the address bus. Outputs q_i from R15 are connected to the inputs d_{0i} of mADDR, and outputs o_i from mSRC2 are connected to the inputs d_{1i} of mADDR. The address bus takes the value from mSRC2 if the instruction works with the memory (codes 1000 and 0000) and the clock signal is at a low level. So, the control signal c_0 of mADDR recieves value $\overline{\mathrm{CLOCK}}\overline{q_{14}}\overline{q_{13}}\overline{q_{12}}$, where q_i are outputs from IR.

The 16x2:1 multiplexer block mDAT determines what will be sent to the output data bus. Its outputs o_i are connected to the output data bus. The outputs o_i from the multiplexer block mSRC1 are connected to the inputs d_{0i} of mDAT, and the corresponding values from the input data bus are connected to the inputs d_{1i} of mDAT. The address bus has the value from mSRC1, if the `STO` instruction is being executed (code 1000) and the clock signal is at a low level. The control signal c_0 of mDAT recieves value $\overline{\mathrm{CLOCK}}q_{15}\overline{q_{14}}\overline{q_{13}}\overline{q_{12}}$.

Writing to the memory occurs only during the `STO` instruction, so the value $q_{15}\overline{q_{14}}\overline{q_{13}}\overline{q_{12}}$ from `IR` is passed to the RW signal on the bus. It is assumed that memory is realized from synchronous D flip-flops ; i.e., that the value can only be written on the rising edge of the clock. If this is not the case, it is necessary to add delay circuits and change the way this signal is generated.

H. Machine instruction cycle

When the master clock signal is high, the value of program counter R15 is set on the address bus with the READ signal. The memory chip responds with the value located at the given memory address, and it is forwarded to the inputs of IR through mIR, unless in case of interrupts.

When the clock signal goes from high to low, the content of the 16-bit memory location representing the instruction enters the IR register.

While the clock signal is low, the multiplexers redirect the data. Values from registers or memory location are processed in ALU and the results are ready to enter the destination registers or memory location.

When the clock signal rises to high, the new value is written into the write enabled destination register or memory location, R15 is updated, and a next cycle starts.

III. Implementation

A. Simulation and synthesis

Gate level simulation: After translating equations into logic gates, the processor has about 7600 elementary multi-input NAND gates if the adders are Ripple Carry, or about 8900 if the adders are based on a four-bit Carry Lookahead Adder. Entering such a scheme in graphic designers and simulators is quite slow and confusing, so we decided to use modular text input. Initially, we described the CPU in structural Verilog, using only modules, wires, arrays of wires and elementary logic gates. Verilator [10] appeared too difficult for installation and usage to be used for teaching, while v2c [11] working at RTL (Register Transfer Level) failed to generate code from our structural Verilog description. The Icarus Verilog [9] successfully compiled and simulated the description of this microprocessor with 64K RAM and 512 bytes of ROM in structural Verilog. However, the simulation was slow to run, since the translated file was about 105 M size.

We improved the simulation performance by building our own digital simulator in C with the help of a C preprocessor. The language used to describe and simulate the circuit resembles structural Verilog, with certain differences that could not be avoided, but they were quickly fixed using search and replace tools. The circuit is however compiled with the standard gcc compiler, like any C program. The corresponding preprocessor macros rename keywords, expand to initialize variables that point to appropriate nodes (represented by array elements) in the digital circuit, or commands that determine new node values by performing elementary logical functions between previous node values, which is also a rudimentary propagation delay simulation. The object module of the simulator created in this way was reduced to only 150 K and run much faster.

ISA level simulation: While simulation at the structure and logic circuits level is important for learning the digital logic and computer architectures, the simulation at machine instruction set architecture level is more interesting for simulating complete computers. Due to the uniform instruction timing and encoding, writing a processor emulator is trivial, so more focus can be given to emulating peripheral devices and writing system software for an imaginary computer with this processor. For this purpose, we have written a cross-assembler for SVEU16 and implemented the programming language Forth (on some systems used as firmware or OS) in SVEU16 assembly.

Synthesis: The processor is thus structurally described and simulated to the level of elementary logic gates. It is possible to convert each logic gate into a set of transistors and realize a processor from discrete or integrated components, but this is difficult, expensive and unnecessary work. The easier and more modern way is to describe the processor at a behavior level. SVEU16 was therefore also described in just 140 lines of RTL Verilog and tested on FPGA using Altera development board. Such a small definition of CPU makes it useful for systems on a chip, which need more cells available for other system parts.

B. Educational usage

SVEU16 was used in two courses at the Faculty of Electrical Engineering of the University of Sarajevo. At one of them, which is an introduction to digital logic, students were given various tasks related to this processor: coding instructions, running the simulator, explaining the operation of individual components, drawing parts of the schematics, determining multiplexer signals from the operational code and evaluating expansion possibilities. In another course, which deals with virtualization and distributed operating systems, students wrote computer emulators with this instruction set, but with different configurations of peripheral devices with minimal software in ROM. The tasks on both courses were mostly successfully completed.

IV. Conclusion

SVEU16 is simple, but a powerful enough implementation of a complete and functional CPU. It was used, by this paper's authors, as an educational CPU at logical and ISA level for students enrolled in two courses at the Faculty of Electrical Engineering, University of Sarajevo. In this paper, CPU instruction set, components and simulations are explained in detail in order to help students to understand how a CPU works, and also to give other teachers and researchers some new ideas and perspectives in basic CPU implementations and simulations.

References

[1] A. S. Tanenbaum, *Structured computer organization*, 3rd ed. Prentice-Hall, Inc., 1989.

[2] D. C. Schuurman, "Step-by-step design and simulation of a simple cpu architecture," in *Proceeding of the 44th ACM technical symposium on Computer science education*, 2013, pp. 335–340.

[3] M. H. H. Ichsan and W. Kurniawan, "Design and implementation 8 bit cpu architecture on logisim for undergraduate learning support," in *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*. IEEE, 2017, pp. 132–137.

[4] A. S. Tanenbaum, *Structured Computer Organization*, 5th ed. Pearson Higher Education, 2006.

[5] Y. N. Patt and S. J. Patel, *Introduction to Computing Systems; Bits & Gates to C C++ and Beyond*. McGraw Hill, 2005.

[6] T. Boscke. (2007) Mproz3 – a 16 bit minimal processor. [Online]. Available: http://www.ikomi.de/pub/mproz/mproz3_e.pdf

[7] D. A. Patterson and J. L. Hennessy, *Computer organization and design: the hardware/software interface*. Elsevier, Inc, 1998.

[8] D. Ellard, D. Holland, N. Murphy, and M. Seltzer, "On the design of a new cpu architecture for pedagogical purposes," in *Proceedings of the 2002 workshop on Computer architecture education, 29th International Symposium on Computer Architecture*, 2002, pp. 6–es.

[9] S. Williams and M. Baxter, "Icarus verilog: open-source verilog more than a year later," *Linux Journal*, vol. 2002, no. 99, p. 3, 2002.

[10] W. Snyder. (2023) Verilator. [Online]. Available: https://www.veripool.org/ftp/verilator_doc.pdf

[11] R. Mukherjee, M. Tautschnig, and D. Kroening, "v2c–a verilog to c translator," in *Tools and Algorithms for the Construction and Analysis of Systems: 22nd International Conference, TACAS 2016, Eindhoven, The Netherlands*. Springer, 2016, pp. 580–586.

Student Perspectives on Source-Code Plagiarism: Case Study of Three Programming Courses

Matija Novak
University of Zagreb / Faculty of Organization and Informatics, Varaždin, Croatia
matija.novak@foi.unizg.hr

***Abstract* - Plagiarism is a serious offence and students plagiarise for various reasons. But sometimes students plagiarise and don't even know that they are doing it. In this research a case study was conducted where questions, which describe different situations, were given to students. Questions are taken from an existing study which was conducted in the UK to get the perspective of academics. Similar questions were later given to students in the UK and China to get the student perspective. In this case, Croatian students from three programming courses (at three different levels) were questioned. In this paper the most interesting findings of student perspectives on plagiarism are presented together with some comparisons to the previous studies.**

***Keywords* - plagiarism; academic offence; student perspective; programming**

I. INTRODUCTION

Every teacher has to deal with plagiarism on a regular basis. There are two main aspects while dealing with plagiarism: detection and prevention. While detection is necessary, the first step in dealing with plagiarism should be prevention. There are many techniques to prevent plagiarism [1] and academic dishonesty in general. One technique is to educate students on that topic. It might seem that it is obvious what plagiarism is but research like [2] shows that this is not the case.

This is especially true when talking about source-code plagiarism which is the focus of this paper. Simple definition of plagiarism, that can be found in dictionaries, is that plagiarism is "the act of taking the writings of another person and passing them off as one's own" [3]. While this is correct there are different aspects that are not covered, especially when talking about source-code plagiarism.

Therefore, researchers try to define source-code plagiarism more precisely. A good definition is given by Cosma and Joy in [4] which is: "*Source-code plagiarism in programming assignments can occur when a student reuses ... source-code authored by someone else and, intentionally or unintentionally, fails to acknowledge it adequately ... , thus submitting it as his/her own work. This involves obtaining ... the source-code, either with or without the permission of the original author, and reusing ... source-code produced as part of another assessment (in which academic credit was gained) without adequate acknowledgement ... The latter practice, self-plagiarism, may constitute another academic offence*".

In this paper the goal is to find out the student perspective on plagiarism. This study is conducted in Croatia on three programming courses that are on three different levels. *The main idea is to compare whether there are any differences in perception between students in these courses.* In addition, the results are compared to existing studies which analyse academic (in UK) and student perspectives (in UK and China).

The rest of the article is structured as follows. Section 2 describes the related work and Section 3 the methodology. Results are presented in Section 4 and in Section 5 discussion is given. Section 6 concludes.

II. RELATED WORK

There are different studies [5,6,7,8] that analyse student and academic perspectives on source-code plagiarism. In 2006, Cosma and Joy [6] conducted a survey about academic perspective on source-code plagiarism. The goal was to find out what constitutes source-code plagiarism. Two issues have received controversial responses regarding source-code reuse and self-plagiarism. The definition of source-code plagiarism given in introduction was partially based on this study.

Later in 2011, Joy, Cosma, Yau and Sinclar [7] conducted a similar survey on student perspectives. Based on the results they divided the students into three groups: the ones that plagiarise deliberately, those that do so negligently and those who genuinely do not understand what plagiarism means. Most misunderstood issues were about source-code reuse from previous assignments, citation, limits of collaboration and when code was converted to a different language. When compared to the academic perception the common issue is code reuse.

In [6,7] the research was done in the UK, while in 2014 a study [8] was made which compared student perspectives between UK and Chinese students. This study showed that, in most areas, there is a common understanding and misunderstanding of what constitutes source-code plagiarism. Although, there were three topics that UK students understand and students from China don't. These were: copying code from a book without acknowledgement, paying someone, and copying and submitting another student's work without consent.

In recent years, while there are studies [9,10] that deal with student perspective on plagiarism focusing on text, there are no new studies focusing on source-code.

III. METHODOLOGY

In this study a survey was conducted on three programming courses in Croatia. Course A is a course at a first year university graduate level. Course B is a course at third year university undergraduate level. Course C is a third year professional undergraduate level. Students in all three courses had already passed some programming courses so it is expected that they know something about source-code plagiarism.

At first the idea was to use the same questions that were used in [7] or [8]. Unfortunately, the original survey and exact questions used were not available. Fortunately, the complete questions used in [6] were found in [11] which is a detailed report of the whole study. While [6] focuses on academics from the papers it is visible that the questions in [7,8] are similar to questions in [6] so that it is possible to compare some results.

A. Survey execution and layout

The survey was performed online over Learning Management System (LMS) Moodle. Access to the survey had only students that enrolled in the course. *The survey was translated to Croatian to eliminate the problem of misunderstanding the questions.*

The survey was divided into five parts. First part was just one yes/no question: "*Do you know what the term source-code plagiarism means?*".

Second part gave the students two definitions of source-code plagiarism. One was the definition presented in the introduction. The second definition was [12]:

"*Plagiarism, in programming assignments, is the act of taking a significant amount of source-code parts (up to the entire source-code) from other students or from the Internet and using it without noting its origin. A 'significant amount' means that the similarity between two solutions of a programming assignment is high enough that an expert (teacher, ethical board, etc.) considers specific student work as sufficiently 'real' plagiarism to accuse the student of plagiarism.*"

These two definitions were considered most accurate on what constitutes as source-code plagiarism. If interested in other definitions there is a systematic literature review paper [13] which gives good overview of definitions about source-code plagiarism. The reason for giving students these definitions was to give them understanding of what was meant by source-code plagiarism. There was a suspicion, before starting the survey, that some students are totally ignorant of what source-code plagiarism means, which was confirmed.

One might argue that these definitions do not include open source-code, but this will depend on the course. If using open source-code is allowed then this can not be considered plagiarism. For example, in the case of the three courses students need to make the assignments themselves from scratch. Now it is not forbidden to gather ideas from others, but if they do use code from someone else they have a form to fill out where they can state which parts of the code were taken and from where.

Third part contained three questions about what source-code plagiarism can involve. *These questions had the answers Agree, Disagree, Neither Agree or Disagree (i.e. don't know)*. Part four contained 15 questions (or better said statements) that describe different situations. *These questions had answers which are: Academic offence, Not an academic offence and don't know*. Here a note was given: "When you answer the following questions, think about whether this would be considered an academic offence or not, in a graded assignment."

Part five contained only two questions which were:

1. What do you think is the minimum percentage that your code needs to be similar to another student so that the teacher starts suspecting you of plagiarism and looks into it in more detail? Answer from 0% to 100%.
2. Here you write any comments if you like.

B. Questions of part three and four

The questions that were given to the students were not all questions that exist in [11]. The reason for this was that some questions were not that relevant since the students do not have these parts in their assignments in these courses. For example: "*Plagiarism in programming assignments can involve documentation of computer programs.*" In courses where the study is done students do not have to write documentation. Also, some questions deal with non-graded assignments which were not the focus of this study. This reduction also had another purpose to limit the number of questions. If the survey is longer, less students want to spend the time to answer it.

The questions in part three were:

1. Plagiarism in Programming assignment can involve the Source-code of a computer program.
2. Plagiarism in Programming assignment can involve the Comments within the source code.
3. Plagiarism in Programming assignment can involve the User-interface of a computer program.

The questions in part four were:

1. A student reproduces/copies someone else source-code without making any alterations and submits it without providing any acknowledgements.
2. A student reproduces/ copies someone else source-code. Adapts the code to his/her own work and submits it without providing any acknowledgements.
3. A student pays another person (other than a student on the same module) to create part or whole of source-code and submits it as his/her own work.
4. A student pays a fellow student on the same module to create part or whole of source-code and submits it as his/her own work.

5. A student steals another student's source-code and submits it as his/her own work without making any alterations.
6. A student steals another student's source-code, edits and submits it as his/her own work
7. For a group assignment, students between different groups exchange parts of source-code with the consent of their fellow group members, and integrate the borrowed source code within their work as if it was that group's own work.
8. For a group assignment, students between different groups exchange parts of source-code without their fellow group members knowing, and integrate the borrowed source code within their work as if it was that group's own work.
9. Assume that students were not allowed to resubmit material they had originally created and submitted previously for another assignment. For a graded assignment, a student has copied parts of source-code that he had produced for another assignment without acknowledging it.
10. Two students work together for programming for a programming assignment that requires students to work individually and students submit very similar source-codes.
11. For an assignment, the student has copied source-code from a book and has intentionally not provided any acknowledgments.
12. For an assignment, the student has copied source-code from a book and has unintentionally not provided any acknowledgments.
13. Providing pretend reference (i.e. references that were made-up by the student that do not exist)
14. Providing false reference (i.e. referenced exist but not match the source-code that was copied)
15. Modifying the program output to make it seem as if the program works.

IV. RESULTS

In total 206 students filled out the survey, the response rate of the survey was as follows:

- Course A: 87.5% response rate, *70 responses* out of 80 students;
- Course B: 70.4% response rate, *95 responses* out of 135 students;
- Course C: 77.6% response rate, *41 response* out of 53 students.

The results were first analysed together and then per course, the results per course are further presented in parentheses. Afterwards, to see if there are any differences in the perception of students between the three courses the statistical analysis was performed using chi-square test.

The hypothesis for each question was: Null hypothesis (H0) – The student perception is not dependent on the course. Alternative hypothesis (H1) – The student perception is dependent on the course.

Regarding question 1 (in the first part of the survey) where students were asked if they know what source-code plagiarism means on average 85% answered yes (Course A - 91.4%, Course B – 86.3%, and Course C – 78%). These results are not unexpected since the students are at least three years into their study program. Based on this question it can be said that at least most of them have heard of the term source-code plagiarism. Based on the responses of other questions it is clear that they do not understand the meaning of source-code plagiarism completely.

A. Results of part three questions

The results of the survey's third part were interesting. First question in this part was that source-code plagiarism can involve source-code. It was expected that everyone would agree, but surprisingly only 71.5% agreed (Course A – 78.6%, Course B – 72.6% and Course C – 63.4%). Less than 10% answered that they don't know. First explanation was that students did not honestly fill out the survey but by analysing other questions from these respondents it does not seem so. Especially, some of the really good comments in the last question came from students that answered this question with disagreement. It might be that they do not know what source-code plagiarism is. But, since definitions were given to them before this part, the only explanation might be they did not understand the question.

Regarding comments, around 38% of students agree that it might be part of source-code plagiarism (Course A – 28.6%, Course B – 45.3%, Course C – 41.5%) and 21% don't know. Regarding user interface, around 52% (Course A – 62.9%, Course B – 56.8% and Course C – 36.6%) agreed that it can involve source-code plagiarism and 16% don't know.

When this is compared to what academics in the UK have answered [11] (comments ~68% agree, user interface ~60% agree) it is not surprising that students have lower agreement. What is interesting is that Course A, which is on a university graduate level, has the lowest agreement for comments (28.6%) and highest for user interface (62.9%). On the other hand, Course C, which is a professional undergraduate, is exactly opposite lowest for user interface (36.6%) and higher for comments (41.5%).

The explanation might be that students in higher years know that comments can be easily deleted and modified without changing the logic. Also they already know that comments are not part of the real functionality while the user interface is. Also, such students build quite complex user interfaces. Therefore, source-code plagiarism highly involves user interface and not so much comments. Students in Course C still use comments to explain code for themself and see it as a possibility an indication of plagiarism if these comments are the same. In addition, these students build quite simple user interfaces so plagiarism is more likely to happen in comments then in user interfaces.

TABLE I. RESULTS OF THE THIRD PART

Q	Percentage (%)							
	Course A		*Course B*		*Course C*		*Total*	
	AO	NO	AO	NO	AO	NO	AO	NO
1	93	6	100	0	95	5	96	4
2	47	37	55	34	68	22	57	31
3	**74**	**13**	**62**	**16**	**68**	**27**	68	19
4	76	11	66	16	68	27	70	18
5	**97**	**3**	**100**	**0**	**83**	**17**	93	7
6	77	10	76	10	73	20	75	13
7	**19**	**54**	**30**	**55**	**44**	**44**	31	51
8	77	16	74	11	73	17	75	14
9	**36**	**50**	**31**	**46**	**54**	**24**	40	40
10	33	44	23	55	24	56	27	52
11	**43**	**34**	**85**	**10**	**90**	**7**	73	17
12	43	34	41	36	51	27	45	32
13	64	13	68	10	76	10	69	11
14	49	16	58	13	54	15	53	14
15	33	43	33	44	24	46	30	44

Q-question, AO- academic offence, NO – not an academic offence
marked bold – significant differences between courses

B. Results of part four questions

This part of the survey had most questions (Q). The results (in percentages) are presented in Table 1 for answers plagiarised (academic offence - AO) and not plagiarised (not an academic offence - NO). The terms plagiarism and academic offence are in this section used interchangeably. While it is clear that they are not the same, in this study there were no answers about other academic offences. The don't know is not given because of the space but can easily be calculated since the three answers add up to 100%. Most questions do not have a significant statistical difference between groups (course). For those that have it will be explicitly stated what the χ^2 and p value are.

For Q1 (copying without modification) there is a clear understanding (96% agree) that it is AO. For Q2 (copying with modification) there is a big decrease in agreement (57% agree) that this is AO. Now the reason is explained in the last question in part five. Students could not decide how big the modification is. And this is something that corresponds to the answers of academics in [11] that it depends on the degree of adaptation. Now in the three courses the rule is that no adaptation is allowed so it makes sense that many answered that it is plagiarism.

For Q3 (paying third person) and Q4 (paying another student) the results are almost equal. Around 70% agree that it is AO. This percentage is almost the same in [11] where ~75% agree. Here the academics say that this is not plagiarism rather some other academic offence so students might have the same reasoning. Although, this was not stated in the comments so we don't know. In this question there is a significant difference in perception between students of the three courses with $\chi^2(4)=9.972, p=0.041$.

For Q5 (stealing without modification) the understanding is clear 93% say it is AO, while for Q6 (stealing with modification) this drops to 75%. Here again the explanation is similar as for the difference between Q1 and Q2 and that is the unknown factor how big the adaptation is. The academics on both questions were around 83% that it is plagiarism. What is interesting is that I Q5 there exists a difference between groups (courses) with $\chi^2(2)=20.561, p<0.001$**.** This is obvious because in Course 3 there is 17% that say this is NO, while in Course A 0% say NO and Course B 3% say NO. The explanation might be that students on the university level are much more ethical. Similar observation is present in Q6, but there is no significant difference between courses.

Regarding group work Q7 (code exchange with knowledge of the group) and Q8 (code exchange without knowledge of the group) there is a big difference. When other team members know that the exchange is taking place only 30% say this is AO. On the contrary when it is done without the knowledge of team members 74% say this is plagiarism. On both questions academics here agree with ~83% that it is plagiarism. From this question it can be said that students do not understand that AO is taking place even when the other side agrees to the exchange. Also, it seems that students don't understand the boundaries of what is allowed. As with Q5 there is a significant difference between groups (courses) in Q7 with $\chi^2(4)=10.5177, p=0.033$. In the two university courses there is a high percentage that it is not an academic offence. In Course A only 19% and in course B only 30% think it is AO, while in Course C 44% think it is AO. It is around 50% for all three courses that it is NO, the difference comes from the don't know answers. The reason for the difference might be that students at the university level have much more courses that have group work and on the professional level it is more individual projects.

Question Q9 goes into the domain of self-plagiarism which academics in [11] don't agree with and only 29% think this is plagiarism. Students here are more strict with 40% saying that it is plagiarism and 20% don't know.

Question 10 (two students working together while they should be working alone) has only 27% say this is plagiarism and 22% don't know, while ~68% academics say it is. This question confirms the statement made in question Q7 that there is a lack of understanding of the boundaries of what is allowed.

Questions 11-14 are about copying from books and citations. In question Q11 (copying from book without citation intentionally) 73% agree that it is AO, but there is a significant difference between courses with $\chi^2(4)=44.657, p<0.001$. The difference is mostly due to Course A where only 43% say it is AO and 23% don't know while in the other two courses it is more than 85% that it is AO.

In question 12 (copying from book without citation unintentionally) 45% say it is AO while ~68% of academics [11] say it is. Here it is shown that students are much more relaxed. In question 13 (copying from book with intentionally putting made up references) 69% think

it is AO and in Q14 (copying from book with intentionally putting wrong references) 53% think it is AO. Q13 and Q14 have lower agreement that it is AO than for Q11 but higher than Q12. When compared to Q11-Q14 it can be concluded that students think it worse when it is intentional the same as academics. Still, academics are much more strict since from their point of view how can one know if it was intentional or not. Also, referencing something (Q13 and Q14) even if it is not correct, more students are uncertain whether this is AO (larger percentage ~25% don't know answers) while when nothing is cited it is more clear to them that it is AO.

Regarding modifying the output 30% think it is AO while 25% don't know. Here academics [11] say ~20% that it is and the reasoning is that it is not plagiarism rather than some other kind of cheating. Some comments from students also state that this question is not plagiarism but rather an incorrect implementation.

C. *Results of part five questions*

In this last part the question was how big the similarity in percentages should be to be considered plagiarism. Answers in Course A ranged from 5% to 90% with an average of 56%. Answers in Course B ranged from 5 to 90% with an average of 59%. Answers in Course C ranged from 5% to 100% with an average of 44%.

In Figure 1 student answers are shown in graph form. The graph shows the frequency of percentages in ranges of 10 percent. From this graph it is visible that Course C is different from Course A and Course B. In Course A and Course B more students gave higher percentages while in Course C more students gave lower percentages. What is visible from this is that there is no agreement and as with academics [11] it depends.

The last question were comments and partially they were mentioned to explain the results of previous questions. While most students did not write anything there were some good observations.

Some comments did go into the direction that they think that students who study together, have the same courses and learn from the same sources will produce the same code. Such a statement is not correct size just by using the same sources for learning it will probably produce higher similarities but not the same. Also, when talking in context of the three courses where projects are relatively large the overall similarity will be smaller than maybe in courses with small assignments. Usually students see only one project, maybe two and do not have the overview of all projects in course as teachers so it is not unexpected that the opinions differ.

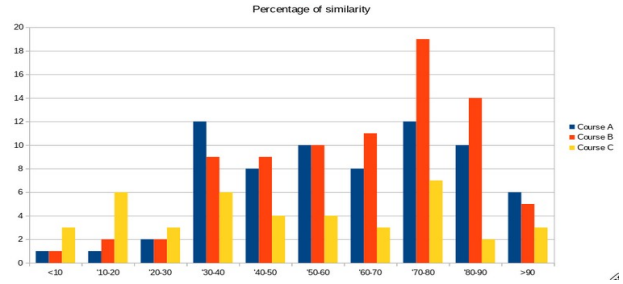


Figure 1. Frequency of similarity percentage in student responses

Other comments got into discussion that just if something is similar it is not necessarily plagiarism. While this is true and agrees with many studies, it is going into the domain of detection and this is not the focus here.

One issue that was pointed out regarding the questions itself was that it was not clear how much copying/adapting are the questions talking about. This is fair comment and such questions could be split, in the future, into multiple questions where the amount is specified.

Some comments talked about code reuse and copying from websites like stack-overflow and how they think it is not needed to be cited. Here again is the problem of how much is necessary to be cited. Another issue is what is the correct way to cite source-code.

Good suggestions in the comments were that students should be educated. Another was that if there is suspicion of plagiarism they should be asked to explain the code written. Explaining the code is a partial solution since it could be that the student has the knowledge to explain but does not know how to write it from scratch. Also, it could be that it was just to make a shortcut to do the assignment faster. Are such situations then acceptable?

V. DISCUSSION

In the previous section the results were discussed on a per question basis in this section a discussion is given based on all results. Also it compares the results to [8]. As it can be seen, questions describe different situations: copying from students, paying someone for doing their assignments, stealing, group work, self plagiarism and modifying output.

Based on the first question the assumption that there are students that do not know what is meant by source-code plagiarism is confirmed. Around 15% of students do not know what is source-code plagiarism. Although 85% say they know what source-code plagiarism means, by analysing other questions it was shown that this understanding is limited.

In the case of copying (Q1) and stealing (Q5) without modification the understanding is excellent and almost all students (above 90%) know this is considered plagiarism. In [8] such situations are understood by 80% of UK students and only 43% of Chinese students.

When modification is involved (Q2, Q4, Q6, Q8) or when talking about paying someone (Q3) there is less understanding, but still many students (approximately 60%) understand that. When looking at [8] situations of modification in the UK understanding is around 70% and in China 57%. While paying someone in the UK 80% of students understand and in China only 16%.

Copying from books (Q11) when no reference is given most students understand (around 73%). On the other hand, in situations where it was copying from books and giving false references (Q14), made up references (Q13) or unintentional not referencing (Q12) the understanding is moderate (between 45-70%). According to [8] regarding copying from books without referencing 81% of UK students understand this is plagiarism but only 26%

from China understand that. On the topic of incorrect references in the UK 11% understand and in China 10%.

The most problematic domain for students in the three courses was collaboration (Q7, Q10) where the understanding is quite low (around 30%). According to [8] the situation of collaboration on individual work is 34% in the UK and 8% in China, while group collaboration is 32% in the UK and 14% in China.

Regarding self plagiarism (Q9) understanding is moderate (around 40%). Based on [8] in the UK the understanding is around 3% and in China around 4%.

Lastly, the modification of program output is considered plagiarism by 30% and 44% do not think it is plagiarism. In [8] such a situation is not considered plagiarism and 80% of UK students and 45% of Chinese students understand.

From these results there is the *indication* that Croatian students are better in understanding self-plagiarism, copying and stealing without modification and collaboration than students from the UK and China. On the other hand they have worse understanding when copying with modification is done or copying from books then students in the UK but better than Chinese students. *These results are not completely accurate since the questions are similar but not exactly the same.* Nonetheless, a good indication is given what to expect.

Overall students in Croatia in the three courses have similar understandings of source-code plagiarism regardless of the study program. The only significant differences appear in cases of: paying someone (Q3), stealing without modification (Q5), code exchange in group work with knowledge of the group (Q7) and copying from a book without citation intentionally (Q11).

Additionally, based on questions in part three in all three courses students have similar level of understanding that source-code plagiarism can involve.

VI. Future work

This survey gave a good first view on students' perspective about source-code plagiarism. In the future it is planned to make additional surveys on other programming courses to see if the results are consistent.

As indicated in the comments, some questions need to be more precise to get more clear responses so in future it is planned to make this improvement.

Another possibility for future work is to give students real source-code which is considered plagiarism by academics and code that is not considered plagiarism and to ask them to make a decision. This would give inside if there is an agreement.

VII. Conclusion

Not all students know what source-code plagiarism means and those that do know, do not understand it completely. Plagiarism is still a highly researched topic and even academics don't agree on all topics what constitutes plagiarism so it is of no surprise that students don't know.

Based on the presented results it can be said that students form Croatia in the three programming courses have a solid understanding of what constitutes source-code plagiarism in most cases. Almost all students understand that copying and stealing is plagiarism but when it comes to modifications the understanding starts to drop. In comparison to students in the UK and China it seems that Croatian students are somewhere in between.

The most misunderstood topic is about collaboration, self-plagiarism and where can source-code plagiarism occur. It can be stated based on this research that students are not adequately educated on the topic of source-code plagiarism. Nonetheless, this is not excuse for allowing plagiarism. While special attention needs to be about what are the boundaries of collaboration on a particular course.

References

[1] T. Hodgkinson, H. Curtis, D. MacAlister, and G. Farrell "Student academic dishonesty: The potential for situational prevention" *J. Crim. Just. Educ.* 27(1), pp. 1–18, 2016 DOI:https://doi.org/10.1080/10511253.2015.1064982

[2] Drisko, J. W. "What Is Plagiarism, How to Identify It, and How to Educate to Avoid It" *Journal of Social Work Education*, pp. 1-12, 2022

[3] Encyclopædia Britannica Inc, "plagiarism" *Encyclopædia Britannica*, 2015. [Online]. Available: http://www.britannica.com/topic/plagiarism. [Accessed: 30-Jan-2023].

[4] Cosma, G., and Joy, M. "Towards a definition of source-code plagiarism" *IEEE Transactions on Education*, 51(2), pp. 195-200, 2008

[5] Yeo, S. "First-year university science and engineering students' understanding of plagiarism" *High Education Research & Development*, 26(2), pp. 199-216, 2006

[6] Cosma, G., and Joy, M. "Source-Code Plagiarism : a Uk Academic Perspective" *Annual Conference of the HEA Network for Information and Computer Science* pp. 116–120, 2006

[7] Joy, M., Cosma, G., Yau, J. Y.-K., and Sinclair, J. "Source Code Plagiarism - A Student Perspective" *IEEE Transactions on Education*, 54(1), pp. 125–132, 2011 DOI:https://doi.org/10.1109/TE.2010.2046664

[8] Zhang, D., Joy, M., Cosma, G., Boyatt, R., Sinclair, J., and Yau, J. "Source-code plagiarism in universities: a comparative study of student perspectives in China and the UK" *Assessment & Evaluation in Higher Education*, 39(6), pp. 743–758, 2014 DOI:https://doi.org/10.1080/02602938.2013.870122

[9] Murugathas, K., and Santharooban, S. (2022). "Students' perspectives on plagiarism detection" *VINGNANAM Research Conference*, Shri Lanka, 2002

[10] Raisyah, N. "Understanding on Plagiarism from English Language Education Students' Perspective", Doctoral dissertation, *Universitas Islam Kalimantan MAB*, 2022

[11] Georgina, C., and Mike, J. "Source-code plagiarism: A UK academic perspective" *Research Report CS-RR-422,* Department of Computer Science, University of Warwick, 2006, Available: https://wrap.warwick.ac.uk/61520/1/WRAP_Joy_cs-rr-422.pdf [Accessed: 30-Jan-2023]

[12] Author, Author. "Title" Doctoral dissertation, *University of xxx. Faculty of xxx,* year

[13] Novak, M., Joy, M., and Kermek, D. "Source-code similarity detection and detection tools used in academia: a systematic review." *ACM Transactions on Computing Education*, 19(3), pp. 1-37, 2019

Utilizing Robotics in Maths in Primary Education: A Case Study

A. Goltsiou, C. Sofianopoulou
Harokopio University/Department of Informatics and Telematics, Athens, Greece
aikgoltsiou@gmail.com, csofian@hua.gr

***Abstract* - The use of robotics and programming help to approach mathematical concepts in an alternative way through differentiated learning. The approach contributes to the cultivation of mathematical thinking through experimentation and code correction, through visualization, as well as the inclusion of students. The aim of the paper is to present how robotics and specifically the Lego Education Spike Essential were integrated into the teaching of Geometry in the 1st Grade of a Greek primary school. The activities were included in three teaching scenarios in the framework of the utilization of the new Greek curriculum. The evaluation of the action was done by investigating the opinions and attitudes of the students and with structured observation by the teacher. The students' formative and final assessment were integrated into Moodle. Also, the programming of students' constructions formed the final evaluation of the implementation. From the evaluation of the action, it emerged that the students understood mathematical concepts, they cultivated collaborative, digital and programming skills in a pleasant collaborative learning environment. It is proposed to utilize robotics from an early age in the teaching of mathematics in the new Greek curricula, as a means of cultivating students' mathematical thinking.**

Keywords - robotics; Lego Education Spike; mathematics; differentiated learning; Moodle

I. INTRODUCTION

The new Greek school curricula are oriented towards learning outcomes through a combination of experiential learning and digital educational tools [1]. The approach is two-fold and achieves the differentiation of learning material and implementation, in accordance with what differentiated pedagogy stands for. There are many individual differences between students, which increase as students progress from one level of education to another [2]. Accepting diversity and moving towards an inclusive school is an international orientation [3]. Gardner [4] with the theory of multiple intelligences supports the different ways of perceiving and understanding the concepts of everyday life by each person individually. Through differentiated pedagogy, a flexible framework is formed, which allows each student to follow his own way of learning in a collaborative learning environment. [5].

Cultivating students' critical thinking is achieved by making them active and engaging with constructions that are meaningful to them. Since the 1990s it has been argued that those students who are involved in programming, who write code are able to understand concepts through the process of experimenting with their ideas. Being able to correct errors in the code fosters critical and mathematical thinking [6].

By utilizing ICT in teaching practice, students' critical thinking is cultivated, through visualization, investigation and the immediate feedback they receive. The participation of each student is immediately visible and motivates the members of the group in cooperative teaching and learning, where teaching can be implemented with blended learning [7].

A. The use of robotics at school-Playing with robots

There are many digital applications used in education. Moodle is a modular learning environment (https://moodle.com/) with module configuration that makes it flexible. It has mainly been used in adult education. In recent years, it has also been utilized by Secondary and Primary Education, contributing to the cultivation of cognitive skills, communication, collaboration, and digital literacy [8, 9, 10]. Among the Moodle resources there are now the interactive .H5P files. The activation of students in teaching is particularly achieved by the use of .H5P files [11, 12] which contribute to the improvement of their performance.

The colored cards of Cody and Roby, created by Prof. Bogliol [13] were exploited in Italy along with other forms of programming learning activities, without the use of computers and tablets. These are unplugged activities. From the action of cultivating digital literacy, it was seen that learning became more interesting and effective, as the students were active in everyday issues. In addition, they improved metacognitive skills as they trying to find a solution. L. Klopfenstein et al. [14] used Cody and Roby cards in an augmented reality game and they emphasized the exciting atmosphere created by the action. Fabrizio Ferrari et al. [15] used Cody and Roby as a pre-organizer to introduce students to coding with Scratch. The game contributed to the development of programming skills and as a cognitive scaffold learning Scratch.

With game-centered learning [10] Vygotsky argues that a greater participation of students is achieved. The Lego Education Spike Essential kit has distance and call sensors. It is programmed by students using either its own block language or the Scratch programming language. Through its use, exploratory learning and understanding of cause and effect is promoted (https://education.lego.com/en-us).

Pinto-Llorente [16] used the Lego Education WeDo 2.0 robotics kits and documented that contributed to the cultivation of children's computational thinking. In addition, engaging with the kits contributed to the activation of the students in primary education, as they had the opportunity to learn how to build 3D models and program them. Veselovská et al. [17] utilized the robotics equipment, Lego Education WeDo 2.0, in 3rd and 4th grade students of Primary Education, who mainly focused on the construction of robotic models, with the aim of integrating robotics into the school curriculum and comparing it with the results of countries with different cultural environments. Wu et al. [18] in their comparative research on Lego Education WeDo 2.0 students from 1st to 4th grade in Taiwan and Slovakia, they concluded that educational robotics contributes to the cultivation of communication, cooperation and engineering skills in students, activating them through inquiry learning.

Petrovič [19] utilized the Lego Education Spike Prime robotics kit in the subject of Physics and suggests extending the use of robotics to other subjects in the curriculum and to mathematics. His suggestions concern the implementation of activities in Physics and mathematics.

Nikolos [20] proposed the use of the Scratch 3.0 language to program the Lego Education WeDo 2.0 robotics kit, for the implementation of a Primary Education teaching scenario. Lego Education WeDo 2.0 kits are leveraged in STEM Methodology as Lytra & Drigas said in systematic review [21].

The aim of this article is to present how robotics and specifically Lego Education Spike Essential were integrated in the teaching of geometry at the 1st Grade of a primary education Greek school. The question is:

Although how can robotics be integrated into the teaching of geometry at the 1st Grade in Primary Education.

II. METHODOLOGY

To investigate the question, a project was designed and implemented in the teaching of geometry during the school year 2022-23. The implementation was done in 15 students, 7 boys and 8 girls, aged 6 years, at the 1st grade class of a Greek primary school. A total of eight (8) teaching hours, two (2) mathematics hours of the program timetable were used for each activity. The activities were part of three teaching scenarios, as extension activities in the daily life of the students. The heterogeneity of the class from a socio-cultural point of view as well as the different learning profiles of the students were taken into account [4] in the context of differentiated teaching [5]. The activities were structured forming a cognitive scaffold for the next activity. At the end of each activity, a formative assessment was done in Moodle, as a differentiation of the students' evaluation method [5]. The action was completed over a two-week period.

Students have had experience in collaborative experiential learning and blended learning in Moodle since the beginning of the school year. They had used the robotics kit, Cody and Roby cards from previous STEM activities with their class teacher. They also knew how to use the Lego Education Spike app on tablets. Students participated in the Lego programming activities in groups and in the plenary in the classroom that means to be attended by all students in the class. Specifically, students collaborated in groups of 5 people, where two students programmed on the tablets taking turns with the others. The rest of the students in the group were assistants and evaluated the programming scenario during its construction. That means that in each activity two students were programmers and the other evaluators. The unplugged activity was implemented in plenary.

During the action they were implemented:

A. Unplugged activity with Cody and Roby cards.

B. Robot programming in straight line, square and rectangle path.

C. Evaluation

- Structured observation by the teacher.
- Student self-assessment in Moodle.
- Investigating students' opinions and attitudes after implementation.

The interactive content (.H5P) files and the quizzes of the Moodle environment were used for the formative assessment. The initial and final evaluation of the action was done with structured observation by the teacher based on the objectives of the activities, investigation with questionnaires of students' opinions and attitudes about the Moodle digital environment and engagement with robotics in mathematics, the activation of the students in the group, the programming scenarios and the constructions completed by the students.

A. *Action flow - implementationοή*

The subject of the action was the consolidation of concepts of geometry with an alternative approach in the context of differentiated learning and with the cultivation of digital skills and programming. Each activity was a cognitive scaffolding for the next one, both at the level of geometry and at the level of programming. After teaching each concept in geometry, the students recalled the characteristic elements of the shapes and programmed the movement of each object with blocks in the application of the tablets to form a line, square or rectangle. The activities:

- 1η Programming robots with straight line motion.
- 2η Unplugged activity with Cody and Roby cards following a specific path.
- 3η Programming robot that moves in a square-shape.
- 4η Programming a robot that moves in a rectangle shape.

The students were engaged in the construction of the robot-car taking turns per group. They chose the construction they would make from the Lego Education Spike library in plenary. The building instructions were projected onto the classroom and groups constructed part of the car while the other groups programmed the construction. So all the groups built and programmed the

same object. When construction was complete, they tested their programming script, improved it and experimented with adding picture and music commands. The Unplugged activity with Cody Roby cards was used to facilitate the students' transition from the hands-on material of the cards to the more abstract of the screen and way of checking the correctness of the commands.

III. WHAT WE HAVE FROM THE IMPLEMENTATION

The students gradually presented the results of the constructions and programming they did to the plenary (Fig. 1) in each activity.



Figure 1. Constructions with Lego Education Spike Essential

The check of the movement of the object was done first in the group and then in a mockup of a route in the city, which built by themselves (Fig. 2). Thus, they connected their construction with everyday life [5], so that they could better understand why they learn to draw shapes.

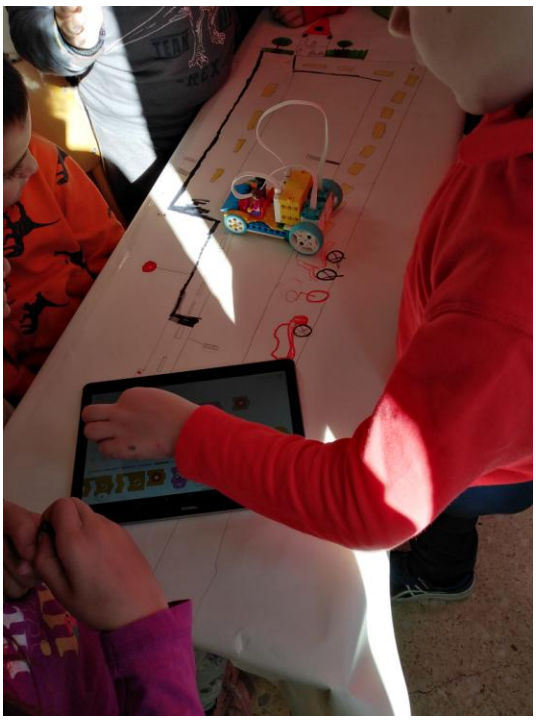


Figure 2. Giving directions

A. *Unplugged activity with Cody and Roby cards*

With the activity Cody and Roby (Fig. 3) a connection of programming with manual material was made. The activity was implemented in plenary by applying peer review.



Figure 3. Giving directions via Cody and Roby cards.

B. *Robot programming in straight line, square and rectangle path*

The unplugged activity served as a scaffold for moving into more complex programming in the Lego Education Spike. The groups created their own script to program the robot that moves in a line, in square-shape and in a rectangle shape (Fig. 4). In fact, they realized that the program would remain the same if they changed the number of steps under each movement command, instead of placing one movement command next to the other. They presented, commented on other groups' scripts, got ideas for their own script and added audio and visual commands.

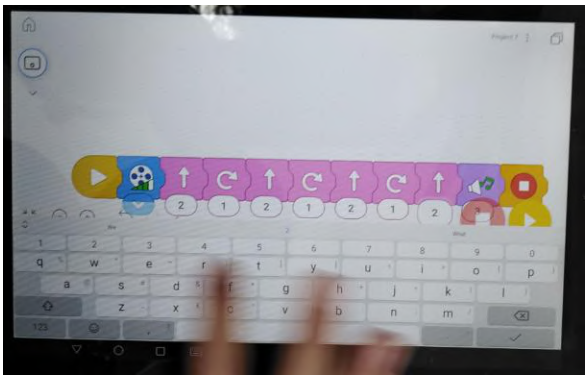


Figure 4. Programming a square

The students perceived the programming of the rectangle as a pattern (Fig. 5), by recalling previous knowledge and using it in a completely different environment.

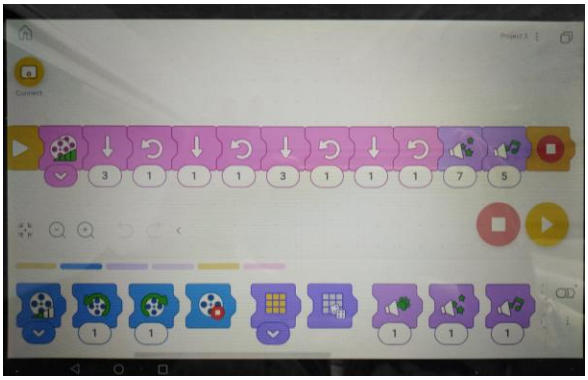


Figure 5. Programming a rectangle

C. Evaluation

The evaluation was about the integration of robotics into the teaching of mathematics. It was conducted with an initial and final questionnaire of the students with structured observation by the teacher based on cognitive goals and cultivating programming skills.

The evaluation of the action was done in combination, through the results of the students' tasks, the quiz, the students' formative assessment tasks and using a questionnaire. The students' opinions, their attitude towards using robotics in the future in the mathematics, the difficulties they had faced and how they were solved, the communication and cooperation in the group were recorded.

From the teacher's initial observation, it emerged that the students had little programming experience with the Lego Education Spike, as they had used the kit with him in a previous STEM activity. The students knew how to program the robot in a straight line. They had cultivated their digital skills due to utilizing the digital classroom since the beginning of the school year. In addition, they knew how to work together experientially. The final rubric observation of the teacher showed that five (5) students of the fifteen (15) students (33.3%) did not identify the error in the programming scenario with the figures. The 46.6% did not know how to correct the error that means seven (7) of fifteen (15) students.

The final evaluation quiz for programming included three (3) questions with pictures, of the matching, multiple choice, true-false type. The matching question had images from the three programs that students had created that was the line, the square and the rectangle. The students had to match the programs with the words straight line, square, rectangle. In the multiple choice questions students had to choose from the arrows what the command meant. In the true-false question students had to decide whether the toy car reached the end of the route or not at the image with Cody and Roby cards. The results showed that 73.3% answered correctly to all the questions.

With an electronic questionnaire, students' opinions and attitudes about robotics mathematics activities and the use of the Moodle environment were investigated. The cooperative activity of the students during the implementation, the completion of the tasks and the interactive content of Moodle constituted the formative assessment. The final evaluation consisted of the students' assignments that were implemented experientially and the final quiz in the digital classroom.

The final investigation of students' opinions and attitudes towards the integration of robotics into the subject of mathematics was recorded in an anonymous electronic questionnaire with dichotomous questions and completed by 14 students. The results are presented in Table 1. Table 1 presents frequency.

TABLE I. POST- QUESTIONNAIR: STUDENTS' OPINIONS

	Students' opinions		
	Questions	***Yes***	***No***
1	I like doing math at school with robotics.	100%	0%
2	I understand geometric shapes better by doing robotics.	100%	0%
3	I see my programming mistakes.	92.9%	7.1%
4	I am correcting my programming mistakes.	85.7%	14.3%
5	I want to continue doing robotics in mathematics.	100%	0%

The final questionnaire highlighted the desire of all the students to continue the involvement with robotics in mathematics. The inability of some students to solve the problem did not deter them from wanting to do math with robotics in the future.

As far as their digital classroom is concerned, 77% of the students say that Moodle is easy to use, helps them learn, they like working on Moodle and want to continue using it.

IV. CONCLUSION

This paper shows how the Lego Education Spike Robotics Kit was used in 1st Grade in primary education to help students learn the shapes square and rectangle better.

The question is how the Lego Education Spike Robotics Kit could be leveraged to help students learn better geometric concepts through programming and cultivate programming skills. During the implementation the teacher observed and recorded the existence of a cooperative climate among the students, a pleasant atmosphere was created and the students learnt geometric concepts with the help of programmable materials. From the questionnaire of students' opinions and attitudes about the involvement with robotics, the students' positive attitude towards the activity emerged. It was observed that students gained programming knowledge through robotics, cultivated critical thinking and problem solving, cultivated 21st century learning skills.

From the teacher's observation rubrics, the investigation of the students' opinions, it emerged that it is possible to utilize robotics in dealing with geometry concepts in primary education. The robotics kit functioned as both a haptic device and a digital tool through its tablet application. The unplugged activity was helpful in understanding application commands. Students benefited, were energized and their creativity and positive attitude towards mathematics was enhanced.

The work supports the studies concerning the use of robotics in mathematical problems and the cultivation of mathematical thinking [13, 14, 15, 16, 17, 18, 19]. It enhances learning through a pleasant atmosphere of collaboration [16, 17, 18, 19, 20] and differentiated [3, 4, 5].

It is proposed to use robotics from an early age, in the teaching of mathematics in the New Greek curriculum, as a means of cultivating students' mathematical thinking. In addition, engaging where possible in the subjects of the rest of the curriculum, as this enhances creativity and mathematical thinking.

V. LIMITATIONS

A Lego Education Spike Essential robotics kit was used to implement the activities in a class of 15 students. It would have been easier to implement if each students group had its own robotics kit. The question arises whether it would be feasible to implement following the same methodology in a more populous class. In addition, it is deemed necessary to train the teachers of each school, in the context of in-school training, with the implementation of laboratory training.

REFERENCES

[1] IEP (2021). http://iep.edu.gr/el/nea-programmata-spoudon-arxiki-selida.

[2] D. Reschley, Learning characteristics of mildly handicapped students:Implications for classification, placement and programming. In: M. Wang, M. Reynolds & H. Walberg (eds). *Handbook of special education. Research and practice. Volume 1: Learner characteristics and adaptive education.*. Pergamon Press, pp. *35-5,*1987.

[3] UNESCO, The Salamanca statement and framework for action on special needs education. Paris: UNESCO, 1994.

[4] H. Gardner, Frames of mind: The Theory of multiple intelligences. New York: Basic Books, 1983.

[5] C. A.,Tomlinson, Grading and Differentiation: Paradox or Good Practice?, *Theory Into Practice, 44:3, 262-269,* 2005

[6] S. Papert, Mindstorms: *Children, Computers, And Powerful Idea,* Basic Books, 1990.

[7] R. Al-Maroof, N. Al-Qaysi, S. A., Salloum & M. Al-Emran, Blended learning acceptance: A systematic review of information systems models. Technology, Knowledge and Learning, 1-36 (2021).

[8] K. Brandl, ARE YOU READY TO "MOODLE"?. *Language Learning & Technology,*19(2), 16-23, 2005.

[9] D. Pérez-Marín, R. Hijón-Neira, A. Bacelo & C. Pizarro, Can computational thinking be improved by using a methodology based on metaphors and scratch to teach computer programming to children? Computers in Human Behavior, 105, 2020.

[10] L. S. Vygotsky, Play and its role in the mental development of the child. Soviet psychology, 5(3), 6-18, 1967.

[11] A. J. Unsworth & M. G Posner, Case Study: Using H5P to design and deliver interactive laboratory practicals. Essays in Biochemistry, 66(1), 19-27, 2022.

[12] R. Singleton & A. Charlton, Creating H5P content for active learning. *Pacific Journal of Technology Enhanced Learning*, *2*(1), 13-14, 2020.

[13] S. Altieri, Recent history of coding in Italy. In *EDULEARN21 Proceedings* (pp. 42-49). IATED, 2021.

[14] L. Klopfenstein, A. Fedosyevev & A. Bogliolo, Bringing an unplugged coding card game to augmented reality. In *INTED2017 Proceedings* (pp. 9800-9805). IATED, 2017.

[15] F. Ferrari, A. Rabbone & S. Ruggiero, Experiences of the T4T group in primary schools. In *The Proceedings of International Conference on Informatics in Schools: Situation, Evolution and Perspectives* (p. 130), 2015.

[16] A. M. Pinto-Llorente, Developing Computational Thinking Using Lego Education WeDo at 4th Grade of Primary Education: A Case Study. In *Research Anthology on Computational Thinking, Programming, and Robotics in the Classroom* (pp. 156-174). IGI Global, 2022.

[17] M. Veselovská, Z. Kubincová, K. Mayerová, Comparison of LEGO WeDo 2.0 Robotic Models in Two Different Grades of Elementary School. In: Merdan, M., Lepuschitz, W., Koppensteiner, G., Balogh, R., Obdržálek, D. (eds*) Robotics in Education. RiE* 2019. Advances in Intelligent Systems and Computing, vol 1023. Springer, Cham, 2020.

[18] P.-J. Wu, F. -Y. Chiu,, K. Mayerová and Z. Kubincová, Educational Robotics at Primary School: Comparison of Two Research Studies, *17th International Conference on Information Technology Based Higher Education and Training (ITHET),* pp. 1-5, 2018.

[19] P. Petrovič, Spike up prime interest in physics. In *Robotics in Education: Methodologies and Technologies* (pp. 146-160). Springer International Publishing, 2021.

[20] Δ. Νικολός, Εισαγωγή στον προγραμματισμό του Lego WeDo 2.0 με το Scratch 3.0. Διεύθυνση Επιμόρφωσης & Πιστοποίησης: *Εκπαιδευτικά Σενάρια από το Υλικό της Εκπαίδευσης Επιμορφωτών Β' επιπέδου ΤΠΕ*, 2018.

[21] N. Lytra & A. Drigas, STEAM education-metacognition–Specific Learning Disabilities. *Scientific Electronic Archives, 14*(10), 2021 .

Web Application for Learning Mathematics

N. Stojkovikj*, L.K. Lazarova * and M. Ananieva**
* Goce Delcev University/Faculty of Computer Science, Stip, North Macedonia
** Goso Vikentiev Vocational School/ Kocani, North Macedonia
e-mail address: natasa.stojkovik@ugd.edu.mk limonka.lazarova@ugd.edu.mk,

***Abstract* – Web application which assists in the process of learning mathematics is presented in the paper. The application and its functionalities will be explained in detail. Also, there will be presented parts from the web application using diagrams such as the E-R diagrams and the UML diagrams.**

Keywords - web application, professor, student.

I. INTRODUCTION

In today's modern way of living, dependent on technology, mathematics has become more important than ever before. Mathematics is a critical tool in the modern world. A deep understanding of mathematics is required in many fields such as science, engineering, finance, medicine, and many other fields. Learning mathematics has many benefits for students in schools, but also in everyday life. Critical thinking, collaboration, communication, and creativity skills are very important skills required in the 21st century and all of them can be acquired by learning mathematics, [1]. Many real-life problems can be solved by using mathematical concepts, so the deep understanding of mathematical concepts can be very helpful. Being able to think logically can help to make analysis and separation of the problem into smaller pieces for better understanding, analysis and finding appropriate solution in detail, [2]. Despite the abstract nature of mathematics, its teaching is scientific thinking among students. The process of teaching should be taken very seriously by the teachers, and they should try to explain all mathematical concepts in an easy and understandable way. With the development of ICT tools created to satisfy the needs in education, the whole education system had changed [3,4]. ICT is a tool that supports the learning process and helps in finding new solutions to all the challenges that education is facing, [5]. Students have many difficulties with learning and understanding mathematical concepts. In [6] the authors have analyzed the types of students' difficulties in learning and understanding mathematical concepts in primary schools. Thus, the integration of the ICT tools in the process of teaching and learning mathematics is necessary in order to decrease and eliminate the difficulties for the students and teachers at the same time, [7]. The author in [8] has given examples of ICT-based activities developed in UK secondary schools. They use readily available software and other ICT tools and have proved inspiring for both students and teachers. In [9] the authors have compared the PISA test scores obtained by students using ICT more intensively with the PISA test scores of an appropriate control group and have showed that show that the number of activities irrespective of the intensity of ICT use, is positively correlated with students' PISA scores in the three domains in the most countries. ICT supports mathematics teachers to improve their designation of lessons, teaching learning tactics, updated to subjective and pedagogical knowledge and expansion of other several relevant skills [10]. The use of ICT makes mathematics teaching healthier and helps to increase the achievement of students [11, 12].

For similar purposes many web applications for teaching and learning mathematics have been developed. The most of them have aim to provide proficiency in mathematics which stands at the top of educational priorities in many countries. Sophisticated web-based learning environments are being developed also for mathematics education, some offering authoring tools for creating courseware, assignments, and exams, some being used for training, assessment, and contests [13, 14, 15].

Many mathematicians, mathematics educators, and researchers have developed many web-based and web-assisted mathematics and mathematics education courses [16]. Web-based Virtual Learning Environments (VLEs) such as Web CT, Blackboard, and eCollege are programs that have been developed to provide online mathematics courses [17]. There are many web applications and educational platforms for teaching and learning mathematics and most of them improve learners' motivation, facilitate their search competence formation, and identifies gaps in basic knowledge of mathematics as a subject. Some of them are interactive [18, 19] and they are designed to help students learn Mathematics, with a focus on high-school algebra and calculus drills, [20]. In [21] the author has selected some mathematics learning applications. In much research it is concluded that the use of web and mobile applications in the educational process for teaching and learning mathematics helped enhanced students' achievement and learning mathematics, [22, 23].

In this paper we have described a web-based application for studying mathematics for students in secondary schools.

II. FUNCTIONALITY OF THE WEB APPLICATION

The developed web-based application is devoted for students in secondary schools who are in the final year of secondary education and who are in the process of preparation for taking mathematics as subject for graduation exam. This web application has only been developed as a tool for learning mathematics and is not available for users to visit and use it. The web application covers the areas covered by mathematics's syllabus for graduation's exam at basic level. It is designed to

facilitate the process of learning mathematics and solving mathematical problems through solved examples, additional exercises for practice and necessary formulas.

On the Fig. 1, the home page of the web application is given. By clicking on the button "For us" , can be read what is the intend of the page. Information for the administrator (teacher who maintains the page) can be found by click on the button "Contact". The login of teachers and login of students can be done by clicking on the button "Login teacher" and "Login student", accordingly.



Figure 1. Home page

A. Funcionality for teacher

The web page is maintained by administrator. The role of the administrator is to full fill the content with solved examples and additional exercises for practicing for every topic which can be found at the graduation exam in mathematics. Also, the administrator gives ID code for every new professor who wants to be registered at the web application and to use it. The ID code of the professor has the following format: *@0000_0000.*

For regular use of the application, the teachers should be registered first. After teachers' registration, they have the same privileges as the other registered teachers (can attach contents, solved examples and problems for exercises). For registration, the professor must enter the following data: Name, surname, age, sex, state, city, address, telephone number (in format (000)000-000), ID code, e-mail address and password. If the registration is successful, the message "Successful registration" will be written, else the message "Please, fill up all necessary fields" will appear. The teachers input email address and password for signing in. If username and password are valid, the professor can used web application (Fig. 2), else the message "E-mail address is not in database. Try again" or "The invalid password is input. Try again" will appear.

At the web application the teachers have the following options:

- to write notes;
- to take student's search;
- to create list of the enrolled students;
- to enter a new student;
- to write statements;
- to answer to the students' questions.



Figure 2. Home page for teacher

The teacher can use the option "Notes" as reminder. By using of the option" Student's search", the teacher will be able to search each student who uses this application. By using of this option, the teacher can do quick and precise searching for every student. If the written student's name exists in the database, the following information for the searched student will be displayed: name, surname, city, mobile phone, e-mail address and password. The option "Student's list" at the web application will show all the students who are in the database, and all data for them.

For registration of a new student, the teacher needs to choose the option "Enter a new student". All needed data for enrolment of a new student are name, surname, date of birth, address and password. The teacher can delete a student if needed. After deleting, the deleted students are not members of web application anymore. For deleting of student, teacher should input his name, surname, and date of birth.

With the option "Student", the teacher will have access to the student's home page. Here, the teacher can read the discussions of the students and the solutions of the exercises. Also, the teacher can give answers to the students' questions.

In option "Statement", the teacher posts a statement to the students. It can be a useful link, information for lessons, additional literature, etc. In the field "Statement Title", the teacher writes the title of the statement. There is a button "Send statement", and by clicking of this button the teacher sends the statement to the enrolled students at the web application.

By choosing the option "Answer the questions", the teacher reads a question from the student in the "Title" field and answers the question in the "Answer" field. After this, the teacher sends the answer to the student by clicking on the "Send answer" button.

B. Funcionality for student

This web application is designed for students who are preparing to take graduation exam in secondary school for the subject mathematics. On this web application, students can easily find materials for learning and exercises for practice. Also, the web application gives the students opportunity to communicate with their teachers (Fig. 3).

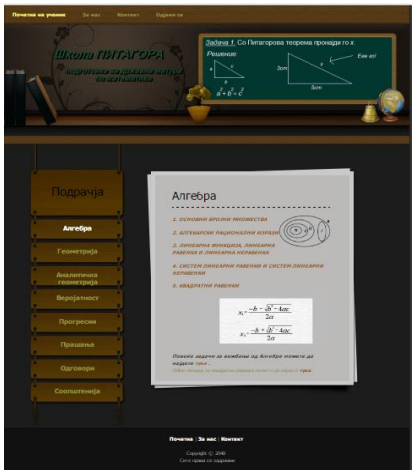


Figure 3. Home page for student

The students can use this web application, only if they are registered by the teacher. For login, the student should write his e-mail and password. After login, the student has access to all topics for preparing the graduation exam in mathematics from a secondary school. By choosing any topic, all the materials in that topic will be opened together with solved examples and additional exercises for practice.

The topics which are included at the web application are the following:

- *Algebra* contains the following topics:
  - Basic number sets.
  - Algebraic rational expressions.
  - Linear function, linear equation and linear inequalities.
  - System of linear equations and system of linear inequalities.
  - Quadratic equations.
- *Geometry* contains the following topics:
  - Geometric figures in a plane.
  - Perimeter and area of 2D figures.
  - Geometric figures in space.
  - Surface area and volume of 3D shapes.
- *Analytic geometry* contains the following topics:
  - Coordinate system.
  - Equations of (straight)line.
  - Relative position of two lines.
- Probability contains:
  - Experiment and event.
  - Statistical probability and random events.
  - Operations with random events.
  - Discrete probability space.
  - Classical definition of probability.
  - Properties of probability.
  - Formula for probability of random events.
- *Progressions* contains:
  - Arithmetic progression.
  - Geometric progression.
  - Infinity geometric progression.

Also, the web application has the following options:

- Questions.
  - In the field “Questions” students can ask questions. The web application does not show which student sends the question. On this way, students have more freedom to discuss for some problems related to the exercises.
- Answers
  - After clicking on the field "Answers" the question from the students and the appropriate answers from the teacher will be shown.
- Statements
  - By clicking on this field, the fields "Title", "Statement" and “Date and time" that are previously filled by the teacher will be displayed.

C. Functional requirements

- Registration of a new member is done by entering the following basic data for the potential member: name, surname, date of birth, country, city, residential address, telephone number, e-mail and password.
- Login of an existing member is done by entering the two basic data username and password. With this data the system will recognize the member.
- Student’ search. The system searches by student's name. So, it is enough for the teacher to write the student’s name. In the table, all students with this name will be displayed.
- Questions and answers. To solve some exercises, the web application allows the student to write the problem in filed "Questions" and send this question to the teacher. In the field “Answers” student can search solution of some previously asked question from the other students.

D. Non-Functional requirements

- **Requirement of simplicity**. This application should be structured in such a way that can easily satisfy the requirements of its users. So, the login and registration fields should be clearly located for easily finding. Also, the option

"Student' search", should take central place. Also, the results should be based without much additional information. On this way the web application can perform the requests precisely and correctly.

- **Requirement for speed.** It is expected that the system will perform the actions with a satisfactory speed, with a maximum response time of a few seconds. It should be considered that the database grows with the increasing number of teachers and students. So, the system needs to work with large capacity (in gigabytes) and according to the size of database and speed's performance.
- **Implementation requirements.** Following the world trends, the largest number of internet users work on the Google Chrome browser. So, the application is expected to have the best functionality and adaptability using this browser.

III. DIAGRAMS

With the use of diagrams, various components of the web-based application will be presented.

A. UML diagrams

The Unified Modeling Language (UML) enables the organization of the design process in a manner that allows for clear understanding and agreement among all parties involved in the development of the software, including the client, analyst, programmer, and others.[24]

Use case diagrams illustrate the connection between an actor and a use case. These diagrams are utilized to represent the system's functions, processes, and the framework in which a solution to a problem is being sought. In Fig. 4, Use Case diagram illustrating the interaction and exchange of information between the student and the teacher through the implementation of a question-and-answer functionality is shown.

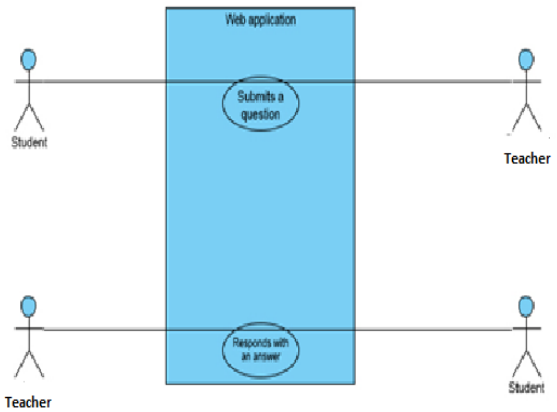


Figure 4. Use Case diagram for question and answer between the teacher and the student.

An activity diagram shows the steps that must be taken (each step is called an "activity") in a specific user scenario to achieve a certain goal or functionality. Activity diagram for functionality user login is shown on Fig. 5.

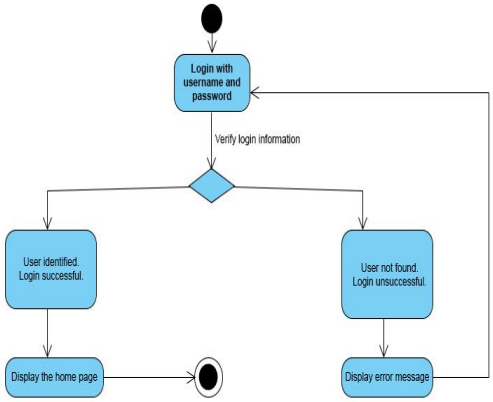


Figure 5.Activity Diagram for the user login functionality

The Entity-Relationship (E-R) diagram define conceptual view of the database and it is utilized in the early phase of database design, known as stage of the conceptual modelling of data.[25] Entity-Relationship (E-R) diagram is shown on Fig. 6.

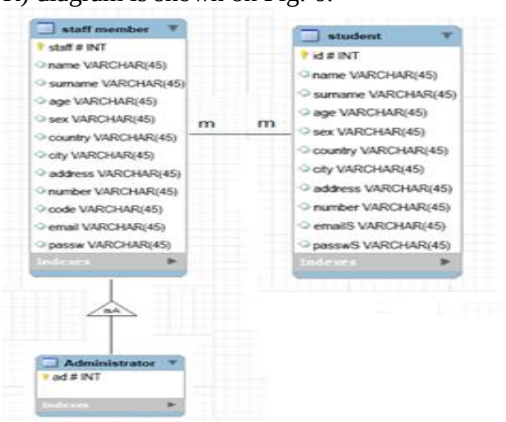


Figure 6. Entity-Relationship (E – R) Diagram

IV. CONCLUSION

In present times, people are gravitating towards utilizing the benefits of new technologies that simplify their daily lives. Computer-aided communication has a crucial place in both individuals' and companies' daily activities. With such advancements, the amount of information and data being stored and exchanged is growing at a big speed rate. Although these technological advancements provide countless advantages, they also pose new challenges to both individuals and organizations. Organizations must continuously evolve

and adapt to these developments in order to maintain their effectiveness.

The Internet, as a global network, holds a crucial place in ensuring the smooth and efficient functioning of various processes. It has revolutionized our lifestyle in every way. Not only has it made various tasks easier, but it has also altered the way people perceive things. The Internet has become an indispensable part of our lives and there is scarcely any task that does not involve modern technology and the Internet in some manner.

The presented web application aims to simplify the work of students through the utilization of the Internet. By utilizing this system, the amount of time and effort required to achieve the same results as with traditional methods is reduced. Also, this web application facilitates the process of learning mathematics which is one of the not very popular subject between the students. This application is integrated with a database that stores information for both students and teachers.

REFERENCES

[1] A. Setyarto and B. Murtiyasa, Sumardi, "Development of 21st century skills in mathematics learning with STEAM in MTs Negeri 2 Wonogiri," Universal Journal of Educational Research, vol.8, no.11, pp.5513 – 5528, 2020, DOI: 10.13189/ujer.2020.081155

[2] L M Rizki and N Priatna, "Mathematical literacy as the 21st century skill", J. Phys.: Conf. Ser., 2019, 1157 042088.

[3] K. Das, "Role of ICT for better mathematics teaching," Shanlax International Journal of Education, vol. 7, no. 4, pp. 19-28, 2019, DOI: https://doi.org/10.34293/ education.v7i4.641.

[4] G.M. Chao, "Impact of teacher training on information communication technology integration in public secondary schools in Mombasa County," Human Resource Management Research, vol. 5, no. 4, pp. 77-94, 2015.

[5] Oduma, C.A and Ile, C.M. "ICT education for teachers and ICT supported instruction: problems and prospects in the Nigerian education system." African Research Review. vol. 8, no. 2, pp. 199-216, 2014.

[6] H. Van Steenbrugge, M. Valcke and A. Desoete, "Mathematics learning difficulties in primary education: Teachers' professional knowledge and the use of commercially available learning packages," Educational Studies vol.36, no.1, February 2010, DOI: 10.1080/03055690903148639.

[7] O. Shestopalova, T. H. Kramarenko, B. Kateryna, "The ICT usage in teaching mathematics to students with special educational needs," Journal of Physics Conference Series, February 2021, DOI: 10.1088/1742-6596/1840/1/012009.

[8] A. Oldknow, "ICT bringing mathematics to life and life to mathematics," Invited papers, 2008.

[9] F. Biagi and M. Loi, "ICT and learning: results from PISA 2009," European Commission JRC Scientific and Policy Report, 2012.

[10] D. Raj Joshi, "Influence of ICT in mathematics teaching," International Journal for Innovative Research in Multidisciplinary Field, vol.3, no.1, pp.7-11, January 2017.

[11] A. Safdar, M. I. Yousuf, Q. P. Malik and G. Behlol, "Effectiveness of information and communication technology (ICT) in teaching mathematics at secondary level," International Journal of Academic Research, vol.3, no.5, I Part, pp.67-72, 2011.

[12] M. Gera, and S. Verma, "Role of ICT in teaching to reduce learned helplessness in mathematics," International Journal of Science and Research, vol.3, no.10, pp.2012–2014, 2012.

[13] R. O. Isidro, J. Sousa Pinto, and A. Batel Anjo, "SA3C - platform of evaluation system and computer assisted learning," WEAS Transactions on Advances inEngineering Education 1:2, pp.1–6, 2005.

[14] FP6-Project "LeActiveMath": language enhanced user adaptive, interactive learning for mathematics, 2004/2006. (http://www.dfki.de/leactivemath/).

[15] J. Sierra, P. Moreno-Ger, I. Martinez-Ortiz, and B. Fernandez-Manjon, "A highlymodular and extensible architecture for an integrated IMS-based authoring system:the <e-Aula> experience," Softw., Pract. Exper. 37:4, pp.441–461, 2007.

[16] D. A. Thomas,and Q. Li, "From Web 2.0 to Teacher 2.0," Computers in the Schools, vol. 25, no. (3-4), pp. 199-210, 2008.

[17] J. Engelbrecht and A. Harding, "Technologies involved in the teaching of undergraduate mathematics on the Web," Journal of Online Mathematics and its Applications, Retrieved August 2, 2011, from http://science.up.ac.za/muti/technologies.pdf .

[18] E. G. Sabirova, T. V. Fedorova, N. N. Sandalova, "Features and advantages of using websites in teaching mathematics (Interactive Educational Platform UCHI.ru)," EURASIA Journal of Mathematics, Science and Technology Education, vol.15, no.5, pp.1-14, 2019.

[19] M. Hossain, R. J. Quinn "Interactive features of Web 2.0 technologies and their potential impact in teaching-learning mathematics," Society for Information Technology & Teacher Education International Conference, Austin, Texas, USA. Mar 05, 2012.

[20] A. P. Tomas1, J. P. Leal and M. A. Domingues, "A Web application for mathematics education", Conference: Advances in Web Based Learning - ICWL 2007, 6th International Conference, Edinburgh, UK, August 15-17, 2007.

[21] I. K. Namukas, "Selection of mathematics learning apps," Conference: OAME 2016At: Barrie, Ontario, Project: Classroom inovation, technologies and online educational curriculum materials (OECM) in mathematics education, May 2016.

[22] J. O. Etcuban, L. D. Pantinople, "The effects of mobile application in teaching high school mathematics," International Electronic Journal of Mathematics Education, e-ISSN: 1306-3030, vol. 13, no. 3, pp. 249-259, 2018, https://doi.org/10.12973/iejme/3906 .

[23] M. Al-Aali, "A study of mathematics web-based learning in schools," American Journal of Applied Sciences, vol.5, no.11, pp.1506-1517, 2008.

[24] G. Booch, J. Rumbaugh and I. Jacobson, Unified Modeling Language User Guide, 2nd ed.,Addison- Wesley Professional, 2017

[25] S. Bagui and R. Earp, Database Desing Using Entity – Relationship Diagrams, 2nd ed., Auerbach Publications, 2011.

Teachers' Expectations for Digivisio 2030 – A Joint Higher Education Digital Transformation Project in Finland

J. Multisilta, T. Mattila and K. Lempinen
Satakunta University of Applied Sciences, Pori, Finland
jari.multisilta@samk.fi, timo.mattila@samk.fi, katja.lempinen@samk.fi

***Abstract* - This paper describes digital transformation (DT) in Finland from the perspective of a medium-sized University of Applied Sciences. Higher education institutions (HEIs) compete with each other, both locally and globally, in terms of student and staff recruitment, government funding and research funding. In addition, networking with companies and other organizations is an important part of HEI strategies and actions. In this paper, we discuss digital transformation in higher education institutions. We also present the results of a questionnaire survey conducted in a medium-sized university of applied sciences in Finland on teachers' attitudes and expectations towards a nationwide digital transformation project (N=82). The data show that teachers' expectations and attitudes range from positive to uncertain. Many believe that in the future digital university, student counselling will play a more important role compared to face-to-face learning. In addition, it is not widely accepted that future students will be self-directed. The role of technology was not seen as a challenge, and teachers expect more training and pedagogical methods to emerge when they collaborate with other teachers from other universities in Finland.**

Keywords – digital transformation; teacher; higher education

I. INTRODUCTION

The Finnish higher education (HE) system is a dual model with two types of higher education institutions (HEIs), namely universities and universities of applied sciences. Universities offer bachelor's, master's and doctoral degrees and conduct both basic and applied research. Universities of applied sciences offer bachelor's and master's degrees and carry out applied research and development in their region. In such the role of the university of applied science is more regional compared to the universities whose role is national. Bot sectors do close cooperation with companies and organizations. The Finnish Ministry of Education and Culture directs the activities of higher education institutes. The higher education system is funded by the government. There are no tuition fees for European students. For research, universities and universities of applied sciences apply for funding from various sources, such as the Academy of Finland, Business Finland, EU programs and private foundations. Funding for higher education institutions is based on an agreement between the Ministry of Education and a higher education institution. The funding is based on a model that allocates funds to an institution according to its performance. The HEI is measured by the number of students it graduates and the amount of external research funding it receives.

In many industry and service sectors in Europe and Finland, the need for a skilled workforce is growing, especially in technology-related fields as well as the social and health sectors. In terms of long-term projections, the decreasing birth rate suggests that Finland's population will begin to decline in 2034 (YLE News, 30.9.2021, Statistics Finland: Falling birth rates cannot maintain population). HEIs are expected to respond to the need for more workforce by providing degree studies and continuing education.

Moreover, due to technological development and the rapid changes in our societies, jobs are constantly changing. For many jobs, workers' skills and knowledge need to be updated systematically. It is not enough to simply have a degree; rather, one must possess the latest knowledge and skills. In addition to job-specific skills, students are expected to learn the so-called future skills such as problem-solving skills, creativity, the ability to act in a self-organized manner, cooperation skills, and social and communication skills, all of which are important in today's changing work life [1, 2].

The contemporary adoption of technologies in HEIs reflects a paradigm shift, whereby technology is seen as facilitating the management of complex learning environments and digital learning [3]. This paradigm shift is known as digital transformation (DT). According to [3], "digital transformation (DT) has become a priority for higher education institutions (HEIs) in this second decade of the 21st century, and this is a natural and necessary process for organizations that claim to be leaders of change and be highly competitive in their domain."

Digital transformation is seen as a "process for increasing efficiency and collaboration, and reducing costs and errors in the management of at-scale training systems" [4]. DT has many implications for higher education institutions. For example, the shift from paper to digital, the automation of processes in HEI's, the increasing use of mobile devices and the cloud have a significant impact on how HEI's will operate in the future [4].

In [5], the researchers discuss the implications of digital transformation. Their main finding is that technology may not be the most important driver of DT, but that technology-enabled teaching and learning processes need to be reformulated. They also conclude that educational content and curricula need to be reformulated.

In [6] there is a report on a study of DT in a small Swedish university. The researchers conclude that the role of the teacher changes in DT and that the changes are reflected in the digital competence and the didactic and pedagogical competence of the teacher. In [6] digital competence is defined to include "the knowledge, skills, understanding and motivation that the individual needs in the process of change that digitalization entails".

In [6] a main finding is that teachers have a good level of digital competence and find it easy to learn new technologies, but they do not use these technologies and methods in their teaching. It can be concluded from [6] that DT should not only take place at the teacher level, but also at the organizational level and as a societal process. In addition, the IT infrastructure, the training of teachers and the management and leadership of the higher education institution are the key drivers of DT [7].

In [8] it is stated that the core processes where DT should be applied are teaching, IT infrastructure, curriculum, management and governance, research, business processes, human resources and marketing. In this sense, DT is a rather holistic process [8].

In their literature review [8], the researchers point out that currently the renewal business model in HEIs is aligned with technological development. However, technology should be seen as a tool to support the redesign of business processes in higher education institutions.

In [8] the elements in the DT process are "people, processes, strategies, structures and competitive dynamics". In addition, based on the survey in [8], an important dimension of DT is the customer experience lifecycle. DT should enhance and replace current products and services and be seen as a way to create "additional and differentiated value" [8].

The survey in [8] also examined the role of teachers in DT. Teachers reflect the need to improve their productivity in teaching through the use of digital tools. In addition, teachers expect that DT will improve communication, collaboration and co-creation of value in higher education.

Based on these definitions, it is clear that DT has a significant impact on HEIs' strategies, teaching, infrastructure, curricula, administration, research, business processes, human resources, and marketing.

In the digital era, learning can also be analyzed using Lee and Hannafin's [9] own it, learn it, and share it (OWL) model. According to the OWL model, students are responsible for their own learning processes and set their own learning goals, thus learning according to their plans via a rich set of cognitive tools and achieving meaningful learning outcomes with a strong link to real-world problems. However, instead of being left to study alone, students should have access to guidance during the learning process. Consequently, student engagement in the learning process is strong in the OWL model.

In this paper we discuss DT in Finnish higher education institutions. As a case study, we discuss the expectations of the staff of a medium-sized Finnish university of applied sciences in relation to digital transformation and the Digivisio project.

II. DIGIVISIO PROJECT

Digivisio 2030 is a joint development program for all Finnish universities and universities of applied sciences. The main aim of Digivisio is to establish Finland as a model country for flexible learning and a global pioneer in higher education in the world. With external funding, Finnish HEIs have formed a consortium for implementing a common model for DT in Finnish higher education. According to Digivisio [10], the aims are as follows:

- Learners are provided with data on their own learning in a secure manner, enabling and supporting their learning throughout their lives.
- In 2030, Finland has an open and recognized learning ecosystem that offers quality, diversity, flexibility, efficiency, and suitability for the life situation and needs, thus generating better learning results.
- The learning ecosystem also provides a platform for research and innovation activities, benefiting society and working life extensively.
- In 2030, the scientific and educational activities of HEIs are key factors in securing the high level of competence of the adult population, national competitiveness, and international impact.

Currently, all HEI's in Finland have their own IT infrastructures and systems, and students need user ID's to each university they would like to study courses. Digivisio 2030 will change this by providing to establish one identity to students. In order to achieve this, all universities must make their systems and services compactible, and they should support Digivisio platform. Students will own their own data in Digivisio platform. In this sense, Digivisio platform supports the OWL model [9].

It is expected that, DT will also change student counseling. HEIs should support both young students and those coming from working life to take part to continuing education. The support should be available regardless of time and place. Counseling can at least partly be implemented with AI solutions [10]. However, it is not easy to support students in complex learning environments and new solutions are needed. According to [11], "there is an urgent need to determine how to support students that have underdeveloped self-regulation skills." Regarding the Digivisio, we can agree with [12]: "With the sudden shift to digital in classrooms around the world, future research may continue to examine how student engagement varies from context to context to understand how educators can better support all students." It can be

concluded that students' needs in Digivisio platform needs more research and implementation of new student-centered guiding and counseling practices. Digivisio represents a major DT that cannot happen without proper change management.

At SAMK, the Digivisio is seen as a continuation of the collective and individual transformations SAMK has been implementing already since early 2000. A comprehensive change in the culture of action does not happen quickly but it needs time. Although faculties and organizational units strive to fulfill their own functions, only the unity of the subsystems enables the organization to achieve its DT goals. Therefore, management should consider the requirements and effects of change on all parts of the system as well as the relationships and dependencies between them [13]. In sum, Digivisio 2030 is a new stage in the transformation that SAMK and other HEI in Finland have been undergoing for many years [14, 15, 16].

III. METHOD AND RESULTS

Between September and October 2022, SAMK faculty members were invited to complete a questionnaire with three background questions and seven open-ended questions about their expectations of the Digivisio project. Prior to the survey, SAMK had organized a briefing on the project, and the status of the project was regularly communicated to the staff in newsletters. The survey was available in either Finnish or English. At the time of the survey, there were 441 employees in SAMK and 84 (19%) responded to the survey. The respondents were from the teaching staff (49%), the research staff (19%) and the administration (32%). The distribution of respondents corresponds to the distribution of staff in SAMK.

In general, 48% of respondents have a positive attitude and 45% have a neutral attitude towards the Digivisio project. Only 6% have a negative attitude and 1% did not answer the question.

Respondents had varying levels of prior knowledge of the Digivisio project. Of all respondents, 38% recognized the project but did not have a clear understanding of it. Only 4% of respondents had no prior knowledge of the project. The majority of respondents had attended project events (48%) or were aware of the project's objectives (11%).

There were seven open-ended questions:

- What do you think about the Digivisio project and its objectives?
- What are your expectations of the Digivisio project?
- How do you think the Digivisio project will affect your current job?
- How do you think you will benefit from the Digivisio project?
- How do you think the Digivisio project will affect the selection and numbers of degree students at SAMK?
- How do you think the Digivisio project will affect student numbers at the Open University of SAMK?
- What kind of information about the Digivisio project do you want and through what means?

The responses to the open-ended questions were merged into text documents and analyzed using thematic analysis. The researcher examined the documents and identified common themes from the data by first coding the data and then grouping the coded data into themes. An inductive approach was used to identify the themes. The themes identified were:

- The goals of the Digivisio project
- Teaching
- Funding and the guidance from the Ministry of Education (MoE)
- Organizational change
- Pedagogy

A. The goals of the Digivisio project

In general, the respondents think that the goals of the project are in the right direction. There were also opinions that reflected the potential outcomes if the project succeeds in its goals.

"The prerequisite for reaching the goal is to re-examine the processes in each higher education institution and think about what we can do differently, and at which points we have built obstacles to the implementation of flexible learning."

There were also fears that the goals may be too ambitious.

B. Teaching

The respondents expect that the Digivisio project will bring flexibility to the learning.

"Digivisio - hopefully - will lead to students being able to smoothly choose the teaching units necessary for developing their own competence."

In contrast to online learning, respondents discussed the importance of face-to-face learning on campus and the role of guidance. Some of the respondents were concerned about the quality of learning in the future and wondered whether students would be self-directed, which would be necessary in digital learning environments.

"I'm a little afraid of what will happen to face-to-face teaching, which, however, is more effective than online teaching in certain subjects."

The quality of the content was also seen as a competitive factor when competing for good students.

"The course offering must be of high quality in order to be able to compete with others for applicants. The course material must also be as new and easily accessible as possible."

C. Funding and the guidance from the Ministry of Education (MoE)

There were not many responses related to funding and guidance from the MoE. The MoE has funded the Digivisio project, and one respondent was concerned that the project will actually be a tool for the MoE to unify

HEI's in Finland. Some respondents questioned how funding for HEI's will be allocated in the future, when digital transformation has changed the way HEI's operate.

"The impact of policy changes on the core funding of organizations is not addressed. It will ultimately determine what will remain permanent."

D. Organizational change

Organizational change seems to be the theme that has drawn many responses. Respondents stated that Digivisio will change the way higher education institutions operate in the future. It was seen that teaching practices and learning opportunities will change because of the digital transformation. The change was also seen as an opportunity through new and flexible service models. In addition, it was seen that lifelong learning and international students will benefit from these new service models.

"New opportunities for learners, universities and their staff. Especially for new, flexible service models."

Competition between HEIs was also reflected in the responses. The Digivisio project was seen as increasing competition between higher education institutions.

"What is SAMK's competitive position in the future? What can we offer online that others don't?"

It was believed that the digital transformation in HEI's will increase the number of students, but it will also increase the dropping out of studies. Some respondents think that the only large HEI's will benefit from DT so that they will get more students. However, the location of the HEI campus might also be a factor in selecting the where to study.

"I think young people will continue to choose where to study based on the city, not the school?"

E. Pedagogy

The respondents emphasized that in Digivisio project the technology should be a tool and the pedagogy should be in focus. The also expected to get new pedagogical methods and more collaboration with other teachers in the future. In addition, training and support for using technology were requested.

The increasing number of IT systems was perceived as stressful, and respondents wished that DT would reduce the number of IT systems. It can be interpreted that teachers experience that many different IT systems create a cognitive load and the time needed to learn a system reduces the time available for regular teaching.

"I wish that at least one system or program would be deactivated, i.e. the functions would be moved to another program or system. "

IV. CONCLUSION

Mostly the data revealed topics related to teaching and management. The role of research did not emerge in our study. So far, the Digivisio project has had a strong focus on teaching, which may explain why research was not mentioned as an important process of the HEI in the data.

The results of the data analysis support previous findings. The role of process development was emphasized, and teachers' technological skills did not raise many questions.

The data shows that teachers' expectations and attitudes vary from positive to uncertain about what to expect. Staff have a clear understanding that Digivisio will change many current processes. Many believe that student counselling will play a more important role in the future digital university compared to face-to-face learning. In addition, it is not widely accepted that future students will be self-directed. The role of technology was not seen as a challenge, and teachers expect more training and pedagogical methods to emerge when they collaborate with other teachers from other HEIs in Finland.

V. DISCUSSION

The COVID 19 pandemic has forced universities to digitise their processes. However, the pandemic is not the only driver of digital transformation in universities worldwide. The shortage of skilled labour, competition for the best students and staff, and student expectations have also forced universities to rethink their offerings.

Teachers and staff may have somewhat contradictory expectations of the digital transformation taking place in universities. Firstly, teachers are usually highly motivated to do their job well. This is why some are apprehensive about the quality of teaching and learning in the digital world. Second, teachers expect that shared digital platforms could bring advanced teaching and learning methods and more collaboration with colleagues in other universities. Although teachers expect to be trained in new technologies, their skills do not appear to be a barrier to driving digital transformation.

In many countries, individual universities are making progress in digital transformation. However, it is not common for a country to develop a national platform for all universities, as Finland has started to do. In Germany, for example, there is the German Forum for Higher Education in the Digital Age (Hochschulforum Digitalisierung), which monitors, shares and acts as a source of information, provides strategic advice and builds key competences of higher education professionals across the country [17]. Finland's advantage is the relatively small number of higher education institutions that share a common vision of higher education.

REFERENCES

[1] Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P., The future of skills: Employment in 2030. Pearson, 2017.

[2] Ehlers, U. D., & Kellermann, S. A., Future skills: The future of learning and higher education. Karlsruhe, 2019 https://www.learntechlib.org/p/208249/

[3] Benavides, L. M. C., Tamayo Arias, J. A., Arango Serna, M. D., Branch Bedoya, J. W., & Burgos, D. (2020). Digital transformation in higher education institutions: A systematic literature review. Sensors, 20(11), 3291. https://doi.org/10.3390/s20113291

[4] C. G. Demartini, L. Benussi, V. Gatteschi and F. Renga, "Education and Digital Transformation: The "Riconnessioni" Project," in *IEEE Access*, vol. 8, pp. 186233-186256, 2020, doi: 10.1109/ACCESS.2020.3018189.

[5] Kamsker, S., Janschitz, G. and Monitzer, S., "Digital transformation and higher education: A survey on the digital competencies of learners to develop higher education teaching," International Journal for Business Education: Vol. 160: No. 1, Article 2, 2020.

[6] Sjöberg, J., Lilja, P., University Teachers' Ambivalence about the Digital Transformation of Higher Education. International Journal of Learning, Teaching and Educational Research. Vol. 18 No. 13 (2019): December 2019.

[7] Bernhard-Skala, C., Organisational perspectives on the digital transformation of adult and continuing education: A literature review from a German-speaking perspective. Journal of Adult and Continuing Education, 25(2), 178–197. https://doi.org/10.1177/1477971419850840, 2019.

[8] Benavides, L.M.C.; Tamayo Arias, J.A.; Arango Serna, M.D.; Branch Bedoya, J.W.; Burgos, D. Digital Transformation in Higher Education Institutions: A Systematic Literature Review. Sensors 2020, *20*, 3291. https://doi.org/10.3390/s20113291

[9] Lee, E., & Hannafin, M., A design framework for enhancing engagement in student-centered learning: Own it, learn it, and share it. Educational Technology Research & Development, 64(4), 707–734, 2016.

[10] Higher Education Institutes' Digivisio 2030. https://Digivisio2030.fi/tiedostot/higher-education-institutes-Digivisio-2030/, 2021.

[11] Edyburn, D., Transforming student engagement in COVID-19 remote instruction: A research perspective. Education Tech Research Dev, 69, 113–116. https://doi.org/10.1007/s11423-020-09919-6, 2021.

[12] Wong, K. M., "A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it" by Lee and Hannafin (2016): An international perspective. Education Tech Research Dev, 69, 93–96. https://doi.org/10.1007/s11423-020-09842-w, 2020.

[13] Mattila, E., Kallio, T. J., & Saru, E., Sivistyksen ja soten yhteistyö kunnissa – askeleita kohti syvempää palveluintegraatiota. Hallinnon Tutkimus, 3/2021, 170–186. pp. 171–172, 2022

[14] Auvinen, P., Ammatillisen käytännön toistajasta monipuoliseksi aluekehittäjäksi? Ammattikorkeakoulun-uudistus ja opettajan työn muutos vuosina 1992–2010, 2004.

[15] Nenonen, M., Tulokset paranevat, miten käy laadun? Tulosperustainen rahoitusmalli koulutuksen tuloksellisuuden ja laadun kehittäjänä ammattikorkeakouluissa, 2020.

[16] Multisilta J., Mattila T., Digital Transformation in Finnish Higher Education: A Perspective from a University of Applied Sciences. In Baptista M. & Isaias P. (2022) Proceedings of International Conference e-Learning 2022, Lisbon Portugal, 19-22.7.2022. IADI Press. pp. 93-110, 2022.

[17] HOCHSCHULFORUM DIGITALISIERUNG: ABOUT US. https://hochschulforumdigitalisierung.de/en Accessed 19.3.2023

Digital Culture and Students' Awareness of Their Participation in This Phenomenon

Leonie Bandić*, Paula Perić*, Tihana Babić**
* VERN' University, Zagreb, Croatia
** Algebra University College, Zagreb, Croatia
lebandic@vernnet.hr, paperic1@vernnet.hr, tihana.babic@algebra.hr

***Abstract* - Digital culture is a complex phenomenon that pervades contemporary digital society. Digital culture can be interpreted through any activity in the digital sphere such as work, communication, creation, recreation, or just entertainment. Modern students are highly connected to and highly dependent on digital culture. For the purposes of this paper, the aim of the conducted research was to investigate whether students are at all familiar with the concept and meaning of digital culture, if they are aware of their participation in digital culture, the impact it has on them, whether they use the content of the digital environment in a targeted and rational manner, and if they feel a connection with it and what feelings it evokes in them in general. The research was conducted with a self-assessment questionnaire that contained questions on emotional, behavioral, and cognitive levels of a sample of students from VERN' University and Algebra University College in the winter semester of the academic year 2022/2023.**

Keywords – digital culture, students, digital culture awareness, communication

I. INTRODUCTION

Digital culture emerged from amplified use of communication technologies, caused by the changes and accelerated prevalence of the electronic broadcast culture. The expeditious development of technologies has changed ways of communication, creation, and culture consumption. Change in societal terms and means, interaction, and alteration of technology usage constitutes digital culture. Synonymous terms used to describe this form of cultural change are cyber culture, internet culture and computer culture.

This research paper seeks to inquire and examine the impact of digital culture on social relationships. Concisely, the study aims to explore the ways digital technologies interfere with interpersonal communication and interaction, and the interference of these changes to our sense of identity, community, and well-being. Through a survey, the study will analyse the ways in which individuals use, participate and associate with digital technologies to maintain and negotiate social connections, as well as the challenges and opportunities that derive from these practices.

This research aims to provide a distinct and analytically grounded understanding of the complex and dynamic relationship between digital culture and social exchange.

II. DIGITAL CULTURE

Digital culture consists of change in societal terms and means, interaction, and alteration of technology usage, caused by the changes and accelerated prevalence of the electronic broadcast culture. Digital culture is marked with newfound set of practices, social norms, and values that modified and shaped our culture. It is disseminated widely and implemented in various industries.

Digitality can be thought of as a marker of culture because it encompasses both the artifacts and the systems of signification and communication that most clearly demarcate our contemporary way of life from others [1]. Cyberspace, which is the basis of digital culture, has different characteristics and fewer boundaries than real space. Digitalization has enabled the process of media convergence to take place and merged the cyberspace with legitimate industries and affect changes that remodel society.

The circulation of media content across different media systems, competing media economies, and national borders depend heavily on consumers' active participation [2]. Arts and entertainment, and the way they have developed, have changed the consumer consumption and media production to their core. With the rapid and constant development of new social media platforms such as TikTok and BeReal, streaming services such as Netflix and Hulu, and platforms for streaming music like Spotify and SoundCloud, consumers are given easy access to participate in their preferred form of entertainment.

III. PARTICIPATION IN DIGITAL CULTURE

Participation as a term refers to actions demonstrating forms of involvement performed by parties within evolving structures of talk [3]. All the same, participation as a noun refers to the action of taking part in something, involvement, engagement, contribution, association, partaking, et cetera.

Taking into consideration how the modern world is immersed in digital surroundings that describe digital culture, it is in favour of the basic human need to socialize with others as humans are social beings by nature. Participating in digital culture, the media could become creators of their own individual worlds, by creating virtual reality with information saturation.

Participating in digital culture gained strength when the feeling of being left out of the virtual world became prominent. (FOMO – fear of missing out) [4]. The need to participate became overwhelming and overbearing, forcing everyone to participate in one way or another.

IV. AWARENESS OF DIGITAL CULTURE

Awareness offers a way of releasing from automatic and useless ways of thinking to reacting. Once individuals are aware of their actions, they can successfully process incoming information from their surroundings and respond adaptively. Unawareness indicates the absence of information processing.

Focused awareness implies an increased focus of attention on experiences happening in the present in an accepting way [5]. The basic assumption of focused awareness is that experiencing the present moment without judgment, with openness and acceptance, enables taking a broader perspective regarding:

- *thoughts* - that are observed only as thoughts that come and go, not the absolute truth,
- *feelings, especially negative* - that are observed as a current experience and not something to fight against,
- *body sensations* - that are viewed as a bodily reaction to feelings [5].

Awareness creates space between perception and responses to situations. Participants of digital culture become able to answer to situations reflexively based on their own choices. The term focused awareness refers to the state of consciousness with attention directed to their own experiences (cognitive, behavioural, and emotional) or to their surroundings in the present moment with the attitude of acceptance [6].

V. RESEARCH METHODOLOGY

A. Research Design

Research in the form of a self-assessment questionnaire has been conducted amongst the students at VERN University and Algebra University College in the academic year 2022 /2023. The research goal was to investigate whether students are at all familiar with the concept and meaning of digital culture and aware of their participation in it.

The self-assessment survey questionnaire contained fifteen claims on emotional, behavioral, and cognitive levels by which the students could express at what level they agree or disagree with the claims developed according to the theoretical framework related to digital culture and adapted according to cognitive, behavioural, and emotional awareness [6]. The questionnaire was composed based on the Likert scale that contained five of the possible answers of agreement (14 statements) and frequency (1 statement). The research was conducted via Google Forms and the total number of respondents was N=160.

B. Research Sample

The research was approached in total by 160 participants, 75.6 % of which are from VERN University and 24.4 % are from Algebra University College.

The highest number of participants that took part in the research are students of undergraduate studies at VERN University:

- Economics and Entrepreneurship (28 %)
- Tourism (14 %)
- Cyber communications (13 %)
- Public relations and media studies (27 %)
- Journalism (2 %)
- Internet of Things (4 %)
- Business informatics (4 %)
- Technical management (2 %)
- Transmedia dramaturgy (2 %)
- Film, television, and multimedia design (4 %).

Students of graduate studies at VERN University:

- Management of sustainable tourism development (9.1 %)
- Business management (27.3 %)
- Accounting and finance (6.1 %)
- Human resources management (24.2 %)
- Business communications management (9.1 %)
- IT management (15.2 %)
- Film and TV production (9.1 %).

Students of Algebra University College:

- Software engineering (15.4 %)
- System engineering (7.7 %)
- Multimedia computing (10.3 %)
- Data science (5.1 %)
- Digital Marketing (28.2 %)
- Design of market communications (33.3 %).

C. Research Results

The results of the research according to all 3 levels will be presented below:

a. Cognitive level of students' awareness of their participation in digital culture

1. FAMILIARITY WITH THE CONCEPT OF DIGITAL CULTURE

Results showed that 30 % (48) of the participants strongly disagree and 31 of the participants disagree with the claim "I am not familiar with the concept of digital culture". The neutral response was given by 26.9 % (43) of the participants. Somewhat fewer participants, or to be exact, 13.1 % (21) agree with the claim and 10.6 % (17) strongly agree with the claim that they are not familiar with the concept of digital culture (as presented in Table 1).

TABLE 1 Response distribution to the claim: *"I am not familiar with the concept of digital culture"* (N=160)

I am not familiar with the concept of digital culture	Frequency of responses	Students in percentages
Strongly disagree	48	30 %
Disagree	31	19.4 %
Neither agree nor disagree	43	26.9 %
Agree	21	13.1 %
Strongly Agree	17	10.6 %

2. THINKING ABOUT PARTICIPATING IN DIGITAL CULTURE

With the claim *"I have never thought about participating in digital culture (ex. How much time I spend, for what purpose I use digital culture)"* 26.3 % (42) of the participants strongly disagree and 20.6 % (33) simply disagree. The biggest number of participants, 26.9 % (43) neither agree nor disagree with the claim so their response is neutral. 13.8 % (22) of the participants agree and 12.5 % (20) strongly agree with the claim (as presented in Table 2).

TABLE 2 Response distribution to the claim: *"I have never thought about participating in digital culture (ex. How much time I spend, for what purpose I use digital culture)"* (N=160)

I have never thought about participating in digital culture	Frequency of responses	Students in percentages
Strongly disagree	42	26.3 %
Disagree	33	20.6 %
Neither agree nor disagree	43	26 %
Agree	22	13.8 %
Strongly Agree	20	12.5 %

3. THINKING ABOUT ONE'S OWN PARTICIPATION IN DIGITAL CULTURE

The results are fairly similar to the former claim, in which 23.1 % (37) of the participants strongly disagree and 26.9 % (43) do not agree with the claim *"I do not wish to think about my participation in digital culture (ex. I do not wish to read my weekly report on using social media)"*. 26.3 % (42) participants of the questionnaire do not have a specific opinion on the claim, so they remained neutral, 12.5 % (20) agree and 11.3 % (18) strongly agree with the claim (as presented in Table 3).

TABLE 3 Response distribution to the claim: *"I do not wish to think about my participation in digital culture (ex. I do not wish to read my weekly report on using social media)"* (N=160)

I do not wish to think about my participation in digital culture	Frequency of responses	Students in percentages
Strongly disagree	37	23.1 %
Disagree	43	26.9 %
Neither agree nor disagree	42	26.3 %
Agree	20	12.5 %
Strongly Agree	18	11.3 %

4. THE EFFECT OF DIGITAL CULTURE

With the claim *"I believe that digital culture does not affect me (ex. The content I consume does not change my opinions, does not change my habits.)"* 51.9 % (44) of the participants strongly disagree and 24.4 % (39) disagree. 26.3 % (42) remain neutral, 14.4 % (23) agree with the claim and 7.5 % (12) strongly believe that digital culture does not affect them (as presented in Table 4).

TABLE 4 Response distribution to the claim: *"I believe that digital culture does not affect me (ex. The content I consume does not change my opinions, does not change my habits)"* (N=160)

I believe that digital culture does not affect me	Frequency of responses	Students in percentages
Strongly disagree	44	27.5 %
Disagree	39	24.4 %
Neither agree nor disagree	42	26.3 %
Agree	23	14.4 %
Strongly Agree	12	7.5 %

5. AWARENESS OF ONE'S OWN PARTICIPATION IN DIGITAL CULTURE

Only 1.9 % (3) of the participants believe that they are completely unaware and 5.6 % (9) are not aware of their participation in digital culture. Furthermore, 21.9 % (35) of the participants do not have a certain opinion. With the claim *"I am aware of my own participation in digital culture (ex. I know what purpose I use it for and how I use it)"* 38.1 % (61) agree and 32.5 % (52) strongly agree (as presented in Table 5).

TABLE 5 Response distribution to the claim*: "I am aware of my own participation in digital culture (ex. I know what purpose I use it for and how I use it)"* (N=160)

I am aware of my own participation in digital culture	Frequency of responses	Students in percentages
Strongly disagree	3	1.9 %
Disagree	9	5.6 %
Neither agree nor disagree	35	21.9 %
Agree	61	38.1 %
Strongly Agree	52	32.5 %

b. Emotional level of students' awareness of their participation in digital culture

6. A SENSE OF CONNECTION TO DIGITAL CULTURE

Most of the participants feel connected, or to be more exact, 40 % (64) agree with the claim, and 25.6 % (41) strongly agree with the claim *"I feel connected to digital culture (ex. Digital culture connects me to the rest of the world – entertainment, work, friends…)"*. 20.6 % (33) of the participants do not have a specific opinion. 10.6 % (17) disagree with the claim and 3.1 % (5) strongly disagree with the claim, they do not feel any connection to digital culture.

TABLE 6 Response distribution to the claim: *"I feel connected to digital culture (ex. Digital culture connects me to the rest of the world – entertainment, work, friends…)"* (N=160)

I feel connected to digital culture	Frequency of responses	Students in percentages
Strongly disagree	5	3.1 %
Disagree	17	10.6 %
Neither agree nor disagree	33	20.6 %
Agree	64	40 %
Strongly Agree	41	25.6 %

7. RELAXATION IN TIME SPENT IN DIGITAL CULTURE

Results show that the largest amount of participants relax in their free time as a part of digital culture, to be more exact, 40 % (64) agree with the claim, and 22.5 % (36) strongly agree with the claim. *"I am relaxed in my time spent in digital culture (ex. It releases me of tension; I participate at the end of a tiring day for the purpose of clearing my mind)".* 6.3 % (10) strongly disagree, 7.5 % (12) disagree and 23.8 % (38) were neutral (Table 7).

TABLE 7 Response distribution to the claim: *"I am relaxed in my time spent in digital culture (ex. It releases me of tension; I participate at the end of a tiring day for the purpose of clearing my mind)"* (N=160)

I am relaxed in my time spent in digital culture	Frequency of responses	Students in percentages
Strongly disagree	10	6.3 %
Disagree	12	7.5 %
Neither agree nor disagree	38	23.8 %
Agree	64	40 %
Strongly Agree	36	22.5 %

8. RATIONAL PARTICIPATION

Only 3.8 % (6) of the participants strongly disagree with the claim *"I participate in digital culture rationally (ex. I use it with a certain purpose/reason, I use targeted content of the digital environment)"* and 15.6 % (25) disagree with the claim. The majority of the participants find that they participate rationally in digital culture, 53 % (56) agree and 18.1 % (29) strongly agree with the claim. 27.5 % (44) of the participants do not have a certain opinion regarding the claim, therefore remain neutral (as presented in Table 8).

TABLE 8 Response distribution to the claim: *"I participate in digital culture rationally (ex. I use it with a certain purpose/reason, I use targeted content of the digital environment)"* (N=160)

I participate in digital culture rationally	Frequency of responses	Students in percentages
Strongly disagree	6	3.8 %
Disagree	25	15.6 %
Neither agree nor disagree	44	27.5 %
Agree	56	35 %
Strongly Agree	29	18.1 %

9. FEELING OF MISSING OUT

Out of 160 participants in the questionnaire, 28.1 % (34) strongly disagree with the claim *"I feel as if I will miss out on something (FOMO – Fear of missing out)"* and 25 % (40) disagree with the claim. 6.3 % (10) strongly agree and 20.6 % (33) agree, whereas 20 % (32) remain neutral on the matter (as presented in Table 9).

TABLE 9 Response distribution to the claim: *"I feel as if I will miss out on something (FOMO – Fear of missing out)"* (N=160)

I feel as if I will miss out on something	Frequency of responses	Students in percentages
Strongly disagree	45	28.1 %
Disagree	40	25 %
Neither agree nor disagree	32	20 %
Agree	33	20.6 %
Strongly Agree	10	6.3 %

c. Behavioral level of students' awareness of their participation in digital culture

10. DIGITAL CULTURE PARTICIPATION FRUSTRATION

31.1 % (50) of the participants disagree and 16.3 % (26) strongly disagree with the claim *"I am frustrated by constant participation in digital culture (ex. I feel restlessness, internal tension)"*. 25.6 % (41) neither agree nor disagree with the claim, whereas 16.3 % (26) agree and 10.6 % (17) strongly agree with the claim (as presented in Table 10).

TABLE 10 Response distribution to the claim*: "I am frustrated by constant participation in digital culture (ex. I feel restlessness, internal tension)"* (N=160)

I am frustrated by constant participation in digital culture	Frequency of responses	Students in percentages
Strongly disagree	26	16.3 %
Disagree	50	31.1 %
Neither agree nor disagree	41	25.6 %
Agree	26	16.3 %
Strongly Agree	17	10.6 %

11. PARTICIPATION IN DIGITAL CULTURE

Only 3.8 % (6) of the participants rated their participation in digital culture as never and 6.9 % (11) consider they rarely participate. 28.1 % (45) believe they sometimes participate in digital culture, 40.6 % (66) claim they often participate and 20.6 % (33) always participate in digital culture (as presented in Table 11 which follows below).

TABLE 11 Response distribution to the claim: *"I participate in digital culture."* (N=160)

I participate in digital culture	Frequency of responses	Students in percentages
Never	6	3.8 %
Rarely	11	6.9 %
Sometimes	45	28.1 %
Often	65	40.6 %
Always	33	20.6 %

12. CONVERSATION ON THE TOPIC

More than half of the participants disagree with the claim *"I do not wish to converse on the topic of digital culture (ex. I dislike the topic "I am reluctant to answer any questions regarding my participation").* 35 % (56) strongly disagree and 31.1 % (50) disagree with the claim. A minority of the participants i.e. 3.8 % (6) strongly agree and 5.6 % (9) agree, whilst 24.4 % (39) do not have a certain opinion on the matter (as presented in Table 12).

TABLE 12. Response distribution to the claim: *"I do not wish to converse on the topic of digital culture (ex. I dislike the topic; I am reluctant to answer any questions regarding my participation)"* (N=160)

I do not wish to converse on the topic of digital culture	Frequency of responses	Students in percentages
Strongly disagree	56	35 %
Disagree	50	31.3 %
Neither agree nor disagree	39	24.4 %
Agree	9	5.6 %
Strongly Agree	6	3.8 %

13. DISTRACTION

With the claim *"Digital culture distracts me while I am focused on something specific (ex. If I receive a notification, I almost immediately check my account)"* 9.4 % of the participants (15) strongly disagree, 12.5 % (20) disagree, and 20 % (32) remain neutral on the topic. 36.9 % (59) agree, whereas 21.3 % (34) strongly agree with the claim (as presented in Table 13).

TABLE 13. Response distribution to the claim: "Digital culture distracts me while I am focused on something specific (ex. If I receive a notification, I almost immediately check my account)." (N=160)

Digital culture distracts me while I am focused on something specific	Frequency of responses	Students in percentages
Strongly disagree	15	9.4 %
Disagree	20	12.5 %
Neither agree nor disagree	32	20 %
Agree	59	36.9 %
Strongly Agree	34	21.3 %

14. DIGITAL CULTURE PARTICIPATORS' OBLIGATION OBSTRUCTION

Considering the claim *"Digital culture obstructs me while I am finishing my obligations (business or academic) and/or other activities (free time spent in the company of friends or relatives)"* the results are as follows: 17.5 % (28) strongly disagree, 20.6 % (33) disagree, 21.3 % (34) remain neutral, 24.4 % (39) agree and 16.3 % (26) of the participants strongly agree with the claim (as presented in Table 14).

TABLE 14. Response distribution to the claim: *"Digital culture obstructs me while I am finishing my obligations (business or academic) and/or other activities (free time spent in the company of friends or relatives)"* (N=160)

Digital culture obstructs me while I finish my obligations	Frequency of responses	Students in percentages
Strongly disagree	28	17.5 %
Disagree	33	20.6 %
Neither agree nor disagree	34	21.3 %
Agree	39	24.4 %
Strongly Agree	26	16.3 %

15. SENSE OF TIME

The majority of the participants in the questionnaire agree with the claim *"I lose sense of time whilst consuming digital cultures (ex. I do not notice how much time I spend preoccupied with digital content – videos, TV shows)",* or to be exact, 31.9 % (51) agree and 24.4 % (39) strongly agree with the claim. 22.5 % (36) neither agree nor disagree with the claim, 11.9 % (19) disagree and 9.4 % (15) strongly disagree with the claim (as presented in Table 15).

TABLE 15. Response distribution to the claim: *"I lose sense of time whilst consuming digital cultures (ex. I do not notice how much time I spend preoccupied with digital content – videos, TV shows)"* (N=160)

I lose sense of time whilst consuming digital cultures	Frequency of responses	Students in percentages
Strongly disagree	15	9.4 %
Disagree	19	11.9 %
Neither agree nor disagree	36	22.5 %
Agree	51	31.9 %
Strongly Agree	39	24.4 %

VI. CONCLUSION

In the research, three levels were applied within the framework of digital culture on which the students, or participants of the said research, were self-questioned on emotional, behavioral, and cognitive levels. Taken into consideration that the participants of the questionnaire were only students at VERN' University and Algebra University College, there is room for further study of digital culture users in a wider spectrum of the population.

In the first group of questions the results have shown that students are familiar with their own participation in digital culture. However, they are unaware of the influence digital culture imposes.

The second group of questions was on an emotional level, and the results have shown that most of the students feel a connection to digital culture and that way of spending free time relaxes them. For the minority of students, it

makes them feel restlessness or creates internal tension. For the time being, the majority of the students do not have a fear of missing out when they are using digital culture.

The results of the behavioral questions have shown that a large group of students often participates in digital culture and that it distracts them from doing various other activities, which leads do losing sense of time. Students are continuous users of the digital environment, although they are aware of the influence it imposes on them and their own participation role in engaging with the content. Taking into consideration the development of the digital environment in the last couple of years, it can be assumed that digital culture is more present in the life of an individual from day to day in various aspects.

REFERENCES

[1] A. Uzelac, B. Cvjetičanin, "Digital Culture: The Changing Synamics", Institute for International Relations, Zagreb, 2008.

[2] H. Jenkins, "Convergence culture: Where old and new media collide", NYU Press, New York, 2006.

[3] A. Duranti, "A Companion to Linguistic Anthropology", Blackwell Publishing Ltd, UK, 2004.

[4] M. Gupta A. Sharma, "Fear of missing out: A brief overview of origin, theoretical underpinnings and relationship with mental health. World J Clin Cases. 2021 Jul 6;9(19): pp. 4881-4889. doi: 10.12998/wjcc.v9.i19.4881. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8283615/#:~:text=Fear%20of%20missing%20out%20(FoMO,to%20maintain%20these%20social%20connections., [accessed 25.4.2023.]

[5] M. M. Linehan, "Cognitive-behavioral treatment of borderline personality disorter", Guilford Press, New York, 1993.

[6] V. Boričević Maršanić and assoc., "Usredotočena svjesnost – mindfulness: primjena u djece i adolescenata", Soc. Psihijat., Zagreb, 2015.

Students' Perception of Influencer Marketing on Instagram

Dominik Blažević*, Tihana Babić**
* VERN' University, Zagreb, Croatia
** Algebra University College, Zagreb, Croatia
doblazevic@vernnet.hr; tihana.babic@algebra.hr

***Abstract* – Influencer marketing on social media is one of the fastest-growing trends in marketing and advertising campaigns. Advertisers are investing more and more funds in that type of marketing because influencers have big and loyal fan bases that they influence, especially when it comes to buying the products that they advertise. Social media platform Instagram is one of the most popular platforms for influencer marketing, given that it is a social media for posting photos and videos, and the captions are usually short (few words and hashtags), so they correspond to what the target audience of most of the advertisers (generation Z) wants. The research presented in this paper was conducted in the winter semester of 2022/2023 at VERN University and Algebra University College, on a sample of 106 students. The purpose of this research was to examine students' perception of influencer marketing, and whether they think that this method of marketing is something that encourages them to purchase a product that is advertised.**

***Keywords* – influencer marketing; influencers; Instagram; social media; digital marketing; college students**

I. INTRODUCTION

Marketing is a profession that is present in every aspect of life in modern networked society. There is a saying that says marketing is everywhere around us, and that statement can be considered true because advertising is present in all segments of life (as ads in physical (real) world or ads on digital media).

Digital media have become more popular than traditional media, given that smartphones are the most popular devices for browsing digital media (social media, web news portals). Smartphones have enabled users to always be online, regardless of where they are, resulting in users being able to access media content whenever they want. Another feature of smartphones in digital media is the fact that smartphones allow users to create their own content, thus allowing users to actively participate in the creation of media content. Some users decide to create content that is not related to what the media publishes, so that they can use the social media platform to express their opinions or to earn money, if they have enough followers. The number of followers has become an important part of digital media because users who have many followers can be considered influencers.

According to Statista, Instagram is the leading platform for influencer marketing on digital media, which is why this platform was chosen as the research topic for this paper. 0 According to Cambridge dictionary, there are two definitions of influencers. The first one is "someone who affects or changes the way that other people behave.", and the second one is "a person who is paid by a company to show and describe its products and services on social media, encouraging other people to buy them".[2]

The two definitions of influencers presented in this paper show the difference between the word influencer and the meaning of influencer as a job in a modern networked society. Although the mentioned definition shows the work of an influencer that is strictly related to marketing, it is not the only work that influencers do, but for some influencers it is the main source of income.

The purpose of research presented in this paper is to examine students' perception of influencer marketing, and their considerations on influencer marketing as a method that potentially encourages them to purchase a product that is advertised.

II. TYPES OF INFLUENCERS

When starting a marketing campaign that involves influencers, advertisers need to be aware that there are different types of influencers. The type of influencer is important because the content should be similar to advertisers' products. Along with influencer content, there is another variable that is important for choosing an influencer, and that is the number of followers.

According to Jana Gross and Florian von Wangenheim, there are four types of influencers. Snoopers, Entertainers, Informers and Infotainers. Snoopers are creators that can be described as social media explorers who publish and share content as a hobby, and their goal is to find audience that has similar interest. Entertainers are creators who entertain their audience with their content. Informers are creators who provide information through their content. Infotainers are creators who combine information and entertainment.[3]

According to SocialMediaToday, five types of influencers are nano, micro, mid-tier, macro, and mega influencers. The distinction between influencers is defined by the number of followers that they have:

1. *nano* influencers - have one to ten thousand followers,
2. *micro* influencers - have ten to fifty thousand followers,
3. *mid-tier* influencers - have fifty thousand to five hundred thousand followers,
4. *macro* influencers - have five hundred thousand to a million followers, and
5. *mega* influencers - have over a million followers.[4]

III. INSTAGRAM

There is no exact definition of Instagram in literature, so it's hard to define it without using words that have already been used by other authors who gave their own definitions of Instagram, but Instagram can be defined as a free social media/network platform designed for publishing photos and videos and for connecting people. As already stated in this paper, Instagram is the most popular social media for influencer marketing campaigns. According to Mircea Fuciu, Instagram is the second largest social media platform, after Facebook. [5]

In addition to the number of followers of the influencer and all previously mentioned divisions, one of the most important things on Instagram is engagement. Engagement can be defined as the overall number of user interactions, according to Anil Narassiguin and Selina Sargent [6]. Engagement can be interpreted as: if an influencer has 10,000 followers on Instagram and has an average of 50 likes and comments per post, that is bad engagement, because this is a low engagement rate (of follower likes). Engagement on Instagram (but also on other social media) is another important factor that advertisers must consider when choosing influencers for marketing campaigns.

Instagram is a platform that is mostly used by younger users; according to Statista, 61,6 % of Instagram users are aged 18 – 34.[7] This statistical data is important for this paper, considering that the focus of this paper is to examine the opinion of students about Influencer Marketing on Instagram. Also, this data shows that students use Instagram the most, because most students belong to that age group. Those users can be considered a part of Generation Z because, according to House buyers 4u, users who were born between 1996 and 2010 are part of Generation Z. [8]

IV. INFLUENCER MARKETING

Influencer marketing is a relatively new type of marketing which is becoming more and more popular every year. In addition to being very popular, influencer marketing is also very profitable for advertisers and organizations. According to the influencer marketing hub, advertisers and organizations earn 5.2 dollars for every dollar spent on influencer marketing.[5] There is no unique definition of influencer marketing because that type of marketing is still new on the market in the context of social media influencers. The previously mentioned definition of the Cambridge dictionary in which the term influencer is defined as a profession, can be used as one of the definitions for influencer marketing on social media.[2]

According to Rong-Ho Lin, Christine Jan, and Chun-Ling Chuang, social media influencers can be considered independent or third-party endorsers who can affect audiences' opinions about a brand or an organization. Influencers create content to raise the number of brand followers and to raise brand awareness. [10] In addition to influencers, users are also an important part of a brand's strategy on social media. Social media allows users to create and share user-generated content, which serves as a source of information for any user who wants to be informed about a product or organization. [11] The above definition explains the importance of content in influencer marketing.

As already stated in this paper, influencers are divided into different types according to content. Another division is the division of the form of content used in the campaigns (how the message is presented to a potential buyer). Therefore, it is important to know what kind of presentation to use for chosen content and how to convey it to the user. According to Engaio digital, there are ten types of influencer marketing campaigns:

1. Giveaways and sweepstakes,
2. Contests and event activations,
3. Social media takeovers,
4. Affiliate marketing,
5. Sponsored posts,
6. Sponsored blogs,
7. Unboxing and reviews,
8. Guest blogs,
9. Brand ambassador programs,
10. Pre-release campaigns. [12]

Sponsored social media posts usually have a hashtag (#) that is used by an influencer in that particular marketing campaign. Influencer marketing hub defines a Hashtag campaign as a marketing campaign in which a specific hashtag is used to promote a brand or a product.[13] Another type of marketing that is closely related to hashtag campaigns is Viral marketing. According to Sprout social, Viral marketing is a method of promotion that expects audience activity (mainly sharing and creating their own content based on the advertised product) [14]. Audience activity in the campaign can be measured by analyzing how much the hashtag of the particular marketing campaign was used by the audience in their own posts.

V. PREVIOUS RESEARCH

Two previous researches that are presented in this part of the paper had similar questions to the ones that were asked in the survey that follows. The first research results showed that the vast majority of the respondents buy a product after seeing an influencer's ad on Instagram. As for the most common use of social media, respondents answered that they use Instagram the most. And the last question that was highlighted was about influencer's ability to shape someone's opinion about a brand. The majority of respondents answered that influencers do in fact change their opinion about a brand or a product. [15]

The second research also had questions that are closely related to the questions asked in the survey conducted for this paper. As to the question concerning influencers and their recommendations on social media, most of the respondents answered that have a neutral opinion about that topic. That means that they don't consider influencer recommendations to be an important part of an influencer marketing campaign. The question that was highlighted was about buying habits. Most respondents answered that influencers do not influence their buying habits. [16]

VI. THE RESEARCH METODOLOGY

A. Research Design

The point of this research was to examine students' opinion about influencer marketing on Instagram. This study wanted to find out if students would buy a product after they see that product advertised by an influencer. This study also wanted to investigate if students think that having an influencer participate in advertising a product is important to encourage them to buy the product.

B. Research sample

The research for this paper was conducted on a sample of 106 students from Algebra University College and VERN University. This research was conducted in the winter semester of academic year 2022/2023. Undergraduate and graduate students were surveyed. 77 students answered that they attend Algebra University College, and 29 students answered that they attend VERN' University. The first five questions in the survey were informational. The questions were year of study, status of the study, gender and already mentioned questions on the level of the study and the University that they attend.

63.2 % of students who participated were female, and 36.8 % were male (N=106). 86.8 % of students who participated were undergraduate students, and 13.2 % were graduate students (N=106).

C. Data collection and analysis

The answers for this paper were collected by an online survey that was created using Google forms. The students had a link on their student portals to voluntarily access the survey. The survey had 10 questions + 5 already mentioned informational questions. All questions were close-ended and had a pre-offered answers. 9 questions were asked on the Likert scale and 1 question had 2 offered answers. Students' answers are presented by gender. Division by Universities is not shown considering that the majority (72.6 %) of surveyed students attend Algebra University College.

VII. THE RESEARCH RESULTS

In this part of the paper, the results of 10 question related to the research are shown. The results are displayed in percentages.

Question 1: ***How often do you buy a product, after seeing an ad on Instagram?***

The respondents' answers to the first survey question show that 72.6 % (of which 38.7 % are female and 33.9 % are male) of surveyed students never or rarely buy products that are advertised on Instagram. (N=106)

These results show that surveyed students do not consider Instagram to be an important advertising "tool" in marketing campaigns.

The summary results of all respondents' answers are presented in Table 1.

TABLE 1. HOW OFTEN STUDENTS BUY A PRODUCT, AFTER SEEING AN AD ON INSTAGRAM (N=106)

FREQUENCY OF BUYING A PRODUCT AFTER SEEING AN INSTAGRAM AD	Students in percentages
Never	36.8 %
Rarely	35.8 %
Sometimes	22.6 %
Often	3,8 %
Always	0,9 %

Question 2: ***What type of content do you prefer in Instagram advertising?***

The second question in the survey was related to the type of content that students prefer in advertising on Instagram. As mentioned earlier in the theoretical part of this paper, content is a very important part of Instagram and influencer advertising. 55.7 % (of which 36.8 % are female and 18.9 % are male) of students answered that they prefer ads with photos and text, and 44.3 % (of which 26.4 % are female and 17.9 % are male) of students answered that they prefer video ads (N=106).

Question 3: ***Do you agree with this statement: "Instagram influencer marketing is a way of advertising that encourages me to buy (a product)?"***

The third question of this survey refers to students' attitudes toward influencer marketing as a way of advertising a product or service. This question was asked to see how many students think that influencers can encourage them to buy a product.

50.9 % (of which 22.6 % are female and 28.3 % are male) of surveyed students answered that they completely or partially disagree with the statement, while 33 % (of which 29.2 % are female and 3.8 % are male) of surveyed students answered that they completely or partially agree with the statement. The summary results of all respondents' answers are presented in Table 2.

The results showed that influencer marketing on Instagram can encourage a third of students, mainly female, to purchase a product.

TABLE 2. STUDENTS' AGREEMENT WITH THE STATEMENT: "INSTAGRAM INFLUENCER MARKETING IS A WAY OF ADVERTISING THAT ENCOURAGES ME TO BUY (A PRODUCT)" (N =106)

FREQUENCY OF INFLUENCER MARKETING ENCOURAGING TO BUY	Students in percentages
I completely disagree	28.3 %
I partially disagree	22.6 %
I neither agree, nor disagree	16 %
I partially agree	22.6 %
I completely agree	10.4 %

Question 4: ***How often do you participate in Hashtag campaigns (Viral marketing)?***

The fourth question of this survey is about the use of hashtags and viral marketing. Brands and influencers often use hashtags (#) in their campaigns. The hashtag is made for users to put that hashtag in their post so advertis-

ers and influencers can see how many users participated in their campaign. 71.7 % (of which 42.5 % are female and 29.2 % are male) of students answered that they never participate in hashtag campaigns.

These results show that the majority of the surveyed students never participate in hashtag campaigns. The summary results of all respondents' answers are presented in Table 3.

TABLE 3. STUDENTS' PARTICIPATION IN HASHTAG CAMPAIGNS (VIRAL MARKETING) (N=106)

FREQUENCY OF PARTICIPATION IN HASHTAG CAMPAIGNS	Students in percentages
Never	71.7 %
Rarely	17.9 %
Sometimes	9.4 %
Often	0.9 %
Always	0 %

Question 5: ***Do you agree with this statement: "It is important to me that an influencer participates in the campaign of a product that is advertised to me on Instagram?"***

The fifth question in this survey refers to students' preferences in advertising on social media. Due to the growth in the popularity of influencer marketing, this question aimed to check whether students think that influencers are an important part of the marketing campaign of the product that they are interested in.

57.5 % (of which 33 % are female and 24.5 % are male) of students answered that they completely or partially disagree with that statement, but 22.7 % agree at least partially. The summary of all respondents' answers is presented in Table 4.

TABLE 4. STUDENTS AGREEMENT WITH THE STATEMENT: "IT IS IMPORTANT TO ME THAT AN INFLUENCER PARTICIPATES IN THE CAMPAIGN OF A PRODUCT THAT IS ADVERTISED TO ME ON INSTAGRAM" (N=106)

IMPORTANCE OF INFLUNCER PARTICIPATION IN A MARKETING CAMPAIGN	Students in percentages
I completely disagree	43.4 %
I partially disagree	14.2 %
I neither agree nor disagree	19.8 %
I partially agree	14.2 %
I completely agree	8.5 %

Question 6**: *Do you support Instagram campaigns involving influencers that are presented negatively by the media (those influencers usually have a large number of followers)?***

Influencers are public figures, given that they address the public, so that is why the media write about them. Popular influencers are usually the subject of media articles, and they are almost always presented in a negative context. This question was asked to see if students are affected by the fact that media writes negatively about an influencer that is a part of a campaign. 51 % (of which 35.9 % are female and 15.1 % are male) of students answered that they completely or partially do not support those campaigns and 38.7 % (of which 19.8 % are female and 18.9 % are male) of students answered that they neither support nor not support those campaigns.

These results show that the majority of surveyed students do not support those campaigns. The summary of all respondents 'answers is presented in Table 5.

TABLE 5. STUDENTS' SUPPORT TO INSTAGRAM CAMPAIGNS INVOLVING INFLUENCERS THAT ARE PRESENTED NEGATIVELY BY THE MEDIA (N=106)

SUPPORTING INFLUENCERS PRESENTED NEGATIVELY BY THE MEDIA	Students in percentages
I completely do not support	36.8 %
I partially do no support	14.2 %
I neither support nor do not support	38.7 %
I partially support	4.7 %
I completely support	5.7 %

Question 7: ***Do you agree with this statement: "Through advertising, influencers focus on a connection with a product (brand), while classic ads are only focused on sales?"***

Advertisers often hire influencers to increase sales of their products, but also to raise brand awareness. This question was asked to examine students' views on the role of influencers in Instagram campaigns.

34 % (of which 18 % are women and 16 % are male) of students answered that they completely or partially disagree with that statement, while 40.6 % (of which 27.4 % are female and 13.2 % are male) of students answered that they completely or partially agree with that statement.

These results are not straightforward, but it can be concluded that surveyed students have very different opinions about this statement, with a larger number of surveyed students who believe that the statement is correct. The summary result of respondents' answers is presented in Table 6.

TABLE 6. STUDENTS' AGREEMENT WITH THE STATEMENT: "THROUGH ADVERTISING, INFLUENCERS FOCUS ON A CONNECTION WITH A PRODUCT (BRAND), WHILE CLASSIC ADS ARE ONLY FOCUSED ON SALES" (N=106)

AGREEMENT WITH THE STATEMENT THAT INFLUENCERS FOCUS ON SALES AND BRAND AWARENESS	Students in percentages
I completely disagree	20.8 %
I partially disagree	13.2 %
I neither agree nor disagree	25.5 %
I partially agree	28.3 %
I completely agree	12. 3%

Question 8: ***How often do you search for product information on Instagram?***

This question was asked to investigate if students prefer being informed about a product on Instagram.

26.4 % (of which 8.5 % are female and 17.9 % are male) of students answered that they never search for product information on Instagram, while 55.7 % (of which 37.7 % are female and 18 % are male) of students answered that they rarely or sometimes search for product information on Instagram. These results show that surveyed students do not consider Instagram to be an important media for finding information about a product.

The summary results of all respondents' answers are presented in Table 7.

TABLE 7. STUDENT SEARCH FOR PRODUCT INFORMATION ON INSTAGRAM (N=106)

FREQUENCY OF SEARCHES FOR PRODUCT INFORMATION ON INSTAGRAM	Students in percentages
Never	26.4 %
Rarely	27.4 %
Sometimes	28.3 %
Often	14.2 %
Always	3.8 %

Question 9: ***How important is it to you that the content that influencers publish is similar to the product that they are advertising?***

Content is a crucial part of an influencer marketing campaign because the advertised product should be similar to the influencer's content. For example, if someone is a sports influencer, that person could not (or rather should not) advertise alcohol. This question was asked to examine what students think about the importance of content in an influencer marketing campaign on Instagram.

57.5 % (of which 42.5 % are female and 15 % are male) of students answered that influencers' and product's content similarity is completely or partially important to them. These results show that surveyed students think content is an important part of an influencer marketing campaign.

The summary results of all respondents' answers are presented in Table 8.

TABLE 8. STUDENT IMPORTANCE ASSESSMENT THAT THE CONTENT WHICH INFLUENCERS PUBLISH IS SIMILAR TO THE PRODUCT THAT THEY ARE ADVERTISING (N=106)

IMPORTANCE OF INFLUENCERS' CONTENT SIMILARITY TO PRODUCT THEY ADVERTISE	**Students in percentages**
Not important at all	17 %
Partially not important	3.8 %
Neither important nor not important	21.7 %
Partially important	29.2 %
Completely important	28.3 %

Question 10: ***How often do you participate in giveaways organized by influencers in collaboration with a brand on Instagram?***

This question was asked to see if students participate in this type of influencer marketing campaigns. 44.3 % (of which 21.7 % are female and 22.6 % are male) of students answered that they never participate in giveaways and 14.2 % (of which 11.4 % are female and 2.8 % are male) of students answered that they often participate in giveaways.

These results show that the majority of surveyed students do not participate in giveaways but there is a solid percentage of those who often participate, so it is hard to provide a correct conclusion about this question.

TABLE 9. STUDENTS' PARTICIPATION IN GIVEAWAYS ORGANIZED BY INFLUNCERS IN COLLABOARTION WITH A BRAND ON INSTAGRAM (N=106)

FREQUENCY OF PARTICIPATION IN INFLUENCER GIVEAWAYS	Students in percentages
Never	44.3 %
Rarely	25.5 %
Sometimes	15.1 %
Often	14.2 %
Always	0.9 %

VIII. DISCUSSION AND CONLUSION

The purpose of this paper was to examine students' perception of influencer marketing, and to find out if that type of marketing is something that encourages them to buy a product advertised by an influencer.

The first question of this research showed the direction in which the answers of the surveyed students will go, considering that 72.6 % of students answered that they do not buy a product after seeing an ad on Instagram. Hashtag campaigns are also a type of influencer marketing campaign which are not popular with surveyed students. 76 % of them answered that they do not participate in hashtag campaigns on Instagram.

And the last question that was very one-sided was the question about participating in giveaways. 69.8 % of students answered that they never or rarely ever participate in giveaways that are organized by influencers in collaboration with the brand. These results show that the majority of surveyed students do not participate in campaigns in which their engagement is expected.

Questions that were strictly related to influencer marketing showed that for a little above half of students influencer marketing is not a way of advertising that encourages them to buy, but also that for a third of them it is.

For more than half of the students that participated in this survey, influencer's participation in marketing campaigns for products that are being advertised to them on Instagram is not important to the students, but for every seventh student their participation is important.

Also, more than a half of surveyed students never or rarely search for product information on Instagram. This can be the case considering that users are probably using

Google search to inform about a product. This research did not have many obstacles, but it would be recommendable to survey a larger number of respondents to reach much more concrete conclusions.

These results show that majority of surveyed students believe that influencers do not influence their decision to buy a product, and more than half of the students answered that they never or rarely search for product information on Instagram.

It can be concluded that the results of previous research do not match the results collected in this research. This fact leads us to the conclusion that the influence of social media is individual, that is, each user uses social media in their way, but also that each user has their way of interpreting the content presented to them. This information is important for people involved in influencer marketing, since the individualization of content is very important for the success of every campaign.

REFERENCES

[1] Statista Research Department: "Influencer marketing worldwide – statistics & facts", Statista, Jan.19, 2023. Available at https://www.statista.com/topics/2496/influence-marketing/#topicOverview [Accessed 6.2.2023.]

[2] Cambridge Dictionary: "Influencer", Cambridge University Press 2023. Available at: https://dictionary.cambridge.org/dictionary/english/influencer [Accessed 6.2.2023]

[3] J. Gross, F. V. Wangenheim: "The Big Four of Influencer Marketing". A Typology of Influencers", Marketing Review St. Gallen, Aug.24, 2018, Vol.2, 30-38. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3230687 [Accessed 12.2.2023.]

[4] M. Walker-Ford: "4 Influencer Marketing Trends that will Dominate in 2020 [Infographic]", SocialMediaToday, Jan. 22, 2020. Available at : https://www.socialmediatoday.com/news/4-influencer-marketing-trends-that-will-dominate-in-2020-infographic/570703/ [Accessed 12.2.2023]

[5] M. Fuciu: "The rise of Instagram – Evolution, Statistics, Advantages and Disadvantages", Revista Economică, 2019. 71:4. Available at http://economice.ulbsibiu.ro/revista.economica/archive/71404fuciu.pdf [Accessed 16.2.2023.]

[6] A. Narassiguin, S. Sargent: "Data Science for Influencer Marketing: feature processing and quantitative analysis", Cornell University, May. 10, 2019. Available at: https://arxiv.org/abs/1906.05911 [Accessed 16.2.2023.]

[7] S. Dixon: "Distribution of Instagram users worldwide as of January 2023, by age group", Statista 2023, Feb.14, 2023. Available at : https://www.statista.com/statistics/325587/instagram-global-age-group/ [Accessed 16.2.2023]

[8] House buyers 4u: "Move over Millennials: Gen Z is the New Focus for the Property market". Available at : https://housebuyers4u.co.uk/blog/move-over-millennials-gen-z-is-the-new-focus-for-the-property-market/ [Accessed 11.3.2023.]

[9] W. Geyser: "What is Influencer Marketing? – The Ultimate Guide for 2023", InfluencerMarketingHub, Jan. 20, 2023. Available at: https://influencermarketinghub.com/influencer-marketing/#toc-1 [Accessed 15.2.2023.]

[10] R.H. Lin, C. Jan, C.L. Chuang: "Influencer Marketing on Instagram", International Journal of Innovation in Management, May. 11, 2019, Vol.7, No.1, 33-41. Available at: https://siim.org.tw/IJIiM/DW/V7N1/IJIiM-19-020.pdf [Accessed 16.2.2023.]

[11] S.S. Lee, E. Kim: "Influencer Marketing on Instagram : How sponsorship disclosure, influencer credibility and brand credibility impact the effectiveness of Instagram promotional post", Journal of Global Fashion Marketing, Routledge, July 2020, Vol.11, No.3. Available at : https://www.researchgate.net/publication/342195205_Influencer_marketing_on_Instagram_How_sponsorship_disclosure_influencer_credibility_and_brand_credibility_impact_the_effectiveness_of_Instagram_promotional_post [Accessed 16.2.2023.]

[12] Engaio Digital : "8 Influencer Marketing Campaign Types, 2023. Available at: https://engaiodigital.com/influencer-marketing-campaign-types/ [Accessed 16.2.2023.]

[13] W. Geyser: "8 Best Hashtag Campaigns on Social Media and Why You Need One", InfluencerMarketingHub, Jan. 4, 2023. Available at: https://influencermarketinghub.com/hashtag-campaigns/#toc-0 [Accessed 11.3.2023.]

[14] B. Barnhart: "What is viral marketing? (And does it work in 2023)?", Sprout social, Feb. 21, 2023. Available at: https://sproutsocial.com/insights/viral-marketing/ [Accessed 11.3.2023.

[15] Da Costa, I. O. M. (2019): "The impact of influencer marketing on consumer purchase intentions and brand attitude: the instagrammers (Doctoral dissertation, ISCTE – Instituto Universitario de Lisboa (Portugal)). Available at: https://repositorio.iscte-iul.pt/bitstream/10071/19035/1/master_ines_milheiro_costa.pdf [Accessed 11.3.2023.]

[16] Kolar, A. (2022): "Odnos Influencera i publike – povjerenje, društveni utjecaj i promjene u ponašanju" (Master's thesis, Sveučilište Sjever, Koprivnica (Croatia)). Available at: https://zir.nsk.hr/islandora/object/unin%3A5094 [Accessed 12.3.2023.]

Tolerance for Disagreement in Communication via Instagram

Dora Njegovan*, Bianka Frigan*, Tihana Babić**
* VERN University, Zagreb, Croatia
** Algebra University College, Zagreb, Croatia
donjegovan@vernnet.hr, bifrigan@vernnet.hr, tihana.babic@algebra.hr

***Abstract* - Instagram represents the most used social media in the Republic of Croatia among the age group between 18 and 34 year olds, which serves them to publish edited photos and videos. Nowadays, Instagram also represents one of the dominant digital communication channels and the ability to communicate successfully as one of the most important life and business skills is equally important in the virtual world. Successful communication requires tolerance, and intolerance leads to a lack of understanding in communication and disagreements that can lead to efficient or unsuccessful cooperation. The research presented in this paper aimed to determine how students self-assess their tolerance for disagreement on Instagram, whether they enjoy when there is a disagreement in communication via Instagram, if they find it amusing, if they try to resolve it, or if they try to avoid the disagreement. The research was conducted on a sample of students from Algebra University College and VERN University in the winter semester of the academic year 2022/2023. The results showed that students consider the differences in opinions to be generally useful, they prefer talking to others who have the same beliefs as them and agree with their opinions.**

Keywords - tolerance; disagreement tolerance; communication; digital communication; Instagram

I. INTRODUCTION

Social media platforms like Instagram allow billions of users to post, create and share different information and content. They also allow users to use them for communication with different people around the world. There are many different social media platforms, and Instagram, along with Facebook, represents the most used social media in the Republic of Croatia, among the age group between 18 and 34 year olds [1],[2].

Nowadays, Instagram also represents one of the dominant digital communication channels, and the ability to communicate successfully as one of the most important life and business skills is equally important in the virtual world.

Successful communication requires tolerance, and intolerance leads to a lack of understanding in communication and disagreements, which can lead to efficient or unsuccessful cooperation. Instagram tends to have a lot of cyber bullyingon the platform, even though the platform does not promote violence on the internet, which mostly comes from disagreeing on different topics. That's why this research was conducted to see how VERN' University and Algebra University College students tolerate disagreement in communication via Instagram.

II. INSTAGRAM

Instagram represents one of the dominant digital communication channels. The studies according to the Paw Research Centre that were conducted in 2018, show that Instagram, together with YouTube, is one of the most popular platforms among adolescents [3].

The last update (28/06/21) shows that the total number of daily Instagram users has come to 1.9 billion users worldwide. Daily, there are more than 500 million Instagram stories posted by active users and the monthly number of users is more than two billion, according to the latest update (22/02/22) [4].

Instagram is a social media platform where people can upload photos and videos and share them with their followers or specific group of people/friends. They can also like, comment, and share other posts with their friends. If a person wants to have an Instagram account, they must be 13 or older. If the accounts represent someone younger than 13, they must state in the account bio that the account is managed by a guardian of the child [5].

Other research has shown that most teens describe the effect of social media as positive (31 %) or negative (24 %), but the largest number of them (45 %) say that the effects are neither positive nor negative [3].

Instagram tends to have a lot of cyber bullyinginvolving violence on the platform, even though the platform does not promote violence on the internet. Violence, in general, comes from disagreeing on certain topics, which leads to discussions and intolerance for others' beliefs [6],[7].

III. COMMUNICATION ON INSTAGRAM

The word "communicate" comes from the word "common", which in Latin means "to share" [8]. Communication is the process of exchanging information, ideas, and feelings between two or more people. Effective communication is crucial in all areas of life, from personal relationships to business and politics. Firstly, communication is essential because it allows individuals to connect and build relationships with others [8].

Tolerance is an essential aspect of effective communication. Communication involves the exchange of ideas, opinions, and information among people, and tolerance allows people to express themselves freely without the fear of discrimination or judgment. "Tolerance does not imply condescension, but benevolence, readiness for respectful dialogue and cooperation" [9], it "is the

ability and skill to tolerate something foreign, to put up with someone else's opinion or character" [9]. "Mastering linguistic tolerance helps to solve the problem of verbal aggression in communication, reduce the level of discourse, reduce the negative impact of discourse, express a tolerant attitude to the object, and ultimately achieve a communicative intention" [9].

It is necessary to consider tolerance in the context of social media, taking into account that very often people do not have the feeling that there are people on the other side, because of "the lack of effective feedback about the damage and pain that their behavior can cause the victims, as well as the opinion that their behavior cannot be considered immoral because there are no visible consequences and the distance from the damage caused by their activities" [10],[11].

IV. RESEARCH METHODOLOGY

A. Research Design

The research was conducted among VERN' University and Algebra University College students in the academic year 2022/2023, during January and February 2022.

The goal of the research was to examine how people react to intolerance on social media platforms such as Instagram.

The survey was conducted on Google Forms, where the participants anonymously answered 13 statements and the students could choose to which level they agree or disagree with each statement. Each statement had five possible answers according to the Likert scale of agreement.

The statements which were used in the research were adjusted according to Gilbert Hoffman's research [12]. Convenience sampling was chosen as a sampling method , and the total number of respondents to all survey questions was 141 (N =141).

B. Research Sample

The participants were female and male students from the age of 18 to 27+. Most were from the age of 18 to 21 (33.3 %), 27 % were from the age of 21 to 24, 19.9 % were from the age of 24 to 27, while 19.9 % of themwere 27+ years old. 63.1 % (89) were female and 36.9 % (52) male students from VERN' University and Algebra University College.

78.7 % were attending the bachelor's degree program and 21.3 % attended the master's degree program.

VERN' University (66 %)

- Economics of entrepreneurship (19.4 %)
- Film, television, and multimedia design (1.1 %)
- Journalism (1.1 %)
- Public relations and media studies (18.3 %)
- Business informatics (3.2 %)
- Technical management (1.1 %)
- Tourism (10.8 %)
- Internet of Things (IoT) (2.2 %)
- Cyber communications and network science (16.1 %)
- IT management (8.6 %)
- Management of sustainable tourism development (4.3 %)
- Entrepreneurial management (1.1 %)
- Accounting and finance (2.2 %)
- Human resources management (8.6 %)
- Management of business communications (2.2 %)

Algebra University College (34 %)

- Software engineering (29.2 %)
- System engineering (8.3 %)
- Multimedia computing (8.3 %)
- Data science (6.2 %)
- Digital marketing (27.1 %)
- Visual communications design (20.8 %)

C. Research results

Students from VERN' University and Algebra University College were first asked if they use Instagram. If they answered "no", the survey ended for them (from a total of 141 participants, 14 (9.9 %) do not use Instagram). If they answered "yes" (91.1 % of the respondents), they were given 13 statements, which included the following:

1. IT IS MORE FUN TO ENGAGE IN A DISCUSSION WHERE THERE'S A LOT OF DISAGREEMENT

The research results have shown that most of the students, ie. 36.5 % of students, don't have an opinion on this given statement *"It's more fun to engage in a discussion on Instagram where there's a lot of disagreement"*. 28.6 % strongly disagree, 20.6 % disagree with the statement, 7.9 % strongly agree and 6.3 % agree with the statement (as presented in Table 1). The total number of student respondents to survey statement 1 was N = 126.

TABLE 1 STUDENTS' RESPONSES TO SURVEY STATEMENT 1 (N = 126)

It is more fun to engage in a discussion on Instagram where there's a lot of disagreement	**Frequency of responses**	**Students in percentages**
Strongly disagree	36	28.6 %
Disagree	26	20.6 %
No opinion	46	36.5 %
Agree	8	6.3 %
Strongly agree	10	8 %

2. ENJOYING TALKING TO PEOPLEON INSTAGRAM WHO HAVE DIFFERENT POINTS OF VIEW

Results have shown that 10.4 % of students strongly agree, and16.8 % agree with the following statement *"I enjoy talking to people on Instagram who have different points of view than mine"*. 22.4 % agree while 22.4 % strongly disagree. 28 % of students don't have an opinion

(as presented in Table 2). The total number of student respondents to survey statement 2 was N = 125.

TABLE 2 STUDENTS' RESPONSES TO SURVEY STATEMENT 2 (N = 125)

I enjoy talking to people on Instagram who have different points of view than mine	Frequency of responses	Students in percentages
Strongly disagree	28	22.4 %
Disagree	28	22.4 %
No opinion	35	28 %
Agree	21	16.8 %
Strongly agree	13	10.4 %

3. PARTICIPATION IN COMMUNICATIONS WITH PEOPLE WHO HAVE DIFFERENT OPINIONS

Results have shown that 16.8 % of students strongly disagree and 24.8 % disagree with the following statement *"I don't like to be a participant in communication on Instagram where people have different opinions",* while 12.8 % strongly agree and 16.8 % agree with the statement. 28.8 % of students don't have an opinion on this statement. The total number of student respondents to survey statement 3 was N = 125 (as presented in Table 3).

TABLE 3 STUDENTS' RESPONSES TO SURVEY STATEMENT 3 (N = 125)

I don't like to be a participant in communication on Instagram where people have different opinions	Frequency of responses	Students in percentages
Strongly disagree	21	16.8 %
Disagree	31	24.8 %
No opinion	36	28.8 %
Agree	21	16.8 %
Strongly agree	16	12.8 %

4. TALKING TO OTHERS WITH THE SAME BELIEFS

The results have shown that most students (36 %) agree and 16 % strongly agree with the statement *"I prefer talking to others who have the same beliefs as mine"*. 8 % disagree, 8 % strongly disagree when it comes to talking with others who have the same beliefs as them, and 32 % of survey respondents don't have an opinion on this statement (as shown in Table 4). The total number of student respondents to survey statement 4 was N = 124.

TABLE 4 STUDENTS' RESPONSES TO SURVEY STATEMENT 4 (N = 124)

I prefer talking to others who have the same beliefs as mine	Frequency of responses	Students in percentages
Strongly disagree	10	8 %
Disagree	10	8 %
No opinion	40	32 %
Agree	44	36 %
Strongly agree	20	16 %

5. DIFFERENCES IN OPINION ARE USEFUL

The fifth survey statement: *"Differences in opinion are generally useful"* has 51.2 % of students agreeing with the statement, 32 % strongly agreeing, 0.8 % of students strongly disagreeing and 2.4 % disagreeing with the statement. 13.6 % of survey respondents don't have an opinion on this statement (as presented in Table 5). The total number of student respondents to survey statement 5 was N = 125.

TABLE 5 STUDENTS' RESPONSES TO SURVEY STATEMENT 5 (N = 125)

Differences in opinion are generally useful	Frequency of responses	Students in percentages
Strongly disagree	1	0.8 %
Disagree	3	2.4 %
No opinion	17	13.6 %
Agree	64	51.2 %
Strongly agree	40	32 %

6. CHANGING THE SUBJECT OF THE CONVERSATION WHEN THERE IS A DISAGREEMENT

Most students, 39.2 % of them, have responded to the statement *"I prefer to change the subject of the conversation on Instagram when there is a disagreement",* that they don't have an opinion. 16 % disagreed and 24 % strongly disagreed, while 11.2 % agreed and 9.6 % strongly agreed with the statement (as shown in Table 6). The total number of student respondents to survey statement 6 was N = 125.

TABLE 6 STUDENTS' RESPONSES TO SURVEY STATEMENT 6 (N = 125)

I prefer to change the subject of the conversation on Instagram when there is a disagreement	Frequency of responses	Students in percentages
Strongly disagree	30	24 %
Disagree	20	16 %
No opinion	49	39.2 %
Agree	14	11.2 %
Strongly agree	12	9.6 %

7. TENDING TO CREATE DISAGREEMENT IN CONVERSATION BECAUSE IT SERVES A PURPOSE

The results have shown that 52 % of students strongly disagreed with the statement *"I tend to create disagreements in Instagram conversations because they serve a useful purpose",* while only 5.6 % strongly agreed. 4 % of them agreed and 21.6 % disagreed with the statement. 16.8 % of students have no opinion when it comes to creating disagreements in conversations on purpose (as shown in Table 7). The total number of student respondents to survey statement 7 was N = 125.

TABLE 7 STUDENTS' RESPONSES TO SURVEY STATEMENT 7 (N = 125)

I tend to create disagreements in Instagram conversations because they serve a useful purpose	Frequency of responses	Students in percentages
Strongly disagree	65	52 %
Disagree	27	21.6 %
No opinion	21	16.8 %
Agree	5	4 %
Strongly agree	7	5.6 %

8. ENJOYING DISCUSSIONS WITH OTHERS ON INSTAGRAM ABOUT TOPICS THEY DISAGREE ON

Most students, 33.9 % of them, strongly disagreed and 16.9 % disagreed with the statement *"I enjoy discussions with others on Instagram about topics where we disagree".* 29 % of students marked the statement with "No opinion", while only 8.9 % strongly agreed and 11.3 % agreed with the statement (Table 8). The total number of student respondents to survey statement 8 was N = 124.

TABLE 8 STUDENTS' RESPONSES TO SURVEY STATEMENT 8 (N = 124)

I enjoy discussions with others on Instagram about topics where we disagree	Frequency of responses	Students in percentages
Strongly disagree	42	33.9 %
Disagree	21	16.9 %
No opinion	36	29 %
Agree	14	11.3 %
Strongly agree	11	8.9 %

9. DISSLIKING TO DISAGREE WITH OTHER PEOPLE

The same percentage of students, 22.4 %, marked "No opinion" and "Strongly agree" on the statement *"I don't like disagreeing with other people on Instagram"*, while 25.6 % "Agree". The same results have shown that 13.6 % of students strongly disagreed and 16 % disagreed with the statement (as presented in Table 9). 22.4 % of students have no opinion. The total number of student respondents to survey statement 9 was N = 125.

TABLE 9 STUDENTS' RESPONSES TO SURVEY STATEMENT 9 (N = 125)

I don't like disagreeing with other people on Instagram	Frequency of responses	Students in percentages
Strongly disagree	17	13.6 %
Disagree	20	16 %
No opinion	28	22.4 %
Agree	32	25.6 %
Strongly agree	28	22.4 %

10. RATHER LEAVE THE CONVERSATION THAN CONTINUE TO DISAGREE

Most students, 25 % of them, strongly agreed with the following statement *"If I could choose, I would rather leave the conversation than continue the disagreement"*, while 15.3 % of them strongly disagreed (as shown in Table 10). The second marked answer to the statement was "No opinion" (24.2 %). 12.9 % of students disagreed and 22.6 % of them agreed. The total number of student respondents to survey statement 10 was N = 124.

TABLE 10 STUDENTS' RESPONSES TO SURVEY STATEMENT 10 (N = 124)

If I could choose, I would rather leave the conversation than continue the disagreement	Frequency of responses	Students in percentages
Strongly disagree	19	15.3 %
Disagree	16	12.9 %
No opinion	30	24.2 %
Agree	28	22.6 %
Strongly agree	31	25 %

11. AVOIDING TALKING TO PEOPLE WITH DIFFERENT OPINIONS

The results show that most students, 33.9 % of them, don't have an opinion. 19.4 % of students disagreed and 17.7 % strongly disagreed with the statement *"I avoid talking to people on Instagram who I don't think will agree with me"*. 12.9 % agreed with the statement and 16.1 % strongly agreed (as shown in Table 11). The total number of student respondents to survey statement 11 was N = 124.

TABLE 11 STUDENTS' RESPONSES TO SURVEY STATEMENT 11 (N = 124)

I avoid talking to people on Instagram who I don't think will agree with me	Frequency of responses	Students in percentages
Strongly disagree	22	17.7 %
Disagree	24	19.4 %
No opinion	42	33.9 %
Agree	16	12.9 %
Strongly agree	20	16.1 %

12. ENJOYING DISAGREEING WITH OTHERS

To the twelfth survey statement: *"I enjoy disagreeing with others on Instagram"*, 51.6 % of students marked their answer as "Strongly disagree" and 15.3 % marked "Disagree", while only 4.8 % of students strongly agreed with the statement (as shown in Table 12). 22.6 % of students have no opinion on the following statement, while 5.6 % agreed with the statement. The total number of student respondents to survey statement 12 was N = 124.

TABLE 12 STUDENTS' RESPONSES TO SURVEY STATEMENT 12 (N = 124)

I enjoy disagreeing with others on Instagram	Frequency of responses	Students in percentages
Strongly disagree	64	51.6 %
Disagree	19	15.3 %
No opinion	28	22.6 %
Agree	7	5.6 %
Strongly agree	6	4.8 %

13. DIFFERENCES ENCOURAGE COMMUNICATION

The largest number of respondents (36.8 %) had no opinion on the (final) 13th statement: *"Differences in opinion on Instagram encourage me to communicate"*. The second largest number of respondents (22.4 %) strongly disagreed and 18.4 % disagreed with the statement (as presented in Table 13). 13.6 % agreed and 8.8 % strongly agreed. The total number of student respondents to survey statement 13 was N = 125.

TABLE 13 STUDENTS' RESPONSES TO SURVEY STATEMENT 13 (N = 125)

Differences in opinion on Instagram encourage me to communicate	Frequency of responses	Students in percentages
Strongly disagree	28	22.4 %
Disagree	23	18.4 %
No opinion	46	36.8 %
Agree	17	13.6 %
Strongly agree	11	8.8 %

V. DISCUSSION AND CONCLUSION

The research conducted for the purposes of this paper was carried out on a sample of students from VERN' University and Algebra University College in Croatia, during the academic year 2022/2023. The students that were surveyed in the research were from different study fields and were between the ages of 18 to 27+.

The goal of this research was to see how students from VERN' University and Algebra University College react to disagreements in communication via Instagram, whether they enjoy when there is a disagreement in communication via Instagram, if they find it amusing, if they try to resolve it, or if they try to avoid it.

The results show that most students who participated in the research think that differences in opinions are generally useful, but the research also shows that they don't like disagreeing with others on Instagram. Also, it shows that they like to hear each other's opinions, but don't like discussing them.

In addition to the stated above, from the answers of the participants, it is not entirely clear why, but it can be noted that a quarter or even a third of students do not have an opinion about any of the statements that were featured in the research.

It can be concluded that even though the surveyed students said that the differences in opinion are generally useful, they prefer talking to those who have the same beliefs as them and agree with their opinions.

This research results gave a closer look at tolerance for disagreement in communication via Instagram, although it was conducted on a small sample of students from VERN' University and Algebra University College in Croatia. Numerous studies have been conducted around the world among young people on Instagram and the impact it has on their quality of life, but as far as communication itself is concerned, more research has been conducted on the subject of cyberbullying than on tolerance for disagreement in communication via Instagram. The focus of this research on the tolerance of disagreement in communication via Instagram is the contribution of this paper, which is manifested in the approach to cyber communication, that is, on Instagram from a different angle; the focus was shifted from harmful consequences to the possibilities of tolerance. The research results can be used for researches to follow, on different social media platforms, or comparisons with students from other universities.

REFERENCES

[1] NapoleonCat, "Instagram users in Croatia", December 2022, Available at https://napoleoncat.com/stats/instagram-users-in-croatia/2022/12/ [accessed: 10. 2. 2023.]

[2] NapoleonCat, "Facebook users in Croatia", December 2022, Available at https://napoleoncat.com/stats/facebook-users-in-croatia/2022/12/ [accessed: 10. 2. 2023.]

[3] M. Anderson, J. Jiang: "Teens, social media & technology", DC: Pew Research Center, Washington, 2018. Available at https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed: 10. 2. 2023.]

[4] S. Aslam: "Instagram by the numbers: Stats, Demographics & Fun Facts", Omnicore, 2018. Available at https://www.omnicoreagency.com/instagram-statistics/

[5] Instagram Help Center, 2023, Available at https://help.instagram.com/424737657584573 [accessed: 16. 2. 2023.]

[6] P. Suciu, "Cyberbullying Remains Rampant On Social Media", Forbes, 2021. Available at https://help.instagram.com/424737657584573 [accessed: 25. 4. 2023.]

[7] D. M. B. Tarigan, L. Riwu, S. Monika, "Use of Cyberbullying Language on Instagram Social Media", ICSS 2022. Available at https://www.shs-conferences.org/articles/shsconf/pdf/2022/19/shsconf_icss2022_01047.pdf [accessed: 25. 4. 2023.]

[8] K. E. Rosengren: "Communication: An Introduction", SAGE Publications Ltd, London, 2000.

[9] J. Zhu: "Manifestation of Tolerance in the Aspect of the Communicative-Pragmatic Approach", Advances in Social Science, Education, and Humanities Research, volume 517, Atlantis Press, 2020. . Available at https://www.atlantis-press.com/article/125951682.pdf [accessed: 20. 2. 2023.]

[10] V. Bilić: "Načini učenja, online aktivnosti i ishodi odgoja net-generacije", Istraživanje paradigmi djetinjstva, odgoja i obrazovanja, Učiteljski fakultet, Zagreb, 2015, pp. 593-607.

[11] T. Babić: „Promjena komunikacijske paradigme u visokom obrazovanju pod utjecajem društvenih medija, Filozofski fakultet Sveučilišta u Zagrebu, Zagreb, 2021. Available at https://www.academia.edu/42565081/Dru%C5%A1tveni_mediji_i_upravljanje_komunikacijama [accessed: 20. 2. 2023.]

[12] G. Hofmann: "Komunacijske kompetencije fizioterapeuta", Filozofski fakultet, Sveučilište u Zagrebu, Zagreb, 2023. Available at https://repozitorij.ffzg.unizg.hr/islandora/object/ffzg:4022 [accessed: 16. 2. 2023.]

Using Drones in Teaching Computer Science

P. Voštinár
Matej Bel University, Department of Computer Science, Banská Bystrica, Slovakia
patrik.vostinar@umb.sk

***Abstract* - More and more teachers from primary and secondary schools are using some type robots, microcontrollers, or some new tools for teaching programming. Thanks to the lower prices of these robotic building blocks and aids, students are better prepared for their future profession. We used four types of drones in the teaching process in primary school and in the extracurricular activity at the Matej Bel University for teaching programming. The contribution deals with describing our experience of using drones in the educational process. In this paper, we introduced a methodic focus on using drones in the educational process.**

Keywords – Airblock; micro:bit; LiteBee Wing; Ryze Tello; programming;

I. INTRODUCTION

Currently, there are many professions in the world that are constantly changing with the development of mankind. New professions are emerging, and some, on the contrary, are disappearing. There are professions that make their work more efficient thanks to new information and communication technologies. For this reason, it is essential that students in primary schools become familiar with new technologies.

Research in our country and around the world shows that the use of ICT technologies increases clarity and efficiency in the teaching process. Thanks to this education, it is possible to use various robotic kits, microcontrollers, and create interesting inter-subject projects, which mostly need to be programmed, which is also why it is important that students learn programming already in the first grade of primary schools. Pupils already use micro:bit educational boards in (not only) computer science classes, they model 3D objects that are then printed on a 3D printer, they use various popular games for teaching not only computer science (e.g. Minecraft: Education Edition), they create mobile applications (e.g. in App Inventor) etc.

Drones are enjoying increasing popularity thanks to their low prices (it's no problem to get a drone with a camera for just a few tens of euros) and modified legislation that allows for their wide, but at the same time, safe use. This is also why it is essential that students in elementary schools already study this technology, so that they know what they can and cannot do with drones, so that they can try to control the drone under the guidance of teachers and not by themselves. Added value is if they know how to programme their own controller for drones or programme the drone autonomously.

Various companies that develop educational hardware and software have also realized the growing availability of drones. This popularity was not missed even by the manufacturers of the most popular robotic construction companies, MakeBlock, Elecfreaks, MakeBlock, DJI, etc.

II. PREVIOUS RESEARCH

Several studies have been devoted to the use of drones in the teaching of computer science at various levels of schools. Velasco and Valente [1] provided a mapping review of the online courses available for drone education. Sanches et al. [2] reviewed some of the most accessible and economical commercial drones for use in various educational environments. Lobo, et al. [3] report in their research an active learning-based instructional approach that prepares students for careers in the drone industry (drone programming using Python, designing and fabricating drones using Computer-Aided Design (CAD)). Their research was implemented in an undergraduate class that engineering and computer science students attended. Quantitative assessments indicated that most students excelled at the technical subjects covered in the class. Phang et al. [4] in their research focused on the use of drones for drone technology for primary school students to learn STEM. Their research described that the use of drones in education was effective for young children. Kinaneva et al. [5] used in their research block-based programming to teach programming with drones. The students who participated in the courses have acquired new knowledge by launching drones to implement different tasks and missions on demand, and at the same time they had fun participating in the classes, which made the educational process more efficient. Yepes et al. [6] analyzed the feasibility of using a set of technologies based on drones, designed based on the theory of significant learning through active methodologies. The study was carried out with 30 high school students and it was found that the workshops with the platform based on drones helped in the understanding, construction, and interpretation of the content covered. Fokides et al. [7] in their study presented the results of a pilot project in which the use of drones to teach primary school students was examined. The target group was 40 fifth-grade primary school students, divided into two groups. The first group was taught with the use of drones, while the second was taught conventionally. The results indicated that the students in the drone group outperformed the students in the conventional teaching group on the Maths evaluation sheet and on all delayed post-tests. In the other two cases (Physics and Geography evaluation sheets), the results were the same. Students' attitudes toward drones were highly positive. Sivanes and Koutromanos [8] investigated the perceptions of 60 in-service teachers' regarding the use of drones through mobile applications. The results showed positive perceptions about using drones through mobile applications. Regarding the facilitation conditions, teachers

indicated the need for teacher training on the pedagogical use of drones, assistance with visual programming, and infrastructure support. Engineering schools have also recognized the need to prepare professionals entering the job market with experience in drone operation, leading to the development of a curriculum focused on this niche [9-15].

III. DRONES SUITABLE FOR THE EDUCATION

The increasing popularity of drones among young people was not missed even by manufacturers of the most popular robotic construction companies, who decided to make drones that can be used in the teaching process.

A. *Airblock drone*

One of the first manufacturers to offer drones for education is MakeBlock, the company behind the popular mBot robots. They created a modular drone called the Airblock, which can also be converted into a hovercraft in water and air. Airblock has won several awards. The Airblock consists of a main module and six additional hexagon-shaped modules that can be attached to the main module in various ways using magnets. Figure 1 shows a picture of all the possibilities of how the drone can be built.



Figure 1. Airblock - modular drone

Airblock has several advantages - it does not contain a camera and thus there is no need to deal with GDPR approvals, it is highly impact resistant (made of high-quality foam) and can be used indoors (Airblock is light, in case of strong wind, we do not recommend using it outdoors, how much the drone can be blown away). We program it using blocks in the mobile application for iOS and Android.

B. *LiteBee Wing*

LiteBee Wing is a drone inspired by the children's favorite bricks building activity - compatible with Lego bricks. Drone could be programmed using a Scratch or Python programming language. For programming is possible to use software for computer or mobile application (iOS and Android).[1]

Figure 2 shows the LiteBee Wing.

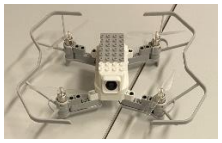


Figure 2. LiteBee Wing drone

C. *Ryze Tech - Tello*

Ryze Tello is one of the cheapest drones sold here. It is not primarily intended for teaching programming, but with the Tello Edu mobile app it can be programmed using blocks. For programming, it is possible to use the commands take off, land, set the speed, take a photo, record a video, use conditions, cycles, variables, basic mathematical operations, work with sensors. Figure 3 shows the Ryze Tello.



Figure 3. Ryze Tello drone with Telle Edu app

D. *Elecfreaks micro:bit dron:bit kit*

This drone uses one of the most popular educational aids for programming, micro:bit. Because of this, it is not possible to use a mobile device to program the drone, but we have to use the website https://makecode.microbit.org/ where we have to add the drone:bit extension to our project. After adding the extension, new blocks will appear with which the drone can be programmed. The advantage of this drone is the possibility to program your own game controller, with which we will program the drone using radio communication - we will set the controller and the drone to the same radio frequency and then we will send commands, what the drone should do (e.g. after pressing the joystick on the controller up, we will send the drone a number 1, which will signal that the drone should climb). Figure 4 shows drone:bit.



Figure 4. Drone:bit

E. *Comparison of drones suitable for teaching programming*

In our opinion, the most suitable drone for teaching in the classroom is the Airblock, because when it hits a wall, we will not damage either the drone or the wall, the drone will fall apart and then we can build it again. LiteBee Wing disappointed us the most, because it lasted the least, we spent the most time on calibration, and after about a month it broke down and we had to send it for a claim. The Ryze Tello and drone:bit drones, which lasted the longest, surprised us with their endurance and the possibility of programming.

[1] https://www.litebee.com/product/liteBeeWing/download/

Table 1 shows a comparison of the drones available in our department.

TABLE I. COMPARISON OF DRONES

	Airblock	**LiteBee Wing**	**Ryze Tello**	**Drone:bit**
Dimensions	230 x 222 x 53	210 x 180 x 50	98 x 92,5 x 41	410 x 410 x 51
Weight	141 g	128g	80 g	509 g
Battery Life	6-8 minutes	10 min	13 min	8 min
Control Distance	10 meters	100m	100 m	100 m
Indoor environment	yes	yes	yes	yes
Outdoor environment	no	no	yes	yes
Charging time	60 min	60 min	90 min	60 min
Camera	no	HD	HD	no
Android, iOS app	yes	yes	yes	no
Programming in PC	yes	yes	yes	yes
Programming languages	block based language	block based language	block based language	block based language, JavaScript, MicroPython
Extensions	no	yes	no	yes
Price	179 €	190 €	109 €	177 €

IV. COMPUTER SCIENCE EXTRACURRICULAR ACTIVITY

Since the academic year 2017/2018, our Department of Computer Science (Matej Bel University in Banska Bystrica) has been organizing free extracurricular activities for students in primary schools. The first year we started with the frequency of one 90 minute lesson once a week. After the first year, based on the great interest of the children, we decided to organized this activity for two groups. The main aim of this activity was to teach children programming to various types of robotics. The secondary objective was to gain more exercises for our teaching trainees of computer science. Each of our students in the department taught 2-3 children in primary school.

During the Covid-19 pandemic (from March 2020) we were teaching online through Microsoft Teams. The content of our activity was changed to software for teaching programming (Minecraft: Education Edition, App Inventor, Scratch, beginning of Python, HTML). During Covid-19, students from our department were preparing teaching materials for the lessons.

In the academic year 2022/2023 we decided to organize this activity only once per week but in two parallel classrooms. The reason was that 9 students are visiting our activity for four years and the rest 15 primary school children started this year for the first time. These two groups are having mainly different content for their lessons. The only similar activity was drone teaching. This year, each group was taught by a student from the department. Figure 5 shows our students from the department in the teaching process.



Figure 5. Teaching drones (our students from department are lectures)

Figure 6 shows the testing of three drones in our hallway.

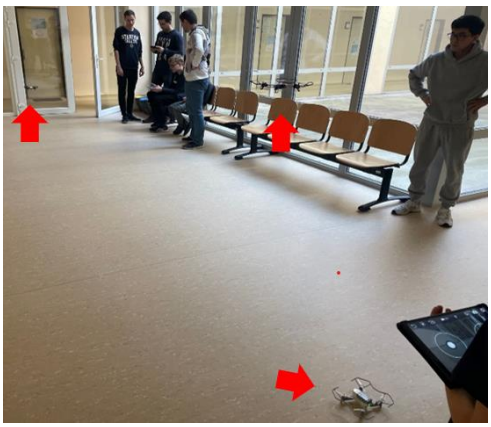


Figure 6. Drone testing.

V. DRONE METHODIC

We prepare various types of tasks to teach programming with drones. In this article, we describe one type of task that was used in extracurricular activities. Table II shows our methodology of using drones.

TABLE II. METHODOLOGY OF USING DRONES

Activity name:	Drones Programming
Target group:	7.-9. year primary school
Expected time:	90 minutes
Goal:	Cognitive goal - the student learns to program using visual programming. Affective goal - The student is able to respond correctly to regulations and rules.
Methods and forms:	• Investigative and research methods, • methods of individual work of students, • narration, description, explanation,

	• question and answer method, • methods of practising and improving skills.
Preparatio, Teaching Aids:	Materials: Computer, iOS or Android mobile device, drone:bit, Airblock drone, LiteBee Wing drone, Ryze Tello drone

Today, we will learn how to use drones and how to programming drones. At the beginning fill out the work list.

Work list

A drone can be an unmanned _____. Regardless of the type of drone, it is characterized by control on _____. The drone can be used, e.g. in (fill in the possible use):
- to the army- _____
healthcare- _____
- police- _____
- agriculture- _____
- companies- _____
- real estate companies - _____

The drone _____ to be used in the vicinity of airports, as there could be a _____ with a plane or helicopter. Drones also _____ to be used to monitor neighbors because it infringes on their _____.
Airblock is a programming drone that can be used (rebuilt) for the following:
- _____ (the air)
- _____ (Water)
- _____ (the earth)

We can program the drone in the Makeblock application. For programming, we use a language that is similar to another block language._____. We will use English blocks to control the drone:
- power on/off- _____
- forward- _____
- turn left/right- _____
- rising/downward - _____
- hover/landing- _____
- shake _____

When we want to set some action that the drone will perform, we set the behavior of this action in the _____ component (also the English name).
If we want the drone to wait, e.g. For 1 second, we assign the command _____ (wait). To set the colors that the drone will display in its control section, we will combine RGB colors: R_____, G_____, B _____.
Before connecting the drone to the mobile device, we need to build the drone. The Airblock drone consists of a control unit and 6 propellers. We connect the propellers to the drone. _____ is used for the connection. To connect the drone, we need to turn on _____ technology in the mobile device. To connect, we have to press the _____ (connect) button.

Task 1:
The drone (it does not matter the specific type) must:
- Wait 5 seconds and take off
- Move 30 centimeters forwards
- Land

Figure 7 shows the solution to 1.

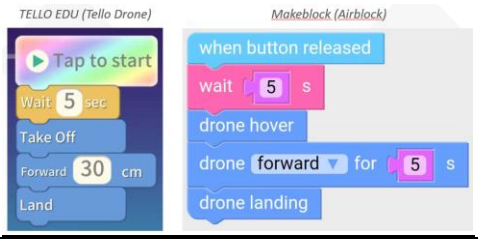




Figure 7. Solution of task 1

Task 2 - square:
The drone must take off and draw square in the air and then land.

Figure 8 shows the solution to task 2.

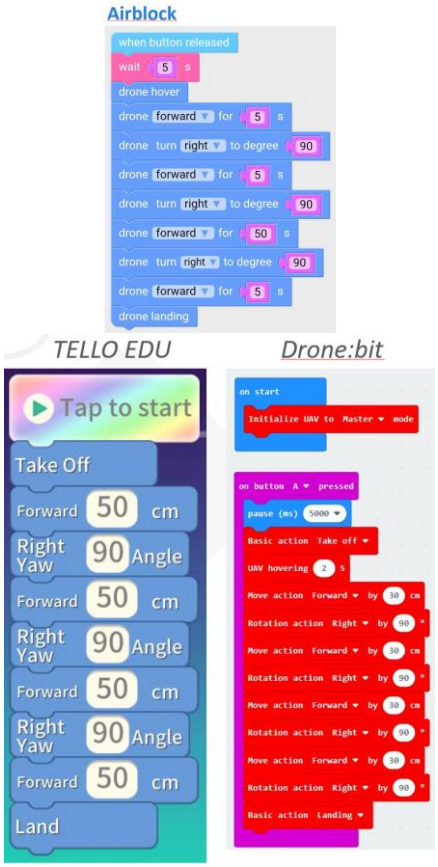


Figure 8. Solution of task 2

Task 3:
The drone must take off and draw:
- drawing 2D geometrical shapes (triangle., rectangle, circle, etc.)
- drawing 3D geometrical shapes (cube, rectangular prism)
- traffic light: drone turns on, shows the red colour for 1 second, wait 1 second, change the colour to orange for 1 second, wait 1 second, change the colour to green for 1 second and hover until 1 metre,
- Record video, take off to a height of 150cm, slowly turn 360 degrees, wait 5 seconds, land, stop video.

VI. METHODOLOGY AND RESEARCH

Our research was similar to other researchers from section 2, we would like to find out if using drones could be interesting, and motivational for students also in the Slovak republic. Our research was conducted in the academic year 2022/23 and still lasts. We were based on our experience in teaching the Airblock drone in the academic year 2017/2018 (the children really liked working with the drone). In our research, we used four types of various drones for learning programming in our extracurricular activity at our university. The main aim of our research was to find out whether we can increase students' interest in studying programming by using drones compared to other robots that we used in our extracurricular activity. A secondary aim was to find out which drone is the best for teaching. We used the questionnaire to measure the data of our research. We asked 24 primary school children several questions in the questionnaire. 22 boys and 2 girls completed the questionnaire. Respondents were students in the fifth and eighth grade (11-14 years old).

The questionnaire consisted of 11 questions. First, two questions focused on gender and the school year they attended.

The next question was aimed at if the children had piloted a drone before. Figure 7 shows the results. Almost all of them said no (20 respondents no, 4 respondents yes).

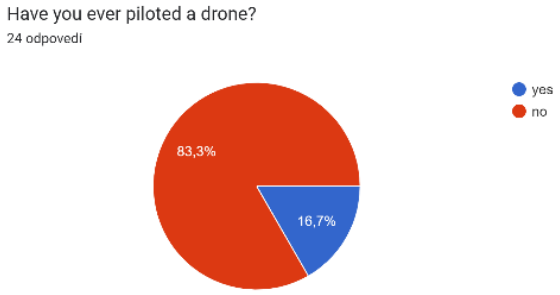


Figure 7. Have you ever piloted a drone?

The third question asked them if they had programmed drones before. All respondents choose not to.

The next question asked if these activities make sense to them. All respondents select yes.

Based on the answers from question Were the classes difficult for you? We can state that the activities were not difficult for the respondents. Figure 8 shows the answers.

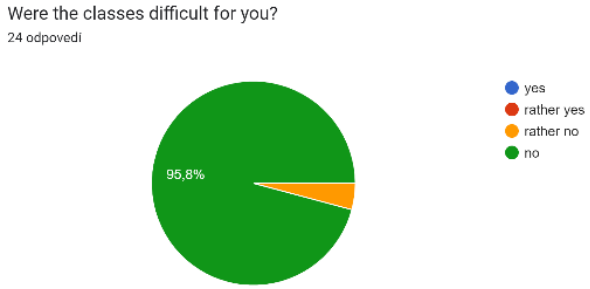


Figure 8. Were the classes difficult for you?

In the question Would you like to continue programming activities with drones? we asked if the students liked the drone activity. Everyone replied that they liked her.

We were curious how they liked working with specific types of drones. Figure 9 shows the answers. The pupils liked the Ryze Tello drone the most and the LiteBee Wing the least.

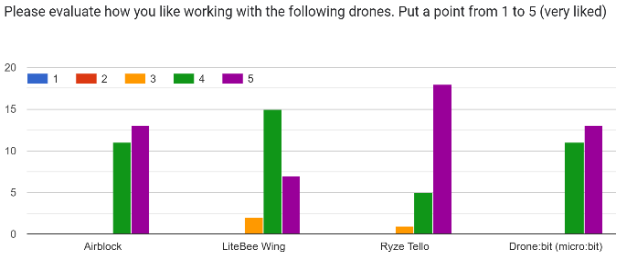


Figure 9. How do children like to work with specific types of drones?

In the next question, we wanted to know which drone the students would take home if they could. Figure 10 shows the answers - children selected drone:bit.

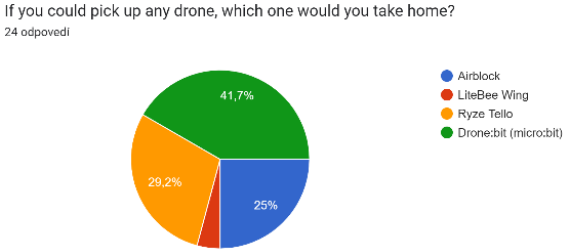


Figure 10. If you could buy any drone, which one would you take home?

In the penultimate answer, we wanted to know which drone programming environment the students liked the most. The answers (Figure 11) revealed MakeBlock, which is used to program the Airblock drone.

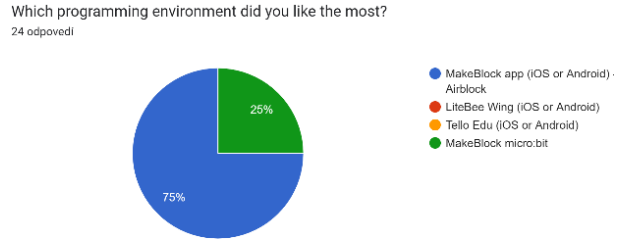


Figure 11. Which programming environment did you like the most?

In the last question, we asked the students which activity they did in the circle they liked the most. The responses (Figure 12) showed that they liked drone programming the most so far.

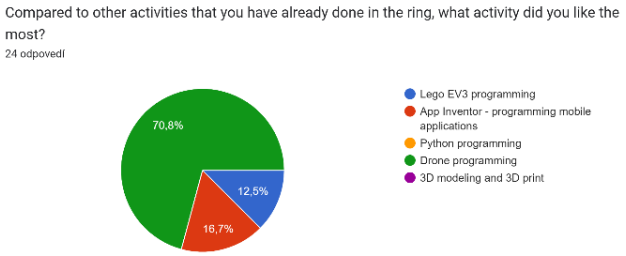


Figure 12. Compared to other activities that you have already done in the ring, what activity did you like the most?

VII. CONCLUSION

In the article, we focus on using drones to teach programming in computer science extracurricular activity. The main purpose was to motivate the children and teach them some programming basics. We described one task as an example of methodological sheets on which the students worked. The questionnaire showed that most of the students had not piloted drones before, none of them had programmed them before. The students liked working with the Ryze Tello drone the most - we assume that it is because of its size (it is the smallest, fastest). Of the drones used, it has the best camera - we assume these are the main reasons why they like it the most. They would most like to take the drone:bit home, perhaps because it lasts the longest and is the largest of them. Among the programming environments, students liked the MakeBlock environment the most - we liked this environment the most, it is intuitive to use. Of the activities in the circle so far, the students liked the activity with the drone the most. Our research confirms similar results from other researchers, using drones can increase interest about studying programming or computer science

ACKNOWLEDGEMENT

This contribution has been processed as part of grant project no. 004TTU-4/2021 Teaching Mathematics and Computer Science Using Interactive Components and project no. 001UMB-4/2023 Implementation of blended learning in the training of professional bachelors and teachers of mathematics and computer science.

REFERENCES

[1] O. Velasco and J. Valente, "Online drone education, a mapping review,' in Proc. IEEE Global Eng. Educ. Conf. (EDUCON), Apr. 2020, pp. 1286–1289.

[2] Searson, Mike, et al. "Panel: Flying High (and Low)! Drones in Education." Society for Information Technology & Teacher Education International Conference. Association for the Advancement of Computing in Education (AACE), 2020.

[3] D. Lobo, D. Patel, J. Morainvile, P. Shekhar and P. Abichandani, "Preparing Students for Drone Careers Using Active Learning Instruction," in IEEE Access, vol. 9, 2021, pp. 126216-126230.

[4] F. A. Phang, J. Pusppanathan, N. D. Nawi and F. K. Ch. Harun, "Drone technology for primary school students in learning STEM", AIP Conference Proceedings, 2022.

[5] A. P. D. Kinaneva, G. Hristov, P. Zahariev and E. J. Raychev, "Play to learn: using drone-aircrafts and block-based programming for improving learning success rates".

[6] I. Yepes, D. A. C. Barone and C. M. D. Porciuncula, "Use of drones as pedagogical technology in STEM disciplines, " in Informatics in Education, vol. 21, pp. 201-233, 2022.

[7] E. Fokides, D. Papadakis, and V. Kourtis-Kazoullis, "To drone or not to drone? Results of a pilot study in primary school settings," in Journal of Computers in Education, vol. 4, pp. 339-353, 2017.

[8] T. Sivenas, G. Koutromanos, "Using Mobile Applications to Interact with Drones: A Teachers' Perception Study," in: Auer, M.E., Tsiatsos, T. (eds) New Realities, Mobile Systems and Applications. IMCL 2021. Lecture Notes in Networks and Systems, vol 411. Springer, Cham.

[9] M. Sadraey, "Unmanned aerial vehicles design education; techniques and challenges," in Proc. ASEE Virtual Annu. Conf. Content Access, 2020.

[10] J. M. Cañas, D. Martín-Martín, P. Arias, J. Vega, D. Roldán-Álvarez, L. García-Pérez, and J. Fernández-Conde, "Open-source drone programming course for distance engineering education," Electronics, vol. 9, no. 12, 2020, p. 2163.

[11] H. Norman, N. Nordin, M. A. Embi, H. Zaini, and M. Ally, "A framework of drone-based learning (Dronagogy) for higher education in the fourth industrial revolution," in Int. J. Eng. Technol., vol. 7, nos. 3–14, 2018, pp. 1–6.

[12] E. S. McDaniel, S. M. Alnaeli, D. A. Juckem, and W. S. Vaz, "Motivating undergraduate computing and engineering research via educational and scientific drones," in Proc. IEEE Int. Conf. Electro Inf. Technol. (EIT), 2019, pp. 536–541.

[13] V. Farr and G. Light, "Integrated STEM helps drone education fly, " in Proc. IEEE Integr. STEM Educ. Conf. (ISEC), 2019, pp. 398–401.

[14] C. S. Maniu, C. Vlad, T. Chevet, S. Bertrand, A. Venturino, G. Rousseau, and S. Olaru, "Control systems engineering made easy: Motivating Students through experimentation on UAVs," in Proc. 21st IFAC World Congr., 2020.

[15] J. Ryu, S. K. LaPaglia and R. Walters, "Idaho Drone League (iDrone) to Stimulate STEM workforce," in Journal of STEM Education, vol. 21, pp. 35–41, 2020.

A Critical Analysis of Students' Cheating in Online Assessment in Higher Education: Post-COVID-19 Issues and Challenges Related to Conversational Artificial Intelligence

G. Bubaš, A. Čižmešija
University of Zagreb, Faculty of Organization and Informatics, Varaždin, Croatia
{gbubas, acizmesi}@foi.unizg.hr

***Abstract* – The latest advances in conversational artificial intelligence (CAI) systems that are available to the general public have intensified the debate on cheating in online assessment in higher education. The issue of cheating in online assessment in the educational context had already been given significant attention owing to school closures (lockdown periods) during the COVID-19 pandemic and the consequent transition from on-site to online teaching and assessment. In our paper we present statistical data on cheating in online courses and academic dishonesty in general, as well as an overview of motives for and methods of cheating reported in the literature, with a focus on the period of the COVID-19 pandemic. Recommendations are also provided regarding the use of specific techniques and tools to prevent cheating in online exams. Since plagiarism is an important problem in academic assessment the potential misuses of CAI in higher education are also briefly analyzed. This paper comprises a synthesis of related review papers and meta studies with reports of empirical research findings, as well as the experiences of the authors of the paper.**

Keywords – e-learning, online assessment, cheating, COVID-19 pandemic, conversational artificial intelligence, review, critical analysis

I. INTRODUCTION

Quality in educational and psychological assessment has been in the focus of professionals for decades [1]. The Standard 6.6. of the *Standards for Educational and Psychological Testing* published in 2014 [2, 116-117] has been labeled "Reasonable efforts should be made to ensure the integrity of test scores by eliminating opportunities for test takers to attain scores by fraudulent or deceptive means." In particular, the following recommendations can be found within this standard: "Test developers should design test materials and procedures to minimize the possibility of cheating."; "Test administrators should follow established policies for dealing with any instances of testing irregularity."; "In any evaluation of work products (e.g. portfolios) steps should be taken to ensure that the product represents the test taker's own work, and that the amount and kind of assistance provided is consistent with the intent of the assessment." In addition, the *International Test Commission*, together with the *Association of Test Publishers*, in 2022 issued their *Guidelines for Technology-Based Assessment* [3]. Two of the guidelines in that document should be especially mentioned, i.e., guideline 8.2. declaring that "A testing organization should continuously analyze the risk of cheating and theft threats and adopt, implement, and maintain appropriate solutions for those threats that carry the highest risk", as well as guideline 8.7. which states that "A testing organization should put in place measures to detect and report cheating or content theft and respond to them as quickly as possible. These measures may include data forensics, monitoring Internet sources for disclosed content, monitoring the test taker during the test, and methods to report test fraud when observed."

In her comments of the *Standards for Educational and Psychological Testing*, Camara [4] emphasizes that the term 'secondary users' of testing can include not only teachers and students, but also policymakers and educational reformers who have influence in shaping or implementing educational policy and programs. Such *secondary users*, who have less expertise in measurement, are more likely to misuse assessment in new and challenging ways in comparison to *primary test users* (e.g. psychologists and educational researchers). It is important that all users, including secondary users, make reasonable efforts to follow the *Standards* [2] where applicable.

A paper by Jenkins et al. published in 2022 [5] reported an *increase of concerns* about cheating in online classes during the COVID-19 pandemic that was especially present in the time of *shutdown* of colleges and their shift to predominantly *online* education and assessment. This study performed on 214 students of psychology at one university in the USA has uncovered that, due to the COVID-19 pandemic, *first time cheating* was on the rise and that cheating in *online classes* was more frequently reported in comparison to cheating in *in-person classes*. The overall reporting of cheating in this study was very high – 74.8% of the surveyed students reported that they had cheated at least with one type of graded material (exams, quizzes, homework, and project/paper). Also, cheating in *online classes* was more present in case of exams, quizzes and homework assignments when compared to *traditionally delivered on-site classes*. Interestingly, only one of the students who participated in this survey reported that he/she had been caught cheating.

In a national survey in Germany [6] performed from November to December 2020 the students (N=1,608) were questioned about their academic dishonesty during the summer semester of the 2019/2020 academic year (this was the first COVID-19 semester with lockdowns at most

academic institutions in Europe). The results of this survey also indicated a much higher self-reported frequency of cheating in *online* versus *on-site* exams. A brief review paper performed by Maryon et al. [7] (published in December 2022) analyzed 16 studies that addressed the issue of *online cheating* during the COVID-19 pandemic and found that, due to the sudden change to online delivery of courses, both educators and researchers recognized *negative effects* on students' academic integrity during the peak of this transition.

In two surveys performed in May 2020 [8] (N=789 instructors) and September 2021 [9] (N=2,868 instructors), a team of experts engaged by the academic publisher and service provider Wiley investigated the instructors' concern about students' cheating in *online* course delivery. In the 2020 survey a total of 93% of the surveyed instructors felt that students were "more likely" (31%) or "significantly more likely" (62%) to cheat *online* than *in-person*. A decrease by 16% in this negative perception was found in the later (i.e. 2021) survey, in which the percentage of instructors who reported that the students were more likely to cheat *online* than *in-person* had declined to 77%. It must be noted that the September 2021 Wiley survey also collected data from students (N=682), most of whom (59%) stated it was "easier" or "significantly easier" to cheat *online*, therefore confirming the previously mentioned instructors' concern about online assessment. For comparison purpose, regarding the period *before* the COVID-19 pandemic, it must be mentioned that in an earlier study [10] (N=659 for students and N=303 for faculty), that was published in 2019, the majority of students (63%) and faculty (56%) at one university in the USA perceived cheating and plagiarism as a greater problem in *online classes*, while only a small percentage of students (5%) and faculty (7%) found cheating and plagiarism to be a greater issue in *traditional courses*.

Assessment is an important component of the education process in higher education institutions (HEI). Accordingly, the recommendations from the *Standards for Educational and Psychological Testing* [2] and *Guidelines for Technology-Based Assessment* [3] should be considered when planning and performing various types and modalities of *online* student evaluation. The previously reported empirical findings regarding the *prevalence of students' cheating in online assessment* before, during and at the time of the easing of the COVID-19 pandemic, emphasize the importance of *quality in online assessment* at higher education institutions.

II. STUDENTS' MOTIVES FOR CHEATING IN ONLINE ASSESSMENT

According to a review paper that investigated *motives for cheating in online assessment* published by Noorbehbahani et al. [11] in 2022, the 38 specific means of cheating derived from the analyzed literature can be grouped into the following *four categories of reasons for cheating* (original formulations are slightly adapted): *teacher related* (e.g. poorly designed exams; restraint in punishing cheaters), *institutional* (e.g. lack of strict rules and policies on cheating), *students' learning characteristics* (e.g. inadequate academic skills and performance; low interest in the course and knowledge acquisition), and *students' personality characteristics* (e.g. laziness; risk proneness).

In a survey [12] performed in the winter semester of 2020/2021 academic year among the students of a technical university in Germany (N=416), 58.4% of the respondents stated that they had the opportunity to cheat during the *written online exams*. Also, 73.2% of the surveyed students affirmed that, in their opinion, the *online* assessment formats provide more opportunities for cheating than *on-site* exams. *Opportunity* can be considered an important factor for developing *intent* to cheat. It is noteworthy that a substantial share of 27.6% of the respondents in this German study confirmed they had engaged in at least partial cheating during the semester in which the survey was performed. One of the conclusions of the author of the study was that the group of students that would be *most likely to cheat* in online assessment consists of (a) those who are *morally willing* to do so and (b) those who find an *opportunity* to cheat.

The first *Wiley survey* performed in September 2021 [9] uncovered various facilitating factors for intent of students (N=682) to cheat in academic assessment: *pressure* (to get good grades; cost of education), *workload* (too much work to meet requirements; hard to balance study with other responsibilities) and *relevancy* (course content for examination is not relevant to the student or his/her major). It can also be concluded from the results of this Wiley survey that the students who are *more likely to cheat* have enrolled in college for *instrumental purpose* (because their parents or family wanted that; to make more money; to find a job they are passionate about). On the other hand, according to the second *Wiley survey*, the students are be *less likely to cheat* if they are *more aware of the negative consequences* (if the consequences are mentioned in the syllabus; if getting caught in cheating lowers grade or implies some other penalty; if proctoring devices are used), as well as if the *honor code had to be signed*. Interestingly, in another study [13] (N=60), when consecutive online exams were performed during summer 2020 (all taken by the students from home), it was found that a lesser level of cheating reduction was associated with the *appeal* not to cheat, students' *approval with non-cheating statement* and *time restriction*, in comparison to a greater level of cheating reduction in case in which a *warning statement of surveillance and administrative action* due to suspicious behavior during online testing was present.

In one of the *highly cited review studies* conducted by Whitley [14] and published in 1998, much before the COVID-19 pandemic (for a bibliometric analysis of most cited related studies see: [15]), the following most influential correlates of students' cheating in academic environment were uncovered: *demographic variables* (cheating was most notably associated with a younger age and being in a marital relationship), *ability* (greater association of cheating was found with a lower ability of performing specific tasks), *dishonesty at earlier levels of education* (e.g. in high school), *attitudes and norms related to cheating*, *personality variables* (morality; procrastination), and *situational variables* (honor codes; perceived workload; competition). In an earlier, also highly cited *empirical study* published in 1993, McCabe and Trevino [16] revealed that the students (N=6,096) were most likely to cheat if (1) they perceive that their peers cheat, (2) their perception of probability of being caught is lower, (3) there is less understanding/acceptance of integrity policy, and (4) there is possibility of a severe penalty. A newer study by Arnett Jensen et al. [17], published in 2002, presented the following *categories of*

academic cheating motives (here they are slightly rephrased and reorganized): *personal academic gain* (necessity to pass the exam; need for a good grade; maintaining class ranking), *conformity* (everyone else is cheating), *unfair treatment by instructor* (perceived unfairness of instructor; exam observed as deliberately too hard to pass), *no harm to others* (it isn't a big deal to cheat; no one suffers from my cheating), *low risk of detection or consequence* (students' assumption that they wouldn't get caught; have cheated before without getting caught; not expect much punishment by instructor if caught), *challenge* (maybe I can get away with it; the class is very competitive), *instrumental* (need to pass to get a job), *relationship or image management* (don't want to disappoint parents), and *situational* (didn't have enough time to study; "froze" during exam; currently depresses or do not have enough energy to prepare for exam). Finally, Jurdi et al. [18], in their paper published in 2011, categorized various predictors of cheating behaviors associated with motivation into *demographic*, *psychosocial*, *academic* and *situational* factors. In their study, performed on Canadian students (N=321), they found that *peers' cheating* had, respectively, a greatest *positive* correlation and *deep learning strategies* a greatest *negative* correlation with their composite measure of engagement in dishonest behavior.

This brief overview of selected studies of *cheating motives* before and during the COVID-19 pandemic, even though some of them are not directly related to *online cheating*, lists many incentives and variables associated with student's misconduct in online exams that need to be taken in account when planning remote assessment at HEIs.

III. VARIETY OF ONLINE CHEATING IN ONLINE COURSES

Janke et al. [6] have listed a considerable number of predominantly *offline* forms or techniques for cheating in academic assessment, while Noorbehbahani et al. [11] created a complementary and more comprehensive list of *online* forms of cheating. When these two lists are combined, the following categories and more specific instances of academic dishonesty during online exams can be proposed:

- *Self-reliance in testing* (examinee is using forbidden material – textbook, notes or cheat sheet; examinee performs internet search or uses offline digital resources on one's computer/smartphone).
- *Other-reliance in testing* (copying other persons' answers; getting help from a consultant or group of peers for performing individual assignment; handing in other persons' work; engaging in forbidden collaboration during testing or oral examination; engaging someone else to falsely impersonate the examinee during online testing or oral examination).
- *Plagiarism, self-plagiarism or falsification in written assignments, essays and seminar papers* (copying content from the internet, a book, or an article without naming the source; submitting the same work as a learning assignment for different courses; referencing sources that one has not read or consulted for writing the submitted text; modifying or making up information from scientific and other sources so that they better fit into the text that is being written).

According to the experiences of the authors of this paper, who have conducted numerous online tests and oral examinations during the COVID-19 pandemic, in combination with the observations of Noorbehbahani et al. [11] and other authors, the following modalities of online assessment and related cheating techniques can be defined (in form of three types of online examination):

(a) *Time limited online tests <u>without</u> surveillance* (cheating can include use of textbooks, cheating sheets, post-it notes, digital documents that can be searched with keywords, use of the internet, help from a consultant/collaborator in the same room or accessed remotely, screen sharing, use of imposters).

(b) *Time limited online tests <u>with</u> video surveillance and browser lock* (cheating can include all of the previous forms, with taking care of the angle of the camera view and more careful whispering by the consultant below the level detectable by microphone, redirecting the webcam, use of multiple devices like a second computer or a smartphone, use of headphones/earphones or a hidden earbud or use of external electronic devices like a miniature camera and spy earpiece).

(c) *Oral examinations using videconferencing tools like Zoom* (cheating by examinees can include all of the previous but with much greater care and control of the angle of the camera view, appropriateness of the movement of hands and direction of gaze, as well as promptness in oral responses to the questions that are being asked).

Having in mind the time given to the students to develop techniques for cheating in online exams (two years of lockdowns at HEIs during the COVID-19 pandemic), alongside with greater *circumstantial incentives* (less time for preparation; test that is difficult to pass; importance of passing the test; inadequate control of the examinees' testing environment; poor instructor skills for use of prevention technology; inadequate organizational culture at a higher education institution; peers that are cheating and pressure on others to also cheat), the sophistication and skill in use of *online cheating techniques* by interested students could gradually evolve to a substantial level, far beyond the ability of an average instructor to adequately confront them.

IV. PREVENTION OF CHEATING IN ONLINE EXAMS

The authors of this paper have performed numerous online assessments in form of *online tests with multiple choice questions*, as well as oral exams using videoconferencing tools like *BigBlueButton*. In case of *online tests* our most commonly used and, according to our experience with more than 20 online testing procedures performed, preferred combination of simultaneous application of various prevention methods was: (a) creation of a new set of questions for every online testing; (b) all students enlisted in a specific course in one generation are accessing the online exam for this course at the same time; (c) at least 35-40 questions in each test; (d) time restriction of test duration to 25-30 minutes for 35-40 multiple choice questions; (e) random order of multiple choice questions in the test for each examinee; (f) presentation in random order of at least 4-6 predefined responses for each multiple choice question; (g) statements of students before taking the test that they will not use unethical means during testing to achieve better results; (h) use of *Safe Exam Browser* which imposes a 'full-screen lock' to the computer of the examinee during the test; (i) video surveillance of the examinees via *BigBlueButton* videoconferencing tool and their use of a webcam or smartphone camera to enable monitoring of their behavior during the whole testing period.

The text of the statement that students agreed with before taking the online tests that were performed by the authors of this paper was: "By taking this online test, I declare and confirm that I will not use illegal and unethical means such as consulting printed or electronic content outside the test,

giving or receiving any help in answering the questions in the test, as well as any other forms of cheating in order to achieve a better result than the one I can achieve only based on my knowledge at the time of gaining access to this test." Video surveillance was organized in groups of up to 25 students per one *BigBlueButton* virtual classroom (it was integrated within the Moodle LMS, where the online exams were placed). The optimal solution for simultaneous online testing of large groups of students was for one instructor to perform video monitoring for a maximum of 2-3 groups with 20-25 students in each group.

The *Safe Exam Browser* (SEB) is an effective solution for preventing cheating on multiple-choice tests since it displays only the necessary navigation in the learning management system and prevents access to internet search engines and other local documents [19] by restricting the number of active applications during the exam period [20]. Because of potential technical problems, ensuring the opportunity for students to take the same online test without SEB is recommended (to discourage false statements about their inability to use SEB, those students who encounter technical problems could be obliged to take a subsequent oral exam). Documentation and instructions for the students on how to download, install and use SEB should also be provided, and a restriction should be made to use only a desktop/notebook computer to access the online test.

In order to mitigate potential negative impacts of online assessment, soon after the beginning of the COVID-19 pandemic the *University Computing Centre* of the University of Zagreb, Croatia, issued recommendations [21] for university instructors on how to hold and deliver distance learning exams. The first recommended step was to prepare the exam carefully and provide students with clear instructions on how to take the exam, including technical requirements, but also to state that any kind of cheating is not allowed. A prior *practice test* that simulates the conditions of the real exam was also advised. Another important guideline, intended to reduce the distribution of questions and answers among students, was to omit feedback for each question after submitting the test so that the correct answers are not revealed. It was also recommended to limit the number of questions per page and disable the return to previously viewed questions. Finally, a large base of well-formulated questions would make it more difficult to share test questions and correct answers.

In case of online exams performed by videoconferencing tools the only suitable (acceptable and inobtrusive) means of preventing cheating that was used by the authors of this paper was to careful monitor students' behavior and proceed to the next question in case the student hesitated with his/her answer to a question for more than a few seconds.

When using video conferencing tools for surveillance a reliable internet connection on the students' side is required, as well as the teachers' ability to monitor all students that are taking the online test (see: Haus et al. [22]). In case large groups of students are taking the online test, it is advised to provide a sufficient number of instructors to monitor online testing and also to complete checking student's cameras at least 5-10 minutes before they are scheduled to start an online test. Haus et al. emphasize the following potential advantages of using smartphone cameras as a surveillance method instead of webcams: (a) easier control of students accessing physical materials (books, devices) or speaking with others in the room during the exam; (b) even when the locked browser apps are used during the exam, the teacher can see students' behavior in real time by zooming in on each student window in the web conference; (c) the teacher can check students randomly or focus on a student who shows suspicious behavior during the exam.

Before using technology for conducting online exams the teachers should develop appropriate competences or be sufficiently instructed on how to perform testing with technology. Also, the students should be provided with adequate guidance and pre-trained, for instance with exercise/trial online testing, and should also be informed in advance about possible technical problems, especially if special software like SEB is used during online testing [23].

More advanced technical methods to prevent students' cheating in online exams include the use of remote *proctoring software*. Automatic online proctoring replaces instructors in monitoring the students while they are taking an online exam and can deter academic misconduct. One of the methods for automatic proctoring implies using software that accesses a student's webcam and monitors the video to identify behavior patterns that could indicate cheating. Besides having proctoring software to automatically control students' behavior during testing, the instructor can also choose to view the video in real time during the exam or review the video recording after the exam. A study conducted before the COVID-19 pandemic by Dendir and Maxwell [24] indicated that online proctoring with webcam recording software can reduce cheating in online courses. The same finding was reported in a study by Arnold [25] in relation to proctored online exams before and during the COVID-19 pandemic (it was conducted each year from 2015 to 2020). This latter study indicated that in each year of the study cheating in *proctored online exams* was not greater than in *proctored face to face on-site settings*. There are various countermeasures in virtual proctored examinations, including the use of 360-degree security cameras [26]. Despite important privacy concerns, the use of a 360-degree camera provides the proctor with similar ability as in the traditional on-site proctoring situation where he/she can oversee the testing place and occasionally walk to the students who are taking the exam [27].

V. PLAGIARISM ISSUES IN RELATION TO CONVERSATIONAL AI

Plagiarism is a form of academic dishonesty that is associated with taking someone else's work and presenting it as one's own. In educational settings it is most frequent in student work in form of homework assignments and essays. The predominant types of plagiarism are related to (a) students' copying or paraphrasing a source without citation, as well as with (b) having someone else do instead of them what is demanded in an assignment and presenting it as their own work. Plagiarism is an important issue in research on higher education from a theoretical [28] and practical perspective, for instance regarding plagiarism detection [29]. Dixon et al. [30] have emphasized that, when text-matching or "similarity detection" software is used for checking student coursework, the occurrence of problematically similar text without proper attribution appears in 25% to above 70% of samples, depending on published study. They also revealed a high percentage of students that were willing to exchange coursework with their peers, which is also considered a form of plagiarism.

Concerns about plagiarism have recently grown with the advance of conversational AI like chatbots or virtual agents. According to data collected in 2019 and presented on the *Statista* portal [31], very high percentages of users' questions

were correctly understood (above 99.8%) and answered (80%-93%) by popular virtual assistants like *Alexa*, *Siri* and *Google Assistant*. Introduction of *ChatGPT* for free public use in November 2022 was a turning point for the academia with numerous observations by educators that it may endanger the academic integrity of students. Since the introduction of *ChatGPT* to the general public, popular news portals like *Insider* [32] have been reporting on the academic exams that *ChatGPT* was able to pass and specialized portals like *TimesHigherEducation* have published commentaries on topics related to *ChatGPT* (for search results see: https://www.timeshighereducation.com/search?search=chatgpt). However, papers in *peer reviewed scholarly literature* that address this issue and discuss, for instance, the opportunities and challenges associated with the use of *ChatGPT*, as well as means for prevention of plagiarism [33] can only recently be found. Some authors have even used *ChatGPT* as a "co-author" to produce articles for scholarly journals about its use in plagiarism (e.g. [34], [35]). The use of *ChatGPT* in education is elaborated in pre-print papers without peer review available on public archive websites like *arXiv* (https://arxiv.org/) or EdArXiv (https://edarxiv.org/discover?q=chatgpt). In peer-reviewed journals there is lack of comprehensive analyses of issues associated with academic integrity and use of Large Language Models (LLMs) tools like *ChatGPT* (see: [36]).

Since the academic community is still adapting to the emergence of conversational AI and tools using LLMs it is too early to derive and present accountable statements about its future impact on higher education. Still, educators should pay much more attention to home assignments and exams/tests performed remotely/online if they are not adequately proctored. The authors of this paper recommend testing not only *ChatGPT*, but also other conversational AI tools like *Bing* chat and *ChatSonic* (https://writesonic.com/chat), as well as similar tools that will soon be available for public use (e.g. Google *Bard*), in relation to how the instructors at HEIs design and deliver online exams and homework assignments. For instance, the authors of this paper have found that *ChatSonic* can produce correct answers to multiple-choice questions in Croatian language that were used in their online exams during the COVID-19 pandemic.

VI. Conclusion

A brief review in the first ("Introduction") part of this paper indicated the fairly high percentage of students who were inclined to cheat in online exams *during* the COVID-19 pandemic, as well as of instructors who were concerned about such academic dishonesty of their students (see: [5], [6], [7], [8]). On the other hand, studies conducted *before* the COVID-19 pandemic also confirm that a greater percentage of students were likely to cheat in *online* assessment in comparison to traditional *on-site* assessment (for instance, see: [10]). In the second chapter of our paper entitled "Motives of students for cheating in online assessment" diverse reasons for academic dishonesty were extracted from the papers published *during* the COVID-19 pandemic (see: [9], [11], [12], [13]), as well as *before* the pandemic (e.g.: [14], [16], [17], [18]). Since our review of the analyzed literature established that the inclination to cheating in online exams was high and the students' motives to do so diverse and numerous, in the third section of our paper ("Variety of online cheating in online courses") we addressed this issue by creating categorizations of means for such academic dishonesty. In the first categorization, the findings of an *empirical study* related to predominantly *offline* cheating [6] were combined with a *review study* associated with *online* cheating [11] by the students in academic assessment. The second categorization in this section of our paper also included the experiences of the authors and addressed means for cheating in (a) *unsupervised* and (b) *supervised online tests*, as well as (c) *oral examinations* with the use of videoconferencing tools like *BigBlueButtom* or *Zoom*. The fourth section ("Prevention of cheating in online exams") presented a practical solution used by the authors of this paper for confronting cheating in online exams by combining several techniques, including students' agreement with an honor statement, video surveillance and use of *Safe Exam Browser*. The use of proctoring applications is also briefly addressed in this section, having in mind both their positive and negative aspects. Finally, in the fifth section or our paper "Plagiarism issues in relation to conversational AI" such tools and services are briefly considered in relation to cheating in assignments which student perform outside monitored *on-site* environments.

The study that is presented here presents a more contemporary analysis of research and addresses in more detail the period of the COVID-19 pandemic when compared to similar reviews of research papers published before the COVID-19 pandemic (i.e.: [37]). However, remote proctoring with the use of video surveillance and the challenges that were recently brought before the academic community with the increasing popularity of the *ChatGPT* conversational AI tool were not extensively analyzed in this paper. The authors of this paper agree that there are numerus controversies associated with remote proctoring in line with the statement of the *ACM US Technology Policy Committee* dated December 16, 2022 [38]. Also, because of the lack of peer reviewed research, it is still too early to make *reliable and accountable statements* about the use of conversational AI or LLM based tools for cheating in academic assessment.

The COVID-19 pandemic has brought academic integrity/dishonesty in online learning into the focus of researchers, policy makers and practitioners (e.g. [39], [40]). The appearance of conversational AI services like *ChatGPT* and *ChatSonic*, as well as their expected improvements and novel services being announced, make it an imperative to confront possible threats to academic integrity of students and critically review the methods of academic assessment that are performed outside proctored environments.

References

[1] Camara, W. J., and Lane, S., "A historical perspective and current views on the Standards for Educational and Psychological Testing", *Educational Measurement: Issues and Practice*, 2007, vol. 25, no. 3, pp. 35–41, Available: https://www.researchgate.net/publication/227836349_A_Historical_Perspective_and_Current_Views_on_the_Standards_for_Educational_and_Psychological_Testing

[2] *Standards for Educational and Psychological Testing*, American Educational Research Association (American Psychological Association / National Council on Measurement in Education), Washington, DC, USA, 2014, Available: https://www.testingstandards.net/uploads/7/6/6/4/76643089/9780935302356.pdf

[3] *Guidelines for technology-based assessment*, International Test Commission, Hemel Hempstead, England, UK / Association of Test Publishers, Washington, DC, USA, 2022, Available: https://www.intestcom.org/upload/media-library/guidelines-for-technology-based-assessment-v20221108-16684036687NAG8.pdf

[4] Camara, W. J., "Issues facing testing organizations in using the Standards for Educational and Psychological Testing", *Educational Measurement: Issues and Practice*, vol. 33, no. 4, pp. 13–15, Available: https://www.researchgate.net/publication/269418376_Issues_Facing_Testing_Organizations_in_Using_the_Standards_for_Educational_and_Psychological_Testing

[5] Jenkins, B. D., Golding, J. M., Le Grand, A. M., Levi, M. M., and Pals, A. M., "When opportunity knocks: College students'

cheating amid the COVID-19 pandemic" [Online first], *Teaching of Psychology*, 2022, https://doi.org/10.1177/00986283211059067

[6] Janke, S., Rudert, S.C., Petersen, Ä, Fritz, T.M., and Daumiller, M., "Cheating in the wake of COVID-19: How dangerous is ad-hoc online testing for academic integrity?", *Computers and Education Open*, 2021, vol. 2, Art. no. 100055, https://doi.org/10.1016/j.caeo.2021.100055

[7] Maryon, T., Dubre, V., Elliott, K., Escareno, J., Fagan, M.H., Standridge, E., and Lieneck, C., "COVID-19 academic integrity violations and trends: A rapid review", *Educational Sciences*, 2022, vol. 12, no. 2, Art. no. 901, https://doi.org/10.3390/educsci12120901

[8] Wiley, *Academic Integrity in the Age of Online Learning: Faculty Survey Results* (Whitepaper), July 2020, Available: https://www.wiley.com/en-us/network/education/instructors/teaching-strategies/academic-integrity-in-the-age-of-online-learning-3

[9] Wiley, *New Insights into Academic Integrity: 2022 Update* (Whitepaper), 2022, Available: https://www.wiley.com/en-us/network/education/instructors/teaching-strategies/new-insights-into-academic-integrity-fall-2021-report

[10] Harton, H.C., Aladia, S., and Gordon, A., "Faculty and student perceptions of cheating in online vs. traditional classes", *Online Journal of Distance Learning Administration*, 2019, vol. 22, no. 4, https://ojdla.com/archive/winter224/hartonaladiagordon224.pdf

[11] Noorbehbahani, F., Mohammadi, A., and Aminazadeh, M., "A systematic review of research on cheating in online exams from 2010 to 2021", *Education and Information Technologies*, 2022, vol. 27, pp. 8413–8460, https://doi.org/10.1007/s10639-022-10927-7

[12] Jantos, A., "Motives for cheating in summative e-assessment in higher education - a quantitative analysis", in *Proceedings of the 13th International Conference on Education and New Learning Technologies (EDULEARN21)*, 2021, pp. 8766-8776, Available: https://www.researchgate.net/publication/353345751_Motives_for_Cheating_in_Summative_E-Assessment_in_Higher_Education_-_A_Quantitative_Analysis

[13] Pleasants, J., Pleasants, J. M., and Pleasants B. P., "Cheating on unproctored online exams: Prevalence, mitigation measures, and effects on exam performance", *Online Learning*, 2021, vol. 26, no. 1, pp. 268-284, http://dx.doi.org/10.24059/olj.v26i1.2620

[14] Whitley, B.E., "Factors associated with cheating among college students: A review", *Research in Higher Education*, 1998, vol. 39, no. 3, 2pp. 35-274, https://www.jstor.org/stable/40196379

[15] Marques, T., Reis, N., and Gomes, J., "A bibliometric study on academic dishonesty research", *Journal of Academic Ethics*, 2019, vol. 17, pp. 169–191, https://doi.org/10.1007/s10805-019-09328-2

[16] McCabe, D.L., and Trevino, L.K., "Academic dishonesty: Honor codes and other contextual influences", *The Journal of Higher Education*, 1993, vol. 64, no. 5, pp. 522–538, https://doi.org/10.2307/2959991

[17] Arnett Jensen, L., Jensen Arnett, J., Feldman, S.S., and Cauffman, E., "It's wrong, but everybody does it: Academic Dishonesty among High School and College Students", *Contemporary Educational Psychology*, 2002, vol. 27, no. 2, pp. 209-228, Available: https://lenearnettjensen.com/wp-content/uploads/2018/10/2002-CEP-Jensen-et-al-Academic-Dishonesty.pdf

[18] Jurdi, R., Hage, H. S., and Chow, H. P. H., "Academic dishonesty in the Canadian classroom: Behaviours of a sample of university students", *Canadian Journal of Higher Education*, 2011, vol. 41, no. 3, pp. 1-35, https://doi.org/10.47678/cjhe.v41i3.2488

[19] Gajewski, R.R., "IT in educational management: Can it support solution of e-cheating problem?". In Brinda, T., Mavengere, N., Haukijärvi, I., Lewin, C., Passey, D. (eds), *Stakeholders and Information Technology in Education*, SaITE 2016, IFIP Advances in Information and Communication Technology, 2016, 493. Springer, Cham, https://doi.org/10.1007/978-3-319-54687-2_10

[20] Lee, K., and Fanguy, M., "Online exam proctoring technologies: Educational innovation or deterioration?", *British Journal of Educational Technology*, 2022, vol. 53, pp. 475–490. https://doi.org/10.1111/bjet.13182

[21] Centar za e-učenje, *Upute za održavanje testa u sustavu Merlin (Moodle) sa studentima online* [In Croatian, "Instructions for Performing an Online Test in the Merlin System (Moodle)"]. E-learning centre, Srce, University of Zagreb, Croatia, Available: https://www.srce.unizg.hr/files/srce/docs/CEU/srce_upute_za_odrzavanje_testova.pdf

[22] Haus, G., Pasquinelli, Y. B., Scaccia, D., and Scarabottolo, N., "Online written exams during COVID-19 crisis", In *Proceedings of the International Conference E-Learning 2020* [Online], July 21-23, 2020, pp. 79-86, Available: https://files.eric.ed.gov/fulltext/ED621620.pdf

[23] von Grünigen, D., de Azevedo e Souza, F. B., Pradarelli, B., Magid, A., and Cieliebak, M., "Best practices in e-assessments with a special focus on cheating prevention", in *Proceedings of the 2018 IEEE Global Engineering Education Conference (EDUCON)*, Santa Cruz de Tenerife, Spain, 2018, pp. 893-899, https://doi.org/10.1109/EDUCON.2018.8363325

[24] Dendir, S., and Maxwell, R.S., "Cheating in online courses: Evidence from online proctoring", *Computers in Human Behavior Reports*, 2020, vol. 2, Art. no. 100033, https://doi.org/10.1016/j.chbr.2020.100033

[25] Arnold, I.J.M., "Online proctored assessment during COVID-19: Has cheating increased?", *The Journal of Economic Education*, vol. 53, no. 4, pp. 277-295, https://doi.org/10.1080/00220485.2022.2111384

[26] Alin, P., Arendt, A., and Gurell, S. (2022): "Addressing cheating in virtual proctored examinations: Toward a framework of relevant mitigation strategies", *Assessment & Evaluation in Higher Education* [Online], https://doi.org/10.1080/02602938.2022.2075317

[27] Turani, A.A., Alkhateeb, J.H., and Alsewari, A.A., Students online exam proctoring: A case study using 360 degree security cameras", *Proceedings of the 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, Dhaka, Bangladesh, 2020, pp. 1-5, https://doi.org/10.1109/ETCCE51779.2020.9350872

[28] Klein, D., "Why learners choose plagiarism: A review of literature", *Interdisciplinary Journal of E-Learning and Learning Objects*, vol. 7, no. 1, pp. 97-110. Available: https://www.learntechlib.org/p/44732/

[29] Foltýnek, T., Meuschke, N., and Gipp, B., "Academic plagiarism detection: A systematic literature review". *ACM Computing Surveys*, 2019, vol. 52, no. 6, Art. no. 112, https://doi.org/10.1145/3345317

[30] Dixon, Z., Whealan George, K., and Carr, T., "Catching lightening in a bottle: Surveying plagiarism futures", *Online Learning*, 2021, vol. 25, no. 3, pp. 249-266, http://dx.doi.org/10.24059/olj.v25i3.2422

[31] Larrichia, F., "Share of questions answered and understood correctly by selected digital assistants as of 2019", Statista, Hamburg, Germany, 2022, Available: https://www.statista.com/statistics/1133702/digital-assistant-intelligence-comparison/

[32] Varanasi, L., "ChatGPT could be a Stanford medical student, a lawyer, or a financial analyst: Here's a list of advanced exams the AI bot has passed so far", *Insider*, 11th February, 2023, Available: https://www.businessinsider.com/list-here-are-the-exams-chatgpt-has-passed-so-far-2023-1

[33] Cotton, D.R.E, Cotton, P.A., and Shipway, J.R., "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT", *Innovations in Education and Teaching International* [Online], https://doi.org/10.1080/14703297.2023.2190148

[34] King, M.R., and ChatGPT, "A conversation on artificial intelligence, chatbots, and plagiarism in higher education", *Cellular and Molecular Bioengineering*, 2023, vol. 16, no. 1–2, https://doi.org/10.1007/s12195-022-00754-8

[35] Frye, B. L., "Should Using an AI Text Generator to Produce Academic Writing Be Plagiarism?", *Fordham Intellectual Property, Media and Entertainment Law Journal* [Online], December 3, 2022, Available: https://ssrn.com/abstract=4292283

[36] Perkins, M., "Academic integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond", *Journal of University Teaching and Learning Practice*, 2023, vol. 20, no. 2, https://doi.org/10.53761/1.20.02.07

[37] Holden, O.L., Norris, M.E., Kuhlmeier, V.A., "Academic integrity in online assessment: A research review", *Fronteers in Education*, vol. 6, Art. no. 639814, https://doi.org/10.3389/feduc.2021.639814

[38] USTPC, "Principles for the Responsible Development of Remote Proctoring Software Issued by ACM US Technology Policy Committee: Equity, Privacy, Security, Accessibility, and Efficacy at Issue", Association for Computing Machinery's US Technology Policy Committee (USTPC), 2023, Washington, DC, USA, Available: https://www.acm.org/binaries/content/assets/public-policy/remote-proctoring-statement.pdf

[39] Gamage, K.A.A., Silva, E.K., Gunawardhana, N., "Online delivery and assessment during COVID-19: Safeguarding academic integrity", *Education Sciences*, 2020, vol. 10, no. 11, Art. no. 301, https://doi.org/10.3390/educsci10110301

[40] Surahman, E., and Wang, T.-H., "Academic dishonesty and trustworthy assessment in onlinelearning: A systematic literature review", *Journal of Computer-Assisted Learning*, 2022, vol. 38, no. 6, pp. 1535–1553, https://doi.org/10.1111/jcal.12

The Development of Smart E-Portfolio Generator System for University students: An SDLC Approach

Ahmed Shuhaiber, Aysha Alkuwaiti, Mera Alremeithi, Hamda Almenhali
College of Technological Innovation (CTI), Zayed University, Abu Dhabi, UAE
Ahmed.Shuhaiber@zu.ac.ae, ayshaalkuwaiti5@gmail.com, meralremeithi@gmail.com, hamdaalmenhali@hotmail.com

***Abstract*— E-portfolio is a creative tool for documenting, analysing and monitoring students' learning achievements and milestones. An e-portfolio is a computerized collection of documents and certifications that reflect a university student, such as demos, resources, and accomplishments. Since earning a university degree is not enough anymore and other soft and technical skills and certification are required, it became hard for students and graduates to collect and display their achievements in one environment, especially ones with imposter syndrome. Therefore, an e-portfolio generator system has been proposed to collect, store and display user's documents and certifications in a dashboard. The methodology followed to develop this system is System Development Life Cycle (SDLC). The E-portfolio generator system is a user-friendly system. It enables the users to upload user documents, live-chat, dashboard creation and receiving alerts. Furthermore, the system will optimize the process of gathering, displaying, and storing user's data in one dashboard. The research is concluded with conclusions and suggestions for further improvement.**

Keywords*— *E-Portfolio, Live-chat, Generator, Users, Dashboard

I. INTRODUCTION

For decades, having a college degree was the "golden ticket" and the best indicator of market readiness. However, that is not the case anymore. The college degree is not the only available path to living and achieving the dream after graduating. Nowadays, the market demands more than just a bachelor's degree. What's more than that, the market wants to see evidence and documentations of what a student carries other than their degree. There are many important criteria employers use for recruitments and the college degree is at the bottom of the list. Apart from this, students may graduate and face a lot of challenges. Students frequently encounter imposter syndrome as they move from the undergraduate level to the graduate and professional level. People with imposter syndrome have a hard time accepting their accomplishments. Instead of appreciating their intelligence and success, they think their accomplishments were solely the product of luck. Anyone can experience imposter syndrome, but college students are more vulnerable. Furthermore, students sometimes believe they do not have valuable learnings and documentations of what they have learnt under their belt when they graduate. However, students achieve more than what they realize by studying throughout their bachelor's journey, yet, they are not aware of it. Reason for that is that students lack the ability to efficiently store and document their achievements properly. This problem needs a solution for a student to thrive in their postgraduate journey.

In response to the previously mentioned hurdles, we propose to create a new method for students and graduates to stand out after graduation and in their job searching process. We suggest creating an e-portfolio builder system, where students can store all their various documents and certificates. This way, students wouldn't have to lose any evidence of their achievements, highlight their best work, track their progress in education, and build a portfolio where they can share it with anyone interested.



Figure 1 – Evidence Documents

In this paper we intend to investigate the subject of gathering the student's achievements and work. The optimal goals of this paper are a) to develop a system where students store their work and achievements b) to allow students to monitor and track their progress through a dashboard and generate a shareable e-portfolio.

II. LITERATURE REVIEW

In order to comprehend the breadth and the technological implementation mechanism of such a system, similar previous research was studied. Each article is summarized by focusing on describing the developed study, purpose of the research, and the findings.

A. *The use of e-portfolios in higher education: From the students' perspective*

Villanueva-Complutense University of Madrid developed this study to ascertain the students' perspectives on the value/utility and significance of the e-portfolio, as well as their personal performance and effort [1]. They were enrolled in three asynchronous undergraduate sections of an elective course on information and communication technologies in education. The approach used by the authors was realistic and quantitative. The findings indicate that students are open to using e-portfolios in the future and do not regard them solely as coursework to be completed for a mark. Respondents recognized the importance of e-portfolios as a tool for analyzing and learning. Students recognize the e-portfolio as a resource for professional progress as well as a tool for job searching. They indicated both competence and job satisfaction. According to findings, the largest push of internal motivation was an appreciation and acknowledgment for the value and use of the e-portfolio, followed by the students' perception of competence [1].

B. *The role of e-portfolios in supporting productive learning*

This research investigates how e-portfolios play these formative roles in order to enhance effective learning [2]. The authors present qualitative findings from interviews with selected first-year undergraduate students at a Hong Kong higher education institution about their experiences creating e-portfolios as assessment tasks. E-portfolios were integrated into three basic courses for first-year students as part of an institutional teaching and learning strategy. The findings reveal that several conditions necessary to foster productive learning were missing in students' experiences: enhanced formative role of e-portfolios through coherent assessment design; positive motivation for students' pursuit of active learning to develop learning interests; engagement of students in reflective and self-regulatory learning as an essential learning process; provision of constructive feedback for sustained learning support; and assistance for students. This study provides implications for utilizing information and communication technology (ICT) to support students' productive learning by explaining how the lack of these conditions hampered students' active participation in e-portfolio tasks and recommending relevant strategies for instructors at the institution in question [2].

C. *VisCa: A Dashboard System to Visualize Learning Activities from E-learning Platforms*

Numerous educational platforms have incorporated data-driven learning and learning analytics into their systems as an emerging data science trend, resulting in more adaptable and individualized learning offerings. Consequently, it is essential to implement a strategy in order to visualize and control the learning experience data. To accomplish this aim, in the proposed study [3] VisCa was developed which is a web-based dashboard system to keep track, save and display learning experiences obtained from e-learning platforms. The Experience API (xAPI) standard is the foundation on which the data model is based when interacting with external platforms. The entire system provides a broad foundation for the learning experience data flow and aids both students and teachers in understanding their current level of learning. The creation of this project will offer a framework for gathering learning activity data that can later be utilized for data-driven learning or learning analytics [3].

D. *E-Portfolio Web-based for Students' Internship Program Activities*

During an internship program, it is essential to keep track of the students' tasks in order to be monitored, assessed, and observed. Students will find it easier to report their work's outcomes to a university or industry supervisors through an e-Portfolio platform to document the whole work procedure. The study [4] developed an e-Portfolio that is easily accessible and suitable for the needs of students and supervisors in recording their monitoring and work process. In this research, the methodology employed is fundamental research. The proposed work focused on implementing the features of internship ePortfolio through displaying them to students who have participated in an internship program [4].

E. *Utilization of the Student E-Portfolio in the Graduate Education: The Case Report of "Nitobe School Program" in Hokkaido University*

A new trans-graduate-school education program called "Nitobe School" was launched in 2015 and was a representative component of the "Top Global University Project" at Hokkaido University, Japan. Graduate education requires evidence-based evaluation and observable proof, such as student portfolios, to make the student's evaluation more objective. Therefore, Nitobe School has initiated the "NITOBE Portfolio". This paper's [5] work mainly focuses on the creation of an e-portfolio system for a unique graduate program that allows students to monitor and track their progress throughout the program and allows instructors to evaluate them based on evidence [5].

III. METHODOLOGY

A. *Planning Phase:*

1) *Technical Feasibility*

For the proposed E-Portfolio Generator System, the main software and hardware to be used are mentioned below:

Hardware:

- Mobile Application/Personal Computer

Software:

- MongoDB
- HTML
- CSS
- JavaScript
- Python

2) Financial Feasibility

The benefits obtained from student's productivity, generating a well-structured e-portfolio for user's satisfaction, and avoiding the possibility of misplacing achievements are the main essence of the tangible benefits. Besides, intangible benefits include a convenient user experience, increased knowledge of achievements and self-satisfied moments, and enhanced job application experiences for users.

Tangible Benfits	Intangible Benefits
The system will reduce the student's time and facilitate the searching process for their important documents.	The system is user-friendly and convenient to use which will improve the user's needs satisfaction.
The system will improve the searching process for the user since unique features will be provided to form a well structured e-profile.	The system will increase the user's knowledge of their achievements since direct access to all of their important documents is provided by the system.
The system will avoid the possibility of losing essential certificate evidence and documents.	The system will enhance the user's experience when applying for a job.

Figure 2 – Tangible and Intangible Benefits

B. Analysis Phase:

1) Requirements gathering

a) Business Requirements

To guarantee the success of the proposed project, the E-Portfolio Generator System should be able to meet the business requirements. The system should:

- Reduce the difficulties faced by users when creating their e-portfolio.
- Help users to have an effective process when searching for a job.
- Facilitate the process of assessing the users' learning progress and accomplishments.
- Allow users to obtain a personalized portfolio of their documentations.

b) User Requirements

- Upload documents and achievements to their e-portfolio dashboard.
- Chat with friends and share their achievements.
- Get alerts when a long interval has passed without new achievement updates.
- View the progress through a well-organized dashboard.
- Edit their e-portfolio.

c) System Requirements

We have listed various requirements and criteria we must satisfy in order to accomplish our goals and guarantee the effectiveness of our proposed system. The system requirements are divided into two categories, functional requirements, and non-functional requirements. These requirements will serve as our source of guidance as we work to establish our system.

Functional System Requirements

Functional requirements cover the aspects of the system that enable end-users to perform their tasks and carry out operations in the system. Input, output, and process are the three areas covered by the functional requirements.

Input

1. The user must sign up in the system with a username and password.
2. The user must provide different information when signing up, such as email address, date of birth, and phone number.
3. The user must access the system with their username and password.
4. The user can upload personal documents and evidence of learning, such as writings, projects, reports samples created throughout their educational journey.
5. The user can upload their documents based on categories, including degrees, certifications, experience letters, and volunteerings.
6. The user can write and put together an "About Me" section.
7. The user can edit and customize their dashboard based on their preferences.
8. The user can choose from the template of dashboards, or customize their own dashboard.
9. The user can organize their dashboard in chronological order or consecutive order.
10. The system gives the user a search bar option to search by any keyword.

Process

1. The system must be able to count the amount of documents uploaded.
2. The system must be able to generate an e-portfolio based on the user's selection of documents.
3. The system must be able to figure out if the document is legitimate, out-of-date, or forged.
4. The system must be able to look for indicators in the documents to make an attempt to ascertain the document's validity.
5. The system must be able to arrange the documents under each category.
6. The system must be able to save edits and changes in documents.

7. The system must be able to save the theme of the dashboard as the user requested and use it every time the user accesses the system.
8. The system must be able to track the user's progress.

Output

1. The system will display a dashboard with the records of documents.
2. The system will greet the user with their name once they access the system.
3. The system will alert the user if the document they are trying to upload is already uploaded previously.
4. The system will provide the user the requested document from the search bar.
5. The system will display the documents based on the order selected by the user.
6. The system will display recent growing trends of certifications and qualifications to give the user suggestions of what to achieve next.
7. The system will present recently accessed documents.

Non-Functional System Requirements

Performance

1. The system must be accessible every day of the week, round-the-clock.
2. The system's response time must be less than 3 seconds.
3. The system must not exceed 3 seconds when loading its pages.
4. The system must support 10,000 users simultaneously.
5. The system's usability must be clear. Users of the system should find it straightforward, simple, and easy to use.
6. The system must be able to withstand portability, which implies that it can be used in various environments. For example, the system can operate efficiently on a Linux machine, iOS, or Windows, in all their versions.
7. The system must be reliable, it should consistently carry out its operations correctly.
8. There are no geographical limitations; anyone can access the system from anywhere.
9. The system must automatically adapt to the screen and viewports in which it is being accessed from. It should be responsive and use flexible layouts.
10. If any failure happens within the system, replicas of the system must be up and running immediately to maintain 24/7 availability.
11. The system provides e-portfolios of high quality, and in clearest and highest resolution.

Control

1. The user must enter a strong password with at least 12 characters.
2. The user's username and password must be confirmed by the system.
3. The system should prohibit users from accessing other user's accounts.
4. The system must use Two-Factor Authentication once the user accesses their account.
5. When resetting passwords, one-time pin code will be sent either through the email address or phone number to verify the user.
6. The system must be able to preclude forged documents and warn the user.
7. All the documents and data in the system will be encrypted using the AES encryption algorithm.
8. The system's databases are regularly backed up for data loss prevention purposes.
9. The system's session timeout is 45 minutes.
10. System's logs are constantly monitored by our SOC team to circumvent incidents.
11. A secure file-sharing service will be used to securely transfer and upload documents

2) Object-Oriented Modeling

The below diagram demonstrates a use case diagram for the proposed system, E-Portfolio Generator. Primarily, the client registers in the system to start utilizing it. Next, the system will update the client's information in the system to manage the subscription. The client can pay for the subscription afterwards and the system will process the payment. Moreover, the accounting department will be notified when the client makes a payment, and the bank will verify it. In addition, the user can request to create personal portfolios and the graphic design department will be responsible to create it for the client. Moreover, the system may send alerts to the client for various reasons. For example, if the user has not updated their profile and exceeds the time limit, the system will send alerts to the client which will be initiated by the admin. Lastly, the admin manages the user accounts and maintains them.

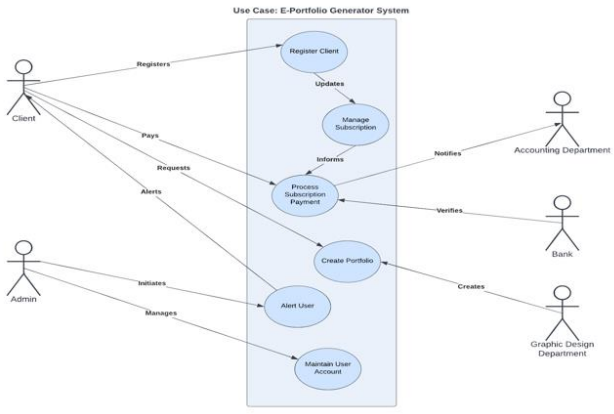


Figure 3 – Use Case

Figure 4 below illustrates the object relationship diagram for the E-Portfolio Generator System. There are nine objects – Person, Client, Admin, Portofolio, Payment, Accounting Department, Transaction, Bank, and Graphic Design Department. Client and Admin objects have an 'is a' relationship to the Person object since both are a person. The client will be able to request for a portfolio and the admin will manage it. The Graphic Design Department will create the portfolio. Therefore, the client will make the payment which will be verified by the Bank. In addition, the Accounting Department will receive the payment and manage the transactions. Lastly, the client will be notified by the Accounting Department about the payment verification and the portfolio.

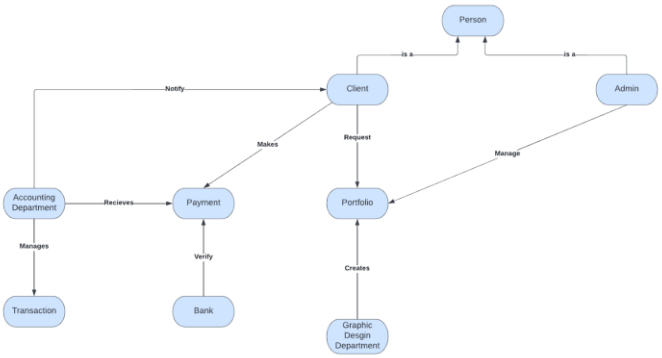


Figure 4 – Object Relationship Diagram

Figure 5 below demonstrates the class diagram of the E-Portfolio Generator System. The main classes of the diagram are – Person, Client, Admin, Portfolio, Payment, Accounting Department, Transaction, Bank, and Graphic Design Department. There is an inheritance relationship between the Client and Admin classes to the Person class since both are a person. The client can request many portfolios or none and the portfolio is requested by one customer. The admin can manage many or no portfolios and each portfolio is managed by one admin. The Graphic Design Department will generate one or many portfolios and each portfolio is created by one Graphic Design Department. There is a composition relationship between the Graphic Design Department and Portfolio classes since the portfolio won't exist without the Graphic Design Department. In addition, the client will make a payment upon the request. The client can make many or no payments and each payment is made by one client. Therefore, the payment is verified by the bank. The bank verifies one or many payments and each payment is verified by one bank. Moreover, the payment is received by the Accounting Department. The department will receive zero to many payments and each payment is received by one Accounting Department. There is a unary association relationship between the two classes since the department contains the payments. Moreover, The Accounting Department can manage zero to many transactions and each transaction is managed by one Accounting Department. Also, one to many transactions will be generated by the payments and each transaction is generated by one payment. There is a composition relationship between Payment and Transaction classes since the transaction won't exist without the payment.

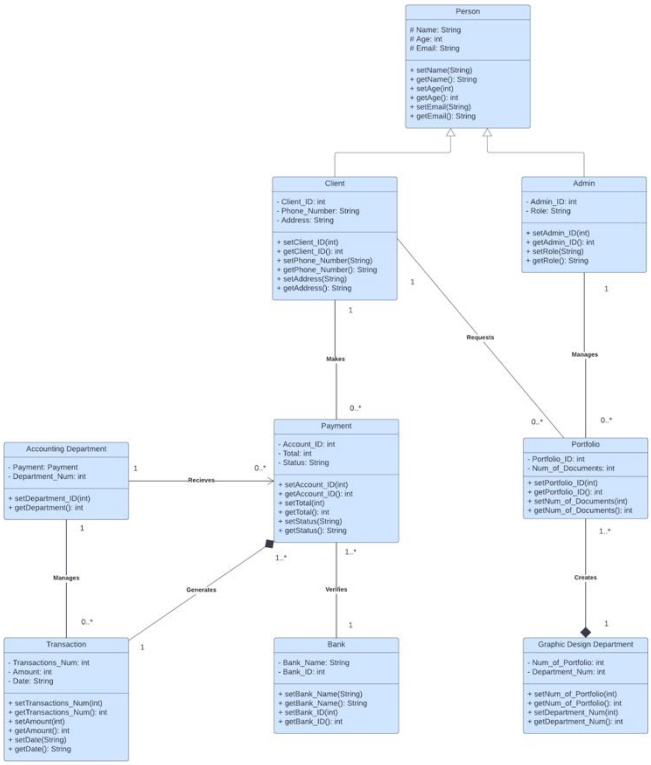


Figure 5 – Class Diagram

C. *Questionnaire Survey*

To further understand user's expectations, a survey has been conducted. As demonstrated in figure 6, 60 percent of the surveyed people responded that they store their certificates in electronic devices. Which illustrates that individuals will strongly benefit from the proposed system as the documents being stored electronically.



Figure 6 – Storage Source

Moreover, another response demonstrated that most of the respondents feel proud when viewing their achievements. As shown in figure 7, the responses indicated that the dashboard within the system will satisfy users in terms of content while viewing their achievements in the dashboard.

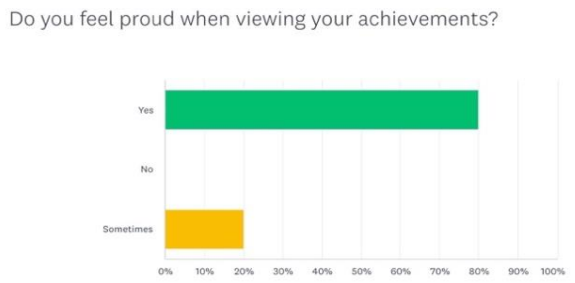


Figure 7 – User Satisfaction

Besides, to gauge the anticipated success of the system, the survey questioned the respondents the amount of satisfaction they'll have when all their achievements are gathered in a single place and displayed on a dashboard. The majority of the respondents as very satisfied and satisfied.

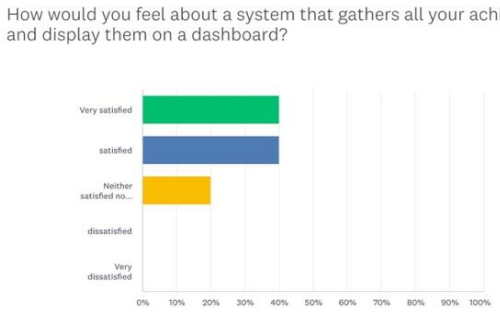


Figure 8 – User Satisfaction

D. Design Phase

1) Business Process design

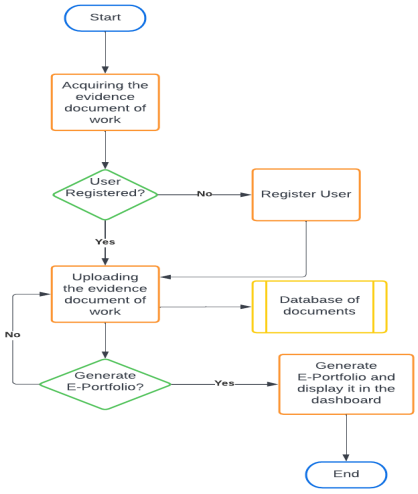


Figure 9 – Flowchart

2) User interfaces (GUI)

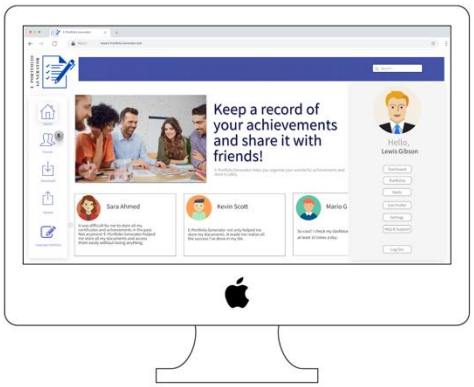


Figure 10 – Home Page

The user homepage includes a headline that briefly informs the users what the system offers. Additionally, it includes reviews of users with their profiles. The menu of the site is placed on the far right vertically. The user can navigate with the buttons of the menu.

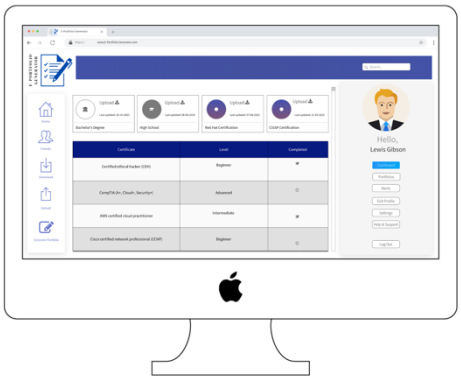


Figure 11 – Dashboard Page

The dashboard of the system includes many functions for a portfolio such as uploading degrees and certifications previously added to the system. Also, a table that a user can create to keep track of achievements and certifications completed.

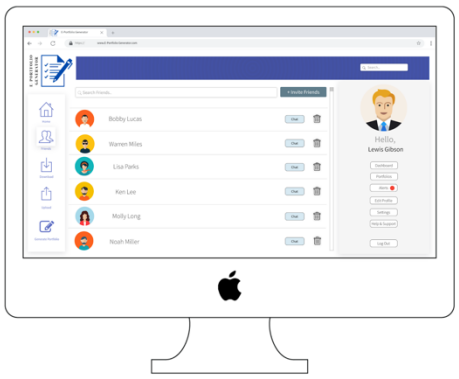


Figure 12 – Network Page

The network page or adding friends includes a search engine that searches for specific usernames within the system. There is a chat option where the user can chat with this user or delete this user. In addition, there is an option to add more friends where the user can add more friends to increase their network within the system.

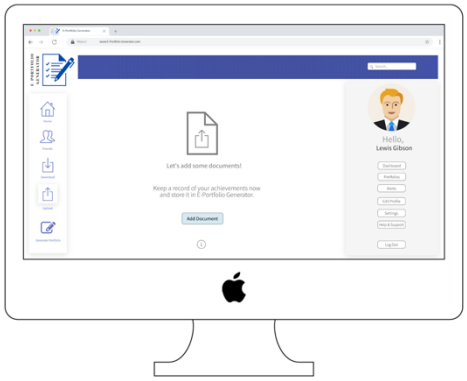


Figure 13 – Uploading Document Page

The uploading document page includes a button that will add documents from the user's device (desktop or files). This page will keep records of the user's documents and add them to the user's dashboard.

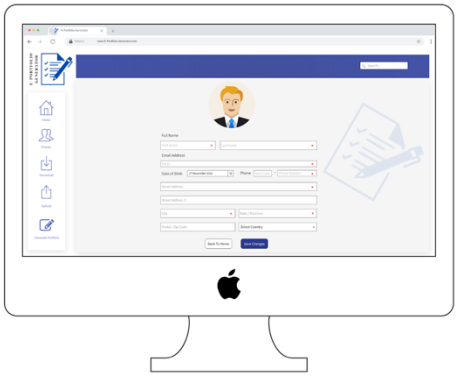


Figure 14 – Editing Profile Page

Moving to the editing profile page, it includes the full name input box (first and last), email address, date of birth, phone number, address, zip code and the country. There is a save changes button to update the user profile.

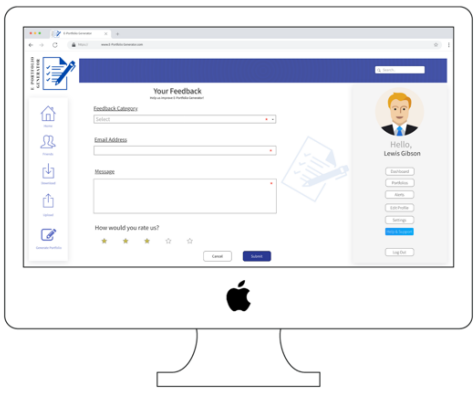


Figure 15 – Feedback Page

The feedback page has four different inputs for a user. The first input is selecting a feedback category from a dropdown list such as customer review or customer complaints. Also, the user must enter their email address to know the source of the feedback. In addition, a user must enter a message to elaborate on their feedback. All of these inputs are mandatory. Lastly, the user can optionally rate the system with a star rating. The user can then submit or cancel the feedback.

IV. DISCUSSION

Indeed, there are many existing platforms that help individuals to highlight their achievements, store them in a central location, and exhibit them to curious parties.

LinkedIn is a professional networking platform created in 2002 and launched in 2003 by Reid Hoffman that aims to connect professionals, facilitate career growth and provide insights on industry trends. It allows users to create a profile highlighting their professional experience, connect with others in their field, and access job opportunities. Additionally, it features company pages and a learning platform for users to expand their knowledge and skills. Furthermore, the E-Portfolio Generator includes certificate validation that is a specialized platform that enables universities and companies to verify student achievements and skills. Furthermore, it is a secure way for students to present their achievements to potential employers. LinkedIn, on the other hand, is a more general social media platform that focuses on professional networking, job searching, and career development. It offers a wide range of features beyond portfolio creation, such as job postings, company pages, and learning resources. Nevertheless, LinkedIn provides a substantial amount of exposure. Perhaps users carry achievements and documents that they would only like to show to interested parties, rather than everyone around the world. LinkedIn profiles can be easily accessed online. On the other hand, the E-Portfolio Generator system assures that only entities with genuinely honest and trustworthy intentions may view others' profiles and documents.

Also, the recruitment platform, 'Nexxt', was founded in 2016 and headquartered in Pennsylvania, USA in which it provides its services to various industries and sectors. Nexxt is considered a recruitment media company that offers a set of recruitment advertising tactics to aid companies in finding and hiring individuals with top talent. The company's main goal is to facilitate the hiring process for both job seekers and employers, additionally, to develop a platform that simplifies creating connections between the two. Although this platform presents a set of recruitment benefits, it only focuses on providing services for enhancing the hiring process. Restrictions on the individual's personalized experience are present. The E-Portfolio Generator System offers greater flexibility in the matter of not only increasing the rate of employment but also allowing individuals to have a more customizable experience. It is evident that Nexxt can

offer a variety of recruitment solutions, however, employers and job seekers are limited in terms of customization in the provided packages. In contrast, the proposed system allows users to customize their portfolio and display it to employers based on their preferences. Users can select from dashboard templates that are provided by the system or create their own. This allows a more personalized experience which will showcase the individuals' documents in a unique format.

Another platform is an online portfolio website builder founded in 2012 by Hamid Abouei and Gustav Degerman called 'Portfoliobox'. Portfoliobox allows users from all around the globe to showcase and share their work with other online users. This is done by creating a website and purchasing a personal domain. That being said, it allows users to develop online portfolios, upload achievements and knowledge documents as images on the website. It also allows them to present them to interested parties.

However, our platform's focal point is not for individuals embracing their achievements and certifications. Instead, it is for individuals creating their own blog of various topics. The E-Portfolio Generator System targets individuals seeking to demonstrate their credible experiences and knowledge. The proposed system accepts acceptable documents uploaded by users. Apart from this, the documents uploaded are also validated by legitimate authorities and entities such as the user's college. This allows them to present genuine and authentic documents that highlight their experience and increase their market demand. Portfolio box has this objective to keep an archive and store individual work in exactly the same way as the e-Portfolio generator system. However, others may use this platform for different purposes and create a blog for other motives.

Considering all of the mentioned elements of the proposed system, the e-portfolio generator system is a highly advanced method of producing portfolios. It is futuristic and progressive. It will target the target audience electronically efficiently. The system will allow the students to be prepared and remove the fear inside them once they graduate. Because, they will gain a direct representation of their achievements. The system will shed light on student's accomplishments, certifications, and experiences. The use of the e-portfolio generator system will be used to assess learning progress and accomplishment which often yields the best results. Portfolios shouldn't just be an inactive repository of student work that is only evaluated at the end of each semester, according to research, portfolios students should incrementally build out portfolios on an ongoing basis. Hence, the system can never be out-dated and will always be in use. It does not only benefit the students by facilitating the process of their portfolio creation, but in addition it will help them to attain different opportunities in various job fields. Furthermore, it can have a significant impact on the students' motivation since they will be more encouraged to participate in workshops, programs, and much more in order to improve their e-portfolio. Additionally, the unique features provided by the system allow the students to obtain a personalized portfolio of their documentations since they can choose the portfolio's format, design and structure. Considering all of these points, the e-portfolio generator will have a significant good influence, as well as enhancing quality and ensuring efficient services to our individuals.

I. CONCLUSION AND FUTURE WORK

In this paper, the planning process, analysis, and design of the E-Portfolio Generator System is provided. The proposed system will allow users to store important documents, facilitate the process for users to gather essential certifications, reduce the probability of losing such documents, users will gain confidence when presenting their e-portfolio, and be prepared when applying for a job. Therefore, developing such a system will guarantee to offer these benefits effectively. Moreover, the survey results demonstrate that the E-Portfolio Generator System is useful to visualize, monitor, and track individual's progress and achievements throughout their academic and career journey. However, future work is still necessary. The future work of the proposed system will include enhancements in the graphical user interface. Also, building and developing an actual system that is functional and usable. Lastly, the system can be eventually integrated with different systems related to the user's desires. For instance, if the user is a student at Zayed University, the user can integrate the E-Portfolio Generator System and their Learning Management System (LMS) into one cohesive system for better experiences and easier transfers of documents.

REFERENCES

[1] M. Ciesielkiewicz, "The use of e-portfolios in higher education: From the students' perspective," *Issues in Educational Research*, vol. 29, *(3)*, pp. 649-667, 2019. Available: https://www.proquest.com/scholarly-journals/use-e-portfolios-higher-education-students/docview/2391970066/se-2.

[2] M. Yang, M. Tai, and C. P. Lim, "The role of e-portfolios in supporting productive learning," *British Journal of Educational Technology*, vol. 47, no. 6, pp. 1276–1286, Jul. 2015, doi: 10.1111/bjet.12316.

[3] Lin, CH., Hu, SS., Lai, HY., Chiang, CF., Tseng, HC., Cheng, YC. (2017). VisCa: A Dashboard System to Visualize Learning Activities from E-learning Platforms. In: Wu, TT., Gennari, R., Huang, YM., Xie, H., Cao, Y. (eds) Emerging Technologies for Education. SETE 2016. Lecture Notes in Computer Science, vol 10108. Springer, Cham. https://doi-org.zulib.idm.oclc.org/10.1007/978-3-319-52836-6_44

[4] A. Juhana et al., "E-Portfolio Web-based for Students' Internship Program Activities," in IOP Conference Series: Materials Science and Engineering, Feb. 2018, vol. 306, no. 1. doi: 10.1088/1757-899X/306/1/012003.

[5] S. Imai, M. Shimamura, A. A. Ravankar, H. Haga, Y. Kobayashi, and Y. Yamanaka, "Utilization of the Student E-Portfolio in the Graduate Education: The Case Report of 'Nitobe School Program' in Hokkaido University," in Proceedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017, Nov. 2017, pp. 147–152. doi: 10.1109/IIAI-AAI.2017.193.

A Smart Desk: A Smart Solution for Young Students

Ahmed Shuhaiber*, Rund Hiyasat**, Amjad Abuelsamen***, Nasreen Ghanem*, Mouza Alhammadi*, Alreem Alghafri*
*College of Technological Innovation (CTI),
**College of Arts and Creative Enterprises,
***College of Interdisciplinary Studies,
Zayed University, Abu Dhabi, UAE
Emails: Ahmed.Shuhaiber@zu.ac.ae, Rund.Hiyasat@zu.ac.ae, Amjad.Abuelsamen@zu.ac.ae, 201800063@zu.ac.ae, 201505910@zu.ac.ae, 202008516@zu.ac.ae

***Abstract* - Boredom while doing homework is a common problem faced by young school students. This can affect their learning progress and achievements. On the other hand, toddlers and young children are becoming more and more experience with technology. This paper aims towards taking advantage of kids' vast experience with technology to better achieve learning outcomes; thereby, developing a smart desk that can smartly check for progress, entertain student while studying, and provide comfort to young users and their parents for daily learning activities. The methodology followed to develop this smart desk is the waterfall SDLC, where a system is developed, analyzed, designed and initially implemented. More precisely, requirements were gathered, modeled, and analyzed, the interfaces and databases were designed, coding and testing took place subsequently. The desk's features include following up with students' homework, measuring students' learning progress based on the scheduled tasks, cheering students when achieving a milestone, alerting for time breaks and more. The smart desk is also a friendly to-use system for juniors. Further, schools and academic institutions can embrace the new technology, thus benefitting from the gains of having a smart desk. The study concluded with implications and future development directions.**

Keywords - Smart Solutions, Smart Desk, Junior Assistant, Kids Tables, System Prototype

I. INTRODUCTION

The current era, also known as the digital age or information age, has brought about a rapid pace of technological advancement and change, however, the traditional methods of learning at home have remained relatively unchanged. This has resulted in a growing number of junior students feeling disengaged and uninterested in their studies. The lack of interaction, hands-on activities, and personalized experiences, combined with the overreliance on technology, has created a monotonous and repetitive learning environment [1]. This, in turn, leads to boredom, decreased motivation, and a lack of attention for many students, thereby affecting their academic progress.

The decline in children's interest in their studies has been a growing concern among parents in recent years, leading to negative impacts on their academic progress and overall learning experience. This trend has been reflected in poor grades and even warnings or dismissals from junior schools, causing alarm in the education sector [2]. The root cause of this issue is complex and multidimensional, with some students feeling bored or disinterested in their studies [3]. This disinterest leads to a lack of attention and motivation, which can have long-lasting impacts on their academic performance and future success [4]. The decline in interest in studies among young students is a worrying trend that has the potential to negatively impact the education sector, particularly for younger students who are just starting their academic journey [2]. If left unaddressed, this trend could reduce the quality of education, affecting the future of these students and the education sector as a whole [4].

The research community, parents, educators, and practitioners are putting efforts into exploring innovative solutions to make learning more enjoyable and comfortable for students. The integration of technology in the field of education has been a game-changer in recent years [5]. With the rapid advancement of technology, there has been a surge of opportunities for innovation in the way students learn and interact with information [3]. The integration of technology in the classroom has been widely recognized as an effective way to enhance the learning process for students of all ages [6]. In addition to providing a more interactive learning experience, technology has also enabled the creation of learning tools and platforms that allow students to access educational resources from anywhere and at any time [6]. One such solution is the development of smart desk systems.

The current research, proposes smart desks for young, school-age kids. It aims towards making high-quality display desks that enhance young children's learning process. It endeavors to enhance and motivate little young students to learn more efficiently by ensuring entertainment while learning. This can be achieved via the reward system proposed in the design.

For this purpose, this article is structured as follows: the following section (section 2) provides an insight to relevant papers from precedent literature. The research methodology and objectives are concisely summarized in Section 3. Next, in Section 4, the discussion and findings showcase the study's significance. While, the study concludes in section 5, limitations and future work recommendations are explained in section 6.

II. LITERATURE REVIEW

The integration of computer technology and multimedia capabilities allows students to interact with digital content in new and innovative ways. Smart Desk systems offer such capability, making the learning process more interactive, engaging, and fun. The concept of Smart Desk Systems can be traced back to the early 2000s when multi-touch technology was first introduced in consumer electronics. A Smart Desk System is a type of interactive desk that incorporates computer technology and multimedia capabilities to provide a collaborative and interactive platform for various activities such as education, meetings, and entertainment [4].

The Smart Desk System is a technological solution designed to meet the needs of the younger generation. It could potentially provide features such as interactive displays, connectivity to devices, and streamlined organization and access to information. These desks have been developed to enhance the learning experience by making it more interactive and engaging for the younger generation with its advanced technology [3, 12]. Thus, The Smart Desk System is an innovative solution that aims to make the learning process for junior students at home more motivating, encouraging, and interactively engaging, while also reducing stress for parents. By using touch and gesture controls, students can interact with digital content in a natural and intuitive way, which can help increase engagement and motivation [5]. With its advanced features and capabilities, the Smart Desk System has the potential to revolutionize the way junior students learn and make remote education a more enjoyable and effective experience [7].

The concept of Smart Desk Systems has been gaining attention in recent years as a potential solution for enhancing the learning experience in both traditional and remote settings. During the COVID-19 pandemic, the shift to remote learning has highlighted the need for technology that can provide a similar experience to that of a traditional classroom's [8]. The Smart Desk System is well suited for this purpose as it can provide an interactive platform for students to engage with digital content, as well as with one another, regardless of their physical presence. consequently, various companies and research institutions have developed their versions of Smart Desk Systems, each with unique features and capabilities. The following provides an insight into some Smart Desk Systems available in the market.

The relevant literature indicated several studies of smart desks and table for educational purposes. For instance, and serving as a coffee table, Kazar tables as in [9] can bear other objects, such as glasses, plates, magazines, etc. and has a wireless charging station. Furthermore, the table also features a speaker system, which can also be linked to devices like home theatres via Wifi, or Bluetooth connections.

Thus, replacing televisions. To maintain high hygiene levels, the table's surface is treated to prevent bacterial growth and spread.

Another instance, but from the educational sector, SMART™ Table. This table can be regarded as a technological tool carrying various benefits for enhancing the students' and teachers' classroom experience [10]. The current study has focused on the key features of the SMART Table. The components of the supply chain process in the development and delivery of the device have been explained. Moreover, the benefits and limitations of the use of the SMART Table in the classroom have been covered. A few of the notable benefits are ease of remembering content, doing homework, and giving presentations. The cost of SMART Table is a key challenge that can limit the integration of the device as a learning tool. Another issue that can create a challenge in the adoption of SMART Table by students and teachers is the lack of readiness to adopt this technology in the classroom. As evident from the findings, a preference for whiteboard and concern towards the lack of surety about the usefulness of SMART Table in communication between teachers and students can limit its applicability in the University of Bridgeport. Nevertheless, the benefit of the SMART Table in the educational sector cannot be overlooked as it supports students and teachers in dealing with complex concepts and content.

Another example is Magnetic resonance wireless power transfer (MWPT) technology can potentially simplify the electrical cable layout and routing for the indoor environment due to its relatively long transmission distance and high efficiency [11]. Targeting at smart table applications, a series-series matching two-coil MWPT-powered LED system operating at 28 cm optimal operating distance is designed and characterized [11]. The effects of the varying operating distances and different blocking materials on the power transfer are studied. In addition, the system performances under different operating frequencies are also measured and analyzed.

III. METHODOLOGY

A. Planning Phase

The project's current resources, including software, hardware, and essential technologies, are analyzed and evaluated in the technical feasibility stage. This analysis provides information on whether the technology and resources needed to develop the project are indeed available to the company. A feasibility study also looks at the know-how of the technical team, the practicality of employing existing technology, the simplicity of upkeep and technological upgrades, among other things.

The smart desk project will be made initially for the little kids in the UAE due to the availability of high advanced technological infrastructure that allows us to develop and allow users to gain access to the system using existing infrastructure technology available to them at their homes for the completion of kids learning needs. Further, manufactures can easily integrate with current/available technologies in the country to the proposed functionalities such as integrating sensors with hardware to detect kid's mood.

This smart student desk must be constantly improved during the designing and implementation phases, and a final prototype must be tested in most of the potential scenarios. The system will be more dependable as a result. The new system may function independently since it will have its own storage and database and does not require integration with other databases. This new system would make learning for kids more enjoyable; as it would entertain kids and transform their lazy and boring traditional learning system to more interactive and user-friendly system. The technical team available with the company are only three staff, who alone will not be able to design and develop smart student desk. Therefore, more technical talents and talents of software development will be hired. We also have an option to outsources the hardware development team.

The desk adopted a flexible design that can best suit the student's individual needs. As children grow taller, the desk's height can be adjusted for a healthier posture. Further, to optimize storage, the desk is equipped with a deep drawer, pencil and stationary storage compartments, and a pull-up tablet and cable storage compartment. The tablet can be placed either horizontally or vertically, thus displaying in landscape, or portrait format, respectively. Additionally, the tablet's angle can also be rotated to fit the user's height and comfort.

Similarly, the desk's surface can be adjusted with a 10° angle. This allows for various kinds of activities to be conducted on the tabletop. An art project, a painting, or a reading assignment, for instance, can be performed on the inclined surface. If the students are copying or writing a piece of text, the desk can be adjusted to a 0° angle to better fit the task. The bottom of the table's top has a grove that prevents pencils and pens from sliding down as the table's inclination is adjusted.

On right-hand-side of the user, lies the control panel, where all buttons can be placed in a closed compartment. From this control panel, users can adjust the desk's height, angle of inclination of the tabletop, and pull the tablet compartment up or down. That side also hosts a set of other compartments that feature pen and stationary storage. Finally, the upper compartment features electric sockets that can be connected to the wall's electricity source. All cables can pass through the desk's wide and hollow legs, to avoid tripping danger. Finally, the tablet's pullup compartment also features shelves with clips to organize cables and avoid messes. The figures below (Figure 1 – 4) show a set of conceptual illustrations of the proposed Smart Table System.

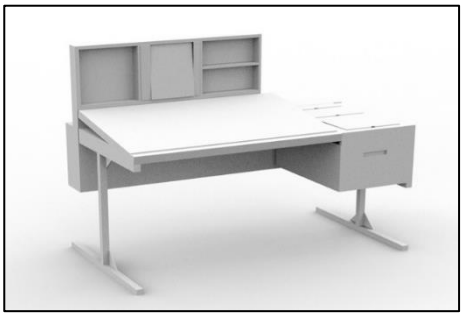


Figure 1 Conceptual model of the proposed table. Source: Authors

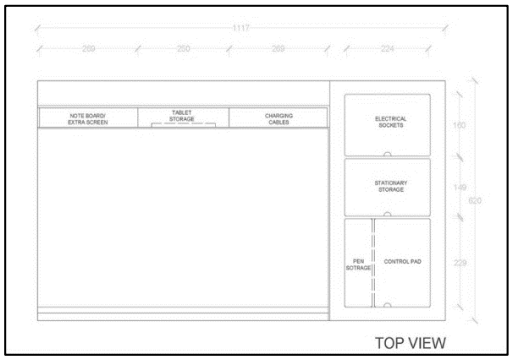


Figure 2. Smart Desk - Top view. Source: Authors

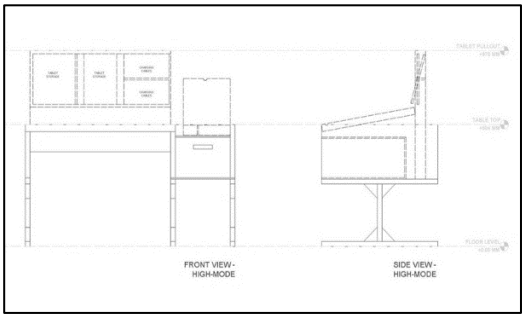


Figure 3. (Left) Desk's Front View (Right) Side View. In Highest mode.

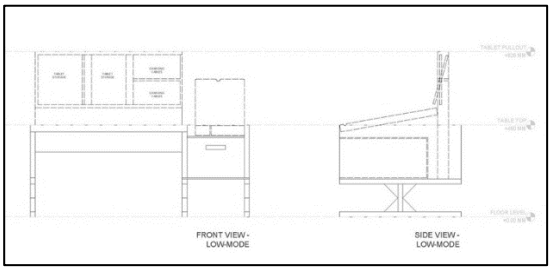


Figure 4. (Left) Desk's Front View (Right) Side View. Both in the lowest mode.

B. Analysis Phase:

1) Requirements gathering

To ensure our product is successful and useful We have explained below the main system requirements to reach our goals. There are two types of system requirements: functional and non-functional system requirements.

Each user mentioned below has an important role in completing the system well and completely because they use the system

regularly and rely on it to improve the relationship between the student child at home and his studies.

Parents

- Faster and easier because they only have to enter their e-mail to receive the child progress reports.
- Able to check the tasks of their children.
- Able to see the child's selection of the weekly reward and put it in the box.

children:

- Able to always use the desk.
- Able to solve tasks using the hologram system
- Able to open the fund every weekend

To ensure our product is successful and useful, we have explained below the main system requirements to reach our goals. There are two types of system requirements: functional and non-functional system requirements.

Functional Requirements

The functional aspect describes what the system is requires in three fronts: input. Output, and process.

Input

- The system will require creating a special account for each child of the family.
- The system will ask to sign in for the child.
- The system will activate the hologram screen.
- The system will ask to choose characters for each child.
- The system will choose the tasks of the day.
- The system will show you options to finish incomplete tasks or start new tasks.
- The system asks the child to scan the book or page that the system is required to help the child with.
- Sometimes the system will scan the child's mood through sensors.
- At the end of all assignments, the system will allow the student to choose his end-of-day rewards.

Output

- The system grants access to students with active account
- The system will welcome the child using the hologram system through different characters.
- If the child adds new tasks, it will be added to the existing tasks that must be completed.
- Alert that there are incomplete tasks.
- The page that was scanned appears in a high-quality display that the child can touch and interact with
- The system sends the end of all tasks to report what the child has achieved in the day.
- Change the display system according to the mood of the child.
- Opening a special box in the desk to place the reward of the child by his parents. The system does not open this box until all the tasks required at the end of each week are completed.

Process

- A student must be logged onto his account before using the desk.
- The hologram system will be turned on and a character speaking to the child will appear.
- The child's choice of today's tasks (completing previous tasks or new tasks).
- Scan the page to view it by hologram.
- Scan the child's mood every half an hour (if he is happy to complete the presentation, but if he is not satisfied, the system will change the character who is talking to the child).

Non-Functional Requirements

This part of the system describe how it operates during performance and control.

A. Performance

- The system will check all tasks.
- The system operates 24 hours a day and seven days a week.
- Charge the desk for a full night after you finish it.
- No system lags.
- The system does not delete tasks on its own.
- It will only take 30 seconds to scan a child's mood.
- The system should send reports to parents at the end of each day.
- The system must save all the tasks of each child in a special file.
- The system tolerates multiple inputs or multitouch’s.

B. Control

- The use of the desk is limited to the student child and school assignments only.
- Only the child himself or one of his parents can access the accounts of the child.
- A child is not allowed to cancel an assignment.

As for the system modelling, we developed the Context Diagram, and the Entity Relations diagram for demonstrating data, as in Figure 5 and 6

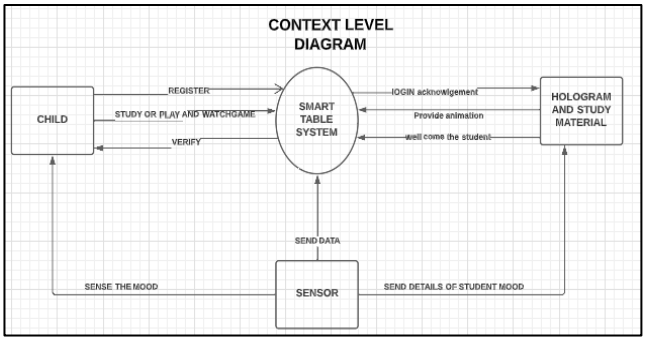


Figure 5. Data Flow Diagram

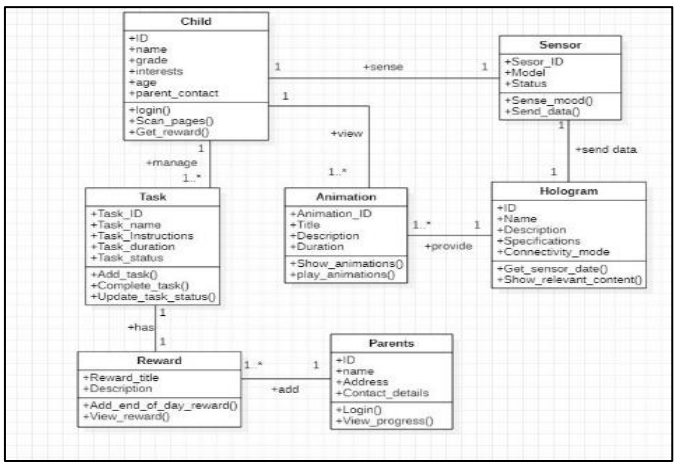


Figure 6. Entity Relationship Diagram

The researchers also considered modelling the decisions and conditions associated with the system, as shown in Figure 7.

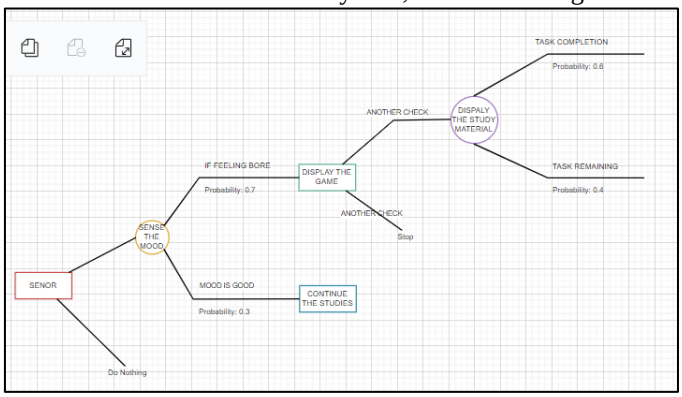


Figure 7. Decision Tree

C. *Design & Implementation Phase*

Also, the figure below (8, 9 and 10) shows samples of the system interfaces, in form of the log-in screen, student's profile and the animation display to all users.

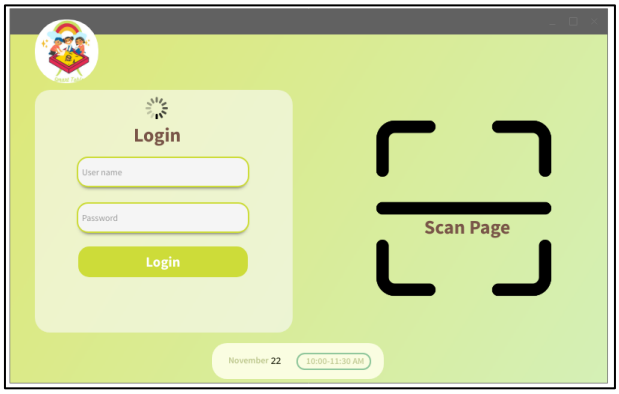


Figure 8. Profile Log-In



Figure 9. User Profile

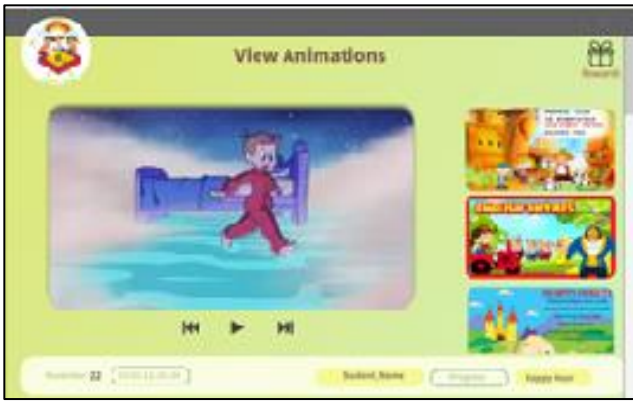


Figure 10. Animation page

D. *Testing Phase*

Alpha testing is the first step in in which developers test the product on their own end. In this stage, we didn't face any structural or functional issues related to the desk. The software was also tested, where users were given system' inputs to try out and observed system' responses according to the requirements. In short, the Alpha testing was successful. In the second phase which was Beta testing, we invited a group of parents along with their kids to test the system, children created

their accounts in the supervision of their parents and fortunately here too our system performed very well. So, the beta test was also successful and fulfilled the user needs.

IV. Discussion

The smart desk has noticeable impact on the studies of the children. The main purpose of smart desk is to engage children in their studies and minimize boredom they feel while studying and lose their interest during studies. Children would feel curious about the smart desk and its Holographic display, which is an innovative idea and kids just think in dreams about these kinds of gadgets. We were also curious about the smart desk that if we are making any mistake because if we talk about the tablets, laptops and smart phones then they gave many advantages but side by side they also distracted kids from their studies and provided them with games and other fun opportunities which is ethically wrong and it also cause of bad performance of students.

Smart desk resolved all these issues and assist children to focus during their educational activities. Smart desk also completes all the objectives and goals which we were expecting from this project and allowed us to explore such a great and innovative idea successfully.

V. Conclusion

In summary, our developed smart desk for kids is designed to make learning and doing homework more fun and interactive than SMART™ and MWPT desks, and more oriented to education than Kazar's table. Our smart desk integrated various technologies that cater to student's needs and interests. Our desk has Adjustable height feature, interactive touch screen, integrated speakers, wireless charging and integrated storage which make the learning process personalized and convenient to students at homes.

Acknowledgement

We are really thankful to the Zayed University to provide us an experienced and best supervision and we would love to express our sincerest gratitude to the University for supporting this project financially.

References

[1] A., Shuhaiber, I., Mashal, & O. Alsaryrah, "Smart homes as an IoT application: predicting attitudes and behaviours". In 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), pp. 1-7, IEEE., 2019.

[2] P., Steurer, & M.B., Srivastava, "System design of smart table". In Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, pp. 473-480. IEEE., 2003

[3] R. El Murr. Smart Desk (Doctoral dissertation, Departments of Computer and Communications Engineering and to the Department of Mechatronics Engineering, American University of Science & Technology), 2018.

[4] D. Rodriguez, S., Prange, L., Mecke, L., & F., Alt, "ActPad–A Smart Desk Platform to Enable User Interaction with IoT Devices". In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-6, 2021.

[5] A., Aryal, B., Becerik-Gerber, F., Anselmo, S.C., Roll, & G. M., Lucas, "Smart desks to promote comfort, health, and productivity in offices: A vision for future workplaces", Frontiers in Built Environment, 76, 2019

[6] J., Muñoz-Benítez, G., Molero-Castillo, & E., Benítez-Guerrero, "Data fusion architecture of heterogeneous sources obtained from a smart desk". In Intelligent Data Sensing and Processing for Health and Well-Being Applications, pp. 23-4, Academic Press, 2018

[7] M. H., Le, T. M., Doan, D. D., Nguyen, & M.S. Nguyen, "Smart Desk in Hybrid Classroom: Detecting student's lack of concentration when studying". In 2022 9th NAFOSTED Conference on Information and Computer Science (NICS), pp. 13-18, IEEE, 2022.

[8] D. D. Nguyen, M. H., Le, X. H., Nguyen, H. T., Ngo, & M. S. Nguyen, "Smart Desk in Hybrid Classroom: Automatic Attendance System based on Face Recognition using MTCNN and ARCFACE". In 2022 International Conference on Multimedia Analysis and Pattern Recognition (MAPR), pp. 1-6, IEEE, 2022

[9] Kazar Digital, Available at: https://www.kazar-digital.com/, Accessed in: January 3, 2023.

[10] W., Hendrick, & L. Casey, "ENGAGE Smart Desk: An API Capable Data Collection and Analysis System for Classroom Behavior", 2021.

[11] T. M., Doan, M. H., Le, D. D., Nguyen, & M. S. Nguyen, "Smart Desk in Hybrid Classroom: Research and Implement a System to detect Inattentive Students". In 2022 International Conference on Advanced Computing and Analytics (ACOMPA), pp. 60-65, IEEE., 2022

[12] M., Yamada, M., Cuka,Y., Liu, T., Oda, K., Matsuo, & L., Barolli, L. "Design of a smart desk for an IoT Testbed: improving learning efficiency and system security". In Innovative Mobile and Internet Services in Ubiquitous Computing: Proceedings of the 11th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS-2017), pp. 27-35, Springer International Publishing, 2017.

Pair Programming Education Aided by ChatGPT

Bojan Banić, Mario Konecki and Mladen Konecki
Faculty of Organization and Informatics,
University of Zagreb, Pavlinska 2, 42000 Varaždin, Croatia
bojan.banic@foi.hr, mario.konecki@foi.hr, mladen.konecki@foi.hr

***Abstract* - Programming education is generally considered as very challenging for both teachers and their students. Students are in many cases more used to memorizing facts than to adopting a new skill that requires intensive and prolonged learning. Also, programming concepts can be quite abstract and hard to perceive by students. In time, many teaching approaches and aiding tools have been identified, but the challenges of programming education have remained one of important research topics. Rapid development of new technology has resulted in new possible approaches in programming education. In this paper, an overview of challenges in programming education is given, and research results about using pair programming in combination with ChatGPT as an educational aiding tool are presented and elaborated.**

Keywords - pair programming; education; ChatGPT

I. INTRODUCTION

Programming is very important part of future information technology professionals because it promotes problem-solving skills [1] and develops computational thinking, which is important for solving complex problems and developing different strategies in areas unrelated to programming [2; 3].

Since the personal computers for home use in the early eighties of the 20th century have started to be used on a wider scale, there has been an idea of teaching students about programming, and many schools have introduced their students to programming in Basic or Logo [4; 5]. However, despite the increasingly available computers and their use for many other purposes most people consider programming a technical activity suitable for only a small part of the population [4; 6].

Possible reason for the loss of interest in programming lies in the fact that programming languages can be quite difficult to learn, and students have problems even with knowing the syntax and semantics of some programming languages [7]. In addition, programming is in many cases presented through problems that are not related to students' interests or experiences, and it is difficult to encourage and motivate students to invest sufficient energy needed to master programming as a skill.

In today's world, the introduction of programming, as well as the application of information and communication technology (ICT), is considered to be of great educational importance, which is why this part of the education is included in many of the most developed European and world educational curricula on the elementary school level [8]. Furthermore, knowledge of programming, as a part of logical reasoning, has become one of the key competencies that all students should possess, as well as employees in a wide range of professions [9].

In order to make learning programming as easy as possible, and to reduce the gap between the level of knowledge and real-world business requirements, new approaches and innovative ways of teaching programming are being developed in schools, colleges, numerous organizations and non-profit associations. One such approach is a combination of pair programming and ChatGPT as an educational assistive tool.

A. Pair Programming

Pair programming refers to the practice in which two programmers work together, while collaborating on the same design, algorithm, code, and testing. The roles of programmers are mutually divided between the person who writes the code, i.e. the driver, and the person who checks it, i.e. the navigator. In addition, both programmers participate equally in designing the solution. It is definitely necessary to prepare the students for the mentioned tasks so that the results are as successful as possible.

Pair programming has become widely accepted since its inception in the 90s thanks to its advantages over individual programming. In secondary schools, this kind of collaborative learning method [10] proved to be successful, and its potential is still a valid research direction. Some of the more significant advantages compared to independent programming are the production of original code with fewer errors, the code is of higher quality and created in a shorter time compared to an individual programmer, programmers learn from each other, build trust and improve teamwork, brain activity is enhanced [11], and programmers feel happier [12; 13]. It has also been proven that pair programming can have a long-term effect on learning outcomes in other areas as well [14]. Pair programming can be applied to simpler tasks for a shorter, but also a longer period of time.

Unfortunately, despite the numerous mentioned advantages, pair programming is not often used in secondary school education, and the reason for this is reflected in certain shortcomings that primarily relate to the

problem of communication, prior knowledge of the partner, and the assignment of appropriate tasks [15].

B. Artificial intelligence and chatbots

Artificial Intelligence (AI) in its advanced form may have skills that resemble the problem-solving skills, communicating in a natural language, learning, and recognizing situations in the same way as humans do [16], but the main difference between AI and other computer programs is its ability to learn.

AI is today important not only for doing business but for education process as well. Because of this, AI is being included in the modern educational programs as a prerequisite to understand the contemporary world. In STEM education it can play an important role in assisting teachers in their roles as facilitators and assessors of learning which is demonstrated by the possibilities to analyze data about the learning process itself [16].

So, it can be said that emerging technologies have also been transforming ways of teaching and learning. For example, chatbots like Chatsonic and ChatGPT (Chat Generative Pre-trained Transformer) are an AI-based programs with technology that recognizes speech and then responds appropriately, providing personalized learning support using computers, mobile phones, or some other devices.

The interaction between chatbots and students may aid learning using a platform for a new educational paradigm in various scientific fields [17]. As a result, because of the teachers shortage Edwards and Cheok [18] have proposed to aid teaching with robots capable of demonstrating artificial intelligence in practice. Furthermore, Zhao et al. [19] claim that the use of AI-based teaching has a positive impact on students' academic achievement, while Topal et al. [17] believe that chatbots can help with science teaching by improving students' performance. Moreover, according to Kim and Kim [20] teachers' use of an AI-enhanced systems may have a positive effect on STEM students' scientific writing.

Despite all, there is a lack of educational perspectives in AI in education research [21], so researchers concern about the absence of educational theories and models, as found in AI-enabled e-learning research published in the last decades.

Following the above, several research question can be identified:

- How to increase student motivation to learn programming?
- Can pair programming promote student interest in programming?
- Can using ChatGPT make programming education more effective?
- etc.

II. RESEARCH METHODOLOGY

In order to conclude about the motivation of students to learn programming, and effectiveness of pair programming aided by ChatGPT the corresponding research was conducted. 45 science and mathematics high school majors participated in the research.

The research has been conducted in two parts. First part of the research aimed to determine the initial motivation for learning programming and attitude towards pair programming and ChatGPT. The second part of the research aimed to determine the motivation for learning programming and attitude towards pair programming and ChatGPT after students' experience with this learning method and technology.

At the beginning of each class, students filled out questionnaires U1, U2 and U3, which tested their motivation for programming, motivation for pair programming, as well as their attitude and habits of using ChatGPT.

All three questionnaires consist of 6 questions and are based on Likert scale where 1 represents the statement "I do not agree" and 5 the statement "I completely agree".

All created questionnaires are more or less based on Technology Acceptance Model and its principles [22; 23].

Questionnaire U1 consisted of the following 6 questions:

- Q1 - Solving mathematical and logical problems is interesting
- Q2 - Programming is a useful skill
- Q3 - Programming is an interesting skill
- Q4 - Programming is difficult to learn
- Q5 - I want to learn programming
- Q6 - Learning programming is interesting

The following 6 questions are part of U2 questionnaire:

- Q1 - Pair programming is fun
- Q2 - Pair programming helps me learn
- Q3 - Pair programming will result in usable program solution in less time
- Q4 - Programming in pairs will result in a higher quality programming solution
- Q5 - During pair programming I can learn from the other team member
- Q6 - I want to program in pair with someone I know

Finally, the following 6 questions are part of U3 questionnaire:

- Q1 - ChatGPT is an interesting programming tool
- Q2 - I learn faster with ChatGPT
- Q3 - I use ChatGPT for learning
- Q4 - I use ChatGPT to create parts of the program code
- Q5 - I use ChatGPT to create the complete program solution
- Q6 - I feel that I don't need to know how to program because I can create programs using ChatGPT

After completing the questionnaires, students were given certain algorithmic tasks in Python determined by the teaching curriculum in informatics. Students were using pair programming, aided by ChatGPT.

Finally, at the end of the class students filled out the same questionnaires again as at the beginning in order to determine a possible change in perception of learning programming, pair programming, as well as using ChatGPT.

To compare the results of the questionnaires, a depended samples t-test was used because to check if there is a statistically significant difference between results at the beginning and at the end of the research.

III. RESEARCH RESULTS

Firstly, the descriptive statistics is given. As it can be seen in Table I, descriptive statistics for questionnaire U1 and all questions at the beginning and at the end of research shows clear difference and shift in students' perception and attitude.

TABLE I. Descriptive statistics of U1 questionnaire

#Q	Valid	Mode	Median	Mean	Std. Error of Mean	Std. Deviation
Q1	45	4.000	4.000	3.644	0.163	1.090
Q2	45	5.000	5.000	4.178	0.163	1.093
Q3	45	5.000	4.000	3.822	0.178	1.193
Q4	45	3.000	3.000	3.156	0.180	1.205
Q5	45	5.000	4.000	3.711	0.200	1.342
Q6	45	3.000	3.000	3.444	0.176	1.179
#Q	**Valid**	**Mode**	**Median**	**Mean**	**Std. Error of Mean**	**Std. Deviation**
Q1	45	4.000	4.000	3.978	0.147	0.988
Q2	45	5.000	4.000	4.222	0.149	0.997
Q3	45	5.000	4.000	3.956	0.168	1.127
Q4	45	4.000	4.000	3.511	0.190	1.272
Q5	45	5.000	4.000	3.911	0.168	1.125
Q6	45	5.000	4.000	3.933	0.186	1.250

Descriptive statistics for U2 questionnaire is shown in Table II.

TABLE II. Descriptive statistics of U2 questionnaire

#Q	Valid	Mode	Median	Mean	Std. Error of Mean	Std. Deviation
Q1	45	5.000	4.000	4.022	0.167	1.118
Q2	45	5.000	4.000	3.800	0.179	1.198
Q3	45	5.000	4.000	4.133	0.161	1.079
Q4	45	5.000	4.000	3.978	0.157	1.055
Q5	45	5.000	5.000	4.489	0.126	0.843
Q6	45	5.000	5.000	4.400	0.133	0.889
#Q	**Valid**	**Mode**	**Median**	**Mean**	**Std. Error of Mean**	**Std. Deviation**
Q1	45	5.000	5.000	4.222	0.155	1.042
Q2	45	5.000	5.000	4.267	0.150	1.009
Q3	45	5.000	4.000	4.178	0.143	0.960
Q4	45	5.000	4.000	4.111	0.146	0.982
Q5	45	5.000	5.000	4.378	0.140	0.936
Q6	45	5.000	5.000	4.422	0.129	0.866

In case of questionnaire U2 the situation is similar as with U1, except regarding the question Q5 where minimal negative difference can be observed at the end of the research.

TABLE III. Descriptive statistics of U3 questionnaire

#Q	Valid	Mode	Median	Mean	Std. Error of Mean	Std. Deviation
Q1	45	5.000	5.000	4.400	0.144	0.963
Q2	45	3.000	3.000	3.022	0.221	1.485
Q3	45	1.000	2.000	2.333	0.220	1.477
Q4	45	1.000	2.000	2.111	0.196	1.318
Q5	45	1.000	1.000	1.933	0.175	1.176
Q6	45	1.000	2.000	1.956	0.174	1.167
#Q	**Valid**	**Mode**	**Median**	**Mean**	**Std. Error of Mean**	**Std. Deviation**
Q1	45	5.000	5.000	4.444	0.129	0.867
Q2	45	5.000	4.000	3.933	0.172	1.156
Q3	45	5.000	4.000	3.622	0.221	1.482
Q4	45	5.000	4.000	3.689	0.217	1.459
Q5	45	5.000	4.000	3.667	0.229	1.537
Q6	45	5.000	4.000	3.489	0.239	1.604

Descriptive statistics of U3 questionnaire is shown in Table III positive difference can be observed in the case of questions Q3, Q4, Q5, and Q6 at the end of research.

The distribution of the difference between the two related groups regarding questionnaire U1 is shown in Table IV.

TABLE IV. Checking normal distribution (U1)

Questionnaire U1

	Beginning res.	End of res.
Valid	6	6
Mean	3.659	3.918
Std. Deviation	0.346	0.230
Skewness	0.049	-0.979
Std. Error of Skewness	0.845	0.845
Kurtosis	0.570	2.850
Std. Error of Kurtosis	1.741	1.741
Shapiro-Wilk	0.989	0.867
P-value of Shapiro-Wilk	0.987	0.216
Minimum	3.156	3.511
Maximum	4.178	4.222

The distribution of the difference between the two related groups regarding questionnaire U2 is shown in Table V.

TABLE V. Checking normal distribution (U2)

Questionnaire U2

	Beginning res.	End of res.
Valid	6	6
Mean	4.137	4.263
Std. Deviation	0.263	0.119
Skewness	0.298	0.241
Std. Error of Skewness	0.845	0.845
Kurtosis	-1.278	-1.315
Std. Error of Kurtosis	1.741	1.741
Shapiro-Wilk	0.945	0.959
P-value of Shapiro-Wilk	0.704	0.812
Minimum	3.800	4.111
Maximum	4.489	4.422

The distribution of the difference between the two related groups regarding questionnaire U3 is shown in Table VI.

TABLE VI. Checking normal distribution (U3)

Questionnaire U3

	Beginning res.	End of res.
Valid	6	6
Mean	2.626	3.807
Std. Deviation	0.958	0.344
Skewness	1.658	1.604
Std. Error of Skewness	0.845	0.845
Kurtosis	2.446	2.612
Std. Error of Kurtosis	1.741	1.741
Shapiro-Wilk	0.792	0.838
P-value of Shapiro-Wilk	0.049	0.124
Minimum	1.933	3.489
Maximum	4.400	4.444

By checking the normality of the distribution in the Table IV, Table V, and Table VI, it is clear that values of skewness are between -2 and +2, but the kurtosis values are not.

Descriptive plots that show the means and standard deviations for each group of results are shown in Figure 1, Figure 2, and Figure 3. It can be observed that all means are higher at the end of the research.

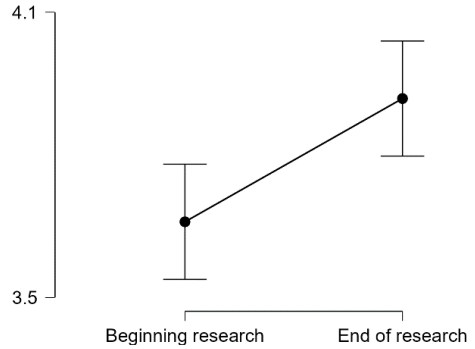


Figure 1. Graphical representation of means in U1 questionnaire

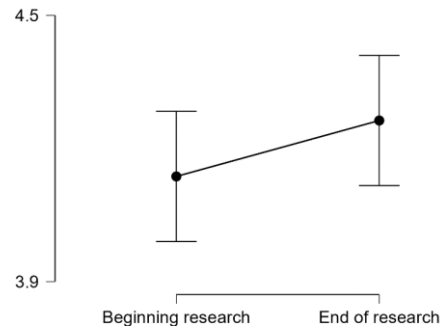


Figure 2. Graphical representation of means in U2 questionnaire

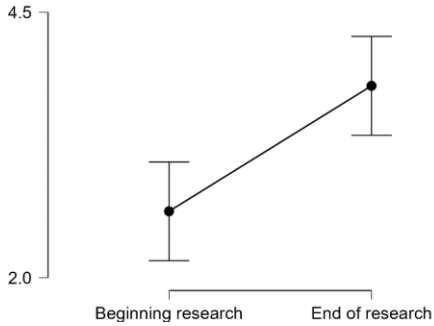


Figure 3. Graphical representation of means in U3 questionnaire

In order to further conclude about obtained research results, a paired samples t-test has been performed.

The respondents on average agreed more (p=0.050) with the statement "Solving mathematical and logical problems is interesting" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.98, SD = 0.988) than before (M = 3.64, SD = 1.090).

The respondents on average agreed more (p=0.015) with the statement "Learning programming is interesting" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.93, SD = 1.250) than before (M = 3.44, SD = 1.179).

The respondents on average agreed more (p=0.003) with the statement "I learn faster with ChatGPT" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.93, SD = 1.156) than before (M = 3.02, SD = 1.485).

The respondents on average agreed more (p=0.000) with the statement "I use ChatGPT for learning" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.62, SD = 1.482) than before (M = 2.33, SD = 1.477).

The respondents on average agreed more (p=0.000) with the statement "I use ChatGPT to create parts of the program code" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.69, SD = 1.459) than before (M = 2.11, SD = 1.318).

The respondents on average agreed more (p=0.000) with the statement "I use ChatGPT to create the complete program solution" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.67, SD = 1.537) than before (M = 1,93, SD = 1.176).

The respondents on average agreed more (p=0.000) with the statement "I feel that I don't need to know how to program because I can create programs using ChatGPT" after engaging in the pair programming and ChatGPT assisted programming education (M = 3.49, SD = 1.167) than before (M = 1.96, SD = 1.167).

IV. CONCLUSION

Programming education is challenging for both teachers and students. The need for programming experts makes programming education an important factor of continuous information systems development.

In this paper, an overview of some of challenges in programing education has been given, and research results of using pair programming and ChatGPT as an assistive educational and programming tool have been presented.

Based on obtained research results it can be concluded that there is a positive difference in students' motivation for programming and using ChatGPT before and after conducted research, but not in the pair programming domain. The reasons can be numerous, and this will part of the further research efforts.

Another part of the future research will be different approaches towards using ChatGPT as an educational tool, and identifying possible challenges that education systems will have to deal with regarding artificial intelligence and other new technologies.

REFERENCES

[1] H.-Y. Wang, I. Huang, and G.-J. Hwang, "Comparison of the effects of project-based computer programming activities between mathematics-gifted students and average students," J. Comput. Educ., vol. 3, no. 1, pp. 33–45, Mar. 2016, doi: 10.1007/s40692-015-0047-9.

[2] M. M. Soysło, "From Algorithmic to Computational Thinking," 2015, pp. 1–1, doi: 10.1145/2729094.2742582.

[3] X. Wei, L. Lin, N. Meng, W. Tan, S. C. Kong, and Kinshuk, "The effectiveness of partial pair programming on elementary school students' Computational Thinking skills and self-efficacy," Comput. Educ., vol. 160, Jan. 2021, doi: 10.1016/j.compedu.2020.104023.

[4] M. Resnick et al., "Scratch: Programming for all," Commun. ACM, vol. 52, no. 11, pp. 60–67, Nov. 2009, doi: 10.1145/1592761.1592779.

[5] S. Popat and L. Starkey, "Learning to code or coding to learn? A systematic review," Comput. Educ., vol. 128, pp. 365–376, Jan. 2019, doi: 10.1016/j.compedu.2018.10.005.

[6] Z. Katai, "The challenge of promoting algorithmic thinking of both sciences- and humanities-oriented learners," J. Comput. Assist. Learn., vol. 31, no. 4, pp. 287–299, Aug. 2015, doi: 10.1111/jcal.12070.

[7] S. I. Malik, G. M. Alfarsi, and M. W. Ashfaque, "A Model for Enhancing Algorithmic Thinking in Programming Education using PAAM ENVIRONMENTAL AND ECOSYSTEM POLLUTION View project A RULE-BASED SYSTEM FOR ADVISING UNDERGRADUATE STUDENTS View project," Artic. Int. J. Interact. Mob. Technol., 2021, doi: 10.3991/ijim.v15i09.20617.

[8] M. Webb et al., "Computer science in the school curriculum: Issues and challenges," in IFIP Advances in Information and Communication Technology, 2017, vol. 515, pp. 421–431, doi: 10.1007/978-3-319-74310-3_43.

[9] K. Umapathy and A. D. Ritzhaupt, "A meta-analysis of pair-programming in computer programming courses: Implications for educational practice," ACM Trans. Comput. Educ., vol. 17, no. 4, Aug. 2017, doi: 10.1145/2996201.

[10] J. Denner, L. Werner, S. Campe, and E. Ortiz, "Pair programming: Under what conditions is it advantageous for middle school students?," J. Res. Technol. Educ., vol. 46, no. 3, pp. 277–296, 2014, doi: 10.1080/15391523.2014.888272.

[11] A. Thapaliya, "Evaluation of brain activity while Pair Programming," in Proceedings - 2020 ACM/IEEE 42nd International Conference on Software Engineering: Companion, ICSE-Companion 2020, Oct. 2020, pp. 104–106, doi: 10.1145/3377812.3382161.

[12] Ö. Demir and S. S. Seferoglu, "A Comparison of Solo and Pair Programming in Terms of Flow Experience, Coding Quality, and Coding Achievement," J. Educ. Comput. Res., vol. 58, no. 8, pp. 1448–1466, Jan. 2021, doi: 10.1177/0735633120949788.

[13] S. Papadakis, "Is pair programming more effective than solo programming for secondary education novice programmers? A case study," Int. J. Web-Based Learn. Teach. Technol., vol. 13, no. 1, pp. 1–16, Jan. 2018, doi: 10.4018/IJWLTT.2018010101.

[14] M. O. Smith, A. Giugliano, and A. Deorio, "Long Term Effects of Pair Programming," IEEE Trans. Educ., vol. 61, no. 3, pp. 187–194, Aug. 2018, doi: 10.1109/TE.2017.2773024.

[15] C. Y. Tsai, Y. F. Yang, and C. K. Chang, "Cognitive Load Comparison of Traditional and Distributed Pair Programming on Visual Programming Language," in Proceedings - 2015 International Conference of Educational Innovation Through Technology, EITT 2015, Apr. 2016, pp. 143–146, doi: 10.1109/EITT.2015.37.

[16] A. Al Darayseh, "Acceptance of artificial intelligence in teaching science: Science teachers' perspective," Comput. Educ. Artif. Intell., vol. 4, p. 100132, Jan. 2023, doi: 10.1016/J.CAEAI.2023.100132.

[17] M. O. Topal, A. Bas, and I. Van Heerden, "Exploring Transformers in Natural Language Generation: GPT, BERT, and XLNet," arXiv Prepr. arXiv2102.08036, 2021, Accessed: Mar. 22, 2023. [Online]. Available at: https://towardsdatascience.com/Transformers-141e32e69591.

[18] B. I. Edwards and A. D. Cheok, "Why Not Robot Teachers: Artificial Intelligence for Addressing Teacher Shortage," Appl. Artif. Intell., vol. 32, no. 4, pp. 345–360, Apr. 2018, doi: 10.1080/08839514.2018.1464286.

[19] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," IEEE Trans. Neural Networks Learn. Syst., vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[20] N. J. Kim and M. K. Kim, "Teacher's Perceptions of Using an Artificial Intelligence-Based Educational Tool for Scientific Writing," Front. Educ., vol. 7, Mar. 2022, doi: 10.3389/FEDUC.2022.755914/FULL.

[21] X. Chen, H. Xie, D. Zou, and G. J. Hwang, "Application and theory gaps during the rise of Artificial Intelligence in Education," Comput. Educ. Artif. Intell., vol. 1, p. 100002, Jan. 2020, doi: 10.1016/J.CAEAI.2020.100002.

[22] M. Y. Chuttur, "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions ," Indiana University, USA . Sprouts: Working Papers on Information Systems, 9(37), 2009.

[23] J. H. Sharp, "Development, Extension, and Application: A Review of the Technology Acceptance Model", Information Systems Education Journal, 5 (9), 2007, ISSN: 1545-679X.

Evaluating Learners' Reactions to an Online Training Programme: A Case Study of Maldives National University

I. Balaban*, J. Badari* i I. Ružić**
*University of Zagreb, Faculty of Organization and Informatics, Varaždin, Croatia; **First Primary School Čakovec, Croatia
igor.balaban@foi.unizg.hr, josipa.badjari@foi.unizg.hr, ivana.ruzic@skole.hr

***Abstract* - In this paper the first level of KirkPatrick's model is used to evaluate the reaction of learners on the online training program conducted at the Maldives National University (MNU). A measurement instrument assessing learners' engagement, satisfaction and relevance of the content was developed and evaluated. A sample of 46 university teachers involved in the training program was used to evaluate the measurement instrument. The results revealed strong positive correlation between all three research constructs. Detailed analysis of participant's responses showed that COVID-19 related lockdown has played a significant role in participants' perception of their engagement and satisfaction with the training program**.

Keywords* - *Kirkpatrick's model; reactions; engagement; satisfaction; relevance; elearning

I. INTRODUCTION

Online training programs have become an unavoidable choice for organizations and individuals looking to enhance their skills and knowledge. The convenience, flexibility, and cost-effectiveness of online training programs have made them an attractive option for learners, and several scientific studies have shown that online training can offer numerous benefits compared to traditional in-person training.

One of the main benefits of online training programs is their accessibility [1]. Online training programs can be accessed from anywhere with an internet connection, making them ideal for learners who are unable to attend in-person training due to geographical, time or other constraints. Additionally, online training programs offer greater flexibility in terms of personal pacing and scheduling [2]. In addition, online training programs can be tailored to the individual needs and preferences of a learner or an institution, providing more personalized and engaging learning experiences.

Finally, online training programs can be more cost-effective than traditional in-person trainings, as they eliminate the need for travel and accommodation and reduce the need for instructor time and resources [3]. This can result in significant savings for both, individuals and organizations, making online training a more attractive option for budget-conscious organizations and financially-challenged individuals.

Overall, the literature suggests that online training programs can offer other numerous benefits over traditional in-person training, including increased accessibility, flexibility, personalization, and cost-effectiveness [4]. These benefits make online training a highly attractive option for organizations and individuals looking to enhance their skills and knowledge.

This paper explores the efficiency of the online training program created for the Maldives National University (MNU) to prepare their teachers to be able to design and conduct online courses due to the increased requirements from their students. Based on that, Erasmus+ project "AMED - Advancing higher education in Maldives through E-learning Development", implemented from January, 2019 to July, 2022 included the co-creation and piloting of a new online study program for professional development of HE teachers and managers focusing on the use of ICT in education.

The design of the AMED study program started with the identification of the training needs in 2019 that targeted students, teachers and institutional readiness for online learning and provided in-depth information on the perception of each target group. The conducted analysis provided details to design and develop the study program curriculum organised in 4 modules: 1/ Foundations of e-learning, 2/ Leadership and management of e-learning, 3/ Designing digital learning, and 4/ The networked teacher. Lecturers learning path involved modules 1, 3 and 4, while the decision makers learning path involved modules 1 and 2. The study program ended with a Capstone project tailored differently for each learning path (lecturers vs decision makers). The study program was developed in English language, implemented within Moodle LMS and successfully piloted in online mode due to the COVID-19 lock-downs at the MNU during 2021 and 2022 among 51 lecturers and decision-makers.

This paper aims to assess the developed AMED study program for lecturers learning path according to the Kirkpatrick's Model, a globally recognized method of evaluating the results of online learning programs [5]. The Model allows program designers and researchers to assess formal and informal training methods and to rate them against four levels of criteria: reaction, learning, behaviour and results. In this paper we will focus on learners' reaction, which is the first level in the Kirkpatrick's Model.

II. RESEARCH GOALS

Considering the fact that Kirkpatrick's model proposes and describes 4 main levels of assessment without implying any kind of measurement instrument or proposing the possible relationships within the instrument, the research aims of this paper are, therefore, as follows:

- To develop a valid and reliable measurement instrument to assess learners' reaction to the AMED study program
- To explore relationships between the main constructs of the instrument
- To interpret the measured assessment results of the AMED study program.

III. RESEARCH CONSTRUCTS

As mentioned in the Introduction, the Kirkpatrick's model proposes the measurement in three dimensions whether the learners found the training 1/ engaging (engagement), 2/ favourable (satisfaction), and 3/ relevant (relevance). Therefore, in this section the research constructs and their importance for the online training programs will be further elaborated.

In order to create an effective educational environment, it is essential for learners to be involved and engaged in the learning process. Learner engagement can be defined as the passion to participate in and accomplish the assigned learning activities, or as learner willingness, need, desire, and compulsion to successfully participate in the learning process [6]. Learner engagement includes three categories: behavioural, cognitive, and emotional engagement [7]. The positive attitude of the learner to participate in the teaching process is called behavioural engagement. Cognitive engagement implies the disposition to think and comprehend one topic or concept, which involves self-regulation. Learners' reactions and feelings during the learning process is known as emotional engagement. Emotional engagement had the greatest impact on learner satisfaction, followed by cognitive and behavioural engagement [8].

Learner satisfaction is a crucial element that contributes to the quality and acceptance of online program. It shows how learners regard learning experience. Learner satisfaction depends on the learner's digital literacy level, social and professional engagement, the learner support system including appropriate academic guidance and the course learning design. There are several factors related to learner satisfaction in online courses which include learning value, instructor enthusiasm, rapport, organisation, interaction, coverage, and assessment [6]. Despite technological advances and efforts to design effective ways to connect in online learning environments, learners in online learning have fewer opportunities to benefit from peer collaboration [9].

Relevance can be defined as the degree to which learners will have the opportunity to use or apply what they have learned. Relevance must be determined by using materials, concepts, examples and language familiar with learners [10]. Relevance can be improved by creating familiarity and good habits: using language and concepts related to the experience of the value of learner life, learning related to something that has already been learned; goal-oriented learning so that learners know the purpose and usefulness of learning; and by applying appropriate strategies regarding the characteristics of learners [11]. Researchers developed relevance interventions that seek to promote learners' beliefs that what they learn connects in meaningful ways to their lives. There are five approaches of the existing relevance intervention research: direct communication, personalization, self-generation, critical reflection, and directed reappraisal [12].

IV. METHODS

The methods section covers the participants (sample), data collection procedure and measurement instrument creation and validation.

A. Participants and data collection procedure

In February 2021, AMED project consortium announced AMED pilot study program to all the academic staff of the MNU. Two learning paths were offered: for lecturers and for decision makers. Invitations were sent to all academic staff to register in the study program using a google form with the description of the pilot program. By the given due date (February 21, 2021), 71 participants registered via the Google form and 3 were registered via personal communication. A total of 51 participants have completed either one of the two learning paths of the study program pilot by the end of project (July, 2022). In this paper we will focus only on the lecturer's learning path (Modules 1, 3 and 4 of the AMED study program, see Introduction). The total of 46 lecturers from the MNU (78% of female and 22% of male) with an average age of 39,8 years participated in the cross-sectional survey in February, 2022. They represent the sample of eligible individuals, i.e. participants who were actively using the training program for several months and have obtained the final certificate. Regarding their previous experience the sample included the balanced number within all groups with max. 30% of the lecturers with 6-10 years of

teaching experience. The sample profile is presented in the Table 1.

TABLE 1. SAMPLE PROFILE (N=46)

Gender	Female	36 (78%)
	Male	10 (22%)
Average age	Male	36,7
	Female	40,6
	Both	39,8
Teaching experience	0 to 5 years	12 (26%)
	6-10 years	14 (30%)
	11-15 years	11 (23%)
	> 15 years	9 (20%)
Motivation to take part in the e-learning change (1-5).	Female	3,5/5
	Male	3,1/5

B. Instrument creation and validation

The instrument development phase was conducted by following the recommendations from Straub and Gefen [13]. We started with the operationalization of research constructs based on the existing measures. Since we used an already developed set of measurement items for the specific constructs, the a-priori content validity was achieved.

As already mentioned, this research uses only the 1st level of Kirkpatrick's model that measures learners' reaction to training programs throughout three (3) main (sub)constructs: engagement, satisfaction and relevance.

Engagement refers to the passion to participate in and accomplish the assigned learning activities, or as learner willingness, need, desire, and compulsion to successfully participate in the learning process [6]. Measurement items used to assess the engagement of learners from Handelsman et al. [14] and Dixon [15] were used in this research.

Satisfaction shows how learners regard learning experience [6]. There are numerous examples found in literature which describe satisfaction. However, in this research we focused on a set of measurement items that best suited our research context. Therefore we focused on measures proposed by Ashby [16], Li et al[17], and Davis [18].

Relevance describes the degree to which learners will have the opportunity to use or apply what they learned [10]. Here we used the measurement items proposed by Hamid & Pihie [19].

The measurement instrument assessing the learners' reaction therefore consisted of 3 main (sub)constructs with a total of 22 items: Engagement (8 items), Satisfaction (11 items) and Relevance (3 items). Participants could record their answers on a 4 - point Likert-type scale.

Collected data was used to evaluate the reliability and validity of the instrument. Some items were removed due to their lower loadings on the posited construct or higher cross-loadings, respectively. In total, 18 items were retained within Engagement (5), Satisfaction (10), and Relevance (3). Once again, the proposed measurement model was evaluated with SmartPLS [20] in terms of validity and reliability (See Table 2).

TABLE 2. INSTRUMENT EVALUATION

Item	Median	Loadings	AVE	CR	CA
Engagement					
ENG1	2	0,640	0,541	0,839	0,790
ENG2	3	0,797			
ENG3	3	0,758			
ENG4	3	0,860			
ENG5	3	0,590*			
Relevance					
REL1	3	0,800	0,762	0,862	0,843
REL2	3	0,911			
REL3	3	0,903			
Satisfaction					
SAT1	3	0,838	0,577	0,935	0,917
SAT2	3	0,613			
SAT3	3	0,594*			
SAT4	3	0,867			
SAT5	3	0,827			
SAT6	3	0,850			
SAT7	3	0,771			
SAT8	3	0,769			
SAT9	3	0,711			
SAT10	3	0,701			

Notes. SD = Standard deviation; Loadings > 0,60; AVE > 0,50; CR > 0,70; CA > 0,70. Scale was a 4-point Likert type.

* item with loading below 0.60 that was retained due to its necessity in the construct

Convergent validity was ensured by examining items' loadings and constructs' Average Variance Extracted (AVE). All items have loaded over 0,60 into their posited construct. The AVE values were all above the threshold of 0,50. Both criteria showed good convergent validity of the model [21]. Values of Cronbach's alpha (CA) and

Composite Reliability (CR) are higher than 0,70 [13] indicating satisfactory internal consistency (reliability).

Final instrument items that were further analysed are presented in Table 3.

TABLE 3. FINAL INSTRUMENT ITEMS IN THREE (SUB)CONSTRUCTS

Engagement
ENG1 I was studying regularly/logging on to the class regularly (Dixon, 2015)
ENG2 I was listening/reading carefully in class/course materials (Dixon, 2015)
ENG3 I applied course materials to my life (Dixon, 2015) (Handelsman, et al., 2005)
ENG4 I really desired to learn the material (Dixon, 2015) (Handelsman, et al., 2005)
ENG5 I participated actively in forums /small group discussions (Dixon, 2015) (Handelsman, et al., 2005)
Relevance
REL1 The content of the courses in this program was current. (Hamid&Pihie, 2004)
REL2 Overall, this study program contributed to the achievement of my wider qualification aim. (Li et al., 2016)
REL3 Overall, I find this study program to be relevant for my work/future tasks.(Li et al., 2016)
Satisfaction
SAT1 Overall, I am satisfied with the quality of the study program. (Ashby, 2011)
SAT2 Overall, I would recommend this study program to other students. (Li et al., 2016)
SAT3 Overall, the study program met my expectations. (Li et al., 2016)
SAT4 Overall, I enjoyed studying this program. (Li et al., 2016)
SAT5 Overall, I was satisfied with the teaching materials provided. (Li et al., 2016)
SAT6 Overall, I was satisfied with the method of delivery of the different teaching materials and learning activities. (Li et al., 2016)
SAT7 Overall, I was satisfied with the assessment during the program. (Li et al., 2016)
SAT8 Overall, I was able to keep up with the workload on this study program. (Li et al., 2016)
SAT9 I was satisfied with the ways to communicate with other peers.
SAT10 I was satisfied with the Learning Management System used in this program.
Note: Answers on 1–4-point Likert-type scale (1 – Strongly disagree; 2 – Disagree; 3 – Agree; 4 – Strongly agree).

V. RESULTS

The Spearman correlation coefficient was calculated for all three constructs in order to explore relationships between the main constructs of the instrument. The results are interpreted according to Evans [22]: 0,00 – 0,19 (very weak), 0,20 – 0,39 (weak), 0,40 – 0,59 (moderate), 0,60 – 0,79 (strong) or 0,80 – 1,0 (very strong). The significance of correlations among the variables is tested at $p < 0,01$. Table 4 indicates strong correlations between all three constructs.

TABLE 4. CORRELATIONS BETWEEN CONSTRUCTS: ENGAGEMENT, SATISFACTION AND RELEVANCE

	Engagement	Relevance	Satisfaction
Engagement	1,000	0,751	0,680
Relevance	0,751	1,000	0,749
Satisfaction	0,680	0,749	1,000

Such findings are in line with the work of Gray and DiLoreto [6] who found that learner engagement mediates the relationship between learner interaction and instructor presence on both perceived learning and satisfaction and that course structure and instructor presence have a significant impact on learner satisfaction. A strong association between satisfaction and engagement was also determined by Rajabalee and Santally [23]. They found that the association between satisfaction and engagement was significant and positively correlated but they also detected a weak but positive significant correlation between satisfaction and engagement with overall performances. In the end, our findings are also supported by Rahim [24] who concluded that improving learner engagement increases learner satisfaction with online learning.

The final step in this research was to interpret the measured assessment results of the AMED study program. From the results shown in Table 2 under construct Relevance we can note the linear representation of the results (Median=3) without any significant deviations. Therefore, deeper analysis was focused on the most deviant items within constructs Engagement and Satisfaction having also proven strong correlations between items.

Within the construct Engagement, the item measuring the regularity of study among participants showed lowest result (Median=2). Within construct Satisfaction the lower result was noticed with the item which measured to which degree the participants were able to keep up with the workload of the study program (Median=3, Mean = 2,500).

Although the program designers and providers were aware of the fact that the learner engagement can be improved by connecting online courses with practice, helping learners in development of specific skills and make them aware of the value of learning and that improved learner engagement increases learner satisfaction with the online learning [24] and despite the fact that they provided the regular feedback and facilitated individual support, the success was limited by the external factors resulting from the COVID-19 related lock-downs in Maldives which, beyond health and psychological factors, included technical factors of which the most important - the instability of

home-internet connections and the lack of the sufficient IT equipment among participants.

Another negative external factor related to the lack of participants' more frequent regularity of study and involvement in study activities was their obligation as higher education teachers to switch their on-site courses to online teaching which was another unexpected and challenging COVID-19 related circumstance demanding their immediate and additional efforts.

Further, it is important to mentiona that the study program was enrolled on the voluntary basis according to the public call and it was expected that the program will attract those MNU teachers that have higher prior content knowledge (as proved by the pre-test) being also the teachers who most frequently and actively got involved in similar trainings in the past. Due to the very high scores in pre-test, their gains in knowledge were rather modest, as shown by the post-test. With that said, they have probably expected a more advanced course than the AMED study program, as it was intended to provide a basic level of knowledge that would allow all MNU teachers to start with e-learning at the MNU. However, the participation of advanced learners in the piloting phase was very much appreciated by AMED study program creators since it provided the stronger criticism and quality feedback for future implementation of AMED study program at the MNU.

VI. CONCLUSION

This paper is focused on assessing the efficiency of the online training program designed to prepare MNU teachers to be able to design and conduct online courses. The first level of Kirkpatrick's model was used to assess the effects of the piloted online training program at the MNU. The measuring instrument that assesses learners' reaction to training programs throughout three main (sub)constructs: engagement, satisfaction and relevance, was developed and validated. The Spearman correlation coefficient indicated strong positive correlations between all three constructs: engagement, satisfaction and relevance. The detailed analysis revealed items that showed lower mean values which can be explained with the effect of negative external factors. For future research, it is suggested to analyse the following three levels of Kirkpatrick's model: learning (Level2), to find out whether learners had acquired the intended knowledge, skills, attitude, confidence and commitment to the training; behaviour (Level3) which describes changes in learners behaviour based on the training they received and results (Level4) which assesses the effects the training program had on the organisational purpose and mission.

REFERENCES

[1] Lucas M, Vicente PN., A double-edged sword: Teachers' perceptions of the benefits and challenges of online teaching and learning in higher education. Educ Inf Technol (Dordr). 2022 Oct 28:1-21. doi: 10.1007/s10639-022-11363-3.

[2] James Brunton, Mark Brown, Eamon Costello & Orna Farrell (2018) Head start online: flexibility, transitions and student success, Educational Media International, 55:4, 347-360, DOI: 10.1080/09523987.2018.1548783

[3] Tang, H. Implementing open educational resources in digital education. Education Tech Research Dev 69, 389–392 (2021). https://doi.org/10.1007/s11423-020-09879-x

[4] Roni Berger & Marilyn S. Paul (2021) Pedagogy vs. Technology: Challenges in Developing Online Courses in Social Work Education, Journal of Teaching in Social Work, 41:3, 275-289, DOI: 10.1080/08841233.2021.1919280

[5] Kirkpatrick, J., Kayser Kirkpatrick, W., (2021). An Introduction to The New World Kirkpatrick Model, Kirkpatrick Partners, retrieved 14 January 2023, URL: www.kirkpatrickpartners.com

[6] Gray, J. A., DiLoreto, M. (2016). The effects of student engagement, student satisfaction, and perceived learning in online learning environments. International Journal of Educational Leadership Preparation, 11(1), n1.

[7] Rojas-López, A., Rincón-Flores, E. G., Mena, J., García-Peñalvo, F. J., & Ramírez-Montoya, M. S. (2019). Engagement in the course of programming in higher education through the use of gamification. Universal Access in the Information Society, 18(3), 583-597.

[8] Deng, R. (2021). Emotionally engaged learners are more satisfied with online courses. Sustainability, 13(20), 11169.

[9] Paulsen, J., & McCormick, A. C. (2020). Reassessing disparities in online learner student engagement in higher education. Educational Researcher, 49(1), 20-29.

[10] Afjar, A. M. et al (2020). J. Phys.: Conf. Ser. 1460 012119

[11] Albrecht, J. R. & Karabenick, A. (2018). Relevance for Learning and Motivation in Education, The Journal of Experimental Education, 86:1, 1-10

[12] Wahyudi, S. et al (2017). Learning Model of Attention, Relevance, Confidence, Satisfaction (ARCS) Supported by Video Tutorial to Improve the Students' Learning Motivation in Vocational High School. Advances in Social Science, Education and Humanities Research (ASSEHR), volume 158.

[13] D. Straub and D. Gefen, "Validation Guidelines for IS Positivist Research," *Commun. Assoc. Inf. Syst.*, vol. 13, no. 24, pp. 380–427, 2004, doi: 10.17705/1CAIS.01324

[14] Mitchell M. Handelsman , William L. Briggs , Nora Sullivan & Annette Towler (2005) A Measure of College Student Course Engagement, The Journal of Educational Research, 98:3, 184-192, DOI: 10.3200/JOER.98.3.184-192

[15] Dixson, M. (2015). Measuring student engagement in the online course: The online student engagement scale (OSE). Online Learning, 19(4). https://doi.org/10.24059/olj.v19i4.561.

[16] Alison Ashby , John T.E. Richardson & Alan Woodley (2011) National student feedback surveys in distance education: an investigation at the UK Open University, Open Learning: The Journal of Open, Distance and e-Learning, 26:1, 5-25

[17] Li, N., Marsh, V., & Rienties, B. (2016). Modelling and managing learner satisfaction: Use of learner feedback to enhance blended and online learning experience. Decision Sciences Journal of Innovative Education., 14(2), 216–242. https://doi.org/10.1111/dsji.12096.

[18] Davis, A. (2017). Measuring student satisfaction in online mathematics courses–RESEARCH. Kentucky Journal of Excellence in College Teaching and Learning., 4(2) Retrieved from: https://encompass.eku.edu/kjectl/vol14/iss/2.

[19] Hamid, J. A., Pihie Z. A. L. (2004). Students' Perception of the Quality of Teaching and Learning in Business Studies Programs, Pertanika J. Soc. Sci. & Hum. 12(1): 71-86, Universiti Putra Malaysia Press, ISSN: 0128-7702

[20] C. M. Ringle, S. Wende, and J.-M. Becker, *SmartPLS 3*. Bönningsted: SmartPLS GmbH, 2015.

[21] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 2nd ed. Thousand Oaks: Sage Publications, 2017.

[22] Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. In *Straightforward statistics for the behavioral sciences*. Thomson Brooks/Cole Publishing Co. module: Implications for institutional e-learning policy. Education and Information Technologies, 26(3), 2623-2656.

[23] Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. Education and Information Technologies, 26(3), 2623-2656.

[24] Rahim, N. B. (2020). Improving student engagement and behavioural outcomes via persistence among distance learners. Akademika, 90(2), 91-102

The Impact of Short and Very Short Videos on the Effectiveness of Teaching and the Principles of Their Development

J. Žufić*, S. Bodiš**
* Jurja Dobrila University in Pula / Faculty of Educational Sciences, Pula, Croatia
** Marčana Elementary School, Marčana, Croatia and Juršići Elementary School, Juršići, Croatia
jzufic@unipu.hr

Abstract **- Short video content for education is increasingly appearing as part of educational content. This was especially evident during the COVID19 crisis. Not only have they persisted, but they are taking an ever-increasing share in classical education Video material not only needs to be developed in accordance with the pedagogical principles, it also needs to keep students' engagement and promote active learning. The length of video material is an important element that needs to be considered in development of teaching contents - how long the video should be. The article outlines active learning guidelines for teaching with video materials and the impact of the length of video materials (short and even shorter) on learning effectiveness. This article reviews previously published research and the results of the research projects conducted in three sixth grade classes in primary schools in Juršići, Marčana and Rovinj. The research included 55 elementary school students and video materials for the subject of Informatics. Students in each class were divided into two groups: One group of students was presented with a short video of 13:27 minutes, and the other group was presented with ten very short videos, ranging from 0:27 to 2:59 minutes.**

Keywords - short videos; very short videos; effectiveness of teaching

I. INTRODUCTION

Videos has become an big part of all levels of education. It is integrated into schools and universities as part of traditional teaching. The question arises as to how long the video teaching material should be. How long is too long and how short is too short? A number of studies have been conducted on how long a video lesson should be in order for viewers to follow it from start to finish. A 45-minute lecture is more than the average student can handle, and not because of shrinking attention spans. It is too hard to listen to someone's monologue for a long time. In addition, video content works in a different way than traditional teaching, and a long video is not only unnecessary, but can also be counterproductive. The majority of users in the conducted research stated that they want educational video content to be shorter than 20 minutes, and they prefer concentrated video content lasting 3 to 6 minutes. As teachers, we know how difficult it is to make a 6-minutes video teaching unit.

A 2018 study [1] "New Research: Video Statistics, Habits, and Trends You Need To Know" found that more than 50% of people prefer videos between three and six minutes long, and just over 25% of them prefer video length of 3 to 4 minutes and the same percentage prefers video recordings of 5 to 6 minutes. The question a video course developer asks is "How long can the viewer's attention be held to deliver the necessary information. Knot [1] states that the reason people stop watching content has nothing to do with the length of the video. They found that 83% of viewers will continue watching the video recording if they are interested in the content's topic.

TED talk [2] has an 18 minute speech rule. According to TED curator Chris Anderson, the 18-minute video is long enough to be serious and short enough to hold attention. Speakers who are used to speaking for 45 minutes think about what exactly they want to say and then shorten their presentation.

Keep in mind that video content is a passive type of learning. How much attention can you expect from students in a one-way conversation? Students want the opportunity to put the acquired knowledge into action. After watching the video, they want to experience what they have learned. Therefore, a 45-minute lecture divided into seven or eight videos will be more effective than a single presentation [2].

Teachers who create teaching material using the method of instructional designers sometimes exaggerate the amount of content they offer students at the expense of greater content value. These are not the same [3]

In creating teaching content, teaching outcomes should always be the main objective. It is important for students to get to the essence of the lesson in as little time as possible [4]

Research conducted on a large number of Massive open online courses found that the most acceptable educational video contents are those lasting up to 6 minutes, and students were most focused on those lasting up to 3 minutes [4].

As many as 95% of the videos on the Boclips platform are shorter than 6 minutes. [5]

A study conducted at the University of Wisconsin-Stout concluded that videos had a positive impact on their learning. Students at the University of Wisconsin-Stout want the videos to last no longer than 15 minutes [6]

The Wistia (Video Hosting and Marketing Tools for Business) web portal hosts over half a million business educational videos, with more than 1.3 billion views. It was found that student engagement initially declines from the first to the third minute; it is constant from the third to the twelfth minute and it drops significantly after the 12th minute mark [7]

YouTube offers many educational videos. It has been measured that the median user's engagement time for videos up to six minutes long was close to 100%, however, with videos of 9-12 minute long, the engagement was 50% and for videos that were 12-40 minute long, the engagement was ~20% [8][9].

The next aspect that needs to be addressed is the question: How video content affects cognitive development, that is, how students remember the material presented through video lessons. Teachers, when creating video content, should try to minimize the external cognitive load, that is, they should take into account the intrinsic (what leads the students to the activity) cognitive load of the students. Two facts are known: first, working memory has a limited capacity and second, working memory must process information to be encoded in long-term memory. Working memory should accept, process and send only important information to long-term memory [10].

Additionally, the cognitive theory of multimedia learning considers that working memory has two channels (visual and verbal) for collecting and processing information [11][12]. Using both channels increases the working memory capacity, but both can become overloaded [13]. Consequently, if a video content creator uses both channels, there is a greater chance for better teaching.

Brame [13] points out that when creating teaching material, it is necessary to take care of three components: a) cognitive load, b) content that encourages student engagement and c) content that will contribute to active learning. Cognitive load is represented by a) Signaling, b) segmentation for sharing information, removing redundant information and c) using audio and visual channels to convey complementary information. Student engagement includes several recommendations: a) keep each video short, b) the teacher when creating the video should use simple language and speak quickly and enthusiastically, and c) should create videos to emphasize relevance to the lesson. Active student learning implies the following elements: a) give students both video and questions to help them learn, b) videos should have features that allow students to have control over the video, and d) videos are a good source of learning material that students will use for their homework.

"Signaling (cueing) [14], is the use of text or symbols in videos to highlight important information." Signaling helps direct students' attention and serves to process information in working memory. Signaling can reduce external load on users and thus helps determine which content is important. Mayer and Moreno [12] and deKoning et al. [14] showed that this approach improves students' ability to learn from animations, and Ibrahim et al. [10] showed that these effects extend to video content as well.

As previously known, active learning has great advantages over passive learning [15] (in [16] and [17]). If video contents promote cognitive activity, teaching effectiveness will increase. Self-regulation is a process that requires students to monitor their own learning. This will allow them to recognize learning difficulties and adapt to the process of active learning. Active learning is aided by the creation of interactive videos. Interactive video, according to Matijević, is a non-linear digital video technology that allows students to pay full attention to educational materials and review each part of the video as many times as they want" [18].

There are two ways of learning interactively: In the first, the student is given feedback after their answer, and in the second, the student is invited to participate in activities that trigger cognitive processes. There are three structures of an interactive video. In the first, students decide at the beginning of the video which video they want to watch. In order for the student to choose another video recording, they must go back to the beginning. In the second structure, the student can choose an-other path at any time, i.e. watch another video. The third structure is one in which the student decides what will happen at each step, that is, there are several possible pathways that the student can choose [18]. The main purpose of interactive visual essays is for students to receive feedback and to move from passive to active learning.

Brame [13] gives recommendations for developing videos:

• Keep videos short and focused on learning objectives.

• It is necessary to use audio and visual elements in teaching to make these elements complementary and not redundant.

• It is necessary to use signaling to highlight important ideas or concepts.

• Use an enthusiastic conversational style in the speech to increase student engagement.

• Videos need to encourage active learning with the help of guiding questions, interactive elements or link them to homework.

The website learndash.com [2] suggests creation of shorter videos. Some of the reasons are:

1 - Learning in *small bites*

2 - Increased student engagement

3 - Makes the most of available learning time

4 - Increases the perceived value of video content (material)

II. RESEARCH

A) *Methods and Participants*

For the purposes of teaching in primary schools and for this research project, an educational material called Maps and Files was created for students of the 6th grade in the subject of Informatics, in the domain Information and digital technology, which covers the teaching unit Creation and basic operations with maps and files, on to the OneDrive service. A video between 6 and 18 minutes long is considered a short video, and less than 6 minutes is considered a very short video. The length of the entire educational video is 13:40 minutes (available on the YouTube service at the link: https://www.youtube.com/watch?v=jelv6KQr8AE). Teaching material contains ten short units. The short video is presented as one unit, while the category of very short videos is presented in 10 short (separate) video units.

Teaching units are:

1. Login to the OneDrive system - 0:21 minutes long (available at https://www.youtube.com/watch?v=BSE22o84BJM),
2. Creating maps - 1:17 minutes long (available at https://www.youtube.com/watch?v=bAxlI-ClsGY),
3. Creating Word (Office 365) file - 0:17 minutes long (available at https://www.youtube.com/watch?v=film3Np7a8o),
4. Editing of Word file - 2:00 minutes long (available at https://www.youtube.com/watch?v=UfnHPONJ4Qw)
5. Renaming – 0:48 minutes long (available at https://www.youtube.com/watch?v=KSN9kvuspaQ),
6. Downloading file on the computer – 1:40 minutes long (available at https://www.youtube.com/watch?v=QsgFRzPLERg),
7. Transferring file – 2:59 minutes long (available at https://www.youtube.com/watch?v=rNiQyvj__bI),
8. Sharing maps or files via usernames – 1:55 minutes long (available at https://www.youtube.com/watch?v=KIauRF6Momg),
9. Sharing via links – 1:34 minutes long (available at https://www.youtube.com/watch?v=m32Xq6dn_t8) i
10. Finding maps and editing files – 0:40 minutes long (available at https://www.youtube.com/watch?v=vR5CUUbEUQw).

Fiftyfive 6th grade primary school students (34 girls and 21 boys) of primary schools in Rovinj (*Vladimir Nazor* Elementary School), Marčana and Juršići participated in the research.

Students were divided into two groups: Group I learned from a short video, one continues teaching unit and Group X learned from 10 very short videos.

Group I consisted of 27 students (14 girls and 13 boys), and Group X consisted of 28 students (20 girls and 8 boys). Group I had an average grade in Informatics from the previous (fifth) grade of 4.50 (girls 4.55 and boys 4.38 - using the *t-test* it was determined that there was no statistically significant difference in the average grade between boys and girls, $t=1.38$, $p<0.05$), and Group X had an average grade of 4.11 (girls 3.86 and boys 4.38 - using the *t-test* it was determined that there was no statistically significant difference in the average grades between boys and girls, $t=1.38$, $p<0.05$). Boys from Group I and Group X had the same average (4.38). The girls from Group I had an average score of 4.55, and the girls from Group X, had an average score of 3.86. Using the *t-test*, it was determined that there is a statistically significant difference in the average grade between the girls of Group I and Group X with $t=2.1$ and $p<0.05$. Girls from Group I have a higher average score than girls from Group X.

Using the *t-test*, it was determined that there was no statistically significant difference in the average grade of all students between Group I and Group X with $t=1.63$ and $p<0.05$.

B) *Student's Testing*

In order to test what students learned, a test consisting of eight practical tasks was created. For each task, students could get from a minimum of 0 to a maximum from 2-5 points (depending on the task). No points were scored for an incomplete task. The entire test could bring a total of 29 points. Testing was done via a remote learning system, which also evaluated the test results.

III. RESULTS AND ANALYSIS OF THE CONDUCTED RESEARCH

Obtained results of the conducted research are presented in Table I.

TABLE I. OBTAINED RESULTS OF STUDENT'S TESTING

	Group I.	**Group X.**
Total number of students [n]	27,00	28,00
Average points, all students [max=29]	22,11	23,19
Standard deviation	8,38	6,87
Average points, girls [max=29]	22,95	22,64
Standard deviation	7,72	7,72
Average points, boys [max=29]	20,00	23,77
Standard deviation	10,11	6,07

Students in Group X achieved slightly better results, but using the t-test it was determined that there was no statistically significant difference in the results of all students between Group I and Group X with $t=0.52$ at the significance level $p<0.05$.

The average results of the boys are slightly different (23.77 against 20.00 in favor of the boys from Group X) between Group I and Group X, but using the t-test it was determined that there is no statistically significant difference in the achieved results of all students between Group I. and Group X. with $t=1.07$, at the level of significance $p<0.05$.

The average results achieved by girls and boys within Group I are similar, and this also applies to Group X.

The average scores of girls between Group I and Group X are the same. With the fact that the initial level of knowledge of girls from Group I was significantly

higher than girls from Group X, it is concluded that there is a certain difference in favor of learning from shorter videos.

The obtained results indicated that there are no statistically significant difference between the effectiveness of short (length from 6 to 18 minutes) and very short (length up to 6 minutes) videos for students aged 12-13 years. This conclusion is limited by a) a relatively small sample, b) only one teaching unit and c) one teaching subject. To create a quality video material, the creator of teaching contents should ask the following questions: Who is my audience? What problem are they trying to solve? What are their goals? What will they need from me to accomplish those goals? What is their skill level? [19] and also be guided by the above guidelines: keep the video short, let it be about elements that influence student engagement and elements that promote active student learning.

IV. CONCLUSION

Many studies and articles have shown that learning from educational videos can be effective, depending on the approach to its development. It is important to take care that the video content is adapted to the age for which it is intended, and it is necessary to use clear and articulate speech and try to make such video content that will encourage student engagement. One long video of 45 minutes can always be divided into videos of up to 6 minutes, so it will be easier for students to adopt the material. Although the research clearly shows that short videos provide more benefits for students and the teachers who make them, sometimes it is justified to choose a video longer than six minutes, even though the attention slowly decreases. However, regardless of the potential of an individual or group to hold their attention, educational video materials should not be about testing people's mental stamina. Student needs or learning outcomes should be the main guidelines when creating video materials.

REFERENCE

[1] R. Knott, Video Length: How Long Should Instructional Videos Be? 2018, https://www.techsmith.com/blog/video-length/ visibled January 10th, 2023.

[2] How Long Should Videos Be for E-Learning? LearnDash Collaborator, 2019, https://www.learndash. com/how-long-should-videos-be-for-e-learning/ visibled January 10th, 2023.

.

[3] What's the Optimum Length for an Instructional Video? And Why is it Important?, 2021, https://www.boclips.com/blog/whats-the-optimum-length-for-an-instructional-video-and-why-does-it-matter visibled January 10th, 2023.

[4] P. J. Guo, J. Kim and R. Rubin How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos, DOI: 10.1145/2556325.2566239, Conference: Proceedings of the first ACM conference on Learning @ scale conference, 2014

[5] What's the Optimum Length for an Instructional Video? And Why is it Important?, 2021, https://www.boclips.com/blog/whats-the-optimum-length-for-an-instructional-video-and-why-does-it-matter visibled January 10th, 2023.

[6] R. Berg, Leveraging Recorded Mini-Lectures to Increase Student Learning, Online Classroom, 2014, vol. 14, n2 https://www.academia.edu/6778520/Leveraging_Recorded_Mini-Lectures_to_Increase_student_Learning visibled January 10th, 2023.

[7] What is the Best Length for an Instructional Video? - eLearning and Professional Studies - UAB The University of Alabama at Birmingham https://www.uab.edu/elearning/news/faculty-best-practices/what-is-the-best-length-for-an-instructional-video visibled January 10th, 2023.

[8] A. Baker, Optimal YouTube Video Length. Content Creator., 2018., https://contentcareer.com/blog/optimal-youtube-video-length/ visibled January 10th, 2023.

[9] J. C. Brame, Effective educational videos. Center for Teaching, (2015). http://cft.vanderbilt.edu/guides-sub-pages/effective-educational-videos/ visibled January 10th, 2023.

[10] M. Ibrahim, P. D. Antonenko, C. M. Greenwood, D. Wheeler Effects of segmenting, signaling and weeding on learning from educational video, Learn Media Technol. 2012; 37:220–235.

[11] R. E. Mayer, Multimedia Learning. New York: Cambridge University Press; 2001.

[12] R. E. Mayer, R. Moreno, Nine ways to reduce cognitive load in multimedia learning. Educ Psychol. 2003; 38:43–52.

[13] J. C. Brame, Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. CBE Life Sci Educ. 2015.

[14] B. deKoning, H. Tabbers, R. Rikers, F. Paas Towards a framework for attention cueing in instructional animations: guidelines for research and design. Educ Psychol Rev. 2009; 21:113–140.

[15] J. K. Knight, W. B. Wood, Teaching more by lecturing less. Cell Biol Educ. 2005; 4:298–310.

[16] D. C. Haak, J. HilleRisLambers, E. Pitre, S. Freeman, Increased structure and active learning reduce the achievement gap in introductory biology. Science. 2011; 332:1213–1216.

[17] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, M. P. Wenderoth Active learning increases student performance in science, engineering and mathematics. Proc Natl Acad Sci USA. 2014; 111:8410–8415.

[18] R. Matijević, Interaktivni videozapisi u obrazovnom procesu. Graduation thesis, Jurja Dobrila University in Pula, 2022.

[19] M. Pierce, The Truth About Instructional Video Length, 2020. https://trainingmag.com/the-truth-about-instructional-video-length/ visibled January 10th, 2023.

Teacher's Motivation for Applying the Unplugged MEMA Method for Early Programming Teaching

S. Babić*, M. Čičin-Šain**
* Polytechnic of Rijeka, Business Department, Rijeka, Croatia
** Cybernetics society Rijeka, Rijeka, Croatia
e-mail: sbabic@veleri.hr; marinacs@efri.hr

***Abstract* - Teachers' motivation is crucial for the application of innovations in their teaching. Therefore, it is important to find out what factors promote teachers' intrinsic and/or extrinsic motivation so that the innovation is accepted and used in the classroom. This paper points out the need to understand the factors that motivate teachers to use the Unplugged MEMA method for early programming teaching. In this context, ways to motivate current and future teachers to use the MEMA method over time are identified, and possible motivating factors are explained using the existing conceptual model of educational innovation adoption. The results of this work may help create a teacher education model for the future application of the MEMA method and the implementation of similar innovations in the classroom.**

Keywords – teacher, teaching programming, motivation, MEMA method, young children, innovation acceptance models

I. INTRODUCTION

Programming is a very useful activity. With programming, a man becomes a computer master. Without the ability to program, a man becomes only a user, using the machine's capabilities, which he knows how to use, and is often frustrated when the computer gives him wrong answers. So why can not we all program, just like we can all read and write?

If in a classroom of any school the question is asked, "Is programming easy?", almost all respondents will answer that programming is difficult. You will get a similar answer if you ask students, adults or retired people. It is amazing that even people who are assumed to have no problems with programming, such as students of computer science and engineering, think similarly. But no, even among these students, many have problems precisely with programming. Therefore, it is logical that we try to answer the question, "Why is programming difficult for most people?"

To write a program, i.e., a set of instructions in a computer language, it is necessary to think like a machine, i.e., to apply computational thinking. Only some find it easy to get into this role, so there are few in the class who have no problems with programming. From teachers' experience, we know that these students often do not need to be good at other subjects. These students find programming so easy that teachers sometimes fear that the students will outperform the teacher. But how can programming be taught to all, not just the talented?

One can try to find the answer in the analogy to reading and writing. Everyone learns to read and write, not only those who are gifted. But it was not always so. A hundred years ago, there were many people who could neither read nor write. How is it that today the situation is quite different and illiteracy has been almost completely eliminated, except in some parts of the world [1].

Illiteracy is eliminated thanks to education. Good education begins in early childhood. What is normal for learning to read is normal for all other skills. Good athletes train in childhood, children learn foreign languages, play instruments, operate various devices - all without problems. Therefore, an early start to programming seems like a good idea.

If we analyze the letters of the Latin alphabet, there are about thirty characters, which in some languages receive some hooks, two dots above the letters or umlauts, tildes, tails below the letters and more. Also, in some languages there are rules on how to read when two or more letters are written together. In each language, the letters of the alphabet are learned very early (usually from the age of six or even earlier) and very gradually (in one year, the letters are learned and over four primary grades, reading and writing are repeated and completed in the mother tongue).

No one expects children in Croatian schools to learn to read in English or in any language other than Croatian, and this is true for other languages as well, so that little Frenchmen learn to read in French, Englishmen in English, and Italians in Italian. They learn to read and write not only in their mother tongue classes, but also in history, nature, geography and mathematics classes, because everywhere something was read and written. Everything is taught to them by their teacher, who knows exactly what the children can do and how much they know.

If children would learn programming step by step in their native language, as one learns letters and reading, without extra lessons, and if their teacher knew how to explain everything necessary, programming could become easy for all, not only for some.

II. MEMA Method As A Proposal For Programming Learning For Young Children

There are many approaches and numerous computer programs designed specifically for students' first encounter with the world of programming. When one of the existing programming languages is used for the child's first encounter with programming, a new set of problems emerge waiting to be solved. Unfortunately, the solutions are neither easy nor cheap. Schools must be equipped with hardware and software to work with these programs, teachers must be trained to use this software on this hardware, and students must learn the new material in addition to other material. Classroom teachers are not capable of teaching such programming, and there is a shortage of computer science teachers because they would rather become programmers, at least those who can program well.

For this reason, the MEMA method (without the use of an electronic computer) is recognized by some as a method for the first steps in programming that certainly alleviates and perhaps even avoids all these problems. The MEMA method uses the MEMA toy - a didactic toy that simulates the operation of a computer's main memory [2]. For the first grade of elementary school (Fig. 1), the MEMA toy consists of 9 drawers of different colors, while in the higher grades it uses the names on labels (Fig. 2) that address the variables. Neither hardware nor software is needed.



Figure 1. Toy MEMA (According to the original author of the Mema method, Ph.D. Marina Čičin-Šain)

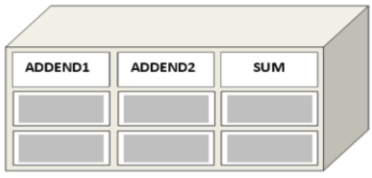


Figure 2. Names of variables on MEMA drawers (According to the original author of the MEMA method, Ph.D. Marina Čičin-Šain; Use the MEMA toy from the first grade, just turn it around - the idea of teacher Divna Bjelanović)

With this method, children in the lower grades of elementary school are taught programming by their teachers, who also teach them other subjects, regardless of whether or not they can program in any of the existing programming languages, which is best for children because they know how young children should be treated.

Children learn programming with minimal extra effort because they repeat along the way what they learn in school anyway, e.g., in first grade they repeat adding and subtracting natural numbers up to 20 [3], in second grade they repeat multiplication tables [4], and in third and fourth grade they practice reading with comprehension and more mathematically complex tasks [5]. Practice has shown that children who have worked with the MEMA method in the lower grades of elementary school master programming in the real programming language more easily [6]. The research was conducted only in a school where only one class worked according to the MEMA method. It was these students who showed statistically significantly better success in programming.

From 2022, the project ERASMUS + will be carried out with the aim of applying the MEMA method and it is expected that the same hypothesis will be confirmed on a much larger sample.

III. Teachers' Motivation

Many years of experience have shown that the greatest problem in introducing a new thing into schools is the lack of motivation on the part of most teachers. As Crosby quotes Tolstoy in his work [7], the teacher's inertia can be described as follows:

- "The teacher always has a tendency to select that method of teaching which is easiest for him;
- that the easier it is for him, the less satisfactory it is for his pupils; and
- that method only is good which gives satisfaction to the pupils. And to give satisfaction to the pupils it is necessary to take account of the differences between them and of their natural aptitudes."

In order to apply the MEMA method in their teaching process, teachers need certain knowledge, skills and attitudes. However, so far, the results of the practice show that the application of the MEMA method in the classroom is problematic, although teachers expressed positive attitudes towards this method during the trainings conducted in the past period.

Therefore, further research needs to be conducted to identify the motivational factors that could encourage teachers to apply the MEMA method in different educational settings.

A. Motivating Current and Future Teachers to use the MEMA Method for Early Learning of Programming

Prof. dr. sc. Marina Čičin-Šain developed the MEMA method for learning programming at an early age more than 40 years ago and has since held numerous lectures and workshops with current and future teachers to this day.

The lectures and workshops have been delivered in various ways: live and online, in small and large groups with current and future teachers in Croatia, Slovenia and

Poland. This can be seen on the Rijeka Cybernetics Society website [8]. However, only a small number of them have used this method in their teaching. As mentioned above, during the training on the MEMA method, teachers regularly have a positive attitude towards this method, and yet most of them do not end up using it.

To date, several studies have been conducted with the goal of obtaining feedback on the usefulness of MEMA in the classroom and the desire of teachers to continue using this method. After lectures about the MEMA method for experienced teachers and prospective teachers, some questionnaires were distributed and some researches were conducted by the authors of this paper.

After a live lecture on MEMA in 2017, a survey was conducted with students – future teachers (N=30) who attended the second year of Teacher Education [9]. The results of this research [9] showed that student-future teachers find programming difficult, and that Mema is useful as an aid for acquiring knowledge about algorithms in mastering numbers, motor skills developing, and learning colors. Interestingly, 58.6% of them think that the MEMA toy is easy to use, 70% think that the MEMA method is interesting, and 68.6% of them would like to learn more about its use in teaching. Also, a large number of future teachers (70%) considered MEMA to be suitable for children aged 7 and 8 years. It should be pointed out that the participants-future teachers pointed out their insufficient IT education in programming before enrolling in college.

At the end of the workshop "Really first steps in learning programming for teachers" on the application of the MEMA method in teaching, which was held within the "All Digital Week 2018" activity, a survey was conducted with teachers (N = 11) who were between 25 and 60 years of age [10]. The results of the research [10] showed that teachers were hesitant in answering the question of whether programming was difficult. Almost all teachers found the MEMA method useful as an aid in mastering numbers, colors, motor skills and learning algorithms. 90.9% of them expressed the opinion that the MEMA method is interesting and that it is easy to use and that they would like to apply the method in their classes. Since only the MEMA method for the first grade was presented at this workshop, the results of the research [10] confirmed that the MEMA method is suitable for children aged 7 and 8 years. It is important to mention that 63.6% of the participants in this study have never learned programming in any of the programming languages (Logo, Qbasic, C++, Python, ...).

From the previously mentioned results, it is evident that current and future teachers believe that the MEMA method enables the development of IT and general content, especially for the development of logical thinking that every child needs regardless of whether or not they will become a programmer or computer scientist.

B. Potential Factors of Teachers' Motivation for Application of MEMA Method

In the literature it is possible to find a large number of models that have in many ways sought to detect the factors, starting from different aspects, which affect the use of innovations by users.

According to Rogers' model the theory of diffusion of innovations [11] the unplugged method MEMA can be seen as an innovation in the teaching process. Acceptance of this innovation by teachers can be associated with the speed of dissemination of MEMA information over some time through a specific communication channel. In doing so, the essential characteristics of MEMA (e.g. price, availability, usability, ease of use, ..) and the education system (e.g. compatibility with curricula of various subjects). It is important to stress that innovation users can make the decision to accept or reject innovation by personal free choice, collectively or commandingly, followed by a phase of using innovation in an environment [11]. The last stage in the process of accepting innovations according to this model [11] is a confirmation in which the user, under the influence of different factors, can even at the final stage make a decision not to accept or accept the innovation. Where the decision to embrace innovation is the result of a command form, the adoption of innovation can still be at different levels. In this case, it is also necessary to motivate users to apply innovation at higher levels. Embracing innovation, according to Rogers' definition of the process of adopting innovation [11] is "an individual's decision to make full use of the innovation available."

From the above, we can assume that the level of adoption of the MEMA method in educational settings will largely depend on teachers as a key stakeholder of the education system and their motivation for the above.

It is known, according to Deci and Ryan [12] and Vallerand [13], that individuals change their behavior under the influence of various factors extrinsic motivation and intrinsic motivation, and amotivation. Extrinsic motivation includes, for example, getting rewards for effort, monetary gain, praise, the possibility of promotion or, in a negative context, avoiding warning and the like. Intrinsic motivation refers to a set of motives that support the self-motivation of an individual to perform an action such as, for example, desire, self-interest, inner satisfaction and the like, while amotivation denotes a type of motivation when individuals depict a lack of motivation. As an example, it is possible that under this model [13] unmotivated classroom teachers do not see some advantage in using the method.

Furthermore, according to the Technology Acceptance Model – TAM [14], one of the most widely used models for technology acceptance, it starts from the claim that the use of technology can be explained by the motivation of the user in the environment in which they are exposed to the actual use of technology. According to this model, two key factors perceived usefulness and expected ease of use affect the user's attitude towards the use of technology, which has a direct impact on its use measure, which directly affects the actual use of technology.

One example of linking the model of acceptance of innovation and technology with the models of competence in the educational environment to understand the factors that influence the motivation of teachers to use technology in teaching is the conceptual model developed by Babić

[15]. According to this model, there are the following significant categories of factors that motivate teachers to apply innovations in teaching: knowledge, skills, and abilities of teachers, attitudes and values, situational factors, institutional factors, and personal factors.

According to the above conceptual model [15] potential factors that can influence the application of completely free *of charge unplugged* MEMA methods in teaching by teachers anywhere in the world are:

- Knowledge of the MEMA method, the skills of its application in teaching, and the ability of teachers to transfer them to the teaching process;
- Positive teaching attitudes about the application of MEMA in teaching with an understanding of the values of the same in teaching;
- Situational factors: characteristics of students and the subject within which the MEMA method is applied (computer science, mathematics, Croatian language, nature, ...)
- Personal factors of the teacher: personal characteristics, innovation, creativity, and self-efficacy.
- Institutional factors: the education system at all levels and possibilities of applying the MEMA method in teaching (time, incentives for effort, praise, the possibility of advancement, ...).

Within the category "situational factors", the role of student characteristics that can influence the use of innovation in the teaching process is emphasized. The practice has shown that the way students accept the MEMA method motivates teachers to design creative tasks using the MEMA method [3].

IV. Conclusion

Although the MEMA method for early programming learning for young children has been proven as useful, free of charge and easy to use, practice shows that it has not been used in teaching on a larger scale so far.

Therefore, this paper discussed the motivation of teachers to accept this method, which is considered a complex process that requires additional research on a larger sample of respondents. On the basis of the existing models of acceptance of innovations in this paper, framework factors of motivation are given that can help in further research of teacher motivation for application of unplugged MEMA method for early programming learning for young children.

The ERASMUS+ project "CybeMEMA: Cybernetics in early childhood education – MEMA as a teaching method" (2022-1-PL01-KA220-SCH-000086129) is currently being implemented, involving partners from Croatia, Poland and Turkey. Within the project, the method will be available in four languages (Croatian, English, Polish, Turkish). In partner countries, the MEMA method will be applied and research will be conducted on its usefulness for learning and teaching programming for young children, involving a larger number of teachers and students.

Acknowledgment

The paper was supported by the ERASMUS+ project "CybeMEMA: Cybernetics in early childhood education – MEMA as a teaching method" (2022-1-PL01-KA220-SCH-000086129).

References

[1] https://nova-akropola.com/covjek-i-svijet/aktualno/pismenost-u-svijetu-i-hrvatskoj/ (accessed 26.2.2023)

[2] M. Čičin-Šain, “Mema - didaktičko sredstvo za simuliranje rada glavne memorije elektroničkog računala u nižim razredima osnovne škole”, Zbornik Pedagoškog fakulteta u Rijeci, 1981.

[3] S. Babić, D. Bjelanović and M. Čičin-Šain, “Programming and Mathematics through Game”,*44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 870-874, 2021. doi: 10.23919/MIPRO52101.2021.9597117.

[4] M. Čičin-Šain, P. Radulović and S. Babić,“Računalno razmišljanje i tablica množenja”, Proceedings of the *45th International Convention on Information and Communication Technology*, Electronics and Microelectronics (MIPRO), p.p. 884-887, Opatija, 2022.

[5] M. Čičin-Šain, S. Babić, „Početnica Mema“, Bilten of Cybernetics society Rijeka, rujan 2014.

[6] M. Čičin-Šain, “Utjecaj predznanja na usvajanje novih informatičkih sadržaja u osnovnoj školi”, Simpozij "Kompjutor na sveučilištu", Cavtat, pp. 1R031-1R034, 1987.

[7] E. H. Crosby, “Tolstoy as a Schoolmaster”, available at: http://playpen.meraka.csir.co.za/~acdc/education/Dr_Anvind_Gupa/Learners_Library_7_March_2007/Resources/books/tolstoy.pdf (accessed: 22.2.2023)

[8] Cybernetics society Rijeka, available at: http://drustvo-kiberneticara.hr/ (accessed: 22.2.2023)

[9] S. Babić, J. Mezak, and M. Čičin-Šain, “Stavovi studenata prema uvođenju računalnog razmišljanja i programiranja pomoću metode Mema za djecu predškolskog i ranog školskog uzrasta”, *Proceedings of the 40th International Convention on Information and Communication Technology, Electronics and Microelectronics* / Biljanović, Petar (ur.). Croatian Society for Information and Communication Technology, Electronics and Microelectronics, pp. 1144-1149, 2017.

[10] M. Čičin-Šain, S. Babić**, “**Zaista prvi koraci u programiranju - radionica unutar aktivnosti All Digital Week”, 41st International Convention on Information and Communication Technology, Electronics and Microelectronics / Skala, Karolj (ur.). Opatija, pp. 980-983, 2018.

[11] E. M. Rogers: Diffusion of Innovations, 4th ed. New York, Free Press, 1995.

[12] E. L. Deci, R. M. Ryan, „Intrinsic motivation and self-determinationin human behavior“, New York: Plenum, 1985.

[13] R. J. Vallerand, „Deci and Ryan's self-determination theory: A view from the hierarchical model of intrinsic and extrinsic motivation“, Psychological Inquiry, vol. 11, no. 4, pp. 312-318, 2000.

[14] F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Quarterly, vol. 13, no. 3, pp. 319-340, 1989.

[15] S. Babić, „The Factors of Teacher's Acceptance of E-learning and Competence for its Implementation at Higher Education Institutions“, Doctoral dissertation, University of Zagreb. Faculty of Organization and Informatics Varaždin, 2016.

Formativno vrednovanje u online okruženju nastave matematike

R. Soldo* i D. Kurtić **
* Strojarska tehnička škola Fausta Vrančića, Zagreb, Hrvatska
** Agencija za odgoj i obrazovanje, Zagreb, Hrvatska
ruzica.soldo@skole.hr, daria.kurtic@azoo.hr

Sažetak

U skladu sa situacijom uzrokovanom pandemijom Covid-19, dolazi do potrebe organiziranja nastave u virtualnim učionicama, koje postaju mjesto uvođenja novih obrazovnih strategija. Nastavnici moraju rekonceptualizirati temeljna pitanja poučavanja, učenja i ocjenjivanja u netradicionalnim, virtualnim, učionicama, suočeni su s mnogim problemima izvođenja i organiziranja nastavnog procesa. Posebno do izražaja dolazi problem vrednovanja, dijelom zbog sumnje u integritet učenika ali i zbog samog izvođenja procesa vrednovanja u *online* okruženju. Ovaj rad ima za cilj ukazati na prednosti i nedostatke formativnog vrednovanja, mogućnosti njegovog integriranja u virtualno okruženje te prikazuje neke metode formativnog vrednovanja provedene u virtualnoj učionici nastave matematike.

Ključne riječi: formativno vrednovanje, virtualna učionica, nastava matematike

I. UVOD

Pandemija Covid-19 je uzrokovala povremeno preusmjeravanje obrazovanja u virtualne učionice te donijela promjene cijelom svijetu u svim sferama života. Kontaktna nastava licem u lice je postala virtualna nastava, sinkrona ili asinkrona, stoga se vrednovanje i ostale sastavnice nastavnog procesa u novim uvjetima poučavanja preispituju i traže se najprikladnije pedagoške metode poučavanja i vrednovanja u online okruženju. Učitelji predmetne nastave osnovne škole i nastavnici srednje škole su organizirali nastavu kroz virtualne učionice u nekom od sustava za komunikaciju i suradničko učenje: *Moodle, Yammer i Microsoft Teams*. Kroz sustav e-učenja potrebno je redovito provoditi vrednovanje aktivnosti učenika i radove domaćih zadaća [1], suradničkog učenja, a formativno vrednovanje može se provoditi kroz diskusije, testove, didaktičke igre, kvizove, istraživačke radove, projekte i prezentacije [2].

A. Vrednovanje

Vrednovanje ili evaluacija je „*sustavni proces prikupljanja, analize i interpretacije informacija o stupnju ostvarivanja ciljeva odgoja i obrazovanja te ciljeva nastave*“[3]. Ocjene su izrazi mjera ostvarenosti i usvojenosti zadataka, a zasnivaju se na podacima do kojih je nastavnik došao ispitivanjem, praćenjem i vrednovanjem. Razvijanjem vještina ocjenjivanja učenika, pruža se široki raspon povratnih informacija i smanjuje opterećenje predavača [4]. Ocjenjivanje je predmet dokimologije, didaktike, pedagogije, psihologije i posebno metodike pojedinih predmeta. Ocjenjivanje je prema Enciklopedijskom rječniku pedagogije [5] postupak kojim se, po utvrđenim propisima, prati odgojno -obrazovni rad i razvoj učenika te određuje stupanj koji je učenik postigao. „*Vrednovanje je sustavno prikupljanje podataka u procesu učenja i postignutoj razini ostvarenosti odgojno-obrazovnih ishoda, kompetencijama, znanjima, vještinama, sposobnostima, samostalnosti i odgovornosti prema radu, u skladu s unaprijed definiranim i prihvaćenim metodama i elementima. Vrednovanje obuhvaća tri pristupa vrednovanju: vrednovanje za učenje, vrednovanje kao učenje, vrednovanje naučenog*“[6].

B. Formativno vrednovanje

Naziv *formativno vrednovanje* počeo se pojavljivati u obrazovnoj literaturi od 1967. godine kad je Michael Scriven prvi put upotrijebio riječ "formativno". Scriven je posebno govorio o razvoju školskog kurikuluma i naglasio da formativno vrednovanje uključuje prikupljanje odgovarajućih dokaza tijekom razvoja i primjene novog kurikuluma na takav način da se revizija kurikuluma može temeljiti na tim dokazima [7]. Formativno vrednovanje prvi put se definira 1971. godine kao upotreba sustavnog ocjenjivanja u procesu izgradnje nastavnog plana i programa, podučavanja i učenja u svrhu poboljšanja bilo kojeg od navedena tri procesa [8]. To je praćenje kvalitete učenikova napretka tijekom cijele godine i dinamike rada učenika. Učitelj može ciljano djelovati na područja nastavnog gradiva kojima je potreban dodatni ili drugačiji rad kako bi pomogao učenicima prebroditi poteškoće u učenju [9], čime spoznaje načine rada koji su najprikladniji za učenike i otkriva njihove talente, a ispravlja slabosti i nedostatke koje usporavaju proces učenja. Učitelj kroz proces formativnog vrednovanja preciznom povratnom informacijom savjetuje i vodi učenika prema željenim postignućima. Formativno možemo vrednovati učenike na dva načina, *vrednovanje za učenje* i *vrednovanje kao učenje*. „*Vrednovanje za učenje služi unapređivanju i planiranju budućega učenja i poučavanja. Vrednovanje kao učenje podrazumijeva aktivno uključivanje učenika u proces vrednovanja te razvoj učeničkoga autonomnog i*

samoreguliranog pristupa učenju. Vrednovanje za učenje i vrednovanje kao učenje ne rezultiraju ocjenom, nego kvalitativnom povratnom informacijom"[6]. Povratna informacija formativnog vrednovanja je kvalitativna te kao takva daje detaljne upute učenicima što trebaju poboljšati, kako bi ostvarili zadovoljavajuće odgojno-obrazovne ciljeve. Formativno vrednovanje daje informaciju o stvarnom stanju učenika pomoću koje se može prilagođavati proces učenja u cilju postizanja očekivanih ishoda, prilagodbom nastavnih metoda te individualiziranim pristupom učenicima [10]. Ovakvo vrednovanje je nenametljivo, integrirano u nastavni proces, ne izaziva stres kod učenika te ih motivira za budući rad.

II. METODE FORMATIVNOG VREDNOVANJA

Vrednovanje u *online* okruženju se sve češće izvodi kao formativni pristup ocjenjivanju, umjesto davanja tradicionalnih testova na kraju poglavlja [11]. Pregled dosadašnjih istraživanja pokazuje kako interaktivno online formativno vrednovanje stavlja fokus na učenika te poboljšava angažman učenika u nastavnom procesu [12]. Prelazak na asinkrono i sinkrono mrežno učenje traži nove načine kako iskoristiti najbolje od postojećih alati za procjenu. Leahy, Lyon, Thompson i Wiliam predstavili su pet pristupa ocjenjivanju u bilo kojoj učionici, tradicionalnoj ili virtualnoj:

1. Upoznati učenike s ishodima učenja i kriterijima vrednovanja
2. Stvaranje učinkovitih rasprava u učionici
3. Pružanje povratnih informacija koje učenike pokreću naprijed.
4. Potaknuti kod učenika odgovornost za vlastito učenje kroz proces samovrednovanja učenika.
5. Uključiti učenike u vršnjačko vrednovanje [13].

Kroz ove preporuke posebna se pozornost stavlja na formativno vrednovanje. Brojne su nove mogućnosti za formativno ocjenjivanje koje nastavnicima omogućava brzo uočavanje značajnih problema i prilagođavanje nastave temeljnim potrebama učenika. U nastavku će se objasniti dva načina formativnog vrednovanja, vrednovanje kao učenje i vrednovanje za učenje.

A. *Vrednovanje kao učenje*

„*Vrednovanje kao učenje podrazumijeva aktivno uključivanje učenika u proces vrednovanja*" [6]. Motiviranjem učenika za rad i njihovim poticanjem za aktivno sudjelovanje u nastavi dolazi do poboljšanja učeničkih postignuća. Učenici kod ovakvog načina vrednovanja preuzimaju odgovornost za vlastito učenje, čime se povećava njihova unutarnja motivacija, što u konačnici vodi prema većem uspjehu učenika [14] i [15]. Vrednovanje kao učenje može se provoditi online testovima korištenjem digitalnih alata Quizlet, Socrative, Google Forms, Hot Potatoes, Kahoot i brojnih drugih.

Prednosti vrednovanja kao učenje u digitalnom okuženju je u mogućnosti pružanja učenicima neograničeni broja pokušaja rješavanja testova ukoliko nisu zadovoljni svojim rezultatima. Primjenom interaktivnih materijala, kvizova i pitalica u nastavi omogućuje se češće provođenje vrednovanja kao učenje uz veću kvalitetu i manje uloženo vrijeme nastavnika [16]. Vrednovanje kao učenje se provodi učeničkim samovrednovanjem ili međusobnim vršnjačkim vrednovanjem, a sve to u cilju boljih postignuća učenika.

B. *Samovrednovanje*

U nastavi se često koriste *online* testovi za samovrednovanje znanja, koje učenici mogu ponavljati više puta i tako kroz proces pokušaja i pogrešaka savladati određene nastavne sadržaje. Istraživanja su pokazala kako je kontinuirano formativno samovrednovanje učenika pozitivno povezano s njihovim postignućima i motivacijom [14]. Samovrednovanje prikazuje učenicima refleksiju vlastitog rada, prikazuje mogućnosti i upozorava na eventualne nedostatke koje treba popraviti. Samoprocjena je ključni proces za pomoć učenicima kako bi razumjeli i shvatili djelovanje i preuzeli odgovornost za svoje učenje. Učitelji u virtualnim učionicama upućuju učenike na korištenje elektroničkog portfolia te sve više koriste *online* testove za samovrednovanje znanja (*Hot Potatoes, Kahoot, Quizzis*). Elektronički portfolio je mapa koja uz primjenu digitalne tehnologije prikuplja dokaze o znanjima i vještinama ispitanika. Elektronički portfolio podržava podatke prikupljene u raznim formatima poput: tekstualnih, grafičkih, vizualni, zvučnih i audiovizualni zapisa [17]. Elektronički portfolio je bitan alat za formativno vrednovanje učenika [18]. Uz korištenje e-portfolia potiču se učenici da stvarajući svoje mape potaknu vlastito kritičko mišljenje i sami vrednuju svoje znanje i postignuća. Postoje mnogi komercijalni alati primjenjivi za izradu e-portfolia kao *FrontPage* i *Dreamweaver*. U radu s učenicima preporuka je koristiti besplatne alate za izradu e-portfolia poput *WordPressa* i *Google Sitesa* [19]. Učenički e-portfolio uz prikupljene dokumente, tekstove, multimedijske sadržaje, sadržava i povratne informacije o uspjehu [17]. E-portfolio je privatan prostor učenika, koji mu daje uvid u vlastiti napredak učenja. Učenik se može odlučiti neke dijelove e-portfolia podijeliti s prijateljima i nastavnicima, čime se osigurava rad u skupinama i potiče socijalizacija. Loša računalna pismenost može učenike obeshrabriti prilikom izrade e-portfolia, dok samoprocjena o vlastitim postignućima može uspješnije učenike učiniti zadovoljnima, a slabije razočarati [20].

C. *Vršnjačko vrednovanje*

Kroz vršnjačko vrednovanje učenicima se pruža mogućnost evaluiranja kolega i praćenja svojih rezultata. Prema istraživanjima [21]. Aautori podržavaju vrijednost suradnje s drugim učenicima u internetskoj nastavi. "Učenici kao partneri sa svojim nastavnicima kontinuirano prate trenutni nivo vlastitog postignuća u odnosu na dogovorena očekivanja kako bi mogli postaviti ciljeve što učiti dalje i upravljati vlastitim napretkom

"[22]. Black i Wiliam ističu u svom istraživačkom radu kako se značajni pozitivni učinci na uspješnost učenika u nastavi matematike mogu postići formativnim vrednovanjem stavljajući naglasak na vršnjačko vrednovanje. S obzirom da je dobivanje povratnih informacija složen proces te zahtjeva određene vještine, učenike se kontinuirano podučavalo pomoću uputa i prakse kako bi razvijali vještine preispitivanja tuđeg rada na formativan način [23]. Učitelj u cijelom procesu ima bitnu ulogu i donosi zaključnu formativnu procjenu nakon što učenici naprave međusobne formativne procjene. Pružaju se nove mogućnosti za formativno vrednovanje koje nastavnicima omogućuje brzo uočavanje nedostataka i prilagođavanje nastave trenutnim potrebama učenika na temelju njihovih potreba. Primjenom zajedničkih digitalnih dokumenata *Google* diska, virtualnih ploča *Linoit i Microsoft Whiteboard* te sličnog, u matematici je omogućeno učenicima stvaranje i rješavanje matematičkih problema pisanom suradnjom i međusobnom komunikacijom. To donekle nadoknađuje socijalnu interakciju među učenicima koju je nudila nastava licem u lice [24]. Vršnjačko vrednovanje u virtualnoj učionici nastave matematike primjenom virtualne ploče *Linoit* uspješno je odrađeno na način da svaki učenik postavi svoje rješenje zadatka na virtualnu ploču, kako bi ga mogao vrednovati drugi učenik. Učenici su uz točnost rješenja trebali napisati komentare i poticajnu poruku za učenika čiji rad vrednuju (Slika 1).



Slika 1. Primjer vršnjačkog vrednovanja na digitalnoj ploči Linoit

D. *Vrednovanje za učenje*

Digitalna tehnologija potiče proces učenja. Pomoću digitalne tehnologije nastavnici lakše i brže formativno prate postignuća učenika kroz stalne aktivnosti u virtualnoj učionici. Neki od alata koje pružaju kontinuiranu formativnu procjenu u internetskom kontekstu vrednovanja i učenja su: *Kahoot, Quizzis, GeoGebra, Testmoz, Mentimeter, LearningApps, Google Forms, Genially* te mnogi drugi. Poznatiji alati za izradu umnih mapa su: *Coggle, Popplet, MindMap i Bubbl.us*. Koristeći digitalnu tehnologiju učenicima se puža mogućnost učenja kroz igru i lakše savladavanje nastavnih sadržaja na zabavan način. Uvođenjem digitalne tehnologije u nastavi veći se naglasak daje procesu učenja, takozvanom vrednovanju za učenje, kako bi se postigli što bolji i kvalitetniji ishodi učenja.

E. *Kvizovi u online okruženju*

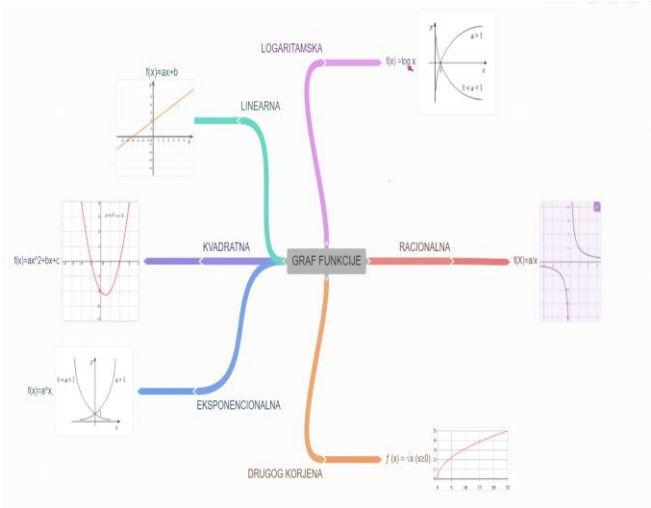
Studije su pokazale poboljšano i aktivnije sudjelovanje učenika u nastavi i učenju uz korištenje kratkih kvizova ([25]; [26]; [15] i [27]). Oni ne pružaju samo informacije učiteljima, već i povećavaju odgovornost učenika. Imaju ograničene mogućnosti unosa i ne mogu primati složeni tekst, pa ne služe za procjenu viših razina razumijevanja jer ne daju nikakav uvid u razmišljanja učenika i razlog njihovog razumijevanja [28]. Kroz nastavu matematike kvizovi su se koristili na početku ili kraju nastavnog sata, da bi se ponovili ključni pojmovi i odnosi, kako bi se učenicima olakšalo razumjevanje budućih sadržaja te stvorila veza s prethodno naučenim sadržajima.

F. *KWL tablica (**K**now-znam, **W**ant to know-želim znati, **L**earned-naučio sam)*

Autorica ove metode formativnog vrednovanja, nastale 1986. godine je Donna Olge. Ona opisuje KWL tablicu kao metodu koja učenika aktivira da poveže prethodno naučeno gradivo s gradivom koje se obrađuje [29]. Učenici na početku sata samostalno zapisuju što znaju o temi koja će se obrađivati, tako da odgovaraju na pitanja u prva dva stupca u KWL tablici. Nakon obrade teme, učenici popunjavaju treći stupac KWL tablice i odgovaraju na pitanje: „*Što smo naučili o temi*?" Primjena KWL tablice u virtualnim učionicama se može ostvariti kroz *Microsoft Whiteboard* ponuđene alate formativnog vrednovanja.

G. *Umne mape*

Umna mapa je vizualni prikaz hijerarhijskih informacija koji uključuje središnju ideju okruženu granama povezanih tema, to je dijagram koji se koristi za vizualno organiziranje informacija [30]. Metoda stvaranja umne mape posebno potiče kreativnost učenika a također je dobro prihvaćena metoda kod učitelja i nastavnika. Umne mape pružaju uvid u proces razmišljanja učenika vezano za određenu temu i razumijevanje određenog pojma. Brojna istraživanja daju dokaze o učinkovitosti umnih mapa u području razmišljanja, učenja, poučavanja, pamćenja, kreativnosti, planiranja, bilježenja, komunikacije i rješavanja problema ([31]; [32] i [33]). Pomoću umnih mapa i digitalnih alata *online* okruženje postalo je zanimljivije i zabavnije učenicima. Poznati digitalni *web* alati za izradu umnih mapa su: *College, Popllet, MindMeister* i drugi. Koristeći dane aplikacije učenici su umnim mapama prikazali zorno matematičke funkcije kroz virtualnu nastavu matematike (Slika 2).



Slika 2. Microsoft Whiteboard kao suradnički alat sa Coggle aplikacijom za izradu umnih mapa-rad učenika

Koristeći digitalne obrazovne tehnologije učitelji i nastavnici mogu formativno vrednovanje jednostavno uključiti u svaki nastavni sat, čime se poboljšava kvaliteta nastave. Nadalje, učenici i nastavnici na ovaj način dobiju povratnu informaciju koji su ishodi nastavnih sadržaja dobro usvojeni, a koje područje nastavnih sadržaja treba ponoviti i poboljšati. Kroz dobro odabrane tehnike i alate, formativno vrednovanje će dati željene učinke u korist usvajanja ishoda i poticanja motivacije učenika uz njihovo aktivno sudjelovanje u nastavnom procesu. Mrežni sustavi formativnih povratnih informacija koji su integrirani u internetski prostor za učenje poboljšavaju angažman i izvedbu učenika ([34]; [35] i [36]).

III. ZAKLJUČAK

Vrednovanje u *online* učionici je složen process. Sumativno vrednovanje često dovodi u pitanje valjanost i pouzdanost ocjenjivanja u internetskom okruženju [37]. S ciljem da dobiju željenu ocjenu, prema mnogim istraživanjima, učenici posežu za nedopustivim metodama varanja na testovima [38] te se postavlja pitanje autentičnosti njihovih radova. Uvođenjem formativnog vrednovanja u *online* učionice izbjegava se navedeni problem jer formativno vrednovanje ne rezultira ocjenom. Primjenom formativnog vrednovanja učenici postaju aktivni sudionici nastavnog procesa, samostalno rade radove, bez straha od loših ocjena i negativnih ishoda vrednovanja ispita. Vršnjačko vrednovanje i samovrednovanje kroz prikladne online alate potiče sudjelovanje učenika u nastavnom procesu. Na taj način učenik je stavljen u središte poučavanja, a učitelj ili nastavnik koordinira i kontrolira cijeli proces vrednovanja i daje konačne rezultate. Kroz formativno vrednovanje učenici dobiju konkretan uvid u vlastiti rad, objašnjenje provedenog rada, što im pruža mogućnost lakšeg razumijevanja područja u kojima trebaju dodatan rad kako bi postigli bolje rezultate. U ovom su radu pokazane samo neke od brojnih metoda formativnog vrednovanja primjenjivih u virtualnoj učionici, kao poticaj učiteljima i nastavnicima da provode formativno vrednovanje u svojoj nastavi koje je obvezno sukladno Pravilniku o izmjenama i dopuni pravilnika o načinima, postupcima i elementima vrednovanja učenika u osnovnim i srednjim školama [6].

REFERENCES

[1] M. Bulić I V. Kostović Vranješ (2019). Utjecaj e-učenja na samoodgovornost učenika pri izvršavanju domaćih zadaća. *Školski vjesnik : časopis za pedagogijsku teoriju i praksu*, vol. *68*, no. pp. 112–126, 2019.

[2] Ministarstvo znanosti i obrazovanja. (2020). *Upute za vrednovanje i ocjenjivanje tijekom nastave na daljinu.* Retrieved from https://mzo.gov.hr/UserDocsImages/dokumenti/Obrazovanje/Upute-zavrednovanje/Upute%20za%20vrednovanje%20i%20ocjenjivanje%20tijekom%20nastave%20na%20daljinu.pdf?fbclid=IwAR3hTiwCR87KUwysOJ0yvhxHrI42u7oTsh7ReG2cza_pVp53tZJRl4vd4b0.

[3] M. Matijević, V. Bilić I S. Opić, *Pedagogija za učitelje i nastavnike*. Zagreb, Školska knjiga, 2016.

[4] T. Anderson, *The Theory and Practice of Online Learning*. Athabasca University Press.2008.

[5] D. Franković, *Enciklopedijski rječnik pedagogije.* Zagreb: Matica hrvatska. 1963.

[6] Narodne novine 2019. Pravilnik o izmjenama i dopuni Pravilnika o načinima, postupcima i elementima vrednovanja učenika u osnovnim i srednjim školama (NN 82/19). Retrieved February 6, 2021, from https://narodne-novine.nn.hr/clanci/sluzbeni/2019_09_82_1709.html

[7] R.W.Tyler, R. M. Gagné i M. Siriven, *Perspectives of curriculum evaluation*. Chicago: Rand McNally, 1976.

[8] B. S. Bloom, J.T. Hastings i G.F. Madaus, "*Handbook On Formative and Summative Evaluation of Student Learning*", New York: McGraw-Hill, 1971.

[9] C. Kivunja, Why Students Don't Like Assessment and How to Change Their Perceptions in 21 st Century Pedagogies. *Creative Education*, vol. 6, no. 20, pp. 2117–2126, 2015.

[10] V. Rajić, Pristupi vrednovanju u obrazovanju. "*School for the Net generation*," pp. 256–276. 2017.

[11] V. Tinto, Taking Student Retention Seriously: Rethinking the First Year of University, 2009.

[12] J. W. Gikandi, D. Morrow I N.E. Davis, Online formative assessment in higher education: A review of the literature. *Computers & Education*, *vol. 57, no. 4, pp*. 2333–2351, 2011.

[13] S. Leahy, C. Lyon, M. Thompson i D. Wiliam, *Classroom Assessment: Minute by Minute, Day by Day*. 5, 2005.

[14] L. Brusać, J. Dadić i T. Kisovar-Ivanda, Učeničkim samovrednovanjem do kvalitetnih učeničkih postignuća. *Magistra Iadertina*, vol. 11, no 1, 2017.

[15] P. S. C. Goh i B. Matthews, Concerns of Student Teachers: Identifying Emerging Themes Through Self-Assessment. In M. M. C. Mok (Ed.), *Self-directed Learning Oriented Assessments in the Asia-Pacific* . Dordrecht: Springer Netherlands. pp. 297–310, 2013.

[16] G. Jugo, I. Matotek, M. Carev i D. Domović, Uporaba Moodle-a 2.0 u vrednovanju znanja. *Medijska Istraživanja: Znanstveno-Stručni Časopis Za Novinarstvo i Medije*, vol. 18, no. 1, 153–163, 2012.

[17] S. Kučina-Softić, *Što je e-portfolio? E-portfolio kao nastavna aktivnost – priručnik za nastavnike*. Sveučilišni računski centar Srce, Svučilište u Zagrebu, 2013.

[18] H. Barrett i J. Carney, Conflicting paradigms and competing purposes in electronic portfolio development. *TaskStream Web Site*, pp. 295–314, 2005.

[19] K. Himpsl i P. Baumgartner, Evaluation of E-Portfolio Software. *International Journal of Emerging Technologies in Learning (IJET)*, vol 4. no. 1, pp. 16–22. 2009.

[20] S. Jarrott i L.E. Gambrel, The Bottomless File Box: Electronic Portfolios for Learning and Evaluation Purposes. 10. 2011.

[21] Q. Wang, W. Chen i Y. Liang, The Effects of Social Media on College Students, 2011.

[22] R. Stiggins, From Formative Assessment to Assessment for Learning: A Path to Success in Standards-Based Schools. *Phi Delta Kappan*, vol. *87*, no. 4, pp. 324–328, 2005.

[23] P. Black i D. Wiliam, Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, vol. *5*, *no.* 1, pp. 7–74, 1998.

[24] S. Kleinman, Strategies for Encouraging Active Learning, Interaction, and Academic Integrity in Online Courses. *Communication Teacher*, vol. 19, no. 1, pp. 13–18, 2005.

[25] Beatty, I. D., & Gerace, W. J. (2009). Technology-Enhanced Formative Assessment: A Research-Based Pedagogy for Teaching Science with Classroom Response Technology. *Journal of Science Education and Technology*, *18*(2), 146–162.

[26] Bennett, K. R., & Cunningham, A. C. (2009). Teaching Formative Assessment Strategies to Preservice Teachers: Exploring the Use of Handheld Computing to Facilitate the Action Research Process. Journal of Computing in Teacher Education, 25(3), 99–105.

[27] Peat, M., & Franklin, S. (2002). Supporting Student Learning: The Use of Computer-Based Formative Assessment Modules. *British Journal of Educational Technology*, *33*(5), 515–523.

[28] Koç, S. (Ed.). (2015). *Assessment in online and blended learning environments*. Charlotte, NC: Information Age Publishing.

[29] Keeley, P. D. (Ed.). (2008). Science Formative Assessment: 75 Practical Strategies for Linking Assessment, Instruction, and Learning (1st edition). Thousand Oaks, CA: Corwin.

[30] Rajapriya, M. (2017). EFFECTIVENESS OF MIND MAPPING IN HIGHER EDUCATION. 7.

[31] Cvetkoska, V. (2017). Student Perceptions Regarding the Mind Map Application in Mathematical Education. *Proceedings of the ENTRENOVA - ENTerprise REsearch InNOVAtion Conference*, *3*(1), 85–91.

[32] Farrand, P., Hussain, F., & Hennessy, E. (2002). The efficacy of the "mind map" study technique. *Medical Education*, *36*(5), 426–431. https://doi.org/10.1046/j.1365-2923.2002.01205.x

[33] Uchenna, E., & Ed, M. (2009). EFFECT OF MIND MAPS ON STUDENTS' INTEREST AND ACHIEVEMENT IN MEASURES OF CENTRAL TENDENCY IN MATHEMATICS. 123.

[34] C. M. Chen i M. C. Chen, Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers & Education*, vol. 52, no. 1, pp. 256–273, 2009.

[35] T. Hatziapostolou i I. Paraskakis, Enhancing the Impact of Formative Feedback on Student Learning Through an Online Feedback System, vol. 8, no. 2, 2010.

[36] T. Van Gog, D. M. A. Sluijsmans, D. Joosten-ten Brinke i F. J. Prins, Formative assessment in an online learning environment to support flexible on-the-job learning in complex professional domains. *Educational Technology Research and Development*, vol. *58*, no. 3, pp. 311–324, 2010.

[37] J. W. Gikandi, D. Morrow i N. E. Davis, Online formative assessment in higher education: A review of the literature. *Computers & Education*, vol. 57, no.4, pp. 2333–2351, 2011.

[38] R. M. Palloff i K. Pratt, *Assessing the online learner: Resources and strategies for faculty* (1st ed). San Francisco: Jossey-Bass, 2009.

Modeliranje linearnom funkcijom u osnovnoj školi

Ivan Nađ
Visoko učilište Algebra, Zagreb, Hrvatska; Fakultet hrvatskih studija, Sveučilište u Zagrebu, Zagreb, Hrvatska;
Osnovna škola Eugena Kvaternika, Velika Gorica, Hrvatska
ivan.nadjzg@gmail.com ; inad@racunarstvo.hr ; inad@hrstud.hr

***Sažetak* - Funkcija je jedan od temeljnih pojmova u matematici. Učenici se već u osnovnoj školi susreću s funkcijama, ponajviše u sedmom razredu i to s linearnom ovisnosti, tj. linearnom funkcijom. Linearna funkcija je obrađena na način na koji se učenici na nastavi susreću s njom, a dani su i različiti primjeri zadataka iz svakodnevnog života koji se mogu modelirati linearnom funkcijom. Pritom se pri modeliranju i analizi koristi alat dinamične geometrije Geogebra.**

Ključne riječi - linearna funkcija; linearna ovisnost; modeliranje; Geogebra

I. FUNKCIJA I MODELIRANJE

Funkcija je jedan od temeljnih pojmova u matematici. U osnovnoj školi, odnosi se na pridruživanje jedne veličine drugoj po nekom pravilu ili pak na ovisnost jedne veličine o drugoj. S obzirom da je pojam (linearne) funkcije nekad teško objasniti učenicima, u osnovnoj školi se ostaje na pojmu (linearne) ovisnosti, tj. ovisnost jedne veličine $f(x)$ o drugoj veličini x.

Konkretnije, što se linearne ovisnosti tiče, to je, prema [1], pridruživanje koje nekom realnom broju x pridružuju realni broj $f(x)$ takav da je $f(x) = a \cdot x + b$. Iako se tako ne imenuju, učenici se s različitim pridruživanjima, tj. funkcijama, učenici susreću već od nižih razreda, a da toga nisu niti svjesni. Primjerice, računanje opsega kvadrata ako je zadana duljina stranice, razni zadaci iz svakodnevnog života kao što su cijene artikala, plaće i slično.

Već ovi nabrojani primjeri su zapravo vrlo prikladni primjeri matematičkog modeliranja, odnosno, prema [1], pronalaska i korištenja matematičkog modela, algoritma ili formule za rješavanje nekog realnog, tj. svakodnevnog problema.

II. ZAŠTO I KAKO MODELIRANJE?

S matematičkim modeliranjem važno je krenuti s učenicima što ranije. Tako učenicima približavamo matematiku kao nešto što nije samo apstraktno, već i nešto što svakodnevno koriste. Učenici tako imaju priliku primijeniti stečena matematička znanja i vještine pri rješavanju nekog stvarnog problema, pogotovo ako su ti problemi njima poznati i bliski, što zasigurno osigurava veću motivaciju za rad.

To je posebno bitno ako pogledamo neka istraživanja koja su analizirala stav učenika prema matematici u osnovnoj školi. Benček, Marenić [2] su 2006. godine na 374 učenika osmih razreda, deset osnovnih škola u gradu Zagrebu ustvrdile da se 81.6% učenika na nastavi matematike dosađuje. Tek 15% učenika je reklo da matematiku uče radi zanimljivosti sadržaja. Uključivanje učenicima zanimljivih sadržaja i njima bliskih problema, a što je kroz modeliranje vrlo lako napraviti, je stoga ključno za motivaciju učenika, a na kojoj leži angažman te, u konačnici, i uspjeh u matematici.

Provesti modeliranje na nastavnom satu može biti vrlo zahtjevno, no ne treba od njega odustati, baš zbog spomenute motivacije. Često sam i sam svjedok bezvoljnosti i nemotiviranosti učenika, pogotovo ako se nastava svede na monotono slušanje i prepisivanje s ploče. Nadalje, nastavnici su često opterećeni količinom sadržaja koje moraju obraditi s učenicima te se često žale da ih nedostatak vremena sputava u ovakvim procesima. Popriličan utjecaj imaju i aktivnosti razreda koje dodatno smanjuju broj nastavnih sati na matematici, kao što su maturalna putovanja, terenske nastave, projektni dani, razni drugi nenastavni dani u nastavi i slično. Zbog svega toga, jako nam je teško planirati nastavu, dostupan fond sati se smanjuje, često smo u nedostatku vremena pa je prirodan strah od svega što traje duže vrijeme, a tu je onda i modeliranje, pogotovo jer se često ovakvi zadaci i aktivnosti oduže i potraju cijeli školski sat, a ishod je, ako niti razred niti mi kao nastavnici nemamo iskustva s ovakvim oblikom nastave, prilično neizvjestan.

Nastavnici se moraju i jako dobro pripremiti i to ne samo što se tiče materijala, nego i samog procesa. Ako vidimo da nam ishod sata neće biti ostvaren, bitno je znati se prilagoditi i krenuti barem prema djelomičnom ostvarenju ishoda sata i završiti sat.

Iako je, koliko god ih dobro poznavali, nekad teško predvidjeti reakciju učenika, bitno je i pokušati prilagoditi ciljeve i aktivnosti učenicima, njihovom uzrastu i sposobnostima, što je više moguće individualno ili ih podijeliti u homogene grupe. Upute učenicima moraju biti jasne i nedvosmislene, a nastavnik kao vođa nastavnog procesa mora vrlo aktivno usmjeravati učenike pri svakom koraku, pogotovo jer su učenici često skloni

prepustiti posao nekome drugome u grupi. Kao što sam već spomenuo, na početku će zasigurno biti problema, takav rad s učenicima će vjerojatno trajati nešto duže, no s vremenom učenici prihvate ovakvo odvijanje nastavnog procesa te sve ide brže i efikasnije. Za nastavnika je vrlo bitna i samoevaluacija, ali i evaluacija od strane učenika kao mjerilo onoga što se i kako htjelo ostvariti, ali i kao pomoć za buduće planiranje ovakvih aktivnosti.

III. LINEARNA OVISNOST U KURIKULUMU

Učenici se u osnovnoj školi, prema Kurikulumu matematike za osnovne škole [3], prvi put s funkcijom (konkretno linearnom ovisnosti) susreću u 7. razredu. S obzirom da i sam ishod učenja vezan uz linearnu ovisnost naglašava da je riječ o primjeni linearne ovisnosti, naglasak tog sadržaja nije na algoritmima i tehnici rješavanja, nego na povezivanju, učenikovom logičkom razmišljanju i sposobnosti rješavanja problema, ponajviše onda i na modeliranju problema. Preporuča se koristiti i programe dinamične geometrije, od kojih je među najzastupljenijima Geogebra pa će kroz rad biti dani i primjeri kako nam upravo Geogebra može pomoći u tome.

IV. ŠTO JE GEOGEBRA?

Geogebra je dinamični matematički alat koji se može koristiti u različitim kontekstima i na različite načine u nastavi. Najčešće se koristi kao alat dinamične geometrije zbog svoje interaktivnosti i prilagodljivosti. Osim geometrije, često se koristi i za izradu raznih tablica, ali i statistiku i diferencijalni račun. U njoj se mogu kreirati i razni kvizovi i ankete. Nudi i mogućnost kreiranja virtualnih učionica. Vrlo je intuitivna i lako se mogu naučiti njezine najosnovnije funkcije. Ima odličnu podršku s vrlo detaljnim vodičima o mogućnostima korištenja svake od funkcija, ali i forume i online zajednice u kojima možete saznati odgovore na većinu svojih pitanja ili pomoći drugim korisnicima s njihovima. Može se koristiti mrežno kroz gotovo sve preglednike, ali i izvanmrežno jer se nudi i kao besplatni računalni program.

V. LINEARNA OVISNOST U NASTAVI

Učenici najprije kreću sa sadržajima koji se temelje na zapisu $f(x) = ax$. Sami zapis vrlo lako možemo otkriti s učenicima popunjavanjem jednostavnih tablica, kao što su, primjerice, Tablica I. i Tablica II. U Tablici I. učenici uočavaju ovisnost duljine u cm o duljini u metrima i zapisuju kao linearnu ovisnost $f(x) = 100x$. U Tablici II., učenici uočavaju ovisnost dubine na koju podmornica može zaroniti o vremenu te također i tu linearnu ovisnost zapisuju formulom, u ovom slučaju $f(x) = -45x$.

Bitno je s učenicima raspraviti što se čemu pridružuje, primjerice duljini u metrima se pridružuje duljina u centimetrima, a vremenu u minutama se pridružuje dubina u metrima. O tome je bitno raspraviti radi sami linearne ovisnosti, tj. razumijevanja učenika što će biti veličina/varijabla x , a što veličina/varijabla $f(x)$, a što je pak bitno radi same primjene formule i modeliranja njome jednom kada do nje dođemo. Pri spomenu naziva linearne ovisnosti, učenici često pitaju što znači „linearna“ i tu možemo komentirati da dolazi od latinske riječi „linea“ što znači pravac/crta.

Duljina (m)	1	2	3	10	50	200	x
Duljina (cm)	100						

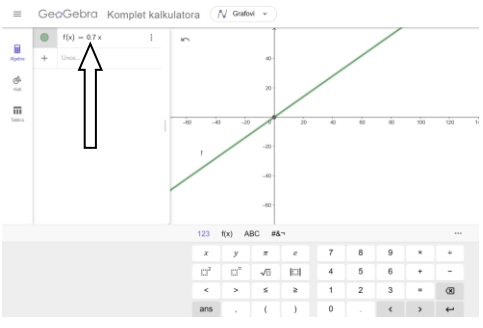
TABLICA I.

Vrijeme(min)	5	1	3	5	8.5	50	x
Dubina (m)	-225						

TABLICA II.

Ovdje je važno s učenicima raspraviti i da je ovdje zapravo riječ o proporcionalnosti dviju veličina, a koju oni u pravilu rade u sedmom razredu prije linearne ovisnosti. Važno je povezati s proporcionalnosti jer se zadaci iz te teme u velikoj mjeri temelje na problemima iz realnog svijeta, a koji su njima bliski pa je prijelaz na zadatke s modeliranjem utoliko lakši.

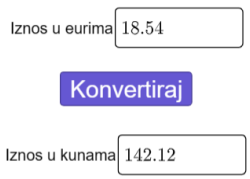
Već je od ovog prvog zapisa bitno s učenicima crtati kako bi učenici i sa samog grafičkog prikaza mogli zaključiti nešto o ovisnosti dviju veličina, primjerice povećava li se jedna veličina povećanjem druge veličine ili se smanjuje, kako nagib pravca utječe na to koliko brzo raste/pada veličina $f(x)$, kako promjena koeficijenta *a* utječe na grafički prikaz i slično. Grafički prikaz također možemo vrlo lako učenicima prikazati u Geogebri. Tu je Geogebra sjajna jer se grafički prikaz automatski prilagođava formuli o kojoj je riječ, a koju je vrlo lako mijenjati kao što je prikazano na Slici 1. Učenici uočavaju da se promjenom koeficijenta a mijenja nagib i da je to njegov utjecaj na grafički prikaz linearne ovisnosti.



Slika 1. Prikaz promjene koeficijenta *a* u alatu Geogebra

Modelirati ovakvim zapisom možemo problem koji je vrlo aktualan u ovo vrijeme, a to je konverzija eura u kune i obrnuto. Učenicima dajemo svakodnevan i njima blizak problem koji će zasigurno biti motivirani rješavati. Možemo krenuti od, primjerice, konvertiranja eura u kune. Najprije ih motiviramo tako da sami predlože

primjere iznosa u eurima koje onda kalkulatorom u Geogebri, kao što je na Slici 2., konvertiramo. U matematici je izuzetno bitna procjena pa ih je dobro prije konverzije pitati za procjenu približne vrijednosti iznosa u kunama. Nakon nekoliko primjera, možemo zajedno s njima doći do formule linearne ovisnosti $f(x) = 7.53450x$.



Slika 2. Kalkulator konverzije u Geogebri

Kako dobar dio učenika želi zarađivati putem društvenih mreža ili kreiranjem i objavom vlastitih sadržaja, zanimljiv bi im sigurno bio i zadatak s računanjem zarade od pregleda videa. Učenike možemo i pitati što misle gdje je tu zapravo zarada (u pregledu reklama). U Geogebri vrlo lako možemo izraditi kalkulatore zarade kao na Slici 3., a koje možemo koristiti prilikom rješavanja zadataka.



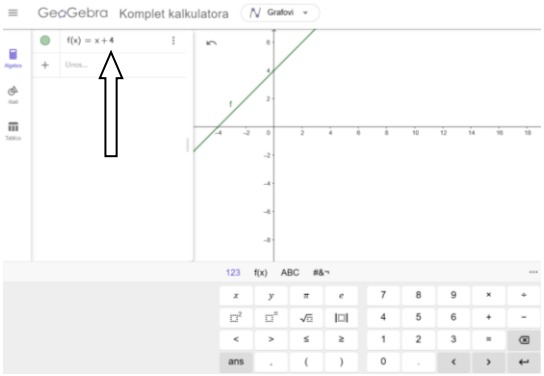
Slika 3. Kalkulator zarade na Youtube-u u kategoriji Gaming, u alatu Geogebra

Nakon zapisa $f(x) = ax$, prirodno je krenuti na $f(x) = x + b$. Do zapisa opet vrlo lako možemo doći popunjavajući tablice kao što je, primjerice, Tablica III. U njoj učenici uočavaju ovisnost iznosa u štednoj kasici o broju proteklih dana (ako svaki dan ubacuje po 1 euro) te zapisuju danu linearnu ovisnost formulom $f(x) = x + 4$.

Broj dana	0	1	2	3	5	20	x
Iznos u eurima	4						

TABLICA III.

Naravno, i ovdje je grafički prikaz bitan kako bi učenici mogli povezati kako promjena koeficijenta b utječe na grafički prikaz. Tu je također Geogebra sjajna kao i u prijašnjem zapisu, a sama promjena zapisa je prikazana na Slici 4. Učenici uočavaju da se promjenom koeficijenta b pomiče gore ili dolje točka u kojoj grafički prikaz presijeca os y te da je to njegov utjecaj na grafički prikaz linearne ovisnosti.



Slika 4. Prikaz promjene koeficijenta *b* u alatu Geogebra

U Geogebri učenici mogu za prethodni primjer sa štednjom izraditi i kalkulator koji bi im prikazivao ušteđeni iznos u odnosu na broj proteklih dana, tjedana, mjeseci ili koju god jedinicu vremena odabrali, pod uvjetom da u svakoj jedinici vremena ubacuju po 1 euro. Primjer takvog kalkulatora je prikazan na Slici 5.



Slika 5. Kalkulator štednje u Geogebri

Učenicima je problematika aktualna jer mogu, primjerice, izračunati koliko bi im jedinica vremena trebalo da uštede za nešto što si žele kupiti. Iako je ovo primjer koji se vrlo lako može i računski riješiti, jako je praktičan za ovladavanje upotrebom Geogebre. I tu je izuzetno bitna uloga nastavnika kao voditelja procesa s ulogom usmjeravanja. Jednom kada izrade ovakav kalkulator, vrlo je lako izraditi kalkulator s uplatama koje nisu samo po 1 euro u jedinici vremena.

Na samom kraju, s učenicima se radi zapis linearne ovisnosti $f(x) = ax + b$. Vrlo prirodno i lako možemo prijeći na ovaj zapis s prošlog ako, već spomenuti prethodni primjer sa štednjom, samo djelomično izmijenimo tako da uplate po jedinici vremena, npr. po danu, više nisu 1 euro nego proizvoljni iznos, primjerice 0.50 eura po danu. Učenici opet mogu popunjavati tablice kao što je Tablica III., no ovaj put primjećuju da je zapis ponešto drugačiji i da glasi $f(x) = 0.50x + 4$. Opet je bitna rasprava, primjerice koliko bi im trebalo da uštede određeni iznos, što da su

uplate tjedne, a ne dnevne, što ako promijenimo koeficijente u zapisu i slično. Sama ta promjena koeficijenata i njen utjecaj na graf se vrlo lako opet može prikazati u Geogebri.

Vrlo česti primjer modeliranja primjenom ovog zapisa je račun cijene taksi prijevoza. Učenicima je dovoljno reći koliko košta početak vožnje, tj. dio koji je u svaki vožnji fiksan, a kolika je cijena po kilometru. Važno je da učenici uoče da je taj fiksni dio, tj. naknada za početak vožnje zapravo koeficijent b, a cijena po kilometru koeficijent a. Učenici mogu isprobati izračunati cijenu vožnje bez korištenja zapisa linearne ovisnosti, ali i korištenjem zapisa. Dobro je ne stati samo na računanju cijene vožnje, nego dodati i, primjerice, pitanje koliku bismo udaljenost mogli prijeći za određeni iznos novca. Kontekst se može promijeniti i ako se promijeni smisao koeficijenta a pa on više ne označava cijenu po kilometru, nego po, primjerice, minuti. S učenicima je onda zanimljivo raspraviti što bi im se više isplatilo u kojem dijelu dana, primjerice za vrijeme velikih gužvi. Zadatak možemo još malo otežati ako u fiksnu naknadu za početak vožnje dodamo i gratis kilometre uključene u cijenu vožnje, ili pak uvećanje cijene u razdoblju povećanje potražnje te dodati raspravu kako bi taj dio utjecao na zapis linearne ovisnosti. Još ako u zadatku iskoristimo stvarne naknade nekog od prijevoznika, učenici mogu koristiti i njihove aplikacije kako bi provjerili točnost svojih rješenja.

Prirodan nastavak toga je i projekt u kojem bi učenici mogli napraviti istraživanje o cijenama svakog od prijevoznika i za svakog od njih doći do zapisa linearne ovisnosti. Možemo uspoređivati i sa cijenama u drugim državama, računati i zaradu vozača, uračunavati troškove vožnje itd., mogućnosti su gotovo pa neograničene.

VI. ZAKLJUČAK

Matematika je često učenicima previše apstraktna i učenici ju često ne razumiju pa se na nastavi, što su i istraživanja pokazala, dosađuju i nedostaje im motivacije za učenje. Stoga je bitno pokušati matematiku što je više moguće približiti učenicima svakodnevnim problemima, problematikom koja je njima zanimljiva. Tu može jako dobro poslužiti matematičko modeliranje jer učenici rješavaju svakodnevne probleme koristeći matematička znanja i vještine. Koristiti možemo i alate dinamične geometrije, kao što je Geogebra, da se odmaknemo od tradicionalnog rješavanja na ploči i u bilježnici. Za takvu nastavu se nastavnik mora dobro pripremiti i očekivati, barem na početku, teškoće s provedbom, no nakon vremena će se učenici priviknuti i modeliranje će biti bolje. Ako već i samo nekoliko učenika iz razreda bude motiviranije za matematiku, cilj je opravdao sredstva.

LITERATURA

[1] https://repozitorij.mathos.hr/islandora/object/mathos%3A413/datastream/PDF/view (preuzeto 30.1.2023.)

[2] A. Benček, M. Marenić, „Motivacija učenika osnovne škole u nastavi matematike“, Školske novine, Zagreb 2006.

[3] Ministarstvo znanosti i obrazovanja, “Odluka o donošenju kurikuluma za nastavni predmet matematike za osnovne škole i gimnazije u Republici Hrvatskoj”, Narodne novine NN 7/2019, Zagreb, 2019.

[4] M.Perić, K.Golubić, D.Vrdoljak, H.Frančešević, S.Davidović , „Linearna funkcija, modeliranje linearnom funkcijom i aritmetički niz” seminar, PMF matematički odsjek, Sveučilište u Zagrebu, mentor : Aleksandra Čižmešija

[5] Bender, E.A., „An Introduction to Mathematical Modeling”, New York: Dover, 1978.

[6] Gershenfeld, N., „The Nature of Mathematical Modeling”, Cambridge University Press, 1998.

Rješavanje jednadžbi u osnovnoj školi uz Geogebru

Ivan Nađ
Visoko učilište Algebra, Zagreb, Hrvatska; Fakultet hrvatskih studija, Sveučilište u Zagrebu, Zagreb, Hrvatska;
Osnovna škola Eugena Kvaternika, Velika Gorica, Hrvatska
ivan.nadjzg@gmail.com ; inad@racunarstvo.hr ; inad@hrstud.hr

Sažetak - **Rješavanje jednadžbi je jedan od izazova koji se često stavljaju pred učenike u nastavi matematike u osnovnoj školi. Računalni alati, kao što je Geogebra, mogu pomoći pri razumijevanju i uvježbavanju rješavanja jednadžbi. U radu je objašnjeno kako koristiti Geogebru za rješavanje jednadžbi, dane su upute za izradu radova i preporuke kako i što kroz te radove prezentirati učenicima.**

Ključne riječi - jednadžbe; Geogebra

I. ŠTO JE GEOGEBRA?

Geogebra je dinamični matematički alat koji se može koristiti u gotovo svakom dijelu nastave matematike, ponajviše kao alat dinamične geometrije zbog svoje interaktivnosti i prilagodljivosti, ali i za izradu proračunskih tablica, statistiku, diferencijalni račun, kreiranje i uporabu virtualnih učionica itd.

Vrlo je intuitivna, svaka naredba sadrži ikonu s crtežom toga što se radi, prevedena je u potpunosti na hrvatski jezik i lako se mogu naučiti njezine najosnovnije funkcije. Znanje programiranja nije potrebno niti za njezine složenije funkcije jer je dostupna odlična podrška s vrlo detaljnim vodičima mogućnosti korištenja svake od funkcija, ali i forumi i online zajednice u kojima možete saznati odgovore na svoja pitanja ili pomoći drugim korisnicima. Može se koristiti mrežno putem svih preglednika, ali i izvanmrežno jer se nudi i kao besplatni računalni program.

Za potrebe ovog rada, korištena je izvanmrežna verzija, no ona se od mrežne razlikuje jedino po rasporedu naredbi na ekranu, funkcionalnosti su apsolutno identične.

II. JEDNAŽBE I PROBLEMI U NASTAVI

S novim kurikulumom [1], linearne jednadžbe se u nastavi matematike ne rade više samo kao zasebna cjelina u 6.razredu, nego su zastupljene vertikalno po svim razinama, od 1. do 8. razreda. U razrednoj nastavi se najčešće svode na pronalazak dijela jednakosti koji nedostaje, kao što je primjerice na Slici 1.

$$586 - \square = 349$$

Slika 1. Primjer linearne jednadžbe u razrednoj nastavi

U predmetnoj nastavi situacija postaje nešto apstraktnija pa se od 5. razreda umjesto „kućica“ uvode nepoznanice, kao što je primjerice na Slici 2., no i dalje se rješavaju na isti način, uporabom veza računskih operacija.

$$586 - x = 349$$

Slika 2. Primjer linearne jednadžbe u 5. razredu

Od 6. razreda, situacija se dodatno apstrahira i učenicima se servira tzv. „prebacivanje“ nepoznanica na jednu stranu i konstanti na drugu stranu, no to je i dalje rješavanje korištenjem veza računskih operacija, no na nešto višoj razini.

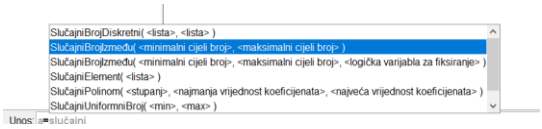
Najčešća preporuka učenicima za usvajanje rješavanja jednadžbi je što više vježbati, no tu često nailazimo na specifične probleme. Školski udžbenici su ograničeni gramažom, tj. brojem stranica pa onda i zadataka te učenicima često nedostaje zadataka za vježbu jer ih često sve iz udžbenika vrlo brzo riješe. Stoga nam u nastavi, ali i radu učenika izvan nastave, izvrsno mogu poslužiti beskonačni generatori jednadžbi, a koje je vrlo lako izraditi u alatu dinamične geometrije Geogebra.

III. LINEARNE JEDNADŽBE U GEOGEBRI

U ovom dijelu će biti detaljno prikazano kako izraditi beskonačne generatore linearnih jednadžbi i njihovih rješenja u Geogebri. Geogebra ima nekoliko različitih mogućnosti prikaza: Algebra, Tablice, Kalkulator vjerojatnosti, različite Grafičke prikaze, Opis konstrukcije i Tipkovnica. Automatski se pri pokretanju prikazuju „Algebra“ (lijevi dio zaslona) i „Grafički prikaz“ (ostatak zaslona). Ova dva automatska prikaza, sa trakom za unos na dnu ekrana, jedino su potrebno za izradu generatora jednadžbi. „Grafički prikaz“ se otvara s prikazanom mrežom i koordinatnim osima što za izradu generatora jednadžbi nije potrebno pa je dovoljno kliknuti desnim klikom bilo gdje unutar ograde „Grafičkog prikaza“ i kliknuti na „Mreža“ i „Koordinatne osi“ kako bi se oni uklonili.

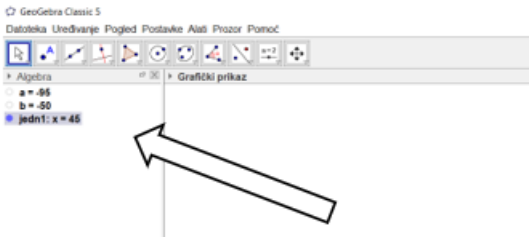
Započinjemo sa zapisom s kojim se učenici i prirodno najranije susreću, već i u razrednoj nastavi matematike, a to je $x + a = b$, gdje su a, b bilo koja dva cijela broja (do 6.razreda osnovne škole se učenici u nastavi matematike ne susreću s brojevima negativnih predznaka). Najprije trebamo definirati varijable a i b, a

definiramo ih slobodnim upisom u traku za unos na dnu ekrana. Definirat ćemo ih tako da u traku za unos upišemo **a=SlučajniBrojIzmeđu(donja granica intervala,gornja granica intervala)** i **b=SlučajniBrojIzmeđu(donja granica intervala,gornja granica intervala)**, kao što je prikazano na Slici 3.



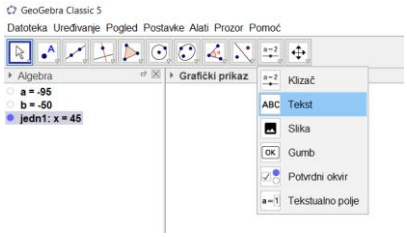
Slika 3. Definiranje varijable a

Zatim trebamo definirati rješenje jednadžbe x, a njega definiramo tako da u traku za unos upišemo $x = b - a$. Rješenje x će se, kao i varijable a i b, kako ih definiramo listati redom u lijevom dijelu ekrana „Algebra“, kao što je na Slici 4.



Slika 4. Varijable i rješenje jednadžbe u dijelu prikaza „Algebra“

Nakon što definiramo varijable i rješenje, trebamo to i grafički prikazati, a to se može napraviti tako da dodamo tekst i unutar njega upišemo jednadžbu. Tekst dodajemo kao što je prikazano na Slici 5.



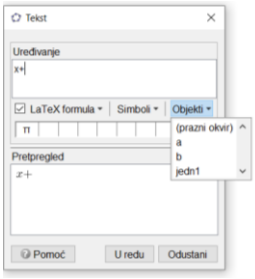
Slika 5. Dodavanje teksta

Nakon što se klikne na „Tekst“, potrebno je kliknuti bilo gdje unutar dijela „Grafički prikaz“, ovisno o tome gdje želimo da tekst, tj. jednadžba, bude prikazan. Nakon klika, otvara se novi prozor kao što je prikazano na Slici 6.



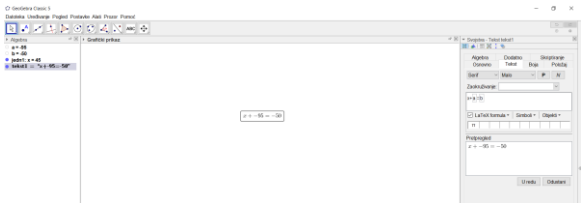
Slika 6. Prozor koji se otvara nakon klika na naredbu „Tekst“ i klika u slobodni dio dijela „Grafički prikaz“

Upis se vrši u dijelu „Uređivanje“. Ovdje je potrebno kliknuti na kvadratić pored kojeg piše „Latex formula“ ako želimo da tekst koji upisujemo u dijelu „Uređivanje“ bude prikazan matematičkim zapisom. Zatim u dio „Uređivanje“ upisujemo „**x +**“ , a umjesto upisivanja varijable a klikamo na „Objekti“, nakon čega se otvara padajući izbornik u kojemu odabiremo prethodno definiranu varijablu a, kao što je prikazano na Slici 7. Tako ćemo povezati jednadžbu koju zadajemo s prethodno definiranom varijablom. Isto radimo i sa varijablom b. Na kraju, na dnu prozora klikamo na gumb „U redu“ i jednadžba će biti dodana u dio prikaza „Grafički prikaz“.



Slika 7. Povezivanje zapisa unutar naredbe „Tekst“ i prethodno definiranih varijabli

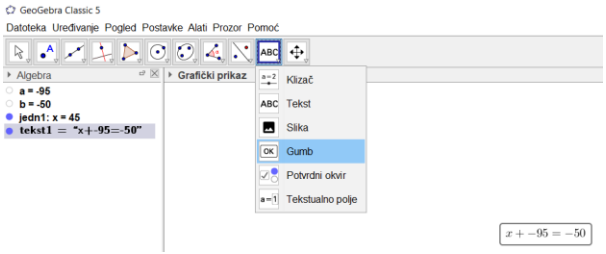
Ako želimo promijeniti, primjerice, boju, font ili veličinu teksta (tj.jednadžbe), dovoljno je desnim klikom kliknuti na jednadžbu i „Svojstva objekta“ pa će se s desne strane prikazati, kao na Slici 8., izbornik u kojemu možemo promijeniti što želimo.



Slika 8. Promjena svojstava objekta „Tekst“

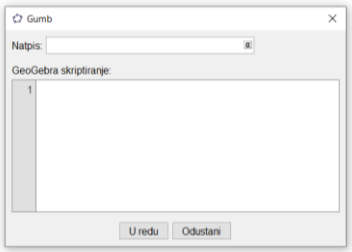
Ako ne želimo da nam se generiraju jednadžbe sa dva računska znaka zaredom, kao što je Slici 8., preporuka je napraviti poseban generator za zapis $x + a = b$, a poseban generator za zapis $x - a = b$, jedino što onda moramo u obje situacije varijablu a prilikom definiranja definirati kao nenegativnu, tj. da je donja granica broj 0.

Za beskonačno generiranje, potrebno je i neko mjesto gdje bismo mogli kliknuti i klikom na koji se svaki put generira nova linearna jednadžba. Za tu potrebu možemo izraditi gumb generiranja jednadžbe. Kako dodati gumb u „Grafički prikaz“, prikazano je na Slici 9.



Slika 9. Dodavanje gumba za beskonačno generiranje u dio „Grafički prikaz“

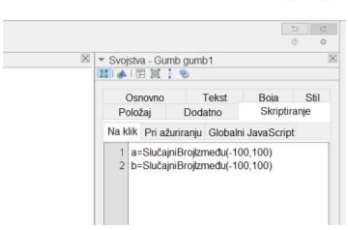
Opet je nakon toga potrebno kliknuti bilo gdje u slobodni dio dijela „Grafički prikaz“, a nakon čega se otvara novi prozor kao što je na Slici 10.



Slika 10. Prozor za definiranje naredbe „Gumb“

U ovom koraku je dovoljno u traci za unos „Natpis:“ zapisati tekst koji želimo da bude prikazan na gumbu, primjerice „Generiraj novu jednadžbu“. Klikom na gumb „U redu“ pri dnu prozora, gumb za generiranje jednadžbe se automatski dodaje u dio „Grafički prikaz“ , na mjesto koje smo prethodno klikom odabrali. Uređivati ga opet možemo desnim klikom na njega i klikom na „Svojstva objekta“, a nakon čega se opet otvara novi izbornik s desne strane.

To još nije dovoljno kako bi nam se klikom na taj gumb stalno generirale nove jednadžbe. Potrebno je još u izborniku s desne strane (do kojeg dolazimo uređivanjem gumba) kliknuti na karticu „Skriptiranje“ i podkarticu „Na klik“. Nakon toga, otvara nam se okvir za unos u koji unosimo što bismo željeli da nam se događa klikom na gumb. Kako mi želimo da se generiraju nove vrijednosti varijabli a i b, tako i upisujemo u taj okvir za unos, iste one definicije varijabli kako smo ih i na početku definirali. Nakon toga je samo dovoljno kliknuti na x u gornjem desnom uglu tog izbornika i gumb za generiranje jednadžbi je definiran.

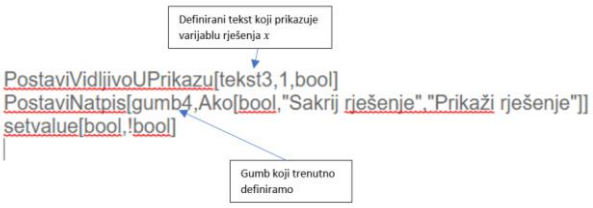


Slika 11. Prikaz definiranja generiranja varijabli a i b u gumbu generiranja jednadžbi

Sada bi se, klikom na gumb „Generiraj novu jednadžbu“, svaki put trebala generirati nova jednadžba.

Ako želimo da ipak rješenje bude sakriveno te se prikazuje ili sakriva po našoj želji, to možemo napraviti dodavanjem novog gumba i opisom što će se svaki put dogoditi klikom na njega. To se može napraviti na nekoliko načina, a opisat ćemo način koji mijenja i tekst gumba ovisno o radnji (ako rješenje nije prikazano, natpis na gumbu je „Prikaži rješenje“, a ako je rješenje prikazano na gumbu piše „Sakrij rješenje“) i koji se zapravo temelji na Booleovoj algebri, a koja je predefinirana u Geogebri.

Najprije u traku za unos upišemo „bool=true“.To radimo samo kako bismo zapisali trenutno stanje (varijabla je u našem slučaju već prethodno dodavanjem teksta prikazana pa smo zato upisali „true“). Svakim klikom na gumb povezan s ovom naredbom, događat će se promjena iz „true“ u „false“ i obrnuto, odnosno prikazivanje i sakrivanje rješenja Mi dodatno i natpis na gumbu u svakoj od tih situacija. Sada dodajemo novi gumb, onako kako smo to prethodno opisali. Opet nam treba kartica „Skriptiranje“ i podkartica „Na klik“, gdje u okvir za unos unosimo što želimo da nam se odvija klikom na gumb. U taj dio, redom po retcima upisujemo kao što je prikazano na Slici 12. Nakon toga je samo dovoljno kliknuti na x u gornjem desnom uglu tog izbornika i gumb je definiran.



Slika 12. Definiranje gumba za sakrivanje/prikazivanje rješenja

Prirodno se, nakon ovih oblika zapisa jednadžbi, prelazi na oblik linearne jednadžbe $ax = b$, a gdje je cijeli ovaj prethodno opisano postupak gotovo identičan, drugačije je jedino definiranje teksta prikaza u oblik linearne jednadžbe koji koristimo i definiranje varijable x koju sada definiramo $x = b/a$. Ne moramo posebno pri definiranju varijabli obraćati pažnju na situaciju kada je varijabla a jednaka nuli jer Geogebra sama prilagođava varijablu rješenja i u prikazu piše ima li ga ili ne.

Na kraju, ostaje još oblik $ax + b = c$ (i, po želji, dodatni oblik $ax - b = c$ ako opet želimo izbjeći 2 računska znaka zaredom), a gdje se samo definiramo dodatnu, slučajno generirana varijabla c i, naravno, drugačije definiranje varijable x koja je sada $x = (c - b)/a$. Ostatak postupka je identičan prethodno opisanom.

Ako želimo unutar iste bilježnice u Geogebri imati zadatke vezane uz više prethodno spomenutih zapisa linearne jednadžbe, onda treba obratiti pažnju da se pri definiranju varijabli one ipak nekako drugačije imenuju, npr. $a1, a2, a3$ itd. jer će se inače klikom na jedan gumb generiranja jednadžbe generirati stalno nove jednadžbe u svim zapisima, a varijabla rješenja x će biti prilagođena samo zadnjoj njenoj definiciji.

Prirodan nastavak toga je i projekt u kojem bi učenici sami mogli napraviti jednostavne beskonačne generatore, ne samo nužno jednadžbi, nego i, primjerice, veličine kuta pa da ostatak razreda mora pogađati naziv kuta (predmetna nastava matematike) ili pak generatore zadataka osnovnih računskih radnji (razredna nastava matematike). Nastavnik kao vođa nastavnog procesa mora vrlo aktivno usmjeravati učenike pri svakom koraku u procesu izrade kako oni ne bi izgubili motivaciju, posebice ako se prvi put susreću s Geogebrom.

IV. GEOGEBRA RUKU POD RUKU S OSTALIM OBLICIMA NASTAVE

Geogebra ima svoje mjesto u nastavi, pogotovo u ovakvim situacijama kada nam jednostavno trebaju dodatni primjeri i dodatni zadatci. Dobro dođe i kao alat učenicima koji često griješe jer je u dijelu „Algebra“, ako to prethodno ne isključimo, uvijek vidljivo kako je varijabla rješenja definirana pa učenici mogu rekonstruirati put do rješenja jednadžbe. Kao učenicima, može poslužiti i nastavnicima pri pripremi materijala za nastavu i uštediti im vrijeme za samostalno rješavanje zadataka i pripremu za nastavu.

No, treba imati na umu i ograničenja. Geogebra, primjerice, ne „namješta“ rješenja i članove jednadžbe da se ne bi potkrali „nezgodni“ razlomci i/ili decimalni brojevi.

Također, učenici su već i sada znatno izloženi uređajima u mnogo aspekata života pa sam i sam u nastavi primijetio da im se ponekad teško koncentrirati na zadatke iz udžbenika i bilježnica jer im takav statičan zapis crnih slova na bijeloj podlozi nije dovoljno pobuđujući za mozak. S druge strane, uređaji za pristup interaktivnim sadržajima su često opremljeni igricama i aplikacijama povezanima s društvenim mrežama pa im konstantne obavijesti i poruke svejedno remete koncentraciju. Zato je bitno ovakve sadržaje dozirati, da oni ne budu prevladavajući.

Uvježbavanje rješavanjem mnoštvom zadataka također nije niti jedini oblik potpunog usvajanja određenih ishoda, pogotovo na višoj razini. U radu s učenicima treba inzistirati na zadatcima primjene, modeliranja stvarnih problema korištenjem stečenih matematičkih znanja i vještina pri rješavanju, pogotovo ako su ti problemi njima poznati i bliski, što zasigurno osigurava veću motivaciju za rad. Primjerice, možemo im dati formulu za preračunavanje eura u kune pa da učenici računaju koliko bi x eura bilo kuna i obrnuto. Učenike uvijek zainteresiraju i zadaci sa zaradama nekih influencera, tj. koliko zarađuju po reklami s obzirom na vrstu reklame pa da učenici računaju koliko se zaradi po pojedinom tipu reklame ili koliko bi kojih reklama trebali „odraditi“ za određenu zaradu, a što još možemo učiniti zanimljivijim učenicima ako diskutiramo koliko vremena i truda treba za koju vrstu reklame te što im se onda najviše isplati raditi. Zadatke možemo povezati i s ostalim školskim predmetima, pogotovo ako se prethodno konzultiramo s kolegama o aktualnoj problematici, tj. što trenutno rade na tom predmetu.

U sve te primjere povezivanja i modeliranja lako možemo uključiti i suradnički rad, rad u parovima i grupama, a gdje uz malo truda možemo osigurati da svaki učenik, primjereno svojim mogućnostima, doprinese zajedno s ostalim učenicima rješavanju nekog većeg i/ili zahtjevnijeg problema što bi mu samom bilo teže, a ponekad i nemoguće.

V. ZAKLJUČAK

Učenicima zbog ograničenja mase i broja stranica školskih udžbenika često nedostaje zadataka za uvježbavanje, pogotovo iz linearnih jednadžbi gdje je to uvježbavanje bitno. Zato u nastavi možemo koristiti alate dinamične geometrije kao što je Geogebra. U Geogebri se, uz vrlo malo znanja i korištenjem njenih osnovnih funkcija, mogu lako izraditi beskonačni generatori zadataka. No, nije se dobro zadržati samo na njima, nego je za potpuno usvajanje ishoda bitno ih koristiti ruku pod ruku s ostalim oblicima nastave, kao što je primjerice modeliranje i suradnički rad.

LITERATURA

[1] Ministarstvo znanosti i obrazovanja, “Odluka o donošenju kurikuluma za nastavni predmet matematike za osnovne škole i gimnazije u Republici Hrvatskoj”, Narodne novine NN 7/2019, Zagreb, 2019.

[2] Priručnici Geogebra (preuzeto 28.1.2023.)

[3] Webinari Geogebra (preuzeto 21.1.2023.)

[4] Tutorijali Geogebra (preuzeto 25.1.2023.)

Industrijski IT certifikati iz područja računalnih mreža

Davorin Valenčić*, Matej Miholjek *, Aleksandar Skendžić *
* Veleučilište Velika Gorica/Velika Gorica, Hrvatska
davorin.valencic@vvg.hr , mmiholjek@yahoo.com , wireless82@gmail.com

Sažetak – Kako bih se držao korak s velikom potražnjom kvalitetnih IT stručnjaka na današnjem tržištu, potrebna su velika ulaganja u njihovo obrazovanje i specijaliziranje. Jedan od najboljih i najbržih oblika kako obrazovati ljude (a pogotovo ih držati u koraku sa trenutnim tehnologijama) je educiranje kroz industrijske certifikate. U ovom radu će se prikazati vrste IT certifikata, razlozi i potreba za certifikacijom i njezine prednosti kod zapošljavanja i za poslodavce i posloprimce. Osim toga će se objasniti važnost industrijskih certifikata u području računalnih mreža za današnje i buduće IT stručnjake i inženjere računalnih mreža. Razmotriti će se i dostupnost industrijskih certifikata od strane najvećih svjetskih proizvođača mrežne opreme.

I. UVOD

Zbog napretka u području informacijskih i komunikacijskih tehnologija (engl. Information and Communications Technology - ICT) došlo je do potrebe za velikom količinom stručnjaka koji će moći raditi u tom području i razvijati nove usluge. Stvaranjem velikog broja novih radnih mjesta došlo je do potrebe za stjecanjem posebnih vještina u svrhu razvoja i opstanka tih stručnjaka u situaciji kada su razvoj tehnologije i konkurencija izuzetno velike. Dakle, stručnjaci moraju sve više preuzimati odgovornost za kontinuirano stjecanje različitih vještina kako bi ostali konkurentni na tržištu. [1]

Novi dinamični načini učenja i stjecanja vještina povećavaju potrebu IT stručnjacima da dokumentiraju svoje ishode učenja novih vještina u obliku priznatih vjerodajnica za razvoj karijere, kao što su diplome i certifikati, te sa time imaju i mogućnost povećanja svoje vrijednost na tržištu rada pojedinca. [2]

Isto tako poslodavci imaju problem utvrditi koje profesionalne vještine nedostaju njihovim postojećim kadrovima i koje profesionalne IT vještine su potrebne za ispunjavanje internih zahtjeva. [2]

Tema ovog rada su industrijski IT certifikati iz područja računalnih mreža. Prvi dio ovog rada je nastao na osnovi diplomskog rada [2], a u nastavku rada je prikazano trenutno stanje na tržištu industrijskih IT certifikata iz područja računalnih mreža.

II. ICT CERTIFIKATI

Certifikati su vjerodajnice koje prepoznaju i potvrđuju nečije znanje i stručnost. Programi certificiranja uključuju određeno razdoblje obrazovanja i osposobljavanja, kada se završi razdoblje certificiranja, polaznik dobiva potvrdu o završetku programa certificiranja. Programi certifikata prikladni su za osposobljavanje pojedinaca za određeni zadatak ili za razvoj kompetencija u određenoj vještini. Certifikati mogu sadržavati diplome polaznika, kao i certifikate iz tehničkog, poslovnog i na primjer područja stranih jezika. [1] [3]

Rast i razvoj ICT tehnologija i njihova posljedica za sve većom potražnjom za radnicima sa specijaliziranim vještinama postavili su značajan zahtjev prema tradicionalnom obrazovnom sustavu da pruži kvalificiranu i održivu radnu snagu.

Kao odgovor na napredak u računalnoj tehnologiji i sporo reagiranje tradicionalnog obrazovanja, IT industrija koristi certificiranje kao način osposobljavanja i akreditacije vlastitih radnika. [4]

Dobivanje certifikata zahtijeva da pojedinac položi uvjete koji se odnose na postavljene kriterije, potvrđujući postizanje određenih vještina. Ocjene IT certificiranja provodi dobavljač ili tvrtka za testiranje, a mnogi dobavljači zahtijevaju da radnici sa već položenim certifikatima idu ponovno na certificiranje (engl. recertification) nakon određenog vremenskog razdoblja kako bi pokazali kontinuiranu sposobnost. Dobavljači certifikata ne zahtijevaju od polaznika da daju dokaz o pripremljenosti prije pokušaja polaganja ispita i za razliku od pripreme kod tradicionalnog obrazovanje na kojem su propisani materijali koji su obavezni za prolazak ispita, izbor materijala je prepušten polazniku ispita. [4]

Priprema za IT certifikacijske ispite može obuhvaćati fizičku opremu, knjige, simulacijski softver, polaganje ispita, nastavne programe sponzorirane od strane dobavljača i obuku pod vodstvom instruktora koja se nudi na formalnim i neformalnim obrazovnim institucijama. Dobavljači IT certifikata mogu predložiti materijale za pripremu ispita ili čak pružiti vlastiti kurikulum za IT uz odgovarajuću naplatu studentima i obrazovnim institucijama.

Skillsoft je objavio svoje godišnje izvješće za 2022. godinu o najpopularnijim IT certifikatima u svijetu. Prema izvješću, 91% stručnjaka u informacijskoj tehnologiji ima barem jednu certifikaciju. Zapravo, ispitanici posjeduju u prosjeku četiri certifikata u svom području, od kojih je 41% stečeno u posljednjih šest mjeseci. [7] [8]

Ovaj rad je podržan od Veleučilišta Velika Gorica, Velika Gorica, Hrvatska.

A. *Prednosti certifikata za posloprimce*

Proces certificiranja ima velike prednosti za posloprimca, kao na primjer: [1][3][5][6]

- Prednost pri zapošljavanju: u današnje vrijeme konkurencija za gotovo bilo koji posao vrlo je jaka, a certificiranje može pokazati da osoba sa certifikatom ima bolje znanje od drugih kandidata.
- Bolja mogućnost za zadržavanje posla: certifikat može pomoći pokazati poslodavcu predanost prema osobnom poboljšanju potrebnih skupa vještina i znanja.
- Napredak u karijeri posloprimca: polaganje i dobivanje novog ili naprednijeg certifikata u određenom području stručnosti može pomoći u napredovanju u karijeri.
- Utvrđivanje profesionalne vjerodostojnosti na tržištu: profesionalna vjerodostojnost jedan je od najvažnijih čimbenika koji pokazuje važnost certificiranja za radnike bez radnog iskustva. Kada radnik posjeduje određenu certifikaciju i naknadne certifikate od istog dobavljača, može pokazati svoju profesionalnu sposobnost poslodavcu.
- Pristup prednostima uključenosti u zajednice: certificiranja pomažu pojedincima da postanu dijelom dviju novih zajednica. Prva zajednica uključuje pojedince koji teže i rade na polaganju certifikata, dok druga zajednica uključuje certificirane profesionalce.
- Pojedinci koji teže certificiranju mogu djelovati kao grupa za podršku među kolegama, te u toj grupi istovremeno mogu doprinijeti sa znanjem, ali i učiti. Dok istovremeno ovlašteni profesionalci i stručnjaci za određenu certifikaciju mogu pomoći sa svojim znanjima i iskustvom u određenim područjima. Zbog toga polaznik može dobiti pouzdane smjernice, savjete za poboljšanje karijere i informacije o izvorima za stjecanje određenog tehničkog znanja.
- Pokretanje osobnog razvoja: postavljanje certifikata kao osobnih ciljevi može pomoći u ostvarenju zadovoljavajućeg iskustva. Certifikati pokreću osjećaj samo-motivacije, discipline i predanosti, što je nematerijalna nagrada stjecanjem certifikatima. To pomaže pojedincu u psihološkom aspektu u napretku kroz karijeru.
- Pokazivanje jedinstvenosti na tržištu: uvođenje novih certifikata specijalnosti od strane dobavljača stvara nove mogućnosti zapošljavanja i napretka u karijeri. Profesionalci se na taj način mogu usredotočiti na dobivanje ovih novih certifikata, koji im mogu pomoći da postanu stručnjaci u novim tehnološkim područjima i smjerovima.
- Relevantnost u današnjem tehnološkom svijetu: neki određeni certifikati zahtijevaju obnavljanje certifikata (engl. recertification). Ponovno certificiranje istaknuti je uvjet za održavanje i unaprjeđivanje znanja kao certificiranog stručnjaka na određenom području. Ponovno certificiranje pomažu profesionalcima da uhvate korak s poboljšanjima i novinama na određenim tehnološkim platformama i održavanju njihovih vještina relevantnima.

Gore navedeno pokazuju i nedavna istraživanja. Osim materijalne nagrade (povećane plaće), oni koji su nedavno stekli IT certifikat pokazuju i povećanu učinkovitost na poslu: [7] [8]

- 56% kaže da im se poboljšala kvaliteta rada,
- 41% kaže da su više angažirani u svom poslu i
- 36% kaže da su brži u obavljanju svog posla

B. *Prednosti certifikata za poslodavce*

Postoje velike prednosti koje certifikacija zaposlenika donosi tvrtkama, kao na primjer: [1][5][6]

- Povećana produktivnost organizacije: ključ bržeg rasta produktivnosti je sposobnost kombiniranja novih tehnologija s odgovarajućim ljudskim kapitalom i promjene poslovanja, procesa i upravljačke prakse tvrtke kako bi se iskoristile nove mogućnosti i poboljšala produktivnost. Jedni od glavni prioriteta koji omogućuju ovakav rast u produktivnosti su obrazovanje, osposobljavanje i usavršavanje radnog kadra.
- Povećanje dostupnih vještina: današnje organizacije se sve više susreću sa manjkom profesionalnih vještina u nekim brzo rastućim tehnološkim područjima, te se zbog toga sve više usredotočuju na zapošljavanje i razvoju tih vještina. Kontinuirano obrazovanje smanjuje jaz u vještinama čak i prije nego što one postanu očite, te s tim poboljšanim vještinama omogućuju radnicima organizacije veću raznolikost u poslovima koje mogu raditi.
- Poboljšana učinkovitost i smanjen rizik: jedan od uzroka smanjenja učinkovitosti je gubitak stručnjaka u organizaciji. Odlazak stručnjaka ima više posljedica na organizaciju, kao na primjer financijski aspekt i troškovi koje nakon odlaska mora potrošiti na pronalazak zamjene. Taj trošak u većini slučaja je veći od plaće toga stručnjaka. Osim toga potrebno je i vrijeme da novi kadar postane produktivan. Kako bi se usporio ili čak i spriječio odlazak stručnjaka i time spriječili financijski gubitci i poslovni rizici, organizacije moraju sve više ulagati u svoje kadrove. Kao jedan od glavnih aspekata ulaganja u radne kadrove i njihovo zadržavanje, certifikacija igra veliku ulogu. Sve veći broj organizacija u zadnje vrijeme je spremno plaćati razne vrste profesionalnog certificiranja svojih zaposlenika.
- Povećanje zadovoljstva zaposlenika: certifikacija gradi samopouzdanje i povećava motivaciju, stvarajući sretne, pozitivne zaposlenike s mogućim stavom orijentiranim na uslugu, što zauzvrat čini organizaciju vjerodostojnijom i profesionalnijom u očima klijenata. Reputacija

tvrtke zbog certificiranja čini je privlačnijom za najkvalificiranije kandidate za nova radna mjesta, kao i za profesionalce za zapošljavanje.

Prema studiji Skillsofta iz 2015. godine postotak tvrtki koje prijavljuju nedostatak profesionalnih IT vještina raste [7].

Isto tako istraživanje iz 2022. godine je pokazalo da: [7][8]

- gotovo 60% svih IT donositelja odluka predviđa nedostatak vještina u sljedećim godinama;
- 41% tvrtki vjeruje da je certificirano osoblje ključno za uklanjanje nedostataka u vještinama;
- više od 96% IT menadžera kaže da certificirano osoblje dodaje vrijednost njihovim organizacijama.

Isto istraživanje je pokazalo da menadžeri za zapošljavanje navode posebne prednosti certificiranog osoblja: [7][8]

- 45% smatra da certifikacija povećava produktivnost;
- 44% smatra da certifikacija zadovoljava zahtjeve klijenata;
- 37% smatra da certifikacija skraćuje vrijeme rješavanja problema
- 35% smatra da certifikacija omogućuje brže dovršavanje projekata
- 35% smatra da certifikacija pruža konkurentsku prednost

Iz tog istraživanja je pokazano i da poslodavci smatraju da su certificirani stručnjaci bolji u uklanjanju nedostataka u vještinama, udovoljavanju zahtjevima klijenata i povećanju produktivnosti, a za to im treba manje vremena. [7] [8]

III. VRSTE IT CERTIFIKATA

IT certifikate razvijaju profesionalci i industrija, kao i neovisni dobavljači, te se oni mogu grupirati u dvije opće kategorije: [1]

- certifikati proizvođača IT tehnologije (engl. vendor-specific certification) i
- certifikati neutralni u odnosu na proizvođače (engl. vendor-neutral certification).

Potvrde i programi proizvođača (engl. vendor) izrađuju se i nadgledaju za određene proizvođače i usredotočuju se prvenstveno na informatičku disciplinu koja okružuje tehnologiju tvrtke, usluge ili liniju proizvoda. Certifikati neutralni u odnosu na proizvođače izrađeni su za određenu grupu stručnjaka iz industrije, javnog i privatnog sektora i usredotočuju se na metodologije i tehnologije oko određene uloge posla.

A. Certifikati neutralni za dobavljače

Certifikati neutralni u odnosu na proizvođače su oni koji se fokusiraju na poslovne vještine demonstriranjem ključnih koncepata kritičnih za trenutne uloge na tržištu. Imaju osnovni cilj, koji se često pojačava sa popularnim hardverskim ili softverskim alatima, ali se više usredotočuju na poslovne vještine, a ne softverske ili hardverske vještine. Certifikati neutralni u odnosu na proizvođače razvijaju i nadgledaju se od strane raznih konzorcija, neprofitnih organizacija i tvrtki specijaliziranih za treninge koje razvijaju standarde i sadržaje korištenjem odbora pojedinaca iz industrije, akademske zajednice i strukovnih udruga ili društava. [2]

Certifikati neutralni u odnosu na proizvođače odražavaju kako tehnologije funkcioniraju usredotočujući se više na koncepte koje se odnose na temeljnu tehnologiju, a ne na proizvod ili tehnologiju određenog proizvođača, te samim time pripremaju pojedinca za rad u tvrtkama sa proizvodima i opremom više proizvođača i različitim tehnologijama. Iako neutralni u odnosu na proizvođače, ispiti su orijentirani na koncepte i usredotočeni su na opće znanje važno za polaznike i odražavaju široko rasprostranjenu prilagodbu određenih tehnologija.

Neke od pogodnosti certifikata neutralnih u odnosu na proizvođače su [1] [2]:

- Neovisnost o proizvođačima omogućava stvaranje programa koji nije vezan uz jednu tehnologiju ili proizvod.
- Programi neutralni u odnosu na proizvođače nude nepristran pogled uz pružanje uravnotežene pokrivenosti teme.
- Neutralnost prema proizvođačima pruža mogućnost rješavanja nedostataka ili problema povezanih s njihovim proizvodima ili tehnologijama.
- Uklanjanje potrebe za promocijom pojedinačnih proizvoda omogućuje da se usredotoče na relevantna tehnološka pitanja toga područja.
- Razvijanje programa koji pokrivaju proizvode više od jednog proizvođača.
- Pružanje mogućnosti certifikacije onima koji još nisu odabrali specifični proizvod.

Jedna od najpoznatijih neovisnih organizacija za dobivanje certifikata neovisnih o proizvođaču u ICT području je organizacija Computing Technology Industry Association (CompTIA). [1]

B. Certifikati proizvođača

Stručnjaci koje su certificirani od proizvođača za specifičnu IT tehnologiju zasnovanu na proizvodima pogoduju tvrtkama zbog smanjivanja troškova tehničke podrške, promicanja vlastitih proizvoda i tehnologija i poticanja polaznika na ponovno certificiranje. Neke od pogodnosti programa certificiranja proizvođača su [1] [4]:

- Programi specifični za proizvođače pružaju detaljne upute o određenim tehnologijama i proizvodima.
- Uživaju prepoznatljivost u cijeloj industriji u odnosu na popularnost samog proizvoda.

- Njihovi programi su usmjereni na trenutne i široko korištene proizvode.
- Više pratećeg materijala kao što su knjige, materijali za obuku su dostupni za programe specifične za proizvođače.
- Neki certifikati proizvođača općenito su priznati na svjetskoj osnovi, dok određene certifikate neutralne prema proizvođačima provode lokalne ili regionalne organizacije.
- Proizvođači mogu pripremiti program za proizvode prije njihovog puštanja, dopuštajući programe koji usko odgovaraju industrijskim trendovima.

Neki od najpoznatijih certifikata proizvođača IT tehnologije prema [1][3] su Cisco, Juniper, Microsoft, Oracle i mnogi drugi.

IV. PROIZVOĐAČI MREŽNIH UREĐAJA

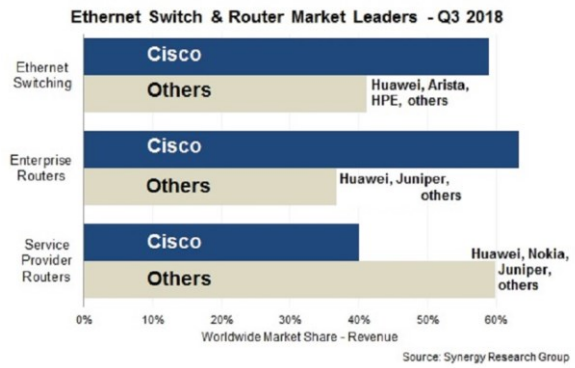
Danas postoji veći broj proizvođača mrežne opreme na svjetskom tržištu. Pod mrežnom opremom danas se uglavnom smatraju uređaji 2. i 3. sloja OSI modela, odnosno usmjernici (engl. routers) i preklopnici (engl. switch), za koje se često koristi skraćenica R/S.

Slika 1. prikazuje stanje 2018. godine na tržištu R/S mrežnih uređaja u svijetu koje je objavila Synergy Research Group. Cisco je najveći proizvođač usmjernika i preklopnika. Po područjima primjene, Cisco ima oko 58% udjela na tržištu u području Ethernet preklopnika u tvrtkama, oko 70% udjela na tržištu usmjernika za tvrtke i oko 40% na tržištu usmjernika za davatelje usluga. Ostali proizvođači (Huawei, Juniper, Hewlett Packard Enterprise - HPE, Arista, Nokia, Alcatel-Lucent, F5, Extreme Networks i drugi) imaju preostali udio. [9]

Slika 2. prikazuje stanje tržišta mrežnih R/S uređaja u svijetu i odnos kompanije Cisco i njegovih konkurenata iz 2021. godine [10]. Odnos između proizvođača Cisco i njegovih konkurenata se promijenio, ali Cisco je zadržao i dalje poziciju lidera u svijetu u području mrežnih uređaja.

V. IT CERTIFIKATI IZ PODRUČJA RAČUNALNIH MREŽA

Svake godine objavljuju se liste najtraženijih i najbolje plaćenih ICT certifikata od strane različitih tehnoloških IT web stranica. Primjeri takvih lista su sljedeće:



Slika 1. Tržište mrežnih uređaja u svijetu 2018. godine [9]

Company	Share of Enterprise & Service Provider Routers	Share of Ethernet Switches
Cisco Systems	35%	45%
Huawei Technologies	31%	10%
Arista Networks	n/a	8%
New H3C Technologies	2%	6%
Hewlett Packard Enterprise	n/a	6%

Slika 2. Tržište R/S uređaja: Cisco i konkurenti 2021. godine [10]

- 15 najtraženijih IT certifikata za 2022. (web stranica www.cio.com) [11]
- 10 najplaćenijih IT certifikata za 2022. (web stranica www.pcmag.com) [12]
- 15 najplaćenijih IT certifikata 2022. (web stranica www.skillsoft.com) [13]
- 15 najplaćenijih IT certifikata u 2022. (web stranica www. techtarget.com) [14]
- 17 najtraženijih IT certifikata za 2023. (web stranica www.indeed.com) [15]
- Najbolji IT certifikati (web stranica www.thebalancemoney.com) [16]

Slično je kompanija Global Knowledge objavila IT certifikate rangirane kao najpopularnije u svijetu [17]. Na Slici 3. se vide i dvije kompanije koje rade certifikate u području računalnih mreža: Cisco i CompTIA.

U nastavku ovog rada zbog tržišnog udjela i duljine prisutnosti na tržištu sa svojim proizvodima i certifikatima, detaljnije će se prikazati način certificiranja za kompanije Cisco i Juniper, a ukratko za ostale proizvođače s većim prisustvom na tržištu (prema slikama 1. i 2.).

Biti će ukratko prikazan i jedini certifikat za računalne mreže od strane organizacija za dobivanje certifikata neovisnih o dobavljaču: CompTIA Network+.

A. CompTIA Network+

CompTIA, jedna od najpoznatijih neovisnih organizacija za dobivanje certifikata neovisnih o proizvođaču, je razvila certificiranje za računalne mreže pod nazivom CompTIA Network+. [18]

CompTIA Network+ potvrđuje tehničke vještine potrebne za sigurnu uspostavu, održavanje i rješavanje problema s ključnim mrežama na koje se tvrtke oslanjaju. CompTIA Network+ nudi fleksibilne mogućnosti obuke uključujući učenje vlastitim tempom, online obuku,



Slika 3 – Najpopularniji IT certifikati za 2022. godinu (Global Knowledge) [17]

prilagođenu obuku i laboratorije za unapređenje razvoja karijere IT stručnjaka u mrežnoj administraciji. [18]

Certifikat CompTIA Network+ potvrđuje da uspješni kandidat ima znanja i vještine potrebne za: [19]

- uspostavljanje mrežne povezanosti postavljanjem žičnih i bežičnih uređaja;
- razumijevanje i održavanje mrežne dokumentacije;
- razumijevanje svrhu mrežnih usluga, osnovne koncepte podatkovnog centra, oblaka i virtualnog umrežavanja;
- praćenje mrežne aktivnosti, identificiranje problema s performansama i dostupnošću;
- provođenje osnovnih tehnika sigurnosti mreže;
- upravljanje, konfiguriranje i rješavanje problema s mrežnom infrastrukturom.

Više informacija o certifikatu CompTIA Network+. može se naći na [18] i [19].

B. Model certificiranja tvrtke Cisco

Kako su dva načina certificiranja dostupna u današnjem svijetu, ona ovisna i neovisna o proizvođaču, tako su se kroz povijest u IT industriji neki certifikati izdvojili i postali poznati širom svijeta.

Jedni od najpoznatijih certifikata proizvođača mrežne opreme potječu od tvrtke Cisco Systems.

Cisco Systems ili skraćeno Cisco je multinacionalna tvrtka koja razvija, proizvodi i prodaje mrežni hardver, softver, telekomunikacijsku opremu i druge usluge i proizvode visoke tehnologije. [20]

Kompanija Cisco je osnovana 1984. godine u San Franciscu, u Sjedinjenim Američkim Državama. [20] Tvrtka se u ranim početcima bavila proizvodnjom i prodajom mrežne opreme a prvi proizvod koji je zapravo pokrenuo uspon tvrtke je bio usmjernik. Prilikom velikog širenja na globalnoj razini u ranim devedesetim godinama prošlog stoljeća, Cisco je kupio nekoliko tehnoloških tvrtki u područjima mrežnih tehnologija kako bi proširio svoj utjecaj na globalnoj razini u mrežnim i telekomunikacijskim tehnologijama [21]. Kako je Internet protokol postao široko rasprostranjen Cisco je postao 2000. godine u jednom trenutku najvrjednija tehnološka kompanija na svijetu. Cisco je u idućim godinama održavao konkurentnost na svjetskom tržištu, ali početkom 2015. godine zbog sve lošijih rezultata na globalnoj razini, Cisco je odlučio krenuti sa ulaganjima u nova područja koja će promijeniti mrežne i telekomunikacijske tehnologije. 2016. godine Cisco [20] je krenuo sa razvojem tehnologija poput softverski definiranih mreža širokog područja, tehnologija zasnovanih na rad u oblaku, te strojnog učenja [21].

Cisco je puno godina imao certificiranje za tehnologije i područja povezana sa mrežnom infrastrukturom, te su certifikati bili bazirani na ta područja i tehnologije. Ali kako svijet mrežnih tehnologija napreduje, tako je Cisco



Slika 4 – Program certificiranja tvrtke Cisco [23]

primijetio važnost programiranja uz mrežne tehnologije. U veljači 2020. Cisco je kompletno promijenio način certificiranja, a novi način je dizajniran tako da uz kombiniranje stručnosti u mrežnoj infrastrukturi i sa snagom programibilnosti i automatiziranja omogućuje prednosti prilikom umrežavanja. [22] Ovaj program certificiranja potvrdit će kritične radne vještine i sposobnost upravljanja sve modernijom infrastrukturom poduzeća. Prikaz novog programa certifikacije je prikazan na slici 4. [23]

Novo Cisco certificiranje se razdvaja na dva glavna područja. Prvo je područje koje je i do sada postojalo, a to je područje mrežne infrastrukture, koje se sastoji od 5 tehnoloških područja: [23]

- mreže poduzeća (engl. Enterprise)
- mreže pružatelja usluga (engl. Service Provider)
- podatkovni centri (engl. Data Centre)
- sigurnost računalnih mreža
- suradnja (engl. Collaboration)

Novo područje je bazirano na spoju mrežnih tehnologija i programiranja, te ima samo jednu tehnološku vrstu (DevNet). U razvoju je i još jedno novo područje: CyberOps.

Sva navedena područja su isto tako podijeljena na 3 osnovne razine koje definiranju razinu znanja i specijaliziranosti, a te razine su [23][24]:

- Suradnik (engl. Associate) - savladavanje osnovnih znanja za pokretanje karijere u području računalnih mreža.
- Profesionalac (engl. Professional) - odabiranje temeljnog tehnološkog puta i stjecanje naprednijeg znanja u tom području.
- Stručnjak (engl. Expert) - stjecanje ekspertnog znanja iz specijaliziranog područja, kako teoretskog, tako i praktičnoga.

Detaljne informacije o Cisco certificiranju mogu se naći na Cisco web stranicama [23] i [24].

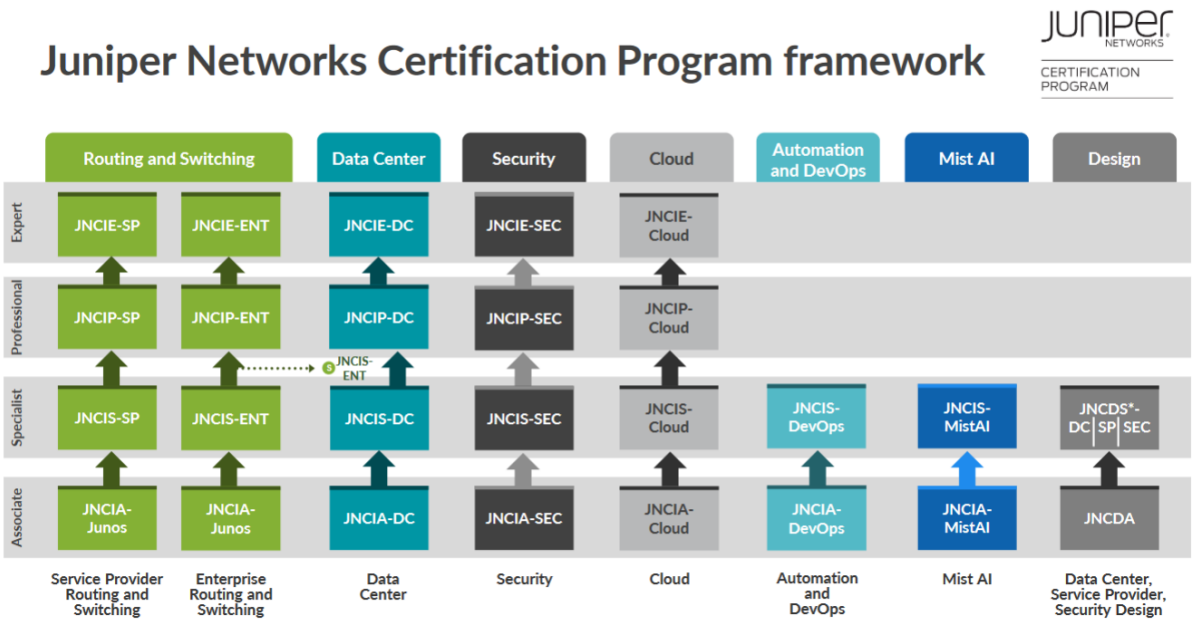
C. Model certificiranja tvrtke Juniper Networks

Juniper Networks se bavi dizajnom, razvojem i prodajom proizvoda i usluga za mreže visokih performansi. Proizvodi tvrtke Juniper Networks odnose se na mrežne zahtjeve za globalne usluge, pružatelje usluga u oblaku, nacionalne vlade, organizacije za istraživanje i javni sektor i druga poduzeća. Tvrtku je osnovao Pradeep S. Sindhu 1996. godine, a sjedište joj je u Kaliforniji, u Sjedinjenim Američkim Državama. [25]

Program certifikacije u tvrtki Juniper je raspoređen na gotovo identičan način kao i kod Cisco certificiranja, postoji nekoliko sličnih tehnoloških područja u kojima se može odabrati karijerni put polaznika. Ta tehnološka područja su raspoređena na četiri razine znanja za određena područja certifikata. Te četiri razine znanja su: [26][27]

- suradnik
- specijalist
- profesionalac
- stručnjak.

Tehnoloških područja su bazirana na usmjeravanju i preklapanju (R/S), podatkovnim centrima, sigurnosti, automatizaciji i mrežnom programiranju, te dizajniranju. Za svaku od ovih razina je potrebno imati certifikat nižeg certifikata, osim razine suradnika. Za sve certifikate potrebno je imati barem osnovna znanja koja su vezana za to područje, dok za certifikate Juniper profesionalca i stručnjaka je potrebno imati pet ili više godina iskustva. Također, isto kao u kod Cisco certificiranja, polaznici prolaze kroz teoretske ispite, ali i kroz praktične ispite na mrežnoj opremi proizvođača na kojima se testira njihovo znanje naučeno kroz teoriju na praktičnim slučajevima. Juniper certifikati isto kao i Cisco certifikati imaju vijek



Slika 5 – Program certificiranja tvrtke Juniper Networks [27]

trajanja od tri godine, nakon kojih je potrebno ponovno izaći na certificiranje. [26][27]

Prikaz kompletnog programa certificiranja kod tvrtke Juniper je prikazan je na Slici 5. Detaljne informacije o Juniper certificiranju mogu se naći na Juniper web stranicama [26] i [27].

D. Modeli certificiranja ostalih proizvođača mrežne opreme

I ostali veći svjetski proizvođači mrežne opreme imaju svoje modele certificiranja koji su po načinu certificiranja, razinama znanja i područjima tehnologija vrlo slični modelima prikazanima za Cisco i Juniper Networks.

Kao primjer se navode najveći svjetski proizvođači mrežne opreme:

- Huawei: model certificiranja je dostupan na [28],
- HPE: model certificiranja je dostupan na [29],
- H3C: model certificiranja je dostupan na [30],
- Arista Networks: model certificiranja je dostupan na [31] i [32],
- Nokia: model certificiranja je dostupan na [33],
- Extreme Networks: model certificiranja je dostupan na [34].

VI. ZAKLJUČAK

Kao što je prikazano u ovom radu, IT certifikati nude brojne prednosti i za zaposlenike i za poslodavce, odnosno poduzeća. Osnovni nedostatak certificiranja su troškovi cijene polaganja ispita certifikata. Pošto validnost većine IT certifikata istječe nakon nekoliko godina (najčešće 3 godine), certifikati se trebaju obnoviti ispunjavanjem određenih uvjeta (najčešće ponovnim polaganjem ispita) što nosi dodatne troškove.

U području računalnih mreža kompanija Cisco je prisutna je od samih početaka umrežavanja računala i nastanka Interneta i još uvijek su tržišni lider. Cisco model certificiranja je prisutan na tržištu već oko 30 godina, nastao je za tadašnje R/S uređaje (usmjernike i preklopnike), a vremenom se razvijao i prilagođavao razvoju već postojećih tehnologija i pojavom novih (na primjer IP telefonija, pružatelji internetskih usluga, podatkovni centri, sigurnost računalnih mreža, i drugi). Ipak osnovni Cisco model je godinama zadržao istu strukturu: sastojao se od tri razine znanja (Associate, Professional i Expert) za početno R/S područje i poslije za nova područja mrežnih tehnologija koja su nastajala (mrežna sigurnost, pružatelji usluga, podatkovni centri, IP telefonija i drugi).

Pojava računalstva u oblaku i mrežnog programiranja unijele su značajne promjene u industriji računalnih mreža, što je kao posljedicu imalo najveću i najznačajniju promjenu u Cisco modelu certificiranja od samog nastanka. Novi Cisco program certificiranja je predstavljen u drugom mjesecu 2020. godine. U novom programu su prisutna, uz tradicionalna područja (R/S, pružatelji usluga, podatkovni centri, sigurnost, …), i potpuno dva nova područja: mrežno programiranje i automatizacija (DevNet) i CyberOps. Osim toga Cisco je svoj certifikat na razini suradnika (CCNA) objedinio i napravio kao osnovu za više tehnoloških područja kako bi pripremio polaznike za sve veću primjenu novijih tehnologija.

I ostali najveći svjetski proizvođači mrežne opreme (Huawei, Juniper, HPE i drugi) nude svoje programe IT certificiranja. Oni su međusobno vrlo slični (i vrlo vjerojatno napravljeni po početnom Cisco modelu, jer je Cisco najdulje prisutan na tržištu računalnih mreža i certificiranja u tom području), započinju na razini suradnika i idu sve do razine stručnjaka u različitim područjima umrežavanja. Isto tako prisutna su nova područja umrežavanja (kao na primjer mrežno programiranje).

Certifikati neovisni o proizvođaču, kao na primjer CompTIA Network+, pokrivaju samo osnovna znanja u području računalnih mreža i ne pokrivaju novije tehnologije.

U radu je prikazano da i dalje na godišnjim listama najtraženijih i najbolje plaćenih ICT certifikata u području računalnih mreža su prisutni uglavnom Cisco certifikati (na primjer CCNA, CCNP ili CCIE za različita mrežna područja).

Cisco CCNA certifikat široko se smatra najboljim mrežnim certifikatom za pokretanje karijere mrežnog inženjera. Većina novih mrežnih inženjera preferira Cisco put certificiranja zbog raspoloživih materijala za učenje i zbog značajne količine dostupnih mrežnih poslova koji zahtijevaju Cisco certifikat i iskustvo.

Slično na višim razinama znanja i iskustva, prevagu stvaraju Cisco certifikati (na primjer CCNP i CCIE) zbog svoje prepoznatljivosti i većih mogućnosti zapošljavanja.

Jedna od prednosti certificiranja u programu drugih proizvođača je ta da za poslove mrežnog stručnjaka za te tehnologije ima manju konkurenciju u odnosu na Cisco stručnjake. Broj dostupnih profesionalaca sa certifikatom drugih proizvođača mrežne opreme puno je manji od onih sa Cisco certifikatom. Tako na primjer kada organizacija zahtijeva skup vještina inženjera za Juniper proizvode, moraju platiti puno više za Juniper stručnjake u usporedbi sa Cisco stručnjacima.[35]

Ipak polaganje certifikata jednog proizvođača mrežne opreme (na primjer Cisco CCNP Enterprise) omogućava puno lakše polaganje odgovarajućeg certifikata drugog proizvođača iz istog područja i iste razine (na primjer Juniper JNCIP-ENT), jer su tehnologije koje se koriste uglavnom iste, a razlike u naredbama i specifičnosti uporabe uređaja se vrlo često brzo svladaju.

Cisco je trenutno i dalje vodeći na tržištu certificiranja za područje računalnih mreža, no putovi certificiranja utemeljeni na Cisco modelu nisu jedina ulaznica za svijet certificiranog mrežnog inženjera.

LITERATURA

[1] Miholjek M.: „Industrijski certifikati za računalne mreže“, diplomski rad na diplomskom studiju Informacijski sustavi, Veleučilište Velika Gorica, 2021. (mentor Valenčić D.)

[2] https://www.pluralsight.com/blog/it-ops/it-certification-101, učitano 22.01.2021.
[3] https://blog.eccouncil.org/the-importance-of-certifications-among-employers-and-professionals/, učitano 22.01.2021.
[4] https://www.computerworld.com/article/2713409/9-great-advantages-of-technical-certifications.html, učitano 22.01.2021.
[5] https://www.globalknowledge.com/us-en/resources/resource-library/articles/10-benefits-of-it-certification-for-you-and-your-employer , učitano 22.01.2021.
[6] https://www.ntinow.edu/blog/computer-certifications-important-employers/, učitano 22.01.2021.
[7] https://www.comptia.org/blog/most-popular-it-certifications, učitano 03.02.2023.
[8] https://www.skillsoft.com/it-skills-and-salary-report, učitano 03.02.2023.
[9] https://www.missioncriticalmagazine.com/articles/91924-report-ethernet-switch-router-market-at-an-all-time-high, učitano 03.02.2023.
[10] https://www.forbes.com/sites/greatspeculations/2022/05/02/a-network-to-rely-upon/?sh=53d6127a66d6, učitano 03.02.2023.
[11] https://www.cio.com/article/193586/top-15-it-certifications-in-demand-for-2021.html, učitano 03.02.2023.
[12] https://www.pcmag.com/news/highest-paying-it-certifications, učitano 03.02.2023.
[13] https://www.skillsoft.com/blog/15-top-paying-it-certifications-of-2022, učitano 03.02.2023.
[14] https://www.techtarget.com/whatis/feature/15-highest-paying-IT-certifications, učitano 03.02.2023.
[15] https://www.indeed.com/career-advice/career-development/top-it-certifications, učitano 03.02.2023.
[16] https://www.thebalancemoney.com/certifications-in-the-tech-industry-2071456, učitano 03.02.2023.
[17] https://www.comptia.org/blog/most-popular-it-certifications, učitano 03.02.2023.
[18] https://www.comptia.org/certifications/network, učitano 03.02.2023.
[19] https://www.comptia.org/certifications/network#examdetails, učitano 03.02.2023.
[20] https://en.wikipedia.org/wiki/Cisco, učitano 03.02.2023.
[21] https://www.cisco.com/c/en/us/about/corporate-strategy-office/acquisitions/acquisitions-list-years.html, učitano 03.02.2023.
[22] https://www.cisco.com/c/en/us/training-events/training-certifications/certifications.html, učitano 03.02.2023.
[23] https://www.cisco.com/c/dam/en_us/training-events/certifications/career-path.pdf, učitano 03.02.2023.
[24] https://learningnetwork.cisco.com/s/certifications, učitano 03.02.2023.
[25] https://en.wikipedia.org/wiki/Juniper_Networks, učitano 03.02.2023.
[26] https://www.juniper.net/us/en/training/certification.html, učitano 03.02.2023.
[27] https://www.juniper.net/content/dam/www/assets/training/us/en/certification-paths-by-credential.pdf, učitano 03.02.2023.
[28] https://e.huawei.com/en/talent/cert/#/careerCert, učitano 03.02.2023.
[29] https://certification-learning.hpe.com/tr/certifications, učitano 03.02.2023.
[30] https://www.h3c.com/en/Training_Certification/Technical_training___certification/Certification_Project/, učitano 03.02.2023.
[31] https://www.arista.com/en/, učitano 03.02.2023.
[32] https://www.rededucation.com/arista/#, učitano 03.02.2023.
[33] https://www.nokia.com/networks/training/src/certifications/, učitano 03.02.2023.
[34] https://www.extremenetworks.com/training/certifications/, učitano 03.02.2023.
[35] https://www.cbtnuggets.com/blog/technology/networking/juniper-vs-cisco-4-reasons-to-go-juniper, učitano 03.02.2023.

Projekt GAMMA: Igranje ili učenje?

Karolina Brleković* i Danijela Ivanović – Ižaković*

* Elektrotehnička i prometna škola Osijek, Osijek, Republika Hrvatska

karolina.brlekovic@skole.hr, danijela.ivanovic-izakovic@skole.hr

Za razliku od drugih predmeta STEM područja, matematika se smatra apstraktnim skupom znanja, izoliranim od stvarnog svijeta. Učenici često izražavaju osjećaj nepovezanosti matematičkih tema s realnim svijetom. Tradicionalne metode podučavanja predmet su kritike jer teško usmjeravaju učeničku pažnju i ne nude autonomiju konstruiranja razumijevanja. Učenje temeljeno na igrama, a posebno učenje temeljeno na digitalnim igrama zbog svoje interaktivne prirode, imaju obećavajući pristup učenju i poučavanju. Ovakvo učenje može biti učinkovito u povećanju motivacije učenika za stjecanje i razvijanje vještina koje nadilaze sadržaje koje nudi trenutni sustav obrazovanja matematike. Dio digitalnih igra zahtijeva od igrača da ovladaju vještinama (npr. strateško i analitičko razmišljanje, rješavanje problema, odlučivanje, kreaciju, brzu prilagodbu promjenama) koje su tražene na današnjem tržištu rada.

Tijekom provedbe projekta, sve igre su pilotirane kroz nastavu matematike u desetak srednjih škola iz Hrvatske, Nizozemske, Grčke i Finske. U pilotiranje je bilo uključeno ukupno 819 učenika. Učenici su igre testirali kroz nastavu matematike, te dali svoj kritički osvrt na njih i sugestije kako igre prilagoditi lakšem učenju učenika i učiniti ih zanimljivijim i vizualno atraktivnijim.

Ključne riječi – Game based learning, igrifikacija, matematika, scenariji poučavanja.

I. UVOD

GAMMA (GAMe-based learning in MAthematics) je dvogodišnji Erasmus + projekt koji se bavi idejom poučavanja matematike koristeći inovativne, kreativne i moderne metode uključujući svrhovitu uporabu digitalne tehnologije odnosno učenja temeljenog na digitalnim igrama (DGBL). Projekt GAMMA ima za cilj iskoristiti vrijeme koje učenici provode u korištenju elektroničkih uređaja, kao i njihovo znanje o digitalnim igrama, u svrhu poboljšanja matematičkih vještina. Glavni ciljevi projekta su izraditi i primijeniti:

- GAMMA priručnik za nastavnike,
- autorski sustav (FAS) za proizvodnja digitalnih igara,
- predložak za izradu scenarija poučavanja,
- 8 scenarija poučavanja temeljenih na digitalnim igrama,
- 5 digitalnih igara za ostvarivanje matematičkih ishoda učenja.

Usporedo s provedbom Erasmus+ projekta, u tijeku je i provedba eTwinning projekta kroz koji se učenike uključuje u kreiranje digitalnih igara, te se proces učenja podiže na višu razinu. Učenici jedni za druge putem digitalnih online alata provode radionice poučavanja o izradi igara na različitim platformama (primjerice Unity, GDevelop, Godot i sl.)

Projekt GAMMA započeo je suradnjom Odjela za matematiku iz Osijeka, Sveučilišta J. J. Strossmayera i Elektrotehničke i prometne škole Osijek u želji za novim pristupom u nastavi matematike koji bi uključivao uporabu digitalnih igara za učenje novih koncepata na drugačiji i učenicima prihvatljiviji način. U suvremenom svijetu postoji sve veća potreba za STEM (znanost, tehnologija, inženjerstvo i matematika) vještinama. To predstavlja veliki izazov između ostalih i za nastavnike matematike. Poučavanje matematike na inovativan, kreativan i suvremen način vodi smanjenju broja učenika s nedostatnim vještinama u području matematike, što je jedan od ciljeva Europske suradnje u obrazovanju i osposobljavanju (ET 2020). Uzimajući u obzir navedeno provedena je anketa među učenicima i nastavnicima, te su na osnovi rezultata ankete definirani glavni ciljevi koji su trebali odgovoriti na izazov uvođenja učenja temeljenog na igri (GBL) u nastavu matematike.

U projektu GAMMA uz koordinatora Odjel za matematiku, Sveučilišta u Osijeku partneri su: Srednja škola Markantuna de Dominisa Rab, Elektrotehnička i prometna škola Osijek, Agencija za odgoj i obrazovanje, First High School of Peristeri Athens, Porin Lyseon lukio, Ethniko Kai Kapodistriako Panepistimio Athens i Satakunnan Ammattikorkeakoulu oy stichting, Hogeschool Van Amsterdam.

Tijekom provedbe projekta, kroz aktivnosti učenja i poučavanja održanih virtualno i na sastancima u zemljama partnerskih organizacija, sudionici su izradili priručnik, učili o načinu korištenja FAS-a, izradili predložak za scenarije poučavanja, scenarije za digitalne igre i osam igara za učenje novih sadržaja iz matematike. Igre i scenariji su pilotirani na nastavi matematike u školama partnerima.

II. PREDSTAVLJANJE ISTRAŽIVANJA O UČENJU MATEMATIKE TEMELJENOM NA IGRAMA

Suvremene metode poučavanja pomiču se od tradicionalnog poučavanja ex katedra i uporabe udžbenika, ka interaktivnim i učenicima prihvatljivijim izvorima znanja. Tradicionalno poučavanje temelji se na konceptu kognitivnog opterećenja obrade informacija.

Primjenom DGBL-a učenici često niti ne promišljaju da igrajući igru uče. Učenje i viši kognitivni procesi odvijaju se na nesvjesnoj razini, a u prvom planu su

koncentracija na ostvarivanje što boljih rezultata u samoj igri i zadovoljstvo učenika [1].

Igre su zabavne aktivnosti koje nas opuštaju i angažiraju na različitim razinama primjerice kulturnim, društvenim, edukacijskim.

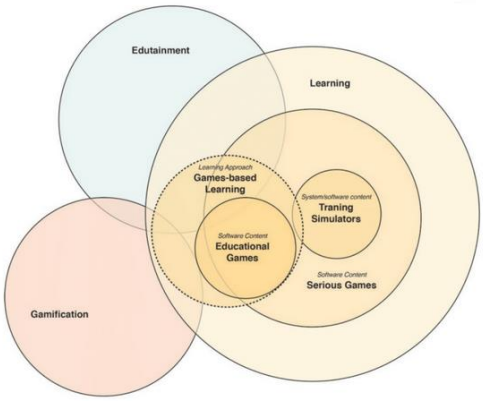
Koja je uloga igara u poučavanju? Igre općenito razvijaju maštu, znatiželju i čuđenje, te pridonose i psihičkom i intelektualnom razvoju pojedinca, posebice onih mlađe životne dobi. One nas potiču da razmišljamo, biramo i odlučujemo jačajući sposobnost koncentracije, samokontrole i fizičke koordinacije.

Želimo li uključiti GBL u nastavi, trebamo se prije svega zapitati što je igra?

- interaktivni izazov,
- niz zanimljivih i smislenih izbora koje igrač napravi u potrazi za jasnim i uvjerljivim ciljem,
- skup aktivnosti koje uključuju jednog ili više igrača,
- sustav u kojem se igrači upuštaju u umjetni sukob, definiran pravilima, koji rezultira kvantificiranim ishodom. [2]

Razlikujemo pojam gamifikacije (igrifikacije) i učenja temeljenog na igri (GBL). Gamifikacija je proces integracije igre u već postojeće aktivnosti ili procese kako bi se povećalo sudjelovanje i angažman sudionika [3]. U gamifikaciji, elementi igre integrirani su u aktivnosti koje nisu igra kako bi svakodnevne aktivnosti, na primjer nastava bile ugodnije, a pri tome se koriste bodovi, pravila igre, statusi i nagrade za poticanje angažmana učenika.

Učenju temeljenom na igri (GBL) nije cilj stvaranje igara kako bi se učenici igrali, to je osmišljavanje aktivnosti učenja, kroz koje se mogu postepeno uvoditi novi koncepti, te voditi učenike prema krajnjem cilju – svladavanju pojedinih ishoda učenja. Struktura GBL-a prikazana je na Slici 1.



Slika 1. Struktura učenja temeljenog na igri [4]

Kod učenja temeljenog na igri (GBL) koristimo određene mehanizme igara poput unaprijed definiranih pravila, strateškog promišljanja, poticanja natjecateljskog duha i sl. kako bismo unaprijedili i/ili ubrzali učenikov proces učenja [5]. Kod učenika se na taj način potiče proces kritičkog mišljenja i razvijaju vještine rješavanja problema.

Zašto GBL?

- Angažira i motivira učenike,
- Kombinira vizualne i slušne inpute, te verbalno i neverbalno → čini učinkovito okruženje za učenje,
- Potiče razmišljanje izvan okvira i stvara osjećaj kontrole, jer igrač ponavlja pokušaje, doživljava neuspjehe, ali i uči na neuspješnim pokušajima→ aspekti učenja kroz rješavanje problema – neuspjeh = prilika za učenje,
- pozitivan neuspjeh može poboljšati povjerenje učenika u matematiku.

Ipak kod uvođenja GBL-a treba razmotriti učinak dodavanja sadržaja za učenje u igru koja će učenike zaista motivirati, te voditi računa o učenicima s malo iskustva kojima će možda biti teško odabrati relevantne informacije u nestrukturiranim i otvorenim okruženjima za učenje.

U projektu GAMMA učenje temeljeno na digitalnim igrama proučavali kroz platforme za izradu edukacijskih igara ChoiCo http://etl.ppp.uoa.gr/choico/ (04.02.2023.)

i MaLT2 http://etl.ppp.uoa.gr/malt2/?equescapegame (04.02.2023.) i već postojeće igre.

Kroz projekt su razvijene i nove digitalne igre:

- E(qua)scape room – sustav linearnih jednadžbi,
- Hot air balloon game – proporcionalnost,
- Function dungeon – linearna funkcija,
- Balloon game – derivacije,
- GeomWiz – geometrijski likovi, površina trokuta, poučak o sinusuma i poučak o kosinusu, Pitagorin poučak,
- JoJo Birdgame – grafovi trigonometrijskih funkcija,
- ProbChallenge – vjerojatnost.

III. IGRANJE ILI UČENJE?

Digitalne igre svakodnevnica su ne samo za učenike nego i za odrasle osobe. Često posežemo za ne pretjerano logički zahtjevnim „slagalicama“ kako bismo relaksirali um. Za učenike su digitalne igre dio njihove „komfor“ zone, u koju u slučaju DGBL-a ulazimo s namjerom da učenje učinimo drugačijim, suvremenijim i opuštenijim. Namjera nam je učenike motivirati za rad i na određeni način im olakšati usvajanje ishoda učenja. Digitalne igre zahtijevaju od igrača vještine kao što su strateško i analitičko razmišljanje, rješavanje problema, odlučivanje, kreaciju i brzu prilagodbu promjenama što su vještine koje se usvajaju i tijekom rada na matematičkim zadacima.

Ukoliko u proces poučavanja želimo uključiti DGBL, nastavni sat treba dobro unaprijed isplanirati i pripremiti:

- provjeriti tehničke mogućnosti prostora i sudionika,
- pripremiti materijale za rad s učenicima,

- odigrati igre koje želimo koristiti kako bismo učenicima mogli pomoći,
- upoznati učenike s novim konceptom rada,
- dobro složiti učeničke timove,
- pratiti, poticati i pomagati učenicima tijekom rada.

Naravno, ovakav način poučavanja od nastavnika traži puno više uloženog vremena i truda u odnosu na pripremu tradicionalnih nastavnih sati.

Za pripremu sata temeljenog na DGBL-a kroz projekt GAMMA razvijen je obrazac za pripremu sata kroz koji se nastavnik može učinkovito pripremiti za izvođenje sata. Nastavnik sam odabire koje od ponuđenih rubrika smatra bitnim za svoju pripremu kao što su primjerice:

- potrebno predznanje učenika,
- vrijeme učenja,
- izvori (računalo, tablet, mobilni uređaj,…),
- korelacija s drugim predmetima,
- aktivnosti (uvod, upute, igranje igre, formativna procjena,….),
- povratna informacija učenika,
- i slično.

Kroz projekt su osim osam digitalnih igara razvijeni i scenariji poučavanja za navedene digitalne igre. Igre i scenariji poučavanja su pilotirani na 819 učenika u desetak škola u četiri zemlje. Kroz proces pilotiranja igara u učionicama provedenih u partnerskim školama u projektu je postalo jasno da su scenariji kod izvođenja nastave temeljene na DGBL bez obzira na iskustvo nastavnika itekako potrebni. Učenike je potrebno usmjeravati, pa čak ponekad, i voditi kroz igru, te voditi računa o pravilima igre. Igre koje koristimo u DGBL-u ponešto se razlikuju od igra koje učenici obično igraju za relaksaciju i to može biti pomalo zbunjujuće. Želja za sakupljanjem bodova i prelaskom na sljedeću razinu igre može biti veća od potrebe da se usvoje novi ishodi. DGBL treba uvoditi postupno i učenike treba naviknuti na drugačiji način rada uz puno strpljenja, jer ostaje pitanje: Igranje ili učenje? Ako sat temeljen na DGBL-u nije pripremljen i vođen na pravilan način, igranje će ostati igranje i neće prijeći u učenje.

IV. AUTORSKI SUSTAVI ZA IZRADU DIGITALNIH IGARA

IV.I. PLATFORMA CHOICO

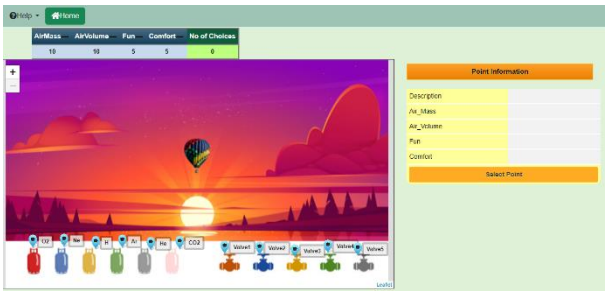
Načela rada platforme razvijene u Grčkoj prije samog projekta, proučavali smo kroz igre postavljene na njoj. Tijekom trajanja projekta, na istoj platformi su osmišljene i kreirane i nove igre. Igre se, kako i sam naziv platforme sugerira, temelje na nizu odabira (engl. *choice*) koji imaju unaprijed definiranje posljedice.

Kroz projekt je na platformi izrađena igra Hot-Air Balloon (Slika 2.) kao i Balloon Game koje se bave konceptima derivacija i proporcionalnosti.

Primjerice, u igri Hot-Air Balloon učenici trebaju dodavati različite plinove ili ih pomoću različitih ventila ispuštati. Cilj je kroz 10 odabira koji čine jednu razinu održati balon u zraku, imajući na umu da svaka odluka utječe na niz parametara.

Ukoliko igrač uspije kroz 10 odabira održati parametre u željenim intervalima, prelazi na slijedeću razinu. Kroz igru igrači uče kemijska svojstva plinova, fiziku i naravno matematiku.

Na početku igre je uputa za igrača (učenika) u kojoj je navedeno u kojim intervalima treba zadržati pojedini parametar.



Slika 2. Sučelje igre ChoiCo Hot-Air Balloon

Svaki od plinova O_2, Ne, H, Ar, HE i CO_2 ima parametre prikazane u tablici koja se otvara klikom na oznaku uz bocu što igraču omogućuje lakši odabir, odnosno predviđanje kako će njegov odabir utjecati na pojedino svojstvo balona.

Korisnici platforme mogu na njoj izraditi svoju vlastitu igru ili *uploadati* igru s vlastitog računala te igrati igre koje su već na platformi i učiti kroz igru.

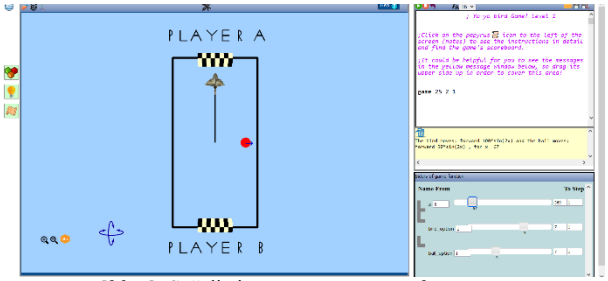
IV.II. PLATFORMA MALT2

Platforma MaLT2 (Slika 3.) temelji se na LOGO-u odnosno kornjačinoj grafici.

Korisnik može pišući programski kod vježbati matematičke koncepte i pomicati vrapca u koordinatnom sustavu, te doći do određenih zaključaka vezanih uz matematičke paradigme. Igre koje su na ovoj platformi razvijene su JoJo Bird game i E(qua)scape room. Iako su obje igre razvijene na istoj platformi, pristup i metodologija koja ih pokreće su drugačije.

JoJo Birdgame je igra za dva igrača u kojoj se kombiniraju potezi zadani trigonometrijskim funkcijama u cilju povezivanja pokreta ptice i jojo-a ovisno o amplitudi i periodu trigonometrijske funkcije. Igra zahtjeva visoku koncentraciju učenika i pogodna je za učenike koji žele više naučiti.

U igri E(qua)scape room platforma služi kao simulator kako bi učenik riješio zadane problemske zadatke iz područja sustava linearnih jednadžbi.



Slika 3. Sučelje igre MaLT2 JoJo Birdgame

V. GAMMA ETWINNING PROJEKT

Zbog boljeg širenja rezultata Erasmus+ projekt GAMMA proširen je na eTwinning projekt „GAMMA za sve“. Provedba eTwinning projekta planirana je kroz period od godinu dana. Kroz projekt ulazimo u dimenziju u kojoj učenici samostalno ili u timovima osmišljavaju pristupe i kreiraju digitalne igre za učenje. U projekt je uključeno više srednjih i osnovnih škola iz Republike Hrvatske, BIH, Republike Srbije, Portugala i Nizozemske.

Kroz projekt provodimo sljedeće aktivnosti:

- Predstavljanje škola, učenika i nastavnika,
- Izrada logotipa projekta,
- Ulazna i izlazna anketa za nastavnike i učenike – kako bismo usporedili ostvarenost planiranih ciljeva,
- Webinari – „Učenik za učenika – kako koristim alate za izradu digitalnih igara“,
- Izrada jednostavnih edukacijskih igara primjerice na platformama za izradu GDevelop i Unity,
- Pilotiranje scenarija i igara izrađenih u GAMMA projektu od studenog do veljače.

Cilj nam je proširiti i nadograditi znanja stečena u projektu GAMMA. Posebice su nam važne povratne informacije koje planiramo prikupiti od učenika koji će izrađivati edukativne igre. Primjerice o njihovim iskustvima tijekom izrade igara s stanovišta povezivanja mehanizama igara s edukacijskim sadržajem, odnosno kako su postavili ciljeve učenja, te mjerili postignuća i napredak učenika kroz igru. S druge strane, također je važno i na koji način su učenici analizirali povratne informacije drugih učenika koji su igre igrali, te kako su i jesu li ih upotrijebili u svrhu poboljšanja igara.

VI. ZAKLJUČAK

Projekt GAMMA sve uključene u projekt uveo je u svijet Game Base Learninga. GBL nema dugu povijest. Veći dio igara koje smo analizirali kroz istraživanje i sam projek,t ili je za niže uzraste, ili je za uvježbavanje već poznatih koncepata. Digitalnih igara za usvajanje novih sadržaja i koncepata, gotovo da nema. Kroz projekt je razvijeno osam digitalnih igara za učenje novih ishoda u matematici sa pripadajućim scenarijima, te priručnik za nastavnike koji žele u svoju nastavu uvesti GBL. Dio igara i scenarija je pilotiran i učenici i nastavnici su anketirani. Dio igara naišao na dobar odaziv i kod učenika i kod nastavnika, dok dio nije.

Najveći izazov kod izrade novih igara je povezivanje digitalnih igara s brzinom i akcijom što je kod igara kreiranih za učenje nespojivo s konceptima učenja. Ipak od ovog pristupa ne treba prebrzo odustati, jer učenicima daje dodatnu motivaciju za rad i istraživanje.

Vjerujemo da se rad na GBL kod nas ali i šire treba nastaviti i da je priručnik koji je razvijen korisna literatura za sve koji žele GBL uvesti u nastavu, te i sami izraditi igre za učenje i poučavanje.

REFERENCE

[1] Boaler. J. (2002.), Learning from Teaching: Exploring the relationship between reform curriculum and equity,

[2] Hébert, C., Jenson, J. (2019). Digital game-based pedagogy: Developing teaching strategies for game,

[3] Spires. A. (2015.) Digital Game-Based Learning,

[4] Martens, A.; Mueller, W. (2016.) Gamification - A Structured Analysis,

[5] Trybus, J. (2015.). Game-Based Learning: What It Is, Why It Works, and Where It’s Going. Miami: New Media Institute, based learning. Journal of Interactive Technology and Pedagogy

Čarobni svijet snjegovića knjigoljubaca – izrada digitalne slikovnice

S. Barbarić, Osnovna škola Milana Langa, Bregana, Samobor, Hrvatska
sbarbaric2@gmail.com

Sažetak **– Autorica rada za učenike nižih razreda osnovne škole osmislila je izradu digitalnih slikovnica kao projekt koji se može realizirati u okviru zadanog kurikuluma nastave informatike. Cilj projekta bio je pokazati koliko pisanje i crtanje na računalu može biti kreativno, a učenje zabavno. Učenici su projektom trebali uočiti kako praktičan rad rukama povezati s informatičkom tehnologijom, programima i aplikacijama. Kao idejni početak projekta, učiteljica je svakom učeniku isplela zimsku kapicu. Zatim bi od prigodnog materijala zajednički izradili maketu grada u kojem žive stanovnici s pletenim kapicama. Plan je bio da će učenici na tu temu pisati priče i crtati na računalu. Učenici su podijeljeni u tri ciljane grupe. Jedna grupa učenika bavila bi se ručnim radom, izradom makete i pletenjem. Druga grupa pisala bi tekst, crtala i slikala uz pomoć aplikacija. Treća grupa, najstarija, imala bi zadatak napraviti digitalni rad od teksta, crteža i fotografija. Digitalni rad mogao je biti video materijal, film ili digitalna slikovnica. Projekt bi završio tako da se medijski materijal objavi na web stranici školskog lista Breganček.**

Kad je učiteljica donijela nove pletene zimske kape za učenike u područnu školu Grdanjci nastalo je veliko veselje. Nakon slikanja i obrade fotografija izradili su video koji su nazvali Kap dobrote = Kapa na dar. Na istom satu video je i objavljen na web stranicama digitalnog školskog lista https://bregancek.net/?p=1213 kako bi bio dostupan svim učenicima i njihovim roditeljima. Slijedeći korak projekta bio je s učenicima matične škole Bregana stvoriti Čaroban grad snjegovića knjigoljubaca u malom. Grad i stanovnike napravili su od drvenih kućica, autića-igračaka, stiropora, platna i ručno pletenih kapica. Autorica je, kao primjer kako treba izgledati digitalni rad, napravila oglednu slikovnicu „Dobro došli u SnjegoGrad“. Objavljena je na web stranici https://bregancek.net/?p=1225. SnjegoGrad je među učenicima bio jako dobro prihvaćen, a film o tome sam su napravili: https://bregancek.net/?p=1225. Svaki razred odabrao je jednu od zadanih tema o snjegovićima, a zajednički su odredili naslov svoje slikovnice. Svaki učenik dobio je mogućnosti osmisliti i urediti jednu njenu stranicu. Slikovnice su vrlo brzo bile završene, uređene i objavljene na https://bregancek.net/?p=1249. Ispisani i uvezeni primjerci slikovnica nalaze se sad u školskoj knjižnici i dostupni su učenicima u tiskanom obliku. Može se zaključiti kako su učenici s veseljem odradili planirani projekt. Osim praktičnog znanja i primjene aplikacija, naučili su kako su pletenje i izrada makete grada vrlo slični programiranju, rade se petlju po petlju, korak po korak, naredbu po naredbu, do konačnog rezultata.

Ključne riječi – digitalna slikovnica, snjegovići, snjegograd, breganček, kape, knjigoljupci, pletenje.

I. UVOD

U kurikulumu nastave informatike za niže razrede osnovne škole navedeni su ishodi učenja tog izbornog predmeta. Tu stoji kako učenici tijekom četiri godine izborne nastave moraju savladati vještinu pisanja i crtanja. Već u prvom razredu trebaju naučiti prepoznati slova na tipkovnici, ponešto nacrtati i motorički raditi rukama i mišem. Zadatak drugog razreda je znati pisati i čitati riječi i rečenice uz pomoć aplikacija na računalu. U trećem i četvrtom razredu u ishodima učenja nalazi se logika programiranja te prepoznavanje i izrada medijskog materijala. Tijekom prva četiri razreda osnovne škole učenici su obavezni naučiti korake i načine kako radi i „misli“ računalo. Moraju znati hoće li „poslušati“ naše želje i naredbe, i što napraviti da bi tako bilo. Isto tako, svi učenici nižih razreda trebaju savladati upotrebu osnovnih aplikacija kao što su Word, Paint i Bojanje 3D. Tijekom četvrtog razreda u nastavnom planu informatike nalazi se PowerPoint aplikacija koja bi sva stečena znanja trebala objediniti.

Tijekom godina učenja informatike djeca praktično nauče koji su sve dijelovi računala korisni i jednostavno se povežu u smislenu cjelinu. Izrađuju različite medijske sadržaje koji se kasnije mogu višekratno upotrebljavati. Cilj je naučiti ih koristiti informatičku tehnologiju za obradu, slušanje i gledanje slika, filma i slušanje muzike. Mnogi od tih pojmova za male dječje glavice su potpuno nerazumljivi. Nakon puno godina rada u nastavi i s djecom pokazalo se da će učenici dobro upamtiti samo ono što vide zorno i u praktičnoj primjeni.

Upravo zato autorica ovog rada je za učenike nižih razreda Osnovne škole Milana Langa u Bregani i područne škole Grdanjci osmislila izradu slikovnice kao projekt. Realizirala ga je zajedno s njima. Pripremila je potrebne materijale kao što su papiri, vuna, drvo, platno, igračke i alati. Učenicima su bili dostupni elektronički uređaji koji se mogu koristiti u radu na projektu. To su računalo, mobiteli, kamera, fotoaparat, pisači, zvučnici i mikrofon, te aparati za plastificiranje i uvezivanje. Da bi učenici mogli sami koristiti uređaje, učiteljica ih je najprije poučila kako svaki pojedini radi i što se njime može. Pustila ih je da pokušaju napraviti nešto svoje. Učenici su pohvaljeni za svaki pokušaj i kreativnu ideju, bez obzira da li je pri tome sve bilo točno odrađeno ili nije. Tako im je približila zadanu temu i informatičku tehnologiju. Dugoročni plan bio je iskoristiti znanje o digitalnoj tehnologiji i izraditi slikovnice uz pomoć računala.

II. ČAROBNI SVIJET SNJEGOVIĆA

Projekt koji je autorica ovog rada provela i radila s učenicima tijekom siječnja i veljače 2022. godine nazvala je „Čarobni svijet snjegovića knjigoljubaca“. U početku je učenicima objasnila osnovni cilj projekta i njegove faze izrade. Prva faza bila je izrada praktičnog zadatka – pletenje kapa i izrada gradića iz mašte. Učiteljica se koristila jednostavnim riječima i primjerima koji su učenicima poznati. Pokazala je fotografije kako bi pletene kape i grad SnjegoGrad mogli izgledati, što bi se tom gradu moglo nalaziti. Nabrojila je moguće likove iz mašte i figurice/snjegoviće koji bi trebali biti stanovnici tog gradića. Svi bi trebali imati i pletene kape na glavi.

Cilj projekta je zainteresirati i potaknuti učenike na pisanje (u aplikaciji Word), crtanje (u aplikacijama Paint, Bojanje 3D). Radom na računalu učenici će pokazati kreativan rad na računalu. U tu svrhu mogu koristiti, fotografije i maketu grada u malom. Učenike treba naučiti kako se nakon pisanja i crtanja jednostavnim alatima i korištenjem samo aplikacije PowerPoint, od autorskog teksta i crteža, može izraditi lijepa digitalna slikovnica. Od fotografija i slika može se napraviti video uradak. Digitalne materijale moguće je staviti na web stranicu kako bi bili dostupni svima. U slučaju škole Milana Langa svi materijali učenika objavljuju se na web stranici digitalnog školskog lista Breganček. Za učenike nižih razreda škole to nije nimalo lak zadatak. Puno je novih pojmova koje su trebali naučiti.

U prvom tjednu projekta autorica je učenike motivirala na rad uz prigodan ručni rad. Za svakog učenika područne škole isplela je zimsku kapicu s pamplonom. U planu projekta i podjeli zadataka učenici područne škole bili su zaduženi za pletenje. Učiteljica je pokazala kako mogu uzeti u ruke igle za pletenje i izraditi bilo kakav koristan predmet od vune. Prvi pokušaj bio je napraviti šal, A malo spretniji učenici pokušali su napraviti kapice za lutke/snjegoviće.

Ideja projekta bila je da će svaka grupa učenika biti zadužena za jedan dio zadatka. Prva grupa učenika će biti zadužena za pletenje, druga za slikanje i fotografiranje, treća za medijsko praćenje, a učenici četvrtih razreda od slika će napraviti video materijal. Tako je na kraju i bilo ostvareno.

Kad je učiteljica donijela ispletene kape učenicima na dar, u područnoj školi više nikoga nije trebalo nagovarati na rad na projektu. Učenici su stavili kapice na glave, u ruke uzeli igle i klupka vune i veselo prionuli zadatku. Pokušali su sami napraviti pletivo. Pri tome koncentrirano su pratili pokrete učiteljice i nastojali ih zorno kopirati. Nastava je bila ogledni primjer nastave uz pomoć metode imitacije. Svakom izrečenom rečenicom učiteljica je naglasila kako je pletenje identično programiranju u bilo kojem programskom jeziku. Programski jezik je za male učenike potpuno strani i nerazumljiv pojam. Ali, dok su učenici u rukama držali igle za pletenje, naučili su kako se plete tako da se radi petlja po petlja. Naučili su da je baš tako i kod programiranja. Radi se korak po korak, naredbu po naredbu, da bi stigli do željenog rezultata. Pokazalo se da među učenicima ima puno talentiranih malih ručica. Nitko od učenika (a kasnije i roditelja ili kolega učitelja) nije mogao zamisliti da se pletenje može povezati s radom na računalu.

Nakon slikanja sa svojim i novonapravljenim kapicama učenici su zajednički obradili fotografije. Od fotografija izradili su video uradak koji je objavljen na stranici: https://bregancek.net/?p=1213 Video su nazvali Kap dobrote = Kapa na dar. Prepoznali su trud učiteljice uložen u izradu ručno pletenih kapa. Kako je to bilo vrijeme početka godine, na satu informatike učenici i učiteljica izradili su prigodne kalendare. Kalendare su zajednički ispisali i plastificirali, te ponijeli kući na radost roditelja.

Još jednom se pokazalo da su djeca i učitelji u područnim školama jedna velika obitelj u kojoj vlada neizmjeran osjećaj pripadnosti, prijateljstva, ljubavi i bliskosti. Kapice i kalendari koji su završili kod učenika kao zajednički rad, cijele godine spominjali su se na svakom satu informatike kao najzabavniji događaj te školske godine.

Nakon toga, za slikovnice zajedničkog naziva „Čaroban grad snjegovića knjigoljubaca“ učenici matične škole Bregana i autorica ovog rada izradili su maketu grada od lijepo obojenih drvenih kućica i autića - igračaka. Učiteljica je napravila snjegoviće knjigoljupce od stiropora, malo platna, nekoliko gumbića i ručno pletenih vunenih kapica s pamplonima. Učenici su šareno bojali kućice i donosili igračke. Glavne zvijezde gradića imena SnjegoGrad bili su snjegovići knjigoljupci.

Kao dodatni uvod u temu učiteljica je smislila priču i izradila prvu slikovnicu pod nazivom „Dobro došli u SnjegoGrad“. Ona je bila i dodatna motivacija učenicima za upoznavanje likova, pisanje i crtanje priča. Učenici su tako iz primjera mogli vidjeti kako digitalna slikovnica treba izgledati. Čitanje slikovnice u kojoj se nalaze poznati izmišljeni likovi kod učenika razvilo je ljubav prema čitanju i pisanju. Htjeli su i oni napraviti nešto slično. A spoznali su da im informatička tehnologija daje priliku kreativno povezati tekst i slike/animaciju. Na web stranici škole breganček može se vidjeti digitalni zapis male slikovnice i mogućnost šetnje virtualnim gradom iz mašte: https://bregancek.net/?p=1225. Igra sa snjegovićima manjem uzrastu učenika bila je jako zabavna. Teško su se odvojili od igračaka koje su mogli rukom opipati. No, kasnije su zato brzo i lagano osmislili vlastite priče o stanovnicima SnjegoGrada. Tema za priče bilo je jako mnogo. Učenici i učiteljica zajednički su odabrali one najbolje koje su našle mjesto u slikovnicama.

Ovim projektom učenici su se trebali baviti tijekom mjesec dana nastave. Tako je bilo planirano. Ali zbog

korone i izolacija, mnogih učenika nije bilo na nastavi uživo i projekt se odužio. Bilo je tjedana kad se u matičnoj školi radilo uživo, a u područnoj školi to nije bilo moguće. I obrnuto. Činilo se da slikovnice nikad neće biti gotove. Svi su bili svjesni toga da će možda i biti gotove, ali ne u zimi, iako je tema slikovnica bila baš zimska. Kako je izgledao rad u razredu vidi se na web stranici: https://bregancek.net/?p=1239

Za uspjeh projekta dobro je bilo što se gradić iz mašte učenicima jako svidio. Priče o njemu nastavili su pisati i kod kuće. Donosili su tekstove i crteže u školu nakon izolacija i izbivanja zbog bolesti. Tekstova je na kraju bilo i previše. Svaki razred odabrao je svoju temu o snjegovićima. Zajednički, glasanjem, odredili su naslov i prigodne priče. Svaki učenik dobio je pravo popuniti tek jednu stranicu prezentacije tj. slikovnice. Tu je mogao staviti autorski tekst i vlastiti crtež. Učenici su koristili alate koje dobro poznaju. Nije bilo poteškoća s aplikacijama. Za uređivanje digitalne slikovnice i skraćivanje teksta te odabir crteža, bila je zadužena učiteljica.

Na kraju, zajedničkim snagama učenika i učiteljice, digitalne slikovnice dobile su zasluženo mjesto na web stranici škole i digitalnom školskom listu Breganček: https://bregancek.net/?p=1249. Učenici su izradili i nekoliko tiskanih primjeraka, plastificirali ih i uvezali spiralama. Tako završene sada krase učionice škole. Napravljeni su i primjerci koji se mogu posuditi u školskoj knjižnici i dostupni su ostalim učenicima.

Čim se završilo s projektom o snjegovićima, a to je potrajalo do prvih dana proljeća, učenici su odmah predložili nastavak priče – novu slikovnicu. Cijela priča oko slikovnica im se dopala. Imali su i ideju o čemu žele pisati u nastavku: o proljeću, cvijeću i kukcima… Htjeli su da učiteljica napiše novu slikovnicu pod nazivom „Dobro došli u KukcoGrad“ i smisli nove likove: kukce, najviše leptire, pčelice i bubamare… Odmah nakon toga oni bi stanovnike Kukcograda maštovito umetnuli u svoje nove priče veseleći se proljeću.



Slika 1. Digitalne slikovnice PŠ Grdanjci, 3.A, 4.A i 4.B



Slika 2. Digitalne slikovnice 3.B, 3.C i 4.C

III. ZAKLJUČAK

Pokazalo se da je tema projekta dobro odabrana. Projekt je zainteresirao i razveselio učenike uzrasta prvih, drugih, trećih i četvrtih razreda. Učenici su mnogo toga novog naučili. Tema im je bila tako zanimljiva da nisu niti primijetili kako uče i usvajaju novo gradivo.

No, znanje koje su oni u dosadašnjem učenju informatike stekli o izradi prezentacije, obradi slika, i izradi video materijala nije bilo dovoljno za samostalnost u radu i obradu podataka. Tu im je trebala pomoć učiteljice. Također, nisu sami mogli stavljati materijale na web stranicu škole. Na tome još u budućnosti trebaju učiti i raditi. Ali sigurno je, s velikim veseljem upijaju sve novo.

Tijekom nastavne godine mnogi učenici bili su bolesni ili u izolaciji zbog korone. Jedini način komunikacije s njima bio je moguć putem informatičke tehnologije. Upravo stoga ovim projektom pokazuje se da pisana i čitana digitalna riječ ima smisla. Može se svakodnevno povezati s mnogo stvari koje su bliske učenicima: igračkama, drvenim kućicama, pletenjem, šivanjem... prirodom.

Sve to potrebno je našim učenicima kako bi im djetinjstvo bilo raznobojna, a ne sumorna stvarnost. Moramo im pokazati da život nije samo ekran računala. Istovremeno trebaju uočiti kako je baš taj ekran jako važan i što se sve s računalom može. Računalo je koristan alat koji olakšava život i donosi nove radosti. Oplemenjuje ga i pomaže na način na koji ljudi to žele i dopuštaju.

U školi odgajamo malu djecu kako bi postali veseli, radoznali, i maštoviti ljudi. Knjigoljupci. Slikari. Informatičari. Kreativci. Dječja mašta nema granica. Baš u tu svrhu nastao je projekt Čaroban svijet snjegovića knjigoljubaca - izrada digitalnih slikovnica. Veslelilo bi da takvih projekata u školama ima više.

IV. ZAHVALA

Zahvaljujem ravnatelju Osnovne škole Milana Langa, Bregana Igoru Matijašiću što mi je omogućio provođenje ovog projekta s učenicima nižih razreda matične škole u Bregani i područne škole u Grdanjcima.

REFERENCE

[1] Godišnji kurikulumi nastave informatike nižih razreda osnovne škole za školku godinu 2021/2022. https://mzo.gov.hr/vijesti/okvirni-godisnji-izvedbeni-kurikulumi-za-nastavnu-godinu-2021-2022/

[2] Upute za pletenje https://pletenje.net/izrada-predmeta-pletenje/

[3] Uputa za izradu makete grada u malom https://www.skolskiportal.hr/sadrzaj/iz-skolskog-svijeta/izradite-maketu-svoga-kvarta-i-mjesta/

[4] Upute za rad s aplikacijama Word, Paint, PowerPoint https://skolazazivot.hr/wp-content/uploads/2020/06/Office365_priru%C4%8Dnik.pdf

[5] Upute za rad s aplikacijom Bojanje 3D https://support.microsoft.com/hr-hr/windows/osnovno-3d-modeliranje-pomo%C4%87u-programa-bojanje-3d-78a27393-4cc5-1c9a-5929-3b4644eb5a94

[6] Upute za rad s uređajem za plastificiranje i uvezivanje https://data.fast.eu/manual/45/45007721/45007721-im-hr.pdf

[7] Web stranica digitalnog školskog lista Breganček Osnovne škole Milana Langa, Bregana https://bregancek.net/

Stavovi i percepcije učenika i studenata prema plagiranju

Marina Mirković i Darko Možnik
Tehnička škola, Požega, Republika Hrvatska
Hrvatsko vojno učilište "Dr. Franjo Tuđman", Zagreb, Republika Hrvatska
marina.mirkovic@skole.hr
darko.moznik@gmail.com

Sažetak

Današnji učenici, budući studenti odrastaju s moćnim tehnologijama od najranije dobi. One im omogućuju pristup, dijeljenje, kopiranje, generiranje i suradnju na kreativnom radu na načine koji se neprestano razvijaju. Pristup webu, velikim bazama podataka dovodi do plagiranja koje postaje ozbiljan problem za izdavače, istraživače i obrazovne ustanove. Ovaj rad analizira istraživanje provedeno putem internetske ankete o procijeni stavova i percepcije prema plagijatu i referenciranju.

U Tehničkoj školi u Požegi anketa je provedena među učenicima trećih i četvrtih razreda, a na Hrvatskom vojnom učilištu "Dr. Franjo Tuđman" u Zagrebu među studentima treće godine prijediplomskog studija smjerova vojno inženjerstvo i vojno vođenje i upravljanje.

Sudionici anketiranja iskazali su svjesnost problema, pokazali nužnost da se ovom temom treba češće baviti i raspravljati. Upoznati su sa alatima za provjeru plagijata, ali nisu svugdje definirane sankcije za plagiranje, odnosno pravilnici koji naglašavaju ovaj problem.

Najbolji način da se spriječi plagijat je educirati učenike o tome kako ispravno provoditi istraživanje, citirati i stvarati jedinstven i originalan rad.

Studenti moraju znati kako komunicirati s kreativnim radom na načine koji su etički i zakoniti. Razumijevanje osnovnih pravila o kreativnom radu bitan je element uspješnog i angažiranog digitalnog građanina. Autorsko pravo postaje bitan element digitalne pismenosti, važno za budućnost kreativnosti i interneta.

Ključne riječi - autorsko pravo, plagijat, citiranje, digitalna pismenost

I. UVOD

Plagijat i varanje učenika oduvijek je jedan od najneugodnijih trenutaka u nastavi s kojima se suočavaju nastavnici. Ovi problematični incidenti često su i savršeni trenuci za poučavanje učenika da uče iz svojih pogrešaka i da ih ne ponavljaju kasnije u životu.

Plagijat je postao globalni problem, događa se u cijelom svijetu. Od lokalnih srednjih škola do prestižnih sveučilišta, bezbrojni oblici varanja gotovo su stalna akademska briga. Načini na koje učenici i studenti varaju na zadaćama, ispitima i u svojim radovima, poprimaju sve jedinstvenije i suptilnije oblike. Uslijed pandemije COVID-19, sve više škola prešlo je na online nastavu, a učenici su postali sve vještiji u varanju i plagiranju svojih radova.

Plagijat može imati nekoliko različitih oblika. Najčešći uključuje „kopiranje i lijepljenje“ tuđih riječi bez navođenja citata ili navođenja autora. Drugi tip uključuje uzimanje nekoliko različitih izvora, njihovo spajanje i predstavljanje kao novih ideja i misli. Ponekad je plagijat očit. Na primjer, nastavnici mogu primijetiti da pisanje u određenom dijelu rada ne odgovara stilu ostalog dijela. Ponekad učenici mogu preformulirati neki tekst tek toliko da zvuči kao njihov izričaj. To još uvijek predstavlja krađu ideja druge osobe. Kako bi internet funkcionirao na slobodan i otvoren način, važno je da građani, a posebno mlade osobe odluče živjeti etički u svojim digitalnim okruženjima [1].

Školski propisane posljedice za nenamjerno plagiranje ovise o pojedinoj ustanovi. Većina Kodeksa časti i pravila o akademskoj čestitosti jasno pokazuju da odgovornost leži na studentima, bez obzira na to jesu li namjerno plagirali ili ne. Imajući to u vidu, neki počinitelji kod prvog prekršaja mogu dobiti blažu kaznu ako se ispričaju i iskreno govore o svojoj zaboravnosti i/ili neznanju. Čini se da pojedine škole postaju vlastiti istražitelji, suci i porota što je dominantan svjetski trend. Ovi se problemi obično rješavaju interno, pri čemu škola odlučuje hoće li učeniku kod kojeg je otkriven plagijat dati suspenziju ili isključivanje iz ustanove. Kako bi procijenili stavove i percepciju učenika i studenata prema plagijatu i referenciranju, provedeno je istraživanje putem internetske ankete.

II. ANALIZA ANKETIRANJA UČENIKA I STUDENATA

Anketa je provedena u siječnju 2023. godine. Sudjelovalo je 90 učenika trećih i četvrtih razreda Tehničke škole u Požegi, među njima 92,2% mladića i 7,8% djevojaka. Na Hrvatskom vojnom učilištu „Dr. Franjo Tuđman“ u Zagrebu u istraživa-

nju je sudjelovalo 57 studenata - 82,5% mladića i 17,5% djevojaka.

Na pitanje „Što je plagijat?“ većina anketiranih navela je: kopiranje tuđih radova bez navođenja autora, krađa tuđeg rada. Manji dio učenika nije znao objasniti i odgovoriti. Isto tako, učenicima je bilo teže odgovoriti koje su posljedice plagiranja. Neki od odgovora su:

- zakonske kazne, akademske kazne, narušen ugled,
- poništavanje rada,
- kazneno odgovaranje i
- novčane kazne.

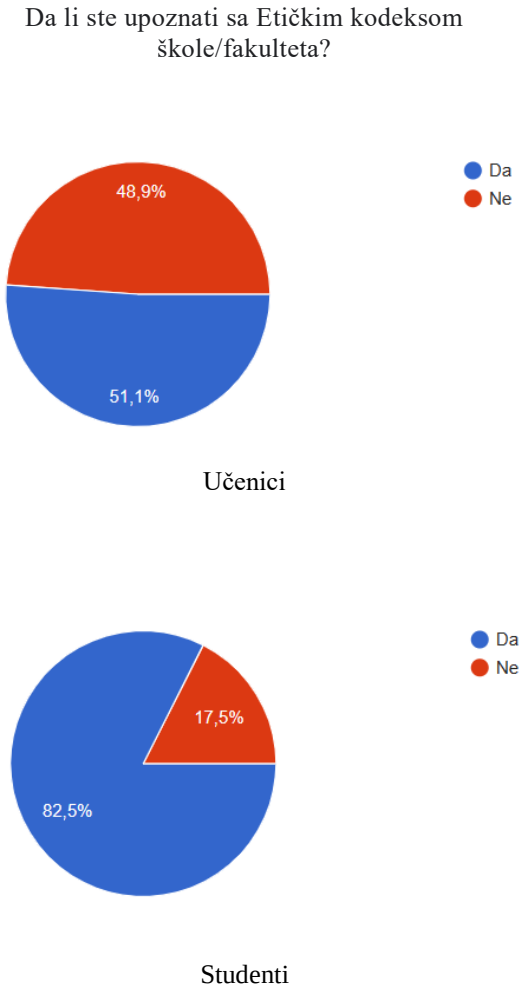
Studenti su bili konkretniji u svojim odgovorima:

- posljedice plagiranja mogu biti kaznene mjere poput odbijanja rada ili ocjene negativno utječu na studijski uspjeh, teškoće u karijeri, izgubljeno povjerenje i reputaciju,
- blaži oblik kazne je smanjivanje ocjene ili odbijanje plagiranog rada od strane profesora, a teži oblik uskraćivanje potpisa što znači da student mora ponoviti kolegij,
- plagiranje može imati negativne posljedice po vašu reputaciju i karijeru. Plagiranje može rezultirati ocjenom "nedovoljan" na zadatku ili ispitu, otkazom iz kolegija ili čak otkazom s fakulteta, te može dovesti do gubitka posla ili štete ugledu i poslovnih odnosa,
- najčešće nikakve ili ne onakve kakve se navode. Stegovna ili kaznena odgovornost i
- ako je plagijator objavom povučenoga rada stekao neke daljnje koristi (napredovanje u poslu, financijsku dobit, neku nagradu i sl.) mogu se pokrenuti i odgovarajući postupci za oduzimanja tako stečenih koristi.

Učenici bi izbjegli plagiranje pomoću autorskih prava, provjerom i navođenjem izvora, stvaranjem samostalnog originalnog djela. Studenti bi pri tome pisali vlastite radove prema pravilima akademskog pisanja, parafrazirali literaturu svojim riječima, citirali i navodili izvore, reference u svojim radovima. Studenti su, za razliku od učenika, naglasili važnost osvješćivanja i educiranja učenika i studenata o ovoj problematici kao i korištenja programa za provjeru plagiranja.

Studenti (75,4%) više od učenika poznaju neki softver za provjeru autentičnosti radova. Najčešće navedeni su PlagScan i Turnitin. Studenti su više od učenika upoznati sa Etičkim kodeksom ustanove na kojoj se obrazuju (slika 1). Kod pisanja radova, 36,8% studenata i 36,7% učenika navodi i potpisuje Izjavu o akademskoj čestitosti.

Učenici i studenti promišljali su o načinima kako smanjiti (spriječiti) plagiranje u učeničkim/studentskim radovima. Neki od odgovora su:



Slika 1. Odgovori učenika i studenata na pitanje “Da li ste upoznati sa Etičkim kodeksom škole/fakulteta?”

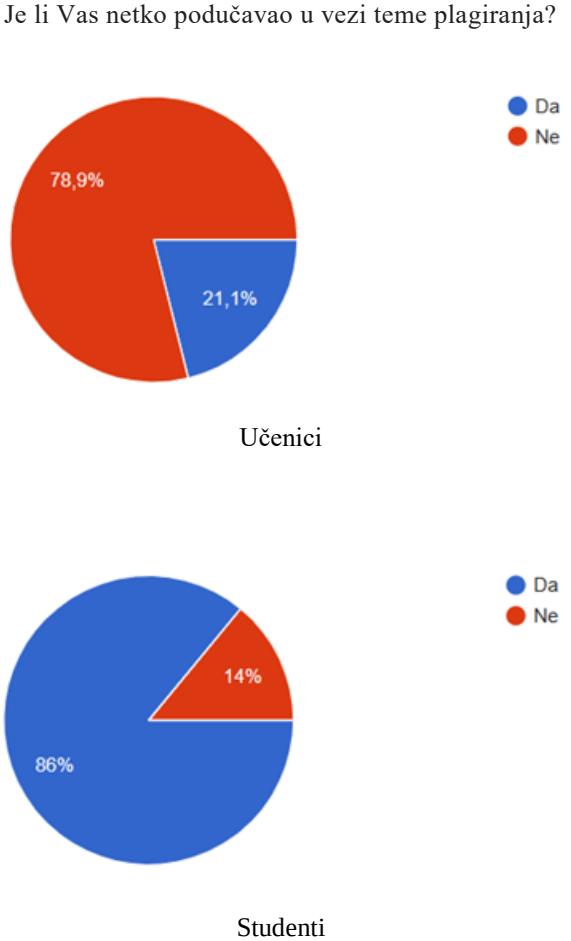
- raditi na tome da učenici/studenti odaberu teme koje ih zanimaju kako ne bi bili vezani time da samo isporuče rad već da se i zainteresiraju za njega te daju svoja kritička mišljenja,
- korištenjem softvera za provjeru autentičnosti te strože kažnjavanje plagijata kod studenata i učenika,
- poučavanjem o pravilnom citiranju i parafraziranju,
- poučavanjem o pravilnom navođenju autora te strogim kaznama ukoliko se utvrdi da je rad plagiran i
- provoditi edukacije o citiranju i intelektualnom vlasništvu, te jasno definirati pravila i kazniti prekršaje.

Motivi (razlozi, uzroci) za stvaranje plagijata koje su naveli učenici su: lijenost, bolja ocjena, kašnjenje u predaji rada, neznanje, manjak vlastitog vokabulara i kreativnosti pisanja, naporan rad, nespremnost učenika, nerazumijevanje

sadržaja, skraćivanje utrošenog vremena. Studenti su dodali i slijedeće motive: loša organiziranost vremena, nemarljivost, neodgovornost, nedostatak vremena za pisanje rada, nedostatak straha od kazne, nezainteresiranost za određenu temu, stjecanja rezultata bez prevelikog truda, nedostatak koncentracije, brzopletost, loše vještine pisanja i izražavanja tekstom.

Većina učenika (92,2%) i studenata (96,5%) upoznati su s pojmovima intelektualno vlasništvo i autorsko pravo. Svi anketirani studenti i 73,3% učenika smatraju da plagiranje nije pošten način stvaranja rada. Studenti u većem broju (75,4%) u odnosu na učenike (45,6%) smatraju da je plagiranje problem u učeničkom/studentskom okruženju. Isto tako plagiranje smatra moralno pogrešnim 86% studenata, a svega 43,3% učenika. Stoga ne iznenađuje rezultat da su gotovo suprotni podaci u odgovorima na pitanje „Je li vas je netko podučavao u vezi teme plagiranja?“ (slika 2). Jasno je da je edukacija o plagiranju važna u što ranijoj fazi obrazovanja.

Seminarske radove 80,7% anketiranih studenata i 66,7% učenika doživljava kao priliku za proširenje vlastitog znanja i dodatno istraživanje dok 21,1% studenata i 33,3% učenika



Slika 2. Odgovori učenika i studenata na pitanje “Je li Vas netko podučavao u vezi teme plagiranja?

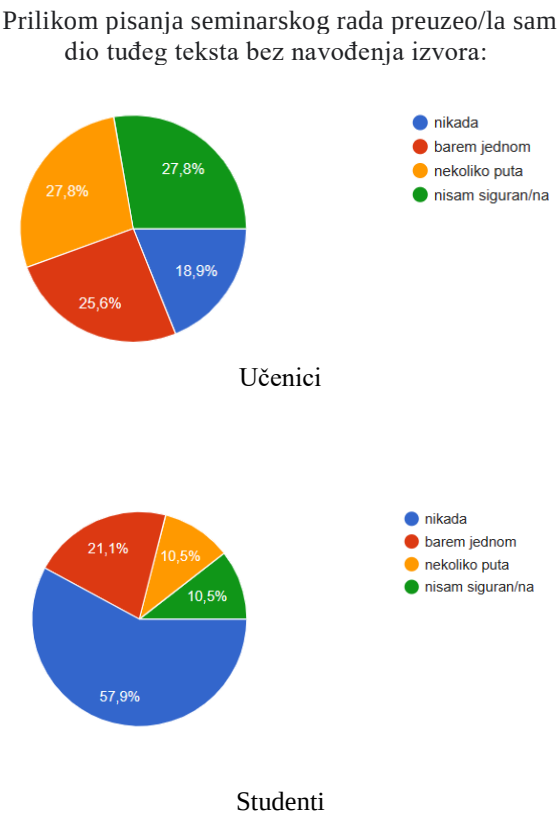
doživljava ih kao nepotrebne zadatke i gubljenje vremena. Među ostalim odgovorima navedeno je:

- razvijanje akademskog pisanja,
- stjecanje vještina u pisanju radova,
- ponekad nepotrebne jer su teme često dosadne,
- za određene predmete kao gubitak vremena, za ostale kao korisno iskustvo u budućem pisanju te općenitoj kvaliteti akademskom pisanja,
- zadatak koji se treba izvršiti i
- dobra ocjena, a usput naučiš gradivo.

Na pitanje „Znam pravila za ispravno citiranje i navođenje izvora“ 45,6% učenika je odgovorilo da zna, 21,1% ne zna, a 33,3% nije sigurno. Kod studenata 94,7% odgovorilo je da zna dok 5,3% nije sigurno.

Studenti znatno manje od učenika, prilikom pisanja seminarskog rada, preuzimaju dio tuđeg teksta bez navođenja izvora (slika 3).

Većina učenika i dio studenata ne poznaju pojam samoplagiranje. Dio anketiranih dao je svoje odgovore: korištenje svojih radova više puta u drugim situacijama, uzimati informacije iz prijašnjih vlastitih radova, recikliranje



Slika 3. Odgovori učenika i studenata na pitanje “Prilikom pisanja seminarskog rada preuzeo/la sam dio tuđeg teksta bez navođenja izvora”

većine svoga rada te objavljivanje na neku sličnu temu, samoplagiranje je kada netko koristi svoj raniji rad bez odgovarajuće reference ili citiranja, kao da je to novi rad.

Na pitanje „Da li ste ikada do sada svjesno plagirali rad?“ potvrdno je odgovorilo 40% učenika i 8,8% studenata. Za kažnjavanje plagiranja znanstvenih, stručnih, završnih, seminarskih i drugih radova izjasnilo se 54,4% učenika i 87,7% studenata. O kažnjavanju takvog postupka, ukoliko se on dokaže, anketirani su dali komentare:

- poništavanje rada/ negativna ocjena/ ponavljanje rada,
- ne treba biti prestroga kazna,
- treba prvo opomenuti, ukoliko se ponovi kazniti
- odbiti rad te studentu ili učeniku dati novu temu i ponovo pisanje rada ili u ekstremnim slučajevima gubitak potpisa,
- treba učeniku dati drugu šansu i konzultirati se s njim,
- izbacivanje iz sustava ili gubitak prava,
- što se tiče studentskih radova mislim da nije potrebno pretjerano kažnjavanje, ali ukoliko netko nastoji zaraditi na osnovu tuđeg rada, mislim da ga je potrebno strogo kazniti i
- potrebno je kazniti studente koji čine takve prijestupe. Međutim, treba kazniti i profesore koji ne primijete plagijat.

Razlozi što kod pisanja zadatka "kopiraju i zalijepe" odlomke s interneta, komentirali su učenici i studenti:

- nemam dovoljno vremena samostalno napisati rad,
- zadatak nije važan,
- moje vještine pisanja su slabe,
- znam da moj profesor neće provjeravati da li je to moj rad,
- moram dobiti dobru ocjenu,
- škola nas uči da su ocjene najbitnije,
- potrebni su vanjski izvori informacija,
- ako bih isto to napisao kopiram tako da ne pišem bespotrebno,
- lijenost i manjak motivacije,
- ako se radi o nekim podjelama ili nešto što ne mogu napisati svojim riječima i
- nemam motivacije truditi se jer znam da profesor neće prepoznati i čini se kao gubljenje vremena.

Komentari na pitanje „Molimo opišite kako bi se osjećali kada bi uvidjeli da je netko koristio Vaše spoznaje u radu bez navođenja Vas kao autora“:

- smetalo bi mi,
- loše,
- osjećao/la bih se polaskano da netko uzima moj rad,
- osjećao/la bih se kao Tesla,
- ako će to nekome pomoći, nemam ništa protiv,
- bio/la bi iznimno uzrujan/a,
- tužio/la bih tu osobu,
- bilo bi mi svejedno,
- bio bi uvrijeđen jer je to ipak moj trud,
- ako će za mene biti loše posljedice, onda bih bio/la ljut/a, ali ako nema nikakvog utjecaja na mene onda mi je sasvim svejedno,
- razočarano,
- ako bi netko stvarno koristio moje podatke, laskalo bi mi,
- kada se osoba posveti stvaranju nečega, teško je gledati da netko uzima dio toga te ga svojata, a ne zna koliko je truda i muke uloženo u kreiranje nekog rada ili dijela. To naravno stvara osjećaj razočarenja kod autora,
- bio/la bih ponosan/a što se koristim ja kao izvor, ali i uvrijeđen/a što se moj rad ne navodi kao izvor i
- ukoliko je seminar ili neki slični zadatak na fakultetu ne bi me toliko smetalo, a ako su u pitanju znanstveni radovi, knjige i slično smetalo bi me zbog nepoštovanja količine uloženog truda.

III. ALATI ZA PROVJERU AUTENTIČNOSTI RADOVA

Etičnost u znanosti i obrazovanju je i danas aktualna tema. Vijeće Europe je 2015. godine donijelo platformu za etičnost, transparentnost i integritet u obrazovanju (ETINED). Ciljevi su i zaštita, razvoj i podupiranje akademske čestitosti u obrazovnim ustanovama, a posebice borba protiv plagiranja.

Za sustavan pristup otkrivanju plagiranja potrebni su alati/programi (softver) za provjeru autentičnosti (izvornosti) radova. Njihova prisutnost razvija i podiže svijest među učenicima i studentima da vode brigu o pravilnom navođenju korištenih izvora u izradi vlastitih radova.

Ovi alati pomažu u postupcima dokazivanja i odlučivanja o izvornosti (autentičnosti) učeničkih i studentskih radova.

Postoji više alata/programa (softver) za provjeru autentičnosti (izvornosti) učeničkih i studentskih radova. U ovom radu kratko su opisana dva alata/programa koja koriste sastavnice Sveučilišta u Zagrebu za provjeru autentičnosti radova, a to su: PlagScan i Turnitin. Svaka ustanova donosi odluku o načinu i modalitetima uporabe navedenog softvera kao i interpretaciju izvješća koje softver generira. Detaljnije o

navedenom softveru za provjeru autentičnosti radova (učeničkih, studentskih i ostalih) nalazi se na poveznicama: www.plagscan.com i www.turnitin.com.

Na primjer, za nabavu licence softvera za provjeru autentičnosti radova Turnitin za javne visokoškolske ustanove u sustavu znanosti i obrazovanja financira Ministarstvo znanosti i obrazovanja RH kroz proračun Srca (Sveučilišnog računskog centra), a Srce provodi proces nabave i koordinira distribuciju licenci i korištenja navedenog softvera. Dostupne licence omogućuju neograničen broj provjera za svakog upisanog studenta, a provjeru radova (od seminarskih do završnih i ostalih) mogu izvršiti nastavnici, ali i sami studenti, po propisanim postupcima i procedurama u skladu sa potrebama obrazovnih ustanova. Za potporu u radu sa softverom za provjeru autentičnosti radova na Srcu postoji Tim za softver za provjeru autentičnosti radova u Centru za e-učenje. Također, postoji i online tečaj Srca: Autorstvo, plagiranje i citiranje: što, kako, zašto? Na Srcu postoje upute i animacije za nastavnike i studente za uporabu softvera za provjeru autentičnosti radova.

Neka najčešće postavljanja pitanja u vezi ovih softvera su npr.:

1. Tko ima pravo koristiti softver?,
2. Koji softver odabrati?,
3. Koje su prednosti korištenja odabranog softvera?,
4. Na koji način radi softver?,
5. Tko može izvršiti provjeru radova?,
6. Kako se pristupa i koristi odabrani softver?,
7. S kojim se izvorima radovi uspoređuju?,
8. Što znači određeni postotak podudarnosti koji softver prikaže nakon analize rada? Kao i ostala pitanja …

Odgovore na ova i slična pitanja u vezi alata/programa (softvera) za provjeru autentičnosti (izvornosti) radova možete naći na web stranicama Srca u rubrici Centar za e-učenje od članova Tima Srca za softver za provjeru autentičnosti radova [2].

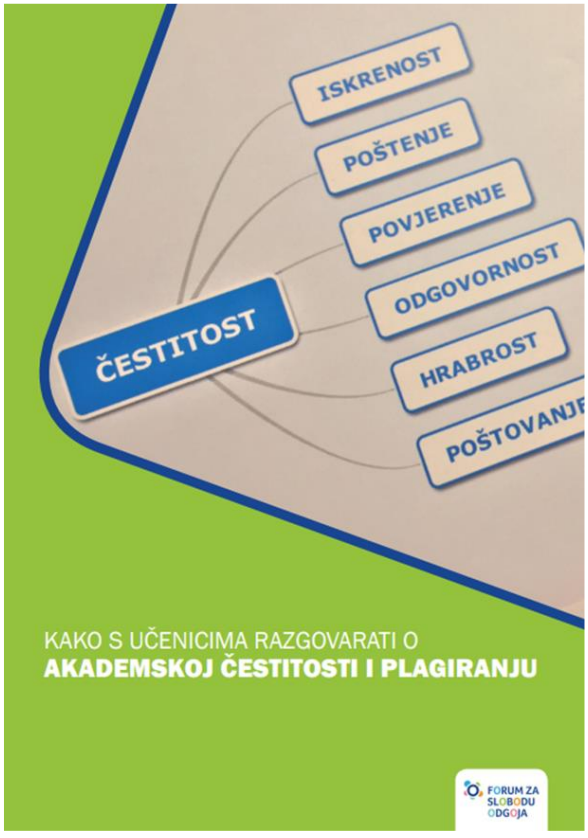
IV. EDUKACIJA

Istraživanje je provedeno na manjem broju ispitanika kojima su autori ovog rada predavači. Planira se buduće istraživanje s više ispitanika na razini županije. Rezultati istraživanja pokazuju da razlikujemo kršenja pravila akademske čestitosti koja su namjerni pokušaj korištenja tuđeg rada i predstavljanja kao vlastitog i kršenja koja su rezultat učeničkog i studentskog nerazumijevanja prakse citiranja i njihovog slabog iskustva u akademskom pisanju. Najbolji način da se spriječi plagijat je educirati učenike i studente o tome kako ispravno provoditi istraživanje, citirati i izraditi jedinstven i originalan rad.

U školama i na fakultetima mora postojati ujednačenost u tome kako se otkrivaju i rješavaju slučajevi plagiranja na cijelom sveučilištu. Ustanove moraju osigurati obuku za akademsko osoblje i temeljito poučiti studente o referenciranju i plagiranju, ali i koristiti osoblje unutar sveučilišta, poput akademskih savjetnika, koji već imaju iskustva u ovom području. Na svakom sveučilišnom odjelu treba identificirati ključno osoblje kojemu se mogu prijaviti sva kršenja politike kako bi se zauzeo jedinstveniji pristup.

Forum za slobodu odgoja objavio je publikaciju pod nazivom „Kako s učenicima razgovarati o akademskoj čestitosti i plagiranju“. U njoj se nalazi niz aktivnosti za nastavnike kako mogu razgovarati s učenicima o akademskoj čestitosti, plagiranju, citiranju i korištenju drugih izvora. Namijenjene su za izvođenje u srednjim školama, kako gimnazijama, tako i u strukovnim školama. Mogu se izvoditi u nizu predmeta: Hrvatski jezik, Etika, Vjeronauk, Sociologija, Filozofija, međupredmetne teme Građanskog odgoja i obrazovanja, kao i u sklopu strukovnih predmeta, s ciljem upućivanja učenika u profesionalnu etiku određene struke [3].

Učenicima će pomoći i „Vodič za učenike pri izradi pisanih radova, plakata i prezentacija kod istraživačkih, projektnih i problemskih zadataka“ u izdanju Školskih knjižničara Republike Hrvatske [4]. Srce ima upute za nastavnike „Pravilno korištenje tuđih i dijeljenje vlastitih obrazovnih sadržaja u e-kolegiju“ kao podsjetnik na što je sve potrebno obratiti pažnju prilikom preuzimanja tuđih i



Slika 4. Publikacija Foruma za slobodu odgoja “Kako s učenicima razgovarati o akademskoj čestitosti i plagiranju”

stvaranja vlastitih sadržaja te njihovog postavljanja u e-kolegij [5].

Državni zavod za intelektualno vlasništvo Republike Hrvatske promiče zaštitu i poštivanje prava intelektualnog vlasništva uz edukaciju građana tečajevima i publikacijama [6]. Brojni dodatni korisni sadržaji, brošure [7], video zapisi, tečajevi koje nalazimo na internetu mogu se koristiti u školama za aktivnosti povodom obilježavanja Dana sigurnijeg interneta. Važno je osvijestiti učenike od osnovne škole da je autorsko pravo sastavni dio njihove digitalne pismenosti [1].

Hrvatska akademska i istraživačka mreža (CARNET) je u sklopu projekta „e-Škole: Uspostava sustava razvoja digitalno zrelih škola“ pripremila webinar i pripadni priručnik „Citiranje u digitalnom okruženju“ koji kroz teoriju i primjere pomaže djelatnicima u školama u izradi vlastitih sadržaja [8].

V. ZAKLJUČAK

Analize na temu plagijata i plagijatorskog ponašanja u Hrvatskoj provedene su i prije istraživanja kojeg su proveli autori ovog rada. Rezultati istraživanja ukazuju na isti problem – studenti nisu dovoljno osviješteni oko problema plagiranja, nedostaju edukacije te je uočljiv manjak motivacije [9].

Plagijat je važno pitanje o kojem se često raspravlja, a koje može imati ozbiljne financijske posljedice za institucije visokog obrazovanja diljem Europe. Pojedina istraživanja pokazala su razlike u odnosu prema plagijatu među različitim europskim zemljama. Jedno istraživanje identificiralo je neke razlike između zapadnih i istočnih zemalja, kao i neke razlike u odnosu učenika i nastavnika prema problemu plagijata [10].

Plagiranje kao praksa dosta je raširena u svijetu zbog nepostojanja univerzalnog pristupa samom konceptu. Različite su kazne za plagiranje u raznim zemljama. U Sjedinjenim Američkim Državama učenici već u školi otkrivaju puno o negativnim posljedicama. Dovoljno rano u svojoj akademskoj karijeri shvaćaju da kršenje pravila tjera ih da prolaze kroz vrlo neugodne situacije u koje su često uključeni roditelji i predstavnici škole [11].

Mnogi studenti, osobito na početku svoje akademske karijere, možda neće u potpunosti razumjeti različite oblike plagijata. U srednjoj školi možda se nije dovoljno pažnje posvetilo ovom problemu. Iako se ti učenici/studenti i dalje moraju smatrati odgovornima, važno je da škole educiraju učenike o pravilnim metodama citiranja i referenciranja na izvore koje su koristili u svojim radovima, odnosno različitim vrstama plagijata i kako ih izbjeći.

Nužno je definirati Kodeks časti za cijelu školu, koji jasno navodi posljedice prijevara i plagiranja s čime moraju biti upoznati roditelji/staratelji i učenici.

Anketiranje je pokazalo da je učenicima i studentima potrebno osigurati dovoljno vremena za dovršavanje seminarskih i istraživačkih radova, odrediti rokove za prikupljanje izvora i pripremu radova. Učenici su skloni plagirati kada shvate da se bliži rok predaje radova.

U svakoj školi/fakultetu potrebno je obučeno osoblje koje razumije da praksa akademskog citiranja nije univerzalna, već nešto što se mora poučavati i prakticirati, a ujedno zahtijeva značajnu promjenu u razmišljanju o poštovanju i autorstvu.

Disciplinska evidencija može biti trag u dokumentaciji studenta/učenika kada se prijavljuje za upis na drugo sveučilište ili školu, što može značajno utjecati na njegov upis. Nadalje, akademska evidencija može biti zatražena ako se osoba prijavljuje za radno mjesto na fakultetu/školi, što će također poslužiti kao njena diskreditacija. Varanje ne utječe samo na to koliko osobi vjeruju školski službenici, već i na ljude u njenom privatnom životu: roditelje, braću i sestre, prijatelje i druge važne osobe. Spremnost osobe na varanje znak je da njena etika ima svoju granicu.

Kako bi bili etički i zakoniti sudionici u svojim internetskim zajednicama, učenici i studenti moraju biti upoznati s osnovnim konceptima intelektualnog vlasništva, autorskim pravima i njihove poštene upotrebe. Nastavnici uz veću suradnju s učenicima, povećanje motivacije učenika, poticanje učenika na čitanje omogućuju da izvorni radovi pridonose vještinama kritičkog mišljenja učenika, obogaćuju cjelokupno znanje, promičući društveni razvoj i napredak.

REFERENCE

[1] Copyright and Creativity for Ethical Digital Citizens, dostupno na https://copyrightandcreativity.org, 27. prosinca 2022.

[2] Softver za provjeru autentičnosti radova, Sveučilišni računski centar, dostupno na https://www.srce.unizg.hr/spa, 25. siječnja 2023.

[3] Kako s učenicima razgovarati o akademskoj čestitosti i plagiranju, Forum za slobodu odgoja, dostupno na https://fso.hr/wp-content/uploads/2017/02/Akademska-cestitost-i-plagiranje_fin.pdf, 12. siječnja 2023.

[4] M. Othman, “Navođenje izvora (citiranje)”, vodič za učenike pri izradi pisanih radova, plakata i prezentacija kod istraživačkih, projektnih i problemskih zadataka, dostupno na http://knjiznicari.skole.hr/upload/knjiznicari/newsattach/474/NAVOENJE_IZVORA_-_vodic_za_ucenike_2.pdf, 12. siječnja 2023.

[5] Pravilno korištenje tuđih i dijeljenje vlastitih obrazovnih sadržaja u e-kolegiju, Sveučilišni računski centar, dostupno na https://www.pmf.unizg.hr/_news/48285/Srce_upute_e-kolegiji.pdf, 24. siječnja 2023.

[6] Autorsko pravo i srodna prava, Državni zavod za intelektualno vlasništvo, dostupno na https://dziv.hr/hr/intelektualno-vlasnistvo/autorsko-pravo/, 18. siječnja 2023.

[7] L. Kaira, “Autorskim pravom štiti se stvaralački rad”, dostupno na http://www.autori.hr/wp-content/uploads/2014/05/brosura-Autorsko-pravo.pdf, 17. siječnja 2023.

[8] Citiranje u digitalnom okruženju, priručnik, CARNET, 2018., dostupno na https://pilot.e-skole.hr/wp-content/uploads/2018/03/Prirucnik_Citiranje-u-digitalnom-okruzenju-1.pdf, 18. siječnja 2023.

[9] P. Marđokić, “Stavovi prema plagiranju kod studenata zdravstvenih studija”, Specijalistički diplomski stručni, Zdravstveno veleučilište, 2019., dostupno na https://urn.nsk.hr/urn:nbn:hr:139:811117, 12. siječnja 2023.

[10] T. Foltýnek, F. Čech, “Attitude to plagiarism in different European countries”, dostupno na https://www.researchgate.net/publication/275419974_Attitude_to_plagiarism_in_different_European_countries, 2. veljače 2023.

[11] A Quick Guide To Plagiarism, dostupno na https://www.aresearchguide.com/6plagiar.html, 3. veljače 2023.

Medijski tehničari – nositelji projekata za mlade

Ljiljana Ille* i Romina Fic **
* Graditeljska škola Čakovec, Croatia, ** Graditeljska škola Čakovec, Croatia
* ljiljana.ille@gmail.com, ** romina.heda@gmail.com

Sažetak – Digitalne kompetencije ključne su za razvoj modernog društva i osnovna je zadaća nastavnika informatike osposobiti učenike za pravilno korištenje ICT tehnologije. Kada je riječ o srednjoškolskim učenicima koji se školuju u obrazovnim programima koji pretpostavljaju svakodnevno korištenje takve tehnologije, kompetencije moraju biti kudikamo šire. U Graditeljskoj školi Čakovec svake se godine provodi nekoliko projekta u kojima učenici, budući medijski tehničari osmišljavaju i realiziraju sve projektne aktivnosti, čime dokazuju sposobnost praktične primjene usvojenih teoretskih znanja. Ove su školske godine provedena dva projekta; Animirani odmor te Mi, srednjoškolci. Učenici su osmislili projekte i realizirali sve planirane aktivnosti u suradnji sa svojim školskim kolegama, vršnjacima iz drugih škola te svojim nastavnicima. To je podrazumijevalo razradu projektne ideje, izradu najave i promocije projekta te komunikaciju s medijima i suradnicima, izradu radijskih emisija, provođenje radionica o uporabi multimedijske tehnologije s vršnjacima iz drugih škola, raspisivanje i provođenje natječaja za učenike u više kategorija te na kraju izradu kataloga svih radova nastalih u projektu, kao i snimanje te montažu video priloga o projektu. Osim praktične primjene stečenih znanja i poznavanja tehnologije, provedenim projektima učenici su pokazali i veliku umješnost u stvaranju novih sadržaja što pak predstavlja najvišu razinu ostvarenosti odgojno-obrazovnih ishoda.

Ključne riječi - digitalne kompetencije; ICT tehnologije; srednjoškolski učenici; projekti; Animirani odmor; Mi, srednjoškolci; praktična primjena znanja; stvaranje novih sadržaja

I. UVOD

Suvremenom društvu nezamisliv je život bez tehnologije i djeca već u ranoj životnoj dobi dolaze u kontakte s njom te stječu kompetencije koje će kasnije proširivati u svom formalnom obrazovanju, ali i tijekom cijelog života. Uporaba digitalne tehnologije (izrada grafičkih i audiovizualnih sadržaja za različite platforme, prezentacija), naprosto je nužna, ne samo u nastavi informatike, nego i u drugim obrazovnim predmetima. Svaki novi stupanj obrazovanja donosi veće zahtjeve u poznavanju i korištenju tehnologije, a kada je riječ o budućim medijskim stručnjacima, izvrsno poznavanje i svakodnevna uporaba tehnologije podrazumijeva se samo po sebi. To znači da svaki korisnik može biti proizvođač sadržaja, ako ima određenu medijsku pismenost i vještine za korištenje digitalne tehnologije i novih medija [1]. Učenici Graditeljske škole Čakovec u programu Medijski tehničar, uz već spomenuta teoretska znanja i vještine vezane uz oblikovanje multimedijskih sadržaja, stječu i praktična znanja potrebna za planiranje i realizaciju projekata. Usvojena znanja i vještine učenici vrlo rijetko pokazuju izvan učionice čime ostaju prikraćeni za realno iskustvo, kao i povratnu informaciju o svome radu od osoba izvan školskog sustava. Isto tako učenici u svojim učionicama ne mogu realizirati sve faze projekta u punom obimu jer su ograničeni vremenom, resursima i ostvarivanjem nastavnih ishoda pojedinih nastavnih sadržaja. Ipak, stječu osnovne kompetencije potrebne za uspješnu realizaciju nekog projektnog zadatka, odnosno usvajaju vještine dobre pripreme i kvalitetne organizacije u provedbi aktivnosti, navike vođenja projektne dokumentacije i njezinog arhiviranja, nauče poštivati zadane rokove, raditi u timu te suvereno i demokratski komunicirati međusobno, ali i s drugim sudionicima pojedinog projekta. Dakako, da je uloga nastavnika u ovom procesu nemjerljiva, ali, da bi učenici izgradili svoje samopouzdanje te stekli iskustvo pisanja, vođenja i realizacije projekata, nastavnik ih mora potaknuti da stečene vještine i znanja primijene na konkretnim zadacima i pritom im biti potpora u cijelom tom procesu. To bi značilo da učeniku ne treba nastavnik predavač, već nastavnik mentor, suradnik i organizator [2]. Budući da na redovnoj nastavi ne mogu ostvariti sve ranije spomenuto, učenici projekte realiziraju u okviru izvannastavnih aktivnosti, posebice u okviru Medijske sekcije koja već dugi niz godina djeluje u Graditeljskoj školi Čakovec. Od početka školske godine 2022./2023. realizirana su dva projekta; *Animirani odmor* i *Mi, srednjoškolci*, o čemu će biti riječi u nastavku, a do kraja školske godine u planu je realizacija još dvaju projekata.

II. PRIPREMNA FAZA PROJEKTA

Prije nego što se uopće krene u pripremnu fazu pojedinog projekta učenici moraju odlučiti o tematici projekta, odnosno što će biti tematska okosnica projekta, a najčešće je riječ o obilježavanju neke značajne obljetnice ili nekog poznatog događaja, kao npr. obilježavanje Svjetskog dana animacije ili Međunarodnog dana srednjoškolaca. Naravno, na izbor teme utječu njihove sklonosti i interesi, kao i neka izravna povezanost s temom.

Nakon što je uspješno odabrana tema kojom će se u projektu baviti, učenici kreću u pripremnu fazu projekta, odnosno u planiranje projektnih aktivnosti (Slika 1) i svih ostalih segmenata koji su nužni za realizaciju projekta u cijelosti. Najprije promišljaju o projektnim aktivnostima koje bi u okviru projekta željeli realizirati i pritom moraju voditi računa o mogućnosti realizacije pojedinih aktivnosti s obzirom na vremenske, prostorne, financijske, tehnološke i ljudske resurse. Neke aktivnosti, zbog spomenutih razloga, u datim uvjetima neće biti izvedive pa ih moraju izostaviti i osmisliti neke druge aktivnosti čija je realizacija moguća. Već u ovom početnom stadiju

pripremne faze do izražaja dolaze učeničke sposobnosti promišljanja i planiranja, ali i njihova snalažljivost i kreativnost.

Pripremna faza:
- Izrada tablice aktivnosti
- Dogovor suradnje:
  - Nastavnici i učenici škole
  - Dječji vrtić Pčelice, Ivanovec; OŠ Ivanovec; ŠAF Čakovec, Radio 1 Čakovec
  - Sponzori:
- Pisanje Opisa projekta
- Izrada prezentacije i prezentacija projekta ravnatelju (termin, aktivnosti, troškovi)
- Raspisivanje natječaja „Animirani odmor" - izrada plakata:
  Natječaj u tri kategorije
  - Animacija (u trajanju do 3 minute)
  - Ilustracija (A4 format)
  - Tekstualni radovi (poezija, proza)
- Izrada Operativnog plana
- Izrada najavnog plakata
- Najava na web stranici škole i fejs stranici škole
- Snimanje radijske emisije Veliki odmor
- Izrada Nagradnog kupona - oslobođenje od odgovaranja
- Prikupljanje učeničkih radova
- Ocjenjivanje i biranje radova
- Printanje tekstova i ilustracija
- Obrada i montaža animiranih formi za objavljivanje i reproduciranje
- Uređenje prostora u **utorak, 26.10.**
- Objesiti na oglasnoj ploči u zbornici popis učenika koji nisu taj dan na nastavi i obavijestiti pojedinačno razrednike *(Učenik sudjeluje u provedbi školskog projekta „Animirani odmor")*

Provedba projekta:
- Program: Predavanje, Izložba, Radionice animacije, Online projekcija i analiza filma i proglašenje i objava najboljih radova.
- Snimanje TV reportaže
- Fotografiranje

Završna faza:
- Vraćanje prostora u prvobitno stanje
- Pisanje teksta za medije + odabir fotografija
- Montaža reportaže (video)
- Izrada Kataloga radova
- Izrada zahvalnica
- Arhiviranje i Razvijanje fotografije za godišnju produkciju
- Pisanje projektnog izvještaja

PRODUKTI:
- Radijska emisija Veliki odmor
- Radovi s natječaja – video
- Radovi s radionica - video
- TV reportaža
- Fotokolaž
- Katalog

Slika 1. Projektne aktivnosti i zaduženja

Učenici zatim pišu opis projekta u kojemu precizno i jasno opisuju projekt i njegove ciljeve te izrađuju vremenik i troškovnik projektnih aktivnosti, a koje će onda predstaviti ravnatelju koji je odgovorna osoba za sva kulturna i javna događanja u školi. Korak koji slijedi od učenika zahtijeva umješnost pregovaranja i dogovaranja jer moraju dogovoriti potencijalne suradnje s kolegama koji su uključeni u planiranje projekta, ali i drugim sudionicima projekta s kojima ostvaruju suradnju u pojedinima segmentima (nastavnici, drugi učenici, suradnici iz drugih odgojno-obrazovnih ustanova, sponzori). U ovoj fazi vrlo su važne i učeničke komunikacijske vještine jer moraju znati predstaviti svoje ideje, ali i zagovarati projekt. Ipak, bez obzira na njihovu komunikativnost i pregovaračke vještine, nastavnička potpora ovdje nikako ne smije izostati, posebice u komunikaciji s ravnateljem, nastavnicima i vanjskim suradnicima na projektu.

Nakon što je projekt predstavljen svim relevantnim subjektima, kreće se u konkretizaciju pojedinih aktivnosti; izrađuje se Operativni plan projekta te najavni plakat i brošura koji se objavljuju na mrežnim stranicama te na Facebook stranici škole, kao i na drugima informacijskim punktovima škole.

Vrlo je važno da konkretne informacije dobiju svi sudionici projekta, i učenici i nastavnici, odnosno ostali sudionici. Učenici raspisuju i nagradne natječaje u različitim kategorijama (Slika 2), ovisno o kojem je projektu riječ, u kojima obavezno navode uvjete natječaja te podatke o načinu i roku dostave radova. U provedenim projektima kategorije natječaja bile su različite za svaki pojedini projekt, od animacije (u trajanju do 3 minute), ilustracije na zadanu temu pa do tekstualnih radova (u prozi ili stihu) te pokretnih slika (u trajanju do 3 minute). Važno je dogovoriti se i imenovati učenika/e koji će biti zadužen/i za provođenje natječaja, od njegovog raspisivanja pa sve do prikupljanja učeničkih radova pristiglih na natječaj po pojedinim kategorijama te obrade i montaže uradaka za prikazivanje. Radovi se pregledavaju, kategoriziraju i ocjenjuju od strane povjerenstava koja su oformljena za svaku pojedinu kategoriju, a čine ih učenici i nastavnici. Svi radovi pristigli na natječaj izlažu se na tematskoj izložbi, a najbolji po kategorijama posebno su istaknuti. Radove pristigle na natječaj potrebno je pripremiti za izlaganje, što znači otisnuti ih (grafički i tekstualni radovi), odnosno obraditi ih i montirati u formi za objavljivanje i reproduciranje (animirane i video forme).



Slika 2. Plakat za nagradni natječaj

Vrlo često je to opsežan posao i iziskuje velik angažman učenika pa se u njega uključuje više učenika. Isto tako svi sudionici moraju se prethodno prijaviti za sudjelovanje u nekoj od ponuđenih aktivnosti, tako da učenici unaprijed mogu pripremiti potrebne resurse za realizaciju aktivnosti. Učenici u ovoj fazi pripremaju i radio emisiju „Veliki odmor" koja se emitira na projektni dan i u kojoj se iznose neke činjenice vezane uz tematiku projekta i projektne aktivnosti.

III. FAZA PROVEDBE PROJEKTA

Faza provedbe projekta najvažnija je jer se sve planirane aktivnosti moraju realizirati u zadanom vremenu

i na zadovoljavajući način pa iziskuje izuzetno dobru pripremljenost i fleksibilnost.

Osim ranije spomenute izložbe, učenici pripremaju i organiziraju prostore za odvijanje pojedinih aktivnosti, vodeći računa o redovnoj nastavi koja se odvija u to vrijeme. Učenici su se unaprijed rasporedili u timove i svatko radi svoju aktivnost, a u prigodni program ili primjerice online analizu filma uključuju se svi.

A. Animirani odmor

Povodom Svjetskog dana animacije 28. listopada realiziran je školski projekt Animirani odmor [3], a aktivnosti su se odvijale tijekom čitavog dana. Započelo se ujutro odlaskom učenika prvog razreda u prostore čakovečke Škole animiranog filma (ŠAF) gdje ih je dočekala voditeljica i upoznala s tehnikama izrade animiranog filma. To su učenici koji nisu izravno uključeni u projekt, ali će to raditi u godinama koje slijede pa su im ovakva iskustva dragocjena i korisna. Slijedila je radionica animacije (Slika 3) s djecom predškolske dobi iz DV Ivanovec. Radionici je prethodilo čitanje s „Pčelicama" koje su proveli školska knjižnica i školska čitalačka skupina, a tema i motivi koje su obradili poslužili su kao osnova za radionicu animacije. Važno je naglasiti da su radionice učenici osmislili i realizirali potpuno samostalno, uz pokoju sugestiju nastavnika. U tim radionicama učenici su primijenili nekoliko tehnika stop motion animacije: objekti, ljudi (piksilacija), glina te kolaž. Radili su po skupinama, ovisno o tome kojom su se vrstom animacije bavili i pritom su usmjeravali predškolsku djecu, tj. davali im upute što i kako moraju u kojem trenutku raditi.



Slika 3. Radionica animacije

U oblikovanju gline u radionicu se uključilo nekoliko učenika iz programa Dizajner keramike koji se također školuju u Graditeljskoj školi. Rezultati radionica bit će kratke animirane forme. Učenici zaduženi za fotografiranje i snimanje sve su bilježili za potrebe dokumentiranja i izrade TV reportaže.

Radionica animacije ponovljena je i u popodnevnim satima, ali s drugim učenicima i drugim sudionicima, odnosnom učenicima predmetne nastave OŠ Ivanovec. U vremenu između radionica animacije učenici su se putem Zoom-a povezali s učenicima Prirodoslovno-grafičke škole Zadar i zajednički pogledali animirani film „Moj život kao Tikvica" te ga analizirali, stručno vođeni od strane nastavnika iz obiju škola. Projektne aktivnosti završile su u večernjim satima proglašenjem najboljih radova u pojedinim kategorijama natječaja i dodjelom prigodnih nagrada učenicima. Dokumentirani su svi događaji na temelju kojih će se u završnoj fazi izraditi TV reportaža te izvještaj za lokalne i regionalne medije.

B. Mi, srednjoškolci

Projekt Mi, srednjoškolci [4] realiziran je 17. studenoga, povodom Međunarodnog dana srednjoškolaca. U ovom projektu provedene su brojne aktivnosti u koje je uključen i veći broj učenika Graditeljske škole Čakovec te susjedne Srednje škole Čakovec pa je organizacija bila zahtjevnija i dugotrajnija. Ideja ovog projekta bila je isticanje učeničkih interesa i sklonosti kroz prezentaciju stečenih znanja i vještina u okviru formalnog srednjoškolskog obrazovanja. Tako su učenici iz različitih obrazovnih programa prezentirali vještine svog budućeg zanimanja; Medijski tehničari prezentirali su vještine fotografiranja i animiranja kroz radionicu animacije (snimanje ispred zelenog platna i montiranje na odabranu pozadinu), Dizajneri keramike prezentirali su rad na lončarskom kolu, Medicinske sestre i tehničari iz Srednje škole Čakovec svim prisutnima demonstrirali su osnove prve pomoći, dok su Fizioterapeutski tehničari izveli nekoliko vježbi za rasterećenje leđnih mišića. Učenici su za sudionike radionica i posjetitelje predvidjeli i nadmetanja u nekoliko sportskih aktivnosti poput šaha i stolnog tenisa, kao i odmjeravanje snaga u pub kvizu. I ovaj projekt završio je kratkim programom s nekoliko glazbenih točki te proglašenjem najboljih radova po kategorijama (Slika 4).



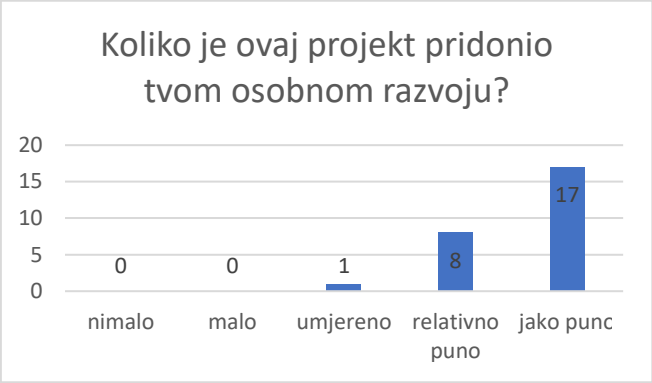
Slika 4. Nagrađeni učenički plakat

Snimljeni su svi događaji od kojih je montirana TV reportaža te izrađen katalog.

IV. ZAVRŠNA FAZA PROJEKTA

U završnoj faza projekta svi prostori u kojima su se odvijale aktivnosti moraju se vratiti u prvobitno stanje. Po povratku u učionicu učenici pišu tekst o projektu, rezultatima projekta i provedenim aktivnostima koji se potom objavljuje na mrežnoj i Facebook stranici Škole,

kao i u lokalnim i regionalnim portalima i glasilima. Učenici izrađuju katalog radova i grafički ga oblikuju, montiraju video reportažu od snimljenih materijala te izrađuju zahvalnice za suradnike u projektnim aktivnostima. Analizira se projekt u cijelosti te se provodi evaluacija projekta.



Slika 5. Pitanje iz upitnika za samoevaluaciju učenika u projektu „Mi, srednjoškolci"

Svi materijali nastali u projektu arhiviraju se, a na samom kraju piše se projektno izvješće koje se također pohranjuje.

IV. ZAKLJUČAK

Kroz rad u projektima učenici stječu nove kompetencije ili proširuju postojeće, uče raditi u timu, poštovati tuđe ideje i stajališta, educiraju svoje vršnjake o načinima korištenja digitalne tehnologije te usvajaju strategije pregovaranja. Osim stečenih spoznaja i iskustava rezultat projekata su i produkti u kojima učenici pokazuju kako usvojena znanja primijeniti za stvaranje novih sadržaja.

Ipak, najvažnija stvar u cijelom procesu je što učenici imaju priliku samoevaulirati svoj kompetencije i rad jer na taj način dobivaju informaciju o svojim sposobnostima i prijeko potrebno samopouzdanje koje će biti pokretač za njihove daljnje aktivnosti.

LITERATURA

[1] Zgrabljić Rotar, N. 2011. "Digitalno doba - Masovni mediji i digitalna kultura", Sveučilište u Zadru i Naklada medijska istraživanja d.o.o., str.48-49.

[2] Matijević, Milan. 2008. Projektno učenje i nastava. Znamen: Nastavnički suputnik 2008./09., str. 188-205.

[3] *Animirani odmor 2021.* Dostupno na: https://online.fliphtml5.com/ywhsn/znwz/?fbclid=IwAR2HH757PC2Frlt39_2JblwUKVIoHbPcrIK5IysyCThL3drbOM3QqpMe8Hg#p=1 (6.2.2023.)

[4] Mi, srednjoškolci 2021.:Dostupno na: https://online.fliphtml5.com/ywhsn/poem/#p=1 (6.2.2023.)

Digitalni alati u inkluzivnom odgoju i obrazovanju

Jelena Bistrović*, Tea Pavičić Zajec**, Tamara Ređep ***

* OŠ „Braća Radić“, Koprivnica, Croatia, jelena.suric@skole.hr
** I. OŠ Varaždin, Croatia, tea232pavi@hotmail.com
*** I. OŠ Varaždin, Croatia, tamara.redep@skole.hr

Sažetak - Porast broja učenika s teškoćama donosi odgojno-obrazovnim djelatnicima i nove izazove. Glavni cilj ovog rada i istraživanja bio je sistematizirati i prikazati važnost i prednosti korištenja digitalnih alata u inkluzivnom osnovnoškolskom obrazovanju. Anketnim upitnikom provjerena je količina i svrha korištenja, nastavno područje te vrsta korištenih digitalnih alata u redovnoj nastavi s učenicima s teškoćama. Istraživanje je provedeno na prigodnom uzorku od 106 učitelja razredne i predmetne nastave u nekoliko osnovnih škola Bjelovarsko – bilogorske, Koprivničko – križevačke i Varaždinske županije. Rezultati su pokazali da većina učitelja koristi digitalne alate u radu s učenicima s teškoćama, a kod većine učitelja svrha korištenja digitalnih alata je povećanje motivacije učenika te bolje razumijevanje apstraktnih nastavnih sadržaja. Učitelji razredne nastave digitalne alate najviše, i više od učitelja predmetne nastave, koriste na satima Hrvatskog jezika, dok učitelji predmetne nastave najviše, i više od učitelja razredne nastave, digitalne alate koriste u STEM predmetima i nastavi stranih jezika. Kao najčešće korištene alate ispitanici navode Kahoot!, Wordwall, Canvu i PowerPoint.

Ključne riječi – digitalni alati, inkluzija, učenici s teškoćama u učenju, osnovnoškolsko obrazovanje

I. UVOD

Svi sudionici odgojno – obrazovnog procesa svjesni su promjena u ponašanju i funkcioniranju osnovnoškolske djece te porasta broja učenika s teškoćama. Jedna od mnogobrojnih uloga učitelja jest osigurati svakom djetetu da se u školi osjeća sigurno, zadovoljno i prihvaćeno. Miljević-Riđički, Maleš i Rijavec [5] u svojoj knjizi navode da škola u svome radu mora primjenjivati načelo jednakopravnosti te poticati primjenu metoda rada koje razvijaju cjelokupnu ličnost djeteta. Mnogobrojne knjige, pravilnici i zakoni koji govore o odgoju, obrazovanju i školstvu prenose poglede ili pravila koji idu u istome smjeru. Nužno je učenicima „osigurati sustavan način poučavanja, poticati i unapređivati njihov intelektualni, tjelesni, estetski, društveni, moralni i duhovni razvoj u skladu s njihovim sposobnostima i sklonostima“ [13]. Prema [8] učenicima s teškoćama u razvoju ograničeno je njihovo puno, učinkovito i ravnopravno sudjelovanje u odgojno-obrazovnom procesu zbog međudjelovanja vlastitih sposobnosti i čimbenika iz okoline.

U povijesnom razvoju zakonskih okvira koji se odnose na učenike s teškoćama primjećuje se promjena u terminologiji. U članu 4. Pravilnika o osnovnoškolskom odgoju i obrazovanju učenika s teškoćama u razvoju iz 1991. godine stoji kako se „potpuna odgojno-obrazovna *integracija* ostvaruje uključivanjem učenika s lakšim teškoćama u razvoju u razredni odjel osnovne škole, u pravilu od I. do IV. razreda.“ Osim potpune integracija, u hrvatskom školstvu postoji i djelomična integracija koja podrazumijeva uključivanje učenika s teškoćama u razvoju na način da se dio nastave odvija u redovnom razrednom odjelu, a dio u posebnom razrednom odjelu. Ministarstvo znanosti, obrazovanja i sporta 2021. godine donosi *Smjernice za rad s učenicima s teškoćama* u kojima se spominje, pojašnjava i usmjerava prema *inkluzivnom* odgoju i obrazovanju. Društvene promjene, promjene u razvojnim karakteristikama učenika te prisutnost sve većeg broja učenika s teškoćama uvjetuju nam nove pristupe i novu terminologiju. „Novi pojam ne uvodi se samo kao promjena imena već i zato što se želi promijeniti pristup, način rada, mišljenje i stavove prema djeci s posebnim potrebama; zapravo želi se naglasiti da je prihvaćanje i poštivanje različitosti (i djece i osoba s posebnim potrebama) jedino pravo rješenje daljnje suradnje među svima nama [4]“. Upravo zbog toga temelj pojma inkluzija podrazumijeva proces, dok je u integraciji naglasak bio na uključivanju. Matilda Karamatić Brčić [3] daje jasan pregled razlika u pojmovima obrazovna inkluzija i obrazovna integracija. *„Uspješnost procesa integracije je ovisila o djetetu, odnosno o njegovoj sposobnosti djelomičnog ili potpunog uključivanja u proces redovnog obrazovanja. Inkluzija se javlja kao puno širi pojam uključenosti svih u odgoj i obrazovanja s naglaskom na cjelokupni školski sustav i odgojno-obrazovni proces. Znači, nije važno samo uključiti dijete, važnije je osigurati kadrovske, materijalne i tehničke uvjete da ono participira u sustavu u kojem se poštuju individualne različitosti i potrebe svih učenika* [3]“.

Poučavanje učenika s posebnim potrebama zahtijeva korištenje različitih strategija umjesto tradicionalnih metoda poučavanja, ali u praksi većina učitelja čini malo prilagodbi zbog ograničenog znanja u pogledu nastavnih prilagodbi koje je potrebno pružiti u inkluzivnom obrazovanju učenicima s posebnim potrebama [2].

II. STRATEGIJE POUČAVANJA I INDIVIDUALIZACIJA POSTUPAKA UČENJA I POUČAVANJA U INKLUZIVNOM OBRAZOVANJU

Inkluzivno obrazovanje pred učitelje postavlja izazov da razvijaju široki spektar strategija poučavanja koje će uključiti učenika s teškoćama u razrednu cjelinu i omogućiti mu napredak. Nacionalno vijeće za specijalno obrazovanje 2010. godine izdaje Pregled literature o načelima i praksi u Inkluzivnom obrazovanju djece s posebnim obrazovnim potrebama [13]. U spomenutom izvješću (prema Davis, Florijan, 2004.) navodi se da postoje rezultati istraživanja koji pokazuju kako kombinacija nastavnih strategija i pristupa proizvodi snažnije učinke od poučavanja samo s jednom strategijom. Inkluzija će zaživjeti u učionicama u svom punom smislu kada će načini poučavanja biti dovoljno fleksibilni i raznoliki kako bi zadovoljili različite zahtjeve svih učenika u razredu. Izvješće Nacionalnog vijeća za specijalno obrazovanje navodi neke moguće strategije diferencijacije u razredu, izvedene iz istraživanja o učinkovitoj nastavi:

- *Koncept skele:* Metafora *koncept skele* potječe iz rada sovjetskog psihologa Vygotskog, a koji je predložio da uz pomoć odrasle osobe učenici mogu izvršiti zadatke koje inače samostalno ne mogu izvršiti. Tijekom učenja i izgrađivanjem znanja zadaci se mogu sve bolje izvoditi, skela se uklanja, a učenici postaju sve samostalniji.
- *Suradničko učenje:* Suradničko učenje najbolje je dokumentiran pristup za promicanje uspješne inkluzije u razrede s raznolikom skupinom učenika koji donosi bolja obrazovna postignuća, društvenu prihvatljivost, bolju razrednu klimu i empatiju.
- *Vršnjačko podučavanje*: Široki raspon studija podupire učinkovitost vršnjačkog podučavanja u kojem učenici iste dobi pomažu u učenju jedni drugima. Iako zahtjeva vrijeme i dobru organizaciju, osim poboljšanja samopoštovanja i društvene interakcije ova metoda „oslobađa“ učitelja dajući mu vremena za pružanje izravnije podrške potrebitim učenicima.
- *Izravna uputa*: Većina učenika s teškoćama u učenju zahtijeva direktnu i eksplicitnu uputu koja povećava njihovu pozornost i akademski angažman. Razina težine zadataka mora biti pažljivo i promišljeno postavljena kako bi se učenicima osigurale šanse za uspjeh. Važna je trenutna povratna informacija učenicima, ispravak i ohrabrenje.
- *Suradničko podučavanje:* Suradničko podučavanje *(eng. Co-teaching)* podrazumijeva oblik redovite suradnje između učitelja razredne ili predmetne nastave i defektologa. Suradničko podučavanje ima više oblika, a najuobičajenija je kombinacija kada jedna osoba podučava, a druga osoba pomaže ili podupire učenje (provjerava razumijevanje, podržava rad, upravlja ponašanjem i sl.).

Smjernice za rad s učenicima s teškoćama [11] učiteljima i nastavnicima daju preporuke za usmjeravanje pažnje tijekom planiranja procesa učenja i poučavanja. U postupku individualizacije potrebno je obratiti pozornost na sljedeće:

- *predstavljanje sadržaja*
- *vremenski okvir potreban za učenje ili izvršavanje zadatka*
- *aktivno uključivanje učenika*
- *načine vrednovanja*
- *osiguravanje primjerenih prostornih uvjeta*
- *prilagodba materijala i asistivne tehnologije* [11].

III. MOGUĆNOSTI DIGITALNIH ALATA U INKLUZIVNOM OBRAZOVANJU

Helena Josjö [2] u svome radu prikazuje postojeća istraživanja o koristi obrazovne tehnologije u učenju učenika s teškoćama (prema Maor i sur., 2011., Peterson-Karlan, 2011., Williams i sur., 2006.) te vlastito istraživanje u kojem rezultati ukazuju na učinkovitost korištenja digitalnih alata za učenike s poteškoćama u čitanju i pisanju i učenike s disleksijom. Josjö prikazuje studiju slučaja iskustava učenika i učitelja u slovenskim školama u primjeni obrazovne tehnologije u inkluzivnom obrazovanju (prema Starcic, 2010.) Implikacije studije su da poboljšani pristup računalima i Internetu u kontekstu školskog rada može poboljšati digitalnu pismenost i e-sudjelovanje učenika s posebnim obrazovnim potrebama te da korištenje digitalnih alata za učenje može poboljšati nastavu usmjerenu na učenika i individualizaciju. Od iznimne je važnosti znanje o korištenju digitalnih alata od strane učitelja kao i pedagoško znanje o samim teškoćama. Govoreći o digitalnim alatima kao podršci inkluzivnom obrazovanju važno je osvijestiti da je za uspjeh učenika s teškoćama u učenju potreban balans tehnologije i pedagogije, odnosno prilagođene nastave.

U ostvarivanju ravnopravnog sudjelovanja svih učenika u odgojno-obrazovnom sustavu može pomoći korištenje dostupnih školskih resursa kao što su besplatni digitalni alati. Raznolika tehnološka rješenja mogu učenicima olakšati praćenje nastave i izvršavanje školskih obveza. Tehnološka rješenja mogu biti naročito korisna u podržavanju različitih načina obrade podataka učenika jer omogućavaju reorganizaciju materijala radi boljeg razumijevanja. Prednosti korištenja digitalnih alata u nastavi višestruke su:

- potiču motivaciju za učenje (nastava postaje zanimljivija)
- korištenje multi senzornih kanala za primanje informacije omogućuje lakše usvajanje novih informacija
- prilagođavanje vrste materijala i tempa učenja individualnim potrebama učenika potiče aktivnu ulogu učenika u procesu učenja
- olakšana je kontrola rezultata učenja ili ponavljanja, i učeniku i učitelju
- povezuje se proces školskog učenja sa svakodnevnim aktivnostima učenika
- omogućava se laka i efikasna grupna izrada projekata [6].

Obradović, Vučetić [7] navode kako interaktivni oblici elektronskog učenja mogu poboljšati vještine mišljenja i rješavanja problema, produbiti kvalitetu naučenog te potaknuti participativno učenje i kreativnost. U svom stručnom radu dalje navode kako Europska komisija (prema European Agency for Development in Special Needs Education, 2011) smatra da je uključivanje digitalnih alata u proces obrazovanja poželjan smjer u borbi protiv diskriminacije i ključno sredstvo za promicanje jednakosti te da odgojno obrazovnim ustanovama prioritet mora biti obuka prosvjetnih radnika u korištenju i izradi digitalnih materijala prilagodljivih učenicima svih sposobnosti i na svim razinama obrazovanja. U školama se ne koriste sve mogućnosti informacijsko-komunikacijske tehnologije stoga se uviđa potreba za tehničkom i društvenom podrškom kako bi se premostila inkluzija učenika s posebnim potrebama [2]. U pregledu postojećih istraživanja Josjö [2] navodi da bi učitelji, defektolozi i stručni suradnici trebali znati kako koristiti informacijsko-komunikacijsku tehnologiju za učenike s posebnom obrazovnim potrebama (prema Anderson i sur., 2009.) što implicira nužnosti cjeloživotne edukacije spomenutih aktera u inkluzivnom obrazovanju.

IV. KAHOOT!, WORDWALL, CANVA, POWERPOINT

U sljedećem poglavlju analizirana su četiri najzastupljenija digitalna alata koja koriste ispitanici istraživanja provedenog u studenom ove godine, u svrhu pisanja ovoga rada, a to su: *Kahoot!, Wordwall, Canva i PowerPoint.*

Kahoot! je interaktivni digitalni alat namijenjen izradi kvizova, diskusija i upitnika. Alat intenzivno koristi elemente učenja igrom u stvarnom vremenu. Sustav bodovanja temelji se na bodovima dobivenima za točan odgovor i vremenu unutar kojeg se daju točni odgovori. Učenici na svojim uređajima kojima pristupaju kvizu imaju prikazane samo odgovarajuće boje i oblike vezane za odgovore (bez teksta odgovora) i odabiru ono polje koje označuje odgovor za koji smatraju da je točan. Nakon svakog odgovora dobivaju povratnu informaciju o tome jesu li odabrali točan odgovor, koliko su bodova osvojili, poredak na ljestvici prema rezultatima ocjenjivanja i ukupan broj bodova koji su do sada osvojili. Na ekranu nastavnika nakon svakog pitanja prikazuje se koliko je učenika odabralo koji odgovor i konačna ljestvica poretka, gdje je prikazano pet učenika s najviše bodova. Kahoot! je jednostavan alat koji omogućuje novi način motiviranja i uključivanja učenika u nastavu. Elementi igre u učenicima mogu potaknuti natjecateljski duh i na taj ih način motivirati da se što više posvete nastavnom sadržaju. Prednost je alata što se može koristiti na bilo kojem uređaju koji ima preglednik, pristup internetu i podržan prikaz HTML5/Java Script sadržaja. Nedostatak je nemogućnost dugotrajnog praćenja rezultata učenika budući da takva funkcionalnost u alatu ne postoji. Rezultat se učenika prati samo na jednom kvizu i to u vrijeme kad se on održava. U konačnici alat je preporučljiv kao novi način za poticanje učenika i uvođenje dinamike u obradi nastavnog sadržaja [9].

Wordwall je online alat za izradu interaktivnih igara u kojima je cilj doći do rješenja kroz razne zadatke u obliku popularnih igara, kao što je kviz, labirint, anagram, i drugi. Krasi ga jednostavnost korištenja i jednostavna mogućnost dijeljenja sadržaja kroz poveznicu i bez potrebe za izradom korisničkih računa od strane učenika. Kroz interaktivne igre i alate koje nudi Wordwall, edukacija i provjere znanja mogu biti zabavnije. Na primjer, na satu geografije učenici mogu pojmove povezati s mjestom na karti. Primjena Wordwall aplikacije na satovima ostalih predmeta također je puna potencijala kroz bogatu ponudu alata. Wordwall aktivnosti mogu se zadati kao interaktivni zadaci i obrazovne igre koje rješava učenik. Kada učitelj postavi igru ili aktivnost kao zadatak, učenici su usmjereni isključivo na tu aktivnost bez da im pažnju ometa odlazak na glavnu stranicu aktivnosti [10].

Sve raširenji besplatni web alat u učionicama je Canva, alat za grafički dizajn. Brzina i jednostavnost alata zasigurno su glavne karakteristike koje ovaj alat čine zanimljivim učiteljima. Osim unaprijed definiranih oko 65000 predložaka koji su podložni promjenama, Canva omogućuje i definiranje vlastitih predložaka u izradi plakata, diploma, brošura, pohvalnica, čestitaka, kalendara, umnih mapa, posjetnica, pozivnica, prezentacija, video zapisa i drugih obrazovnih materijala. Dizajni su svrstani prema spomenutim kategorijama što omogućava brzo i lako snalaženje. Canva pruža i mogućnost timskog rada na obrazovnim sadržajima, odnosno dijeljenje materijala s više korisnika. Za korištenje ovog web alata potrebna je registracija e-mail adresom ili koristeći svoj Facebook ili Google korisnički račun.

Microsoft PowerPoint svjetski je poznati, dostupan i omiljeni među učiteljima zbog svoje jednostavnosti korištenja. Osim osnovnih mogućnosti, koje učitelji većinom koriste (oblikovanje pozadine, odabir fonta, boja, naslova, teksta i natuknica), PowePoint nudi i brojne multimedijske mogućnosti izrade digitalnih sadržaja poput ugradnje slike, pokreta, glazbe i videa. Nadalje, u PowerPoint prezentacije mogu se ugraditi i povezati grafički, auditivni i audio-vizualni elementi različitih formata. Hipertekstualne poveznice i gumbi pružaju odmak od linearnog prikazivanja sadržaja te omogućuju korisniku individualno kretanje kroz sadržaj [3].

V. METODOLOGIJA ISTRAŽIVANJA

Budući da je cilj ovog istraživanja bio ispitati i prikazati korištenje digitalnih alata u redovnoj osnovnoškolskoj nastavi tijekom rada s djecom s teškoćama u učenju anketnim upitnikom anonimnog karaktera ispitan je odnos učitelja predmetne i razredne nastave prema digitalnim alatima u obrazovnoj inkluziji djece s teškoćama u učenju. Na osnovu toga, postavljen je

cilj istraživanja: Prikupiti informacije o korištenju digitalnih alata u radu s učenicima s teškoćama u učenju, saznati u kojem nastavnom području učitelji koriste alate i u koju svrhu, provjeriti koji su najčešće korišteni alati.

Istraživanje je željelo provjeriti odnos učitelja razredne i predmetne nastave prema digitalnim alatima u inkluzivnom obrazovanju stoga polazi od sljedećih hipoteza:

H1: Većina učitelja koristi digitalne alate u radu s učenicima s teškoćama u učenju.

H2: Učitelji digitalne alate većinom koriste u STEM predmetima (matematika, fizika, kemija) zbog nastavnih sadržaja koji zahtijevaju kritičko i logičko zaključivanje i razmišljanje.

H3: Digitalne alate učitelji koriste zbog povećanja motivacije učenika.

VI. UZORAK

Istraživanje je provedeno školske godine 2022./2023. na uzorku od 106 učitelja razredne i predmetne nastave, od toga 59 (55,7%) učitelja predmetne nastave i 47 (44,3%) učitelja razredne nastave (Tablica I.). Uzorak je prigodan.

TABLICA I. PRIKAZ UZORKA - BROJ UČITELJA RAZREDNE (RN) I PREDMETNE (PN) NASTAVE PREMA GODINAMA RADNOG STAŽA

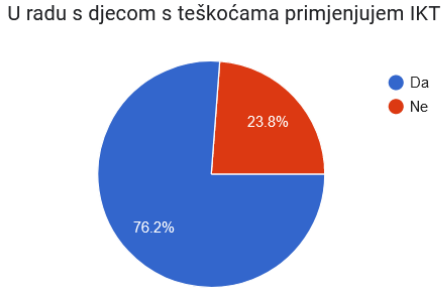
Godine staža	PN[1]	RN[2]	UKUPNO
0-5 godina	8	5	13
5-10 godina	15	4	19
10-20 godina	15	14	29
Više od 20 godina	21	24	45
Ukupno	59	47	106
%	55,66 %	44,34 %	100%

1. predmetna nastava
2. razredna nastava

Odnos učitelja prema korištenu digitalnih alata u inkluzivnom obrazovanju ispitan je anketnim upitnikom. Upitnik se sastojao od 6 pitanja. Pet pitanja su zatvorenog tipa, a jedno nudi mogućnost samostalnog davanja odgovora. Ispitivanje su provele autorice ovoga rada u dostupnim osnovnim školama Koprivničko – križevačke, Bjelovarsko-bilogorske i Varaždinske županije.

VII. PRIKAZ REZULTATA

Sukladno očekivanjima, prije provedenog istraživanja, analizom anketnih upitnika utvrđeno je da većina učitelja (76,2%) primjenjuje digitalne alate u radu s djecom s teškoćama (Graf I.) ($\chi2$ (1, n = 105) = 28,810, p<0,001).



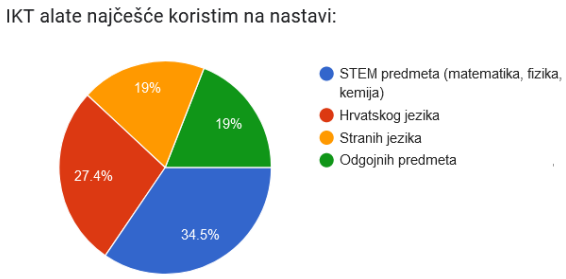
Graf I. Prikaz odgovora na pitanje o korištenju IKT (N=105)

Istraživanjem je provjereno utječu li godine radnoga staža na (ne)primjenu digitalnih alata u radu s učenicima s teškoćama obzirom na uvriježeno mišljenje da „stariji“ učitelji digitalnu tehnologiju koriste u manjoj mjeri u odnosu na mlađe generacije. U Tablici II. vidljiv je prikaz (ne)korištenja digitalnih alata u postotcima po određenoj kategoriji radnoga staža. Iako je postotak nekorištenja digitalnih alata najveći u kategoriji *Više od 20 godina*, odnosno kod najstarijih učitelja, daljnjim statističkim analizama utvrđeno je da ova razlika nije statistički značajna ($\chi2$ (3, n = 105) = 3,652, p=0,302).

TABLICA II. PRIKAZ KORIŠTENJA DIGITALNIH ALATA PREMA GODINAMA RADNOG STAŽA

Godine staža	DA N (%)	NE N (%)	UKUPNO N
0-5 godina	12 (92,3)	1 (7,7)	13
5-10 godina	15 (78,9)	4 (21,1)	19
10-20 godina	23 (79,3)	6 (20,7)	29
Više od 20 godina	30 (68,2)	14 (31,8)	44
Ukupno	80	25	105

Analiza odgovora na pitanje o područjima primjene digitalnih alata (Graf II.) pokazala je da najviše učitelja (34,5%) koristi digitalne alate na nastavi STEM predmeta (matematika, fizika, kemija). Veliki broj ispitanika digitalne alate koristi na nastavi Hrvatskoga jezika (27,4%), a podjednaki broj učitelja (po 19% u svakoj grupi predmeta) digitalne alate koristi u nastavi stranih jezika i odgojnih predmeta (Likovna kultura, Glazbena kultura, Tjelesna i zdravstvena kultura). Uočena razlika nije statistički značajna ($\chi2$ (3, n= 84) = 5,619 , p=0,132-). U Tablici III. dobiveni su zanimljivi, ali i statistički značajni ($\chi2$ (3, n = 84) = 12,485 (p=0,006)) rezultati koji pokazuju da učitelji razredne nastave digitalne alate najviše, i više od učitelja predmetne nastave, koriste na satima Hrvatskoga jezika, dok učitelji predmetne nastave najviše, i više od učitelja razredne nastave, digitalne alate koriste u STEM predmetima i nastavi stranih jezika.



Graf II. Prikaz nastavnih područja u korištenju IKT (N=86)

TABLICA III. PRIKAZ PODRUČJA KORIŠTENJA DIGITALNIH ALATA U RAZREDNOJ, ODNOSNO PREDMETNOJ NASTAVI

Nastavno područje	**PN[1]**	**RN[2]**	**UKUPNO**
Missing (nedostajući odgovori)	***11***	***11***	***22***
Hrvatskog jezika	7	16	23
Odgojnih predmeta	8	8	16
STEM predmeta	20	9	29
Stranih jezika	13	3	16
Ukupno	59	47	106

1. predmetna nastava
2. razredna nastava

Analiza odgovora na tvrdnji *U radu s učenicima s teškoćama digitalne alate koristim zbog* (Tablica IV.) pokazala je da najviše učitelja (30,2%) digitalne alate koristi kako bi povećali motivaciju učenika. Veliki broj učitelja (29,1%) digitalne alate koristi zbog boljeg razumijevanja apstraktnih nastavnih sadržaja ($\chi 2$ (5, n = 86) = 40,977, p<0,001).

TABLICA IV. DESKRIPTIVNA STATISTIKA POJEDINIH TVRDNJI O SVRSI KORIŠTENJA IKT U RADU S UČENICIMA S TEŠKOĆAMA

Tvrdnja	**PN[1]**		**RN[2]**		**Ukupno**	
	N	%	N	%	N	%
Veće motivacije učenika	14	16,28	12	13,95	26	30,23
Boljeg razumijevanja apstraktnih nastavnih sadržaja	13	15,12	12	13,95	25	29,07
Bržeg i učinkovitijeg rješavanja zadataka	5	5,81	2	2,33	7	8,14
Trenutne povratne informacije za učenike	2	2,33	3	3,49	5	5,82
Zornog prezentiranja nastavnih sadržaja	15	17,44	6	6,98	21	24,42
Nešto drugo	1	1,16	1	1,16	2	2,32
Ukupno	**50**	**58,14**	**36**	**41,86**	**86**	**100**

1. predmetna nastava
2. razredna nastava

Zbog pretpostavke da učitelji koriste digitalne alate u radu s učenicima s teškoćama u učenju ispitanici su upitani koje alate najčešće koriste u nastavi (N=50). Najviše ispitanika navodi kako koristi Kahoot! (30%) i Wordwall (28%). Canvu koristi 14% ispitanika, a PowerPoint 12% ispitanika. Osim spomenutih alata pojedini ispitanici navode sljedeće alate: Quizziz, Liveworksheet, Plikers, Mathduel, Learning apps, Genially, Geogebru i drugo.

VIII. ZAKLJUČAK

Dvije neosporne činjenice su da se broj učenika s teškoćama u učenju u školama povećava te da su učitelji dužni svakom učeniku omogućiti učinkovito i ravnopravno sudjelovanje u odgojno-obrazovnom procesu. Ovim radom željela se prikazati važnost, ali i navika korištenja digitalnih alata u inkluzivnom obrazovanju.

Cilj istraživanja bio je prikupiti informacije o korištenju i svrsi korištenja digitalnih alata u osnovnoškolskom inkluzivnom obrazovanju te provjeriti kojim se sve digitalnim alatima ili platformama učitelji služe u radu s djecom s teškoćama u učenju.

Iako postoje potencijalna ograničenja za interpretaciju rezultata, a to su veličina uzorka, objektivnost ispitivača i nepoznavanje trenutnih materijalnih uvjeta rada ispitanika, istraživanje donosi sljedeće zaključke. Rezultati istraživanja sugeriraju zaključak da gotovo većina ispitanika koristi digitalne alate u radu s učenicima s teškoćama u učenju (76,2%) čime se potvrđuje prva postavljena hipotezu (H1). Ispitanici digitalne alate najčešće koriste u STEM predmetima (34,5%) i nastavi Hrvatskoga jezika (27,4%) no uočena razlika nije statistički značajna stoga se odbacuje druga hipoteza (H2). Nadalje, rezultati potvrđuju posljednju hipotezu (H3): Digitalne alate učitelji koriste zbog povećanja motivacije učenika. Iz cjelokupnog rada vidljivo je da učitelji uviđaju mogućnosti i prednosti korištenja digitalnih alata u inkluzivnom obrazovanju.

Prikazana istraživanja ukazuju na važnost prilagodbe nastave za učenike s teškoćama korištenjem digitalnih alata na način da informacijsko-komunikacijska tehnologija postane most za kvalitetno provođenje obrazovne inkluzije. Da bi se ova ideja realizirala neizostavno je redovito educiranje i međusobna suradnja odgojno – obrazovnih aktera (učitelja, defektologa i stručnih suradnika) o tome kako koristiti digitalne alate za sve učenike, uključujući i učenike s teškoćama. Važno je naglasiti, iako informacijsko-komunikacijska tehnologija može imati mnoge koristi za učenike s teškoćama ona ne može zamijeniti učitelja.

Analizirani rezultati potiču na sljedeća promišljanja i/ili daljnja istraživanja: Zašto još veći broj učitelja ne koristi prednosti lako dostupnih digitalnih alata? Na koji način se digitalni alati mogu još više implementirati i u druga nastavna područja, poput odgojnih predmeta ili stranih jezika? Kakva je percepcija studenata Učiteljskog studija o digitalnim alatima kao alatima za obrazovnu inkluziju?

REFERENCE

[1] R.A. Berk, "Research on PowerPoint®: From Basic Features to Multimedia," *International Journal of Technology in Teaching and Learning*, vol 7, 1, pp. 24-35, 2011.

[2] H. Josjö, "ICT and inclusion, Teachers' perceptions on the use of information and communication technology for students with special educational needs in general educational settings", UMEA Universitet, 2012.

[3] M. Karamatić Brčić, "Svrha i cilj inkluzivnog obrazovanja", ACTA IADERTINA, vol 8, pp. 39-47, 2011.

[4] S. Kobešćak, "Što je inkluzija?, Dijete, vrtić, obitelj", Časopis za odgoj i naobrazbu predškolske djece namijenjen stručnjacima i roditeljima, vol. 6, 21, pp. 23-25, 2000.

[5] R. Miljević-Riđički, D. Maleš, M. Rijavec, Odgoj za razvoj, 2nd ed., Zagreb: Alinea, 2001., pp. 27.

[6] S. Obradović, D. Bjekić and L. Zlatić, "Obrazovanje nastavnika za rad u e-okruženju sa učenicima sa specifičnim smetnjama u učenju," in Zbornik radova, vol. 1, M. Danilović, D. Golubović, S. Popov, Eds., Čačak: Tehnički fakultet, 2011, pp. 316-325.

[7] S. Obradović, M. Vučetić, "IKT u inkluzivnom brazovanju", 5th International Conference, Faculty of Technical Sciences Čačak, 2014.

[8] Pravilnik o osnovnoškolskom i srednjoškolskom odgoju i obrazovanju učenika s teškoćama u razvoju, Narodne novine, 24/ 2015.

[9] T. Rađep, T. Pavičić Zajec, B. Marčinkovič, "Ukrotimo kompetence 21. stoletja", Profil Klett, Zagreb, 2019., http://slo-erasmus-osbos.splet.arnes.si/files/2019/10/Priru%C4%8Dnik_Ukrotimo-kompetencije-2kor.pdf [pristupljeno: 3.2.2023.]

[10] M. Režić, "Wordwall – učimo igranjem, e-laboratorij, https://e-laboratorij.carnet.hr/wordwall-ucimo-igranjem/ [pristupljeno: 2.1.2023.]

[11] Smjernice za rad s učenicima s teškoćama, Ministarstvo znanosti i obrazovanja, 2021.

[12] Zakon o odgoju i obrazovanju u osnovnoj i srednjoj školi, Narodne novine, 68/2018.

[13] E. Winter, P. O'Raw, "Literature Review of the Principles and Practices relating to Inclusive Education for Children with Special Educational Needs", The National Council for Special Education, 2010.

Suvremene metode poučavanja robotike i umjetne inteligencije

Ivana Ružić
I. osnovna škola Čakovec, Čakovec, Hrvatska
ivana.ruzic@skole.hr

***Sažetak* – Ubrzani razvoj tehnologije uveo je primjenu robotike i umjetne inteligencije u gotovo sve sfere ljudskog života. Obrazovanje često zaostaje za primjenom napredne tehnologije, a još češće zaostaje s poučavanjem i praktičnom primjenom i stvaranjem učeničkih radova u području najnovije tehnologije. Napredna tehnologija, u ovom primjeru robotika i umjetna inteligencija mogu učenje učiniti interaktivnim i privlačnim za djecu i mlade koji često imaju smanjen raspon pažnje. Budući da najnovija tehnologija ima futurističku dimenziju, ona razvija znatiželju kod djece i mladih, potiče veću motivaciju, koncentraciju, angažman učenika, logično razmišljanje, kritičko promišljanje, vještine rada u timu, pozitivan stav prema istraživanju i rješavanju problema. Rad prikazuje primjenu robotike i umjetne inteligencije u poučavanju učenika 5. – 8. razreda osnovnih škola koristeći suvremene metode poučavanja poput istraživačkog učenja i učenja temeljenog na problemima iz stvarnog života.**

Ključne riječi – robotika, umjetna inteligencija, obrazovanje djece i mladih

I. UVOD

Tehnologija oblikuje sadašnjost i budućnost čovječanstva, a obrazovanje omogućuje razvoj tehnologije. Manje od 2% ukupnog broja ljudi na svijetu, razvija, gradi i upravlja tehnologijom. Više od 98% ljudi su „obični“ korisnici koji ne znaju ništa o unutarnjoj arhitekturi bilo koje tehnologije, oni je samo kupuju i koriste.

Jedna od trenutno najpopularnijih tehnologija jesu robotika i umjetna inteligencija (UI). Robotika i umjetna inteligencija su dva različita područja koja mogu koegzistirati. S razvojem sve sofisticiranijih sustava umjetne inteligencije, povećava se i primjena umjetne inteligencije u robotici.

Korištenje robotike i umjetne inteligencije u svakom području ljudskog djelovanja ubrzano raste iz dana u dan. U području obrazovanja postoje velika potreba i potencijal za primjenu, poučavanje i stvaranje djece i mladih u području robotike i umjetne inteligencije.

Robotika i umjetna inteligencija ne samo da mogu potaknuti razvoj vještina programiranja i računalnog razmišljanja, već i poboljšati inovativnost i vještine kritičkog mišljenja. Robotika i umjetna inteligencija STEM koncepte čine interaktivnim i privlačnim za djecu i mlade koji često imaju smanjen raspon pažnje. Budući da robotika i umjetna inteligencija imaju futurističku dimenziju, one razvijaju znatiželju kod djece i mladih, potiču veću koncentraciju, logično razmišljanje i pozitivan stav prema rješavanju problema. Uz pomoć različitih vrsta robota i računalnih programa za robotiku i umjetnu inteligenciju, učitelji, nastavnici i stručni suradnici mogu provoditi suradničke aktivnosti koje olakšavaju razvoj vještina vođenja i suradnje, socijalizaciju i timski rad kod djece i mladih.

II. UMJETNA INTELIGENCIJA U ROBOTICI

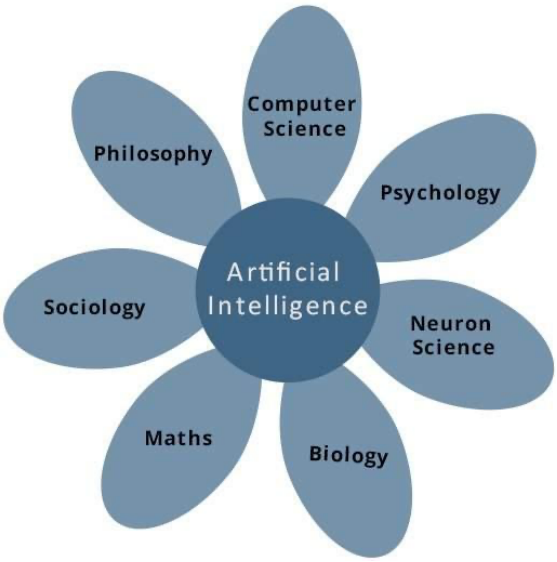
Umjetna inteligencija je grana računalstva koja uključuje razvoj računalnih programa sa ciljem izvršavanja određenih zadataka, koji bi inače iziskivali ljudsku inteligenciju. Algoritmi koji se zasnivaju na umjetnoj inteligenciji mogu učiti, percipirati, obavljati određene zadaće, prevoditi i logički zaključivati. Prema ocu umjetne inteligencije, Johnu McCarthyju, umjetna inteligencija je:“ grana znanosti i inženjerstva koja stvara inteligentne strojeve, a posebice inteligentne računalne programe”[1].

Može se reći da je umjetna inteligencija način stvaranja računala i/ili robota koje kontroliraju programi koji inteligentno razmišljaju, na sličan način na koji razmišljaju ljudi. Umjetnu inteligenciju predstavljaju sofisticirani sustavi koji oponašaju ljudski um kako bi učili, rješavali probleme i donosili odluke bez potrebe za posebno programiranim uputama.

Ciljevi umjetne inteligencije:

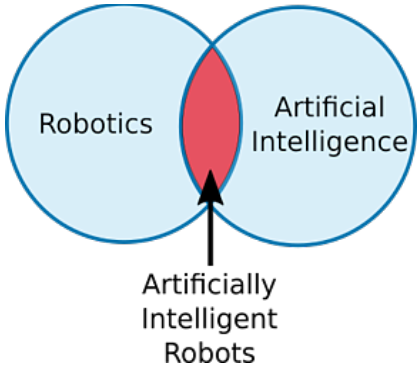
- stvoriti ekspertne sustave – sustave koji pokazuju znakove inteligentnog ponašanja, uče, demonstriraju, objašnjavaju i podučavaju njihove korisnike,
- implementirati ljudsku inteligenciju u strojeve − stvaranje sustava koji razumiju, razmišljaju, uče i ponašaju se kao ljudi.

Umjetna inteligencija je grana znanosti i tehnologije koja se temelji na različitim područjima znanosti poput računalstva, biologije, psihologije, lingvistike, matematike i inženjerstva[2]. Umjetna inteligencija nastoji razviti računalne funkcije povezane s ljudskom inteligencijom, primjerice razmišljanjem, učenjem i rješavanjem problema. Slika 1. prikazuje različita područja znanosti koja mogu doprinijeti izgradnji inteligentnog računalnog sustava.



Slika 1. Različita područja znanosti koja doprinose razvoju umjetne inteligencije[2]

Robotika i umjetna inteligencija su dva zasebna područja znanosti koja se minimalno preklapaju[3]. Slika 2. Vennovim dijagramom prikazuje usko područje preklapanja robotike i umjetne inteligencije. Budući da dolazi do preklapanja, korisnike često zbunjuju ovi pojmovi.



Slika 2. Vennov dijagram robotike i umjetne inteligencije[3]

Robotiku možemo opisati kao granu inženjerstva i računalnih znanosti koja razvija strojeve za izvršavanje programiranih zadataka bez daljnje ljudske intervencije.

Robotika se sastoji od mehaničkog i električnog inženjerstva i računalstva, koje se bave dizajniranjem, stvaranjem i primjenom robota. Roboti su umjetni agenti koji sudjeluju u stvarnom svijetu. To su strojevi koje se programira, koji obično autonomno ili polu-autonomno obavljaju niz radnji. Svrha robota je da manipuliraju predmetima tako da ih percipiraju, dižu, premještaju, mijenjaju njihova fizička svojstva i uništavaju. Time zamjenjuju ljudsku radnu snagu od izvršavanja ponavljajućih radnji bez umora i osjećaja dosade.

Aspekti robotike:

- roboti imaju mehaničku konstrukciju i oblikovani su kako bi izvršili određeni zadatak,
- roboti imaju elektroničke komponente koje napajaju i kontroliraju stroj te ugrađeni računalni program koji određuje vrijeme i vrstu aktivnosti koje robot izvršava.

Spoj umjetne inteligencije i robotike čine roboti koji koriste umjetnu inteligenciju, to su roboti koji su specijalizirani za obavljanje složenih zadataka, a njima upravljaju inteligentni programi[3]. Umjetna inteligencija pomaže robotima obavljati važne zadatke, omogućuje im „ljudski pogled“ s pomoću kojeg pronalaze i prepoznaju različite predmete. Roboti se razvijaju na principu strojnog učenja i pomoću velikog broja podataka koji se koriste za unaprjeđenje sustava. omogućuju im bolje prepoznavanje predmeta i izvršavanje radnji. „Vid računala“ je proces unutar kojeg računalo percipira slike i videozapise u digitalnim formatima. Umjetna inteligencija u robotici pomaže robotu da izvrši određeni zadatak, ali i omogućava da inteligentnije postupa u različitim scenarijima.

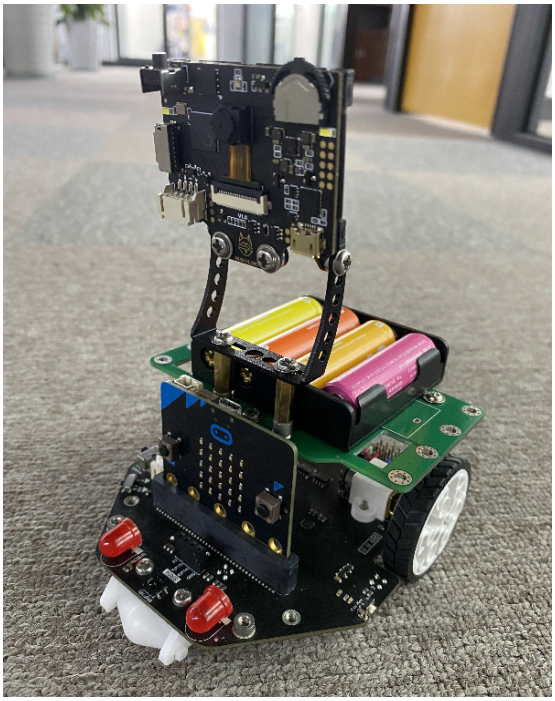
Danas se roboti s umjetnom inteligencijom koriste u zdravstvu, agrikulturi, industriji automobila, skladištima itd. Među najpoznatijim primjerima umjetne inteligencije i robotike jesu roboti Sophia, Digit, Pepper, Atlas, Spot, NAO i mnogi drugi.

III. UMJETNA INTELIGENCIJA I ROBOTIKA U OBRAZOVANJU

ARTIEbot je pokretan robot s kamerom koji koristi umjetnu inteligenciju, a osmišljen za projekt *„Erasmus+ ARTIE: Artificial Intelligence in Education – challenges and opportunities of the new era: development of a new curriculum, guide for educators and online course for students“*. ARTIEbot je izgrađen od micro: Maqueen Plus i HuskyLens UI kamere. Riječ je o naprednom STEM obrazovnom robotu za micro:bit koji ima velik broj mogućnosti: optimizirano upravljanje napajanjem, napajanje baterijama velikog kapaciteta i potpuno je kompatibilan s HuskyLens UI kamerom, što ga čini pristupačnim nastavnim alatom. Micro:Maqueen Plus ima visoko kvalitetne značajke i nudi više priključaka za različita proširenja. Nije pogodan samo za nastavu u učionici, već se može koristiti i za izvannastavne aktivnosti i razna robotička natjecanja.

Funkcionalnosti ARTIEbota:

- pronalaženje lica,
- prepoznavanje lica,
- praćenje lica,
- pronalaženje predmeta,
- praćenje predmeta,
- praćenje linije.



Slika 3. ARTIEbot

Kurikulum, metodički priručnici za učitelje i online tečaj za učenike (namijenjen učenicima od 5. do 8. razreda osnovne škole) koji su rezultat navedenog projekta, razvijeni su na način da učenici samostalno, pojedinačno ili u timu, uz pomoć učitelja sastavljaju svog robota i stvaraju programe za pronalaženje, prepoznavanje i praćenje lica i različitih predmeta, stvaranju govora, upravljanje glasovnim uputama i sl. koristeći se blokovskim programskim jezicima.

IV. SUVREMENE METODE POUČAVANJA UMJETNE INTELIGENCIJE I ROBOTIKE

Suvremene metode poučavanja temelje se na promjeni paradigme poučavanja u kojoj je učenik onaj koji je u procesu aktivnog učenja, kroz istraživanje i interakciju s vršnjacima, usvaja nove spoznaje i razvija svoje vještine. Promjenom paradigme učenja učitelj poprima ulogu onoga koji vodi i savjetuje, a učenik pojedinačno ili češće u timu, kritički razmišlja, istražuje i povezuje. Najčešće suvremene metode poučavanja jesu istraživačko učenje (engl. *Inquiry-Based learning*) i učenje temeljeno na problemu (engl. *Problem-Based learning*).

Istraživačko učenje predstavlja oblik nastave u kojem učenici uz podršku učitelja vode nastavni proces te istraživanjem razvijaju sposobnost razumijevanja novih sadržaja, a najčešće se provodi po 5E modelu (engl. *Engagement, Exploration, Explanation, Elaboration, Evaluation*)[4].

Učenje temeljeno na problemu je metoda poučavanja u kojoj se složeni problemi iz stvarnog svijeta koriste kao sredstvo za promicanje učenja koncepata i načela učenika za razliku od izravnog predstavljanja činjenica i koncepata[5].

Nastavne aktivnosti s odgovarajućim ishodima koje primjenjuju načela istraživačkog učenja i učenja temeljenog na problemu, a koriste robotiku i umjetnu inteligenciju organizirane su u kurikulumu: *„Artificial Intelligence in Education: A new curriculum“*, opisane u metodičkim priručnicima za učitelje: *„Challenges and opportunities of the new era: a handbook for teachers“* i organizirane kao online tečaj za učenike: „Twinspace – an open window to other educational *realities: Online course for students“*.

Nastavne aktivnosti osmišljene i pripremljene u projektu *„Erasmus+ ARTIE: Artificial Intelligence in Education – challenges and opportunities of the new era: development of a new curriculum, guide for educators and online course for students“* uključuju slijedeće aktivnosti:

- samostalno sastavljanje robota ARTIEbota od zadanih dijelova (micro:bit, Maqueen Plus i HuskyLens UI kamere),
- programiranje robota ARTIEbota koji se kreće u okruženju na autonoman način, izbjegava prepreke, prepoznaje objekte na stazi, obavlja radnje ovisno o prepoznatim objektima,
- stvaranje programa koji omogućuje verbalnu komunikaciju s ljudima: prepoznavanje govora i pretvaranje u tekst (eng. *Speech-To-Text, STT*),
- stvaranje programa za računalno stvaranje govora (eng. *Text-To-Speech, TTS*),
- stvaranje programa koji prepoznaje jezik i obavlja računalnu analizu sentimenta,
- stvaranje programa koji omogućuje neverbalnu komunikaciju s ljudima (prepoznavanje i praćenje lica, prepoznavanje emocija i gesti),
- stvaranje programa koji prepoznaje objekte,
- stvaranje programa za glasovno upravljanje.



Slika 4. Primjer programa za prepoznavanje govora, jezika i analizu sentimenta

Temeljem usvojenih znanja i razvijenih vještina programiranja, učenici mogu stvarati vlastita programska rješenja za prepoznavanje zaštićenih biljnih i životinjskih vrsta, prepoznavanje i sortiranje različitih vrsta otpada,

prepoznavanje emocija i gesti, glasovnim uputama upravljati ARTIEbotom u obavljanju zadatka i dr.

Kurikulum, metodički priručnici za učitelje i online tečaj za učenike testirani su u pet različitih škola u Hrvatskoj, Poljskoj i Portugalu, u nastavi ih je koristilo osamnaest učitelja, nakon čega je provedeno vrednovanje. Anonimni online upitnik za vrednovanje ponuđenog kurikuluma, metodičkih priručnika i online tečaja ispunilo je osamnaest učitelja. Za vrednovanje ponuđenih projektnih rezultata korištena se Likertova ljestvica u rasponu od 1 – 5 (*Potpuno se slažem, Slažem se, Nemam mišljenje, Ne slažem se, Uopće se ne slažem*). Rezultati vrednovanja kurikuluma, metodičkih priručnika i online tečaja pokazuju slijedeće:

- 91% ispitanika navelo je da su metodički priručnici i online tečaj tehnički i vizualno atraktivni (inovativni, uključuju multimediju, interaktivnost i suvremene metode poučavanja) (odgovori: *Slažem se* i *Potpuno se slažem*), a 9% nije imalo mišljenje,
- 90% ispitanika navodi da metodički priručnici i online tečaj na smislen i organizirani način nude aktualne i atraktivne obrazovne sadržaje koji nisu dostupni u udžbenicima i drugim nastavnim pomagalima (odgovori *Slažem se* i *Potpuno se slažem*), no 5% ispitanika je suprotnog mišljenja (*Ne slažem se*), dok 5% ispitanika nema mišljenje,
- 89% ispitanika smatra da online tečaj za učenike ima transformativnu funkciju i mijenja paradigmu poučavanja (odgovori *Slažem se* i *Potpuno se slažem*), a 11% ispitanika se ne slaže s tim mišljenjem (*Ne slažem se*),
- 70% ispitanika navodi da online tečaj pomaže učenicima u sagledavanju, formuliranju i rješavanju problema odgovarajućim rasporedom sadržaja, pitanja i zadataka, podiže motivaciju i angažman učenika (odgovori *Slažem se* i *Potpuno se slažem*), 10% ispitanika ima suprotno mišljenje (*Ne slažem se*, *Uopće se ne slažem*), a 20% nema mišljenje,
- 90% ispitanika navodi da online tečaj motivira učenike za stjecanje novih znanja, razvoj njihovih interesa, usavršavanje vještina potrebnih za kritički promišljanje i suradničko učenje (odgovori *Slažem se* i *Potpuno se slažem*), no 10% ispitanika je suprotnog mišljenja (*Ne slažem se*),
- 80% ispitanika navodi da online tečaj učenicima omogućuju usvajanje znanja i razvoj vještina, provjeru i vrednovanje napretka (odgovori *Slažem se* i *Potpuno se slažem*), 8% ispitanika je suprotnog mišljenja (*Ne slažem se, Uopće se ne slažem*), a 12% nema mišljenje.

Nakon primjene i evaluacije kurikuluma, metodičkih priručnika i online tečaja u partnerskim ustanovama, učinjene su korekcije u skladu s preporukama učitelja koji su sudjelovali u testiranju i vrednovanju. Svi projektni rezultati javno su dostupni i slobodni za korištenje svim zainteresiranim učiteljima, nastavnicima i stručnim suradnicima na mrežnom mjestu projekta: https://www.erasmus-artie.eu.

V. ZAKLJUČAK

Rad predstavlja primjenu napredne tehnologije (robotike i umjetne inteligencije) u obrazovanju djece i mladih uzrasta od 5. do 8. razreda osnovne škole. Praktična primjena napredne tehnologije omogućava razumijevanje iste, što je preduvjet za budući razvoj vještina programiranja i računalnog razmišljanja, razvoja inovativnost, vještina rada u timu i kritičkog mišljenja kod djece i mladih potrebnih za njihovo buduće obrazovanje, buduće poslove i cjeloživotno učenje.

Rad predstavlja praktične primjere koji čine projektne rezultate međunarodnog partnerstva pod nazivom „Erasmus+ ARTIE: Artificial Intelligence in Education – challenges and opportunities of the new era: development of a new curriculum, guide for educators and online course for students“, iskustva njihovih korištenja i rezultate vrednovanja primjene novog kurikuluma, metodičkih priručnika i online tečaja za učenike.

LITERATURA

[1] McCarthy, J. 1997. What is Artificial Intelligence?, dostupno na: http://www-formal.stanford.edu/jmc/whatisai/whatisai.html

[2] Artificial Intelligence Tutorial, dostupno na: https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm

[3] A. Owen-Hill, What's the Difference Between Robotics and Artificial Intelligence?, srpanj 2021., dostupno na: https://blog.robotiq.com/whats-the-difference-between-robotics-and-artificial-intelligence

[4] O. Bijelonić, Suvremene metode poučavanja, 28.2.2022., dostupno na: https://www.skole.hr/suvremene-metode-poucavanja/

[5] Center for Innovation in Teaching & Learning, Teaching and Learning, dostupno na: https://citl.illinois.edu/citl-101/teaching-learning

Motivacija nastavnika za usvajanje znanja vezanih uz nove tehnologije

A. Tihomirović, S. Šišić
Tehnička škola Zagreb, Zagreb, Hrvatska,
andrej.tihomirovic@skole.hr

Sažetak **- Nastava je dvosmjeran proces za koji je ključna motivacija obje strane. Izvođenje nastave korištenjem novih tehnologija ovisi o nizu faktora, a kvaliteta naučenog ovisi o percepciji učenika. U većini slučajeva današnji učenici su nisko motivirani, „bombardirani" sadržajima koji im odvlače pozornost, te je poučavanje korištenjem novih tehnologija izuzetno zahtjevan i intelektualno intrigantan posao. Motivacija nastavnika je intrinzična, jer ni na koji način nije financijski vrednovana. Može se uočiti i konstantni trend povećavanja broja maturanata koji se izvan redovne škole pripremaju za maturu. Na tržištu rada, nije moguće pronaći zainteresirane predavače iz STEM grupe predmeta, a satnica iz pojedinih predmeta se smanjuje svakom novom reformom, čime se još više umanjuje motivacija nastavnika iz STEM područaja za usvajanjem novih znanja vezanih uz nove tehnologije, čime se spirala negativnih događanja za društvo sve više ubrzava i nepovratno usmjerava u negativnom smjeru. Stručno usavršavanje nastavnika ne informatičara trebala bi ići u dvije kategorije, za nastavnike koji su zainteresirani i one koji nisu. Prije organizacije edukacije trebalo bi provoditi testiranje o kompetencijama i vještinama nastavnika, te bi poslije provedene edukacije trebalo raditi certifikaciju polaznika na način kako se to radi u realnom sektoru, a novo dobiveni certifikati bi trebali financijski utjecali na primanja predavača.**

Ključne riječi – nastava; nove tehnologije; motivacija; STEM; reforme školstva; manjak stručnjaka; certifikacija

I. UVOD

Sustav obrazovanja u RH podijeljen je u obvezno obrazovanje u trajanju od 8 godina, za koje su zadužene osnovne škole. Srednjoškolsko obrazovanje dijeli se na strukovno i gimnazijsko, pri čemu duljina obrazovanja varira između tri do pet godina, te uz uvjet prolaska na državnoj maturi moguće je nastaviti školovanje na višoj odnosno visokoj razini. Na visokoj razini školovanja moguće je završiti formalno školovanje na stručnoj kao i na znanstvenoj razini. Fokus dosadašnjih reformi školstva bio je usmjeren na gimnazijsko obrazovanje, čime se iz fokusa izgubila većina populacije koja pohađa škole u RH. Posljedice takvog načina rada danas rezultiraju činjenicom da je nemoguće pronaći kvalitetnog majstora keramičara, vodoinstalatera, električara ili nekog drugog tipa majstora. Ovaj rad bavi se uzrocima ovakvog stanja i posljedicama koje će tek nastupiti, kroz prizmu srednjoškolskog obrazovanja u struci tehničar za računalstvo.

II. TRENUTNO STANJE

Učenici upisuju srednju školu ovisno o uspjehu u osnovnoj školi bez vanjskog vrednovanja, čime dolazi do poplave super odlikaša. Prema istraživanju Instituta za društvena istraživanja u Zagrebu[1] [1] u RH to je 4,9 % generacije od petog do osmog razreda. Ta djeca prema istoj studiji obitavaju u većim gradovima i većinom upisuju gimnazijske programe, a tek 0,9 % te populacije upisuje srednjoškolske programe iz STEM područja. Kako bi super odlikaši u osnovnim školama postojali u tom obimu, očita je „prilagodba" kriterija ocjenjivanja. Učenici koji upisuju strukovno zanimanje tehničar za računalstvo (zanimanje koje je u fokusu ovog rada) nisu u kategoriji super odlikaša. Prosjek njihovih ocjena iz osnovne škole su tek nešto niži. U razgovoru s velikim brojem polaznika ovog smjera u periodu zadnjih pet godina, u dvije različite županije (Zagrebačka i grad Zagreb), utvrđeno je da je veliki broj ovih učenika u periodu od petog do osmog razreda osnovne škole promijenio dva ili više nastavnika matematike, jednog od osnovnih iz STEM područja. Različiti kriteriji različitih nastavnika, zajedno s potrebom produkcije super odlikaša, uz smanjenje satnice iz matematike, kroz različite reforme, za posljedicu imaju postojanje nehomogenih skupina učenika čije vještine i kompetencije iz matematike u osnovnim školama, za potrebe struke nije dostatno. Također, učenici su u osnovnom školstvu, koje je obvezno, navikli dobiti „nešto za ništa", što je direktna posljedica sve većeg broja učenika s individualiziranim i prilagođenim programom. Uz to, radne navike vezane uz učenje nisu izgrađene kroz sustav osnovnog obrazovanja. Sve do sada navedeno, stavlja nastavnike STEM područja u srednjem školstvu u izrazito neugodan položaj. Plaće u realnom sektoru u STEM području su bitno veće nego plaće nastavnika u školama. Ulazni parametri učenika u srednjim strukovnim područjima iz STEM područja su nedostatni, učenici nemaju izgrađene radne navike. Očekivanja učenika i njihovih roditelja vezanih uz srednju školu su uglavnom visoka, odnosno, u većini slučajeva očekuje se da će učenik ovog smjera nastaviti školovanje poslije mature na višoj ili visokoj razini. Nastavnika iz STEM područja,

[1]B Jokić i Z Ristić Dedić: Analitičko izvješće SUPERODLIKAŠI http://os-bukovac-zg.skole.hr/upload/os-bukovac-zg/images/static3/2254/attachment/ANALITICKO_IZVJESCE-Superodlikasi-IDIZ.pdf (31/01/2023)

bez stalnog zaposlenja, nema na tržištu rada. Postojeći nastavnici u srednjim školama su sve stariji, što za svoju posljedicu ima veće opterećenje za one nastavnike koji nisu kronično bolesni. Tako je opterećenje na pojedine stručnjake iz STEM područja bitno veće od propisane tjedne norme u školi. Zahvaljujući globalnoj pandemiji, školski sustav je bio prisiljen koristi mogućnost rada na daljinu korištenjem novih tehnologija. To je možda i najvažniji pomak školstva u RH unatrag zadnjih dvadesetak godina. U tom periodu nastalo je iznimno puno kvalitetnih materijala za poučavanje u digitalnom obliku, ali velikim dijelom nastavničke populacije provlači se teza da je kvaliteta nastave na daljinu bila mala i da je količina znanja koju su učenici usvojili, ispodprosječna. S druge strane, postoje i drugačiji primjeri koji pokazuju da su djeca koja su se pripremala za natjecanje korištenjem novih tehnologija postigla iznadprosječne rezultate za svoju dob [2] [2]. Metodom razgovora je utvrđeno da učenici srednjih škola provedu na internetu između 48 i 72 sata na tjednoj bazi, pregledavajući multimedijalne sadržaje, koji nemaju nikakvih dodirnih točaka sa školskim aktivnostima, stoga ne čudi da je u velikom broju slučajeva učinkovitost korištenja novih tehnologija kod velikog dijela školske populacije iznimno niska. Prilikom završavanja srednjoškolskog obrazovanja, s godinama sve veći broj učenika se nesamostalno priprema za polaganje mature pohađajući različite tečajeve za pripremu mature, pri čemu troškove snose njihovi roditelji. Kroz razgovor je utvrđeno je da veliki broj polaznika tih tečajeva tek tada počne kontinuirano raditi na usvajanju novih znanja. Sve do sada navedeno upućuje na nastavak negativnog trenda vezanog uz srednjoškolsko STEM obrazovanje.

III. ŽELJENO STANJE

Profesorica Šišić tečajeve na Loomenu/Moodle s učenicima provodi petnaestak godina. Učenici smatraju da ih je postojanje e - tečaja dodatno motiviralo za rad, korištenje e - tečaja im je olakšalo savladavanje gradiva, a izrada radnih zadataka im je pomogla da uvježbaju određeno gradivo te bi ovakav način učenja preporučili i drugima. Navedeni odgovori su iz pred pandemijskog vremena. Odgovori poslije nastave na daljinu se svode na rečenicu: ”Vratite nas u školu.” Iz istraživanja profesorice Šišić, koje nije formalizirano, vidljivo je da većina učenika želi dolaziti u školu, ali cilj nije usvajanje novih znanja već socijalizacija. Ono što treba postići kroz taj oblik socijalizacije je osjećaj pripadnosti, koji se može realizirati kroz rad. S promjenama u strukturi školovanja, koja bi uključivala vanjsko vrednovanje na kraju svake školske godine iz svakog predmeta za svaki razred strukovnih škola, moglo bi se utjecati na motivaciju učenika, povećanje kvalitete njihovog znanja, uspješnost nastavnika i škola, te bi se na temelju rezultata pojedinih generacija moglo utvrditi što treba mijenjati u pojedinom predmetnom kurikulumu. Na taj način bi se stvorila učinkovita negativna povratna veza koja bi sustav držala u ravnoteži, a istovremeno bi imali učinkovit alat za analizu trenutnog stanja, odnosno temelj za donošenje odluka koje bi bile usmjerene k boljitku društva u budućnosti.

IV. MOTIVACIJA NASTAVNIKA

Nastavnik početnik ima obavezu polaganja stručnog ispita unutar dvije godine od početka rada u školi. Uvjet za polaganje stručnog ispita je pedagoški minimum, odnosno položeni ispiti iz psihologije, pedagogije, didaktike i metodike. Navedeni uvjeti, te financijska naknada za rad u sustavu srednjeg školstva, glavni su razlog nedostatka kadra među predavačima. Maslowljeva hijerarhija potreba i ostale motivacijske metode [3] [3] navedene u radu M. Varga, nisu primjenjive u sustavu školskog obrazovanja, zbog manjka financijskih sredstava koje osnivači izdvajaju za odvijanje srednjoškolske djelatnosti. Osiguravanje potreba nastavnika iz STEM područja na primjeru Maslowljeve piramide u sustavu školstva eventualno je dostatno za osiguravanje fizioloških potreba, najniže od pet razina. Navedeno objašnjava nedostatak nastavničkog kadra u STEM području u srednjoškolskom obrazovanju. Postojeći nastavnički kadar u srednjoškolskom obrazovanju uglavnom ima osigurane dvije najniže razine hijerarhije u modelu Maslowljeve piramide, te se njihova motivacija za rad u sustavu srednjoškolskog obrazovanja razlikuje od motivacije nastavnika početnika. Postojeći nastavnici iz STEM područja su uglavnom populacija 40+ godina koji su materijalno zbrinuti i nemaju ambiciju za napredak na društvenoj ljestvici. Određeni broj nastavnika tijekom školske godine koristi bolovanje zbog različitih kroničnih bolesti. To je ujedno i glavni razlog zašto ne prelaze u realni sektor. Time izravno utječu na količinu posla koju za njih obavljaju njihove radne kolege. Ukoliko se bolovanje pojedinih kolega produlji na više tjedana, zbog nemogućnosti pronalaženja privremene zamjene na tržištu rada, dolazi do prenaprezanja radnih kolega i smanjenja motivacije za rad. Navedeno je uzrok promjene radnih mjesta i većih problema s kadrom u pojedinim srednjim školama. Po definiciji STEM područje pokriva znanost, tehnologiju, inženjerstvo i matematiku, tako da nastavnici iz ovog područja nisu isključivo informatičari. Gledajući ovu grupu nastavnika izdvojeno, kod njih prevladava intrinzična motivacija, odnosno njihova znatiželja i način rješavanja problema su ugrađeni u njihov karakter, te kao takvi drugačije funkcioniraju u odnosu na ostatak populacije. Ova populacija nastavnika, željna je novih znanja, ne treba im posebna vanjska motivacija, iako bi im uvelike bila od koristi jer bi na taj način mogli jednostavnije provesti svoje ideje u djelo. Taj stav najbolje karakterizira jedna izjava koju u šali znaju reći stariji nastavnici, a glasi: „Većina ljudi krade stvari s posla da bi nešto imali kod kuće, samo nastavnici kradu

[2] A Pocrnić i A Tihomirović, Analiza uspjeha na natjecanjima iz Osnova informatike u pandemijskom vremenu, 8th International Conference“Vallis Aurea“ 2020, pp 0039-0349

[3] M Varga: Upravljanje ljudskim potencijalima kroz motivaciju, https;//hrcak.srce.hr (31/01/2023)

od kuće da bi imali na poslu“. No, za usvajanje novih znanja i vještina vezanih uz nove tehnologije potrebno je vrijeme dokolice i promišljanja, što većina nastavnika iz STEM područja nema iz prethodno navedenih razloga. Stoga su i edukacije koje se trenutno provode neučinkovite.

V. TRENUTNI SUSTAVI STRUČNIH USAVRŠAVANJA

U Republici Hrvatskoj stručno usavršavanje odgojitelja, učitelja, nastavnika, stručnih suradnika i ravnatelja su pravo, obveza i odgovornost svih odgojno-obrazovnih djelatnika. Ono se temelji na Zakonu o odgoju i obrazovanju u osnovnoj i srednjoj školi[4][4], Zakonu o predškolskom odgoju i naobrazbi, Pravilniku o napredovanju učitelja i nastavnika u osnovnom i srednjem školstvu i Državnom pedagoškom standardu osnovnoškolskoga sustava odgoja i obrazovanja te Zakonu o Agenciji za odgoj i obrazovanje. Međutim, odredbe navedenih zakona su vrlo općenite, pri čemu su pojedinosti i postupci preciznije definirani odgovarajućim pravilnikom, no on trenutno nije u potpunosti usklađen sa zakonima, a trebao bi definirati postupke i uvjete za olakšavanje provedbe. Pravilnikom o normi rada nastavnika u srednjoškolskoj ustanovi[5][5] regulirano je stručno usavršavanje u kategoriji ostalih poslova kao obveza svakog nastavnika. Jedini način provjere povođenja stručnog usavršavanja je Pravilnik o napredovanju učitelja i nastavnika u osnovnom i srednjem školstvu [6][6] i to za one nastavnike koji žele napredovati. U Pravilniku o napredovanju učitelja i nastavnika u osnovnom i srednjem školstvu koristi se termin profesionalni razvoj i definiran je kao “kontinuirana aktivnost koja uključuje različite procese kao što su edukacija (online ili uživo), praksa, primanje i davanje povratnih informacija pri čemu odgojno-obrazovni radnici preuzimaju odgovornosti za vlastito učenje i cjeloživotni razvoj kompetencija”. Istim pravilnikom predviđen je Okvir nacionalnoga standarda kvalifikacija za učitelje u osnovnim i srednjim školama koji je krovni dokument za razvoj učiteljske profesije u Hrvatskoj, a “sadrži skupove ishoda učenja organizirane prema glavnim kompetencijama nužnima za kvalitetno obavljanje ključnih poslova i aktivnosti u svakodnevnome radu odgojno-obrazovnih radnika te za njihov profesionalni razvoj i koristi se kao alat za provjeru složenih kompetencija u postupku napredovanja tijekom karijere za učitelje, nastavnike, stručne suradnike i ravnatelje” [7] [7]. Stručno usavršavanje nastavnika provodi se u organizaciji MZO, AZOO, ASOO, CARNET, AMPEU, NCVVO, sveučilišta, strukovnih udruga, izdavača i ostalih. Konkretno za matematiku, osim navedenih, Hrvatsko matematičko društvo - nastavnička i inženjerska sekcija i PMF. Osim navedenih organizatora danas se na raznim platformama mogu pronaći mnogi MOOC-ovi, online tečajevi i slični oblici stručnih usavršavanja koji su sada objedinjeni na Europskoj platformi EU Academy. „Osnovna motivacija nastavnika pristupanju edukaciji trebala bi biti motivacija za postizanjem kvalitete u poučavanju, motivacija za učenjem odnosno stručnim usavršavanjem te za osobnim razvojem. No, to nije baš uvijek tako. Motivacijski elementi mogu se standardno podijeliti na ekstrinzična (vanjska motivacija) − pod stalnim utjecajem poticaja i pritisaka izvana, usmjerena na ispunjavanje obveza (licenciranje, mogućnost napredovanja….) te intrinzična (unutarnja motivacija) − odražava osobni cilj i interes za određeno područje, te težnju za osobnim kompetencijama (nova znanja i spoznaje, razmjena iskustava, težnja za informiranošću, radoznalost)“ [8] [8]. Vrlo često organizirano stručno usavršavanje ovisi o financijskim mogućnostima škole zaposlenika te je ograničavajući faktor zainteresiranom pojedincu. „Temeljne su zadaće stručnoga usavršavanja: − osposobiti radnike u odgojno-obrazovnim ustanovama za provedbu okvirnoga kurikuluma; − osposobiti radnike u području odgoja i obrazovanja za razvijanje osam ključnih polja kompetencija učenika; − osposobiti nastavnike, stručne suradnike i ravnatelje u srednjim školama za provođenje državne mature; − omogućiti radnicima u odgoju i obrazovanju trajni profesionalni razvoj u stručno-predmetnom, pedagoško-didaktičko-metodičkom, psihološkom, organizacijskom, komunikacijsko-refleksivnom i informatičkom području.“[9][9]. Vrlo često se događa da ponuđeni sadržaji ne odgovaraju temi skupa i ne zadovoljavaju očekivanja sudionika te da se na ključne probleme dolazi na kraju skupa koji je vremenski ograničen. Kompetencije i vještine sudionika su različite te bi pristup voditelja edukacije trebao biti diferenciran za različite razine, jer inače dolazi do osipanja za vrijeme skupa. Stručno usavršavanje u informatičkom području organizirano je od strane CARNET-a i to su standardizirane radionice za sve sudionike, čak i za

[4] Zakon o odgoju i obrazovanju osnovnoj i srednjoj školi, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2008_07_87_2789.html (30.01.2023)

[5] Pravilnik o normi rada nastavnika u srednjoškolskoj ustanovi, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2010_07_94_2630.html (30.01.2023)

[6] Pravilnik o napredovanju učitelja, nastavnika, stručnih suradnika i ravnatelja u osnovnim i srednjim školama i učeničkim domovima, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2010_07_94_2630.html (30.01.2023)

[7] Okvir nacionalnog standarda kvalifikacija za učitelje u osnovnim i srednjim školama, http://nvoo.hr/?page_id=391 (30.01.2023)

[8] T Dubrović i B Smojver: Motivacija nastavnika za usavršavanje i profesionalni razvoj, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 49- 56, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

[9] R Pažin-Ilakovac i M skelac: Stručno usavršavanje u organizaciji Agencije za odgoj i obrazovanje – opis stanja, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 12-15, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

nastavnike informatike i računalstva čiji je konačni produkt bitno različit od sudionika do sudionika te se provodi na način da se što više skrati jer su edukacije organizirane u radne dane nakon održane nastave. Najviše praktičnih primjera i uputa o korištenju digitalnih alata održava se na Županijskim vijećima, a mogu se pronaći i u online stručnim usavršavanjima domaćih i stranih edukatora, no to je izbor pojedinca. „Pri organizaciji stručnih usavršavanja savjetnicima se nameću sljedeća pitanja: 1. Jesmo li poticatelji promjena ili, iz bilo kojih razloga, zadržavamo status quo? 2. Što želimo promijeniti? 3. Koliko svoje nastavnike vidimo kao multiplikatore onoga što mi radimo? 4. Koliko nastavnika znamo koji su sudjelovali u stručnom usavršavanju, a onda su to primijenili u praksi? 5. Jesmo li spremni „nametnutu promjenu“ redefinirati s ciljem poticanja „usmjeravane promjene“? 6. Jesmo li se i koliko spremni pozabaviti onim nastavnicima koji ne odobravaju ono o čemu im govorimo i/ili nisu spremni to o čemu govorimo odmah i prihvatiti? 7. Jesmo li i koliko spremni slušati sve nastavnike koji su nam „povjereni“? 8. Razumijemo li tvrdnju koja govori: „da bi se promjena dogodila, ljudi na nju moraju biti spremni“? 9. Jesmo li sami spremni s obzirom na promjenjivost svijeta u kojemu živimo stalno i iznova tragati i osmišljavati, čak i onda kada nas nitko izvana na to ne tjera?“[10][10]. O ovim pitanjima bi zaista trebalo dobro razmisliti i u ovisnosti o odgovorima prilagoditi način održavanja stručnih usavršavanja. Dosta savjetnika je davno izašlo iz razreda, a promjene se neminovno događaju. I savjetnici bi se nakon nekog vremena provedenih u agenciji trebali vratiti u škole kako bi bolje razumjeli nastavnike.

VI. STRUČNA USAVRŠAVANJA KAKVA BI TREBALA BITI

U završnom sveučilišnom radu Nikolausa Vesko na poslijediplomskom specijalističkom studiju na sveučilištu u Zadru, dobro je opisano upravljanje ljudskim potencijalima u obrazovnom sustavu Republike Hrvatske. Organizirana i sustavna stručna usavršavanja nastavnika općih predmeta događaju se s uvođenjem novog kurikuluma u eksperimentalne škole, Škola za život. Osnovni problem je što su nastavnici smatrali ovo stručno usavršavanje nametnutim, iako nije postojao način provjere dovršenosti, osim za one koji žele napredovati. Stručno usavršavanje organizirano je korištenjem platforme Loomen, po temama koje su imale određeni vremenski rok s određenim zadacima koje je trebalo u datom roku završiti. Teme su bile različite, od stručnih, pedagoških, ICT, itd. Drugi problem koji se pojavio za vrijeme stručnog usavršavanja je pandemija. Nastavnici su tada mnogo vremena provodili za računalom da bi pripremili i održali online nastavu. No, tada su zapravo mogli uvidjeti prednosti stručnog usavršavanja i materijala koji su bili dostupni, među kojima su i upute za korištenje digitalnih alata. Svakako bi na nivou škole trebalo dogovarati stručna usavršavanja na svim razinama, što i nije slučaj u svakoj obrazovnoj ustanovi. U istraživanju koje je proveo V. Nikolaus, na pitanje oko provedbe formalnih razgovora oko usavršavanja na području matematike, najčešći odgovor je bio da se provodi samo s nekima (55,92%), zatim da se redovito provode strukturirani intervjui sa svima (37,14%), te je najmanji broj njih izjavio da se uopće ne provode takvi razgovori (6,94%). Također, većina škola ima okvirni plan usavršavanja nastavnika matematike (53,06%), dio ih redovito i detaljno provodi plan usavršavanja (25,31%), dok ostatak ne provodi ili provodi usavršavanje bez konkretnog plana (21,63%) [11] [11]. Iz perspektive profesora matematike, učenici smjera tehničar za računalstvo, trebaju dio svog školovanja provesti u realnom ICT sektoru. Na taj način dobili bi informacije o matematičkim alatima koji se koriste u struci, kao i konkretnu primjenu alata. Iz te perspektive, primjena tih matematičkih metoda i alata učenicima bi bila puno prihvatljivija kroz nastavu. Prije pristupanja nastavnika stručnom usavršavanju, posebno za korištenje IT tehnologije, trebalo bi anketirati prijavljene sudionike o njihovoj motiviranosti i očekivanjima, o (pred)znanjima potrebnim za stručno usavršavanje i na taj ih način grupirati u različite skupine kako bi efikasno postizali ciljeve. Stručna usavršavanja treba organizirati u obliku radionica, kako bi se u toku radionice dobio očekivani finalni proizvod te provesti kolegijalno vrednovanje uradaka i vrednovanje predavača. Certifikati (potvrde) bi trebali biti izraženi u postotku kako bi svaki pristupnik dobio povratnu informaciju o svom radu.

VII. ZAKLJUČAK

Da bi odgojno-obrazovni radnici bili kompetentni razviti kod učenika kompetencije potrebne na globalnom tržištu rada u 21. stoljeću i sami se moraju cjeloživotno profesionalno usavršavati. No, za kvalitetno stručno usavršavanje koje potiče profesionalni razvoj svakog pojedinca potrebni su preduvjeti, a to su nastavnici koji su svjesni odgovornosti za svoj profesionalni razvoj. Zemlje koje su u prvom redu ulagale u profesionalni razvoj nastavnika i te imale zakonski regulirano njihovo usavršavanje najbrže su ostvarile napredak u kvaliteti obrazovanja, a time pomogle i napretku zemlje u cjelini.

Najučinkovitiji model stručnog usavršavanja nastavnika iz STEM područja treba biti baziran na praktičnom radu u

[10] M Barbaroša-Šikić: Izazovi u osmišljanju i upravljanju programima usavršavanja, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 16-21, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

[11] N Vesko; Upravljanje ljudskim potencijalima na području matematike u odgojno obrazovnim ustanova Republike Hrvatske, Sveučilište u Zadru, Zadar 2019, https://repozitorij.unizd.hr/islandora/object/unizd%3A3323/datastream/PDF/view (30.01.2023)

realnom sektoru. Nastavnik bi tada predavao jednoj generaciji učenika u vremenu trajanja njihovog strukovnog obrazovanja (između tri do pet godina), te bi potom imao obavezu raditi minimalno godinu dana u realnom sektoru. Po isteku vremenskog perioda od godinu dana, mogao bi birati želi li nastaviti predavati u srednjoj školi ili ostati u realnom sektoru. Kako bi školski sustav imao na raspolaganju najbolje stručnjake iz pojedinog strukovnog područja, plaća nastavnika treba biti viša za minimalno 50% u odnosu na plaću stručnjaka iz realnog sektora.

Na ovaj način bi se osigurala konstantna opskrbljenost školstva sa relevantnim znanjima iz strukovnih područja, nastavnici bi bili visoko motivirani, a transfer znanja bi bio bitno učinkovitiji jer bi znanje bilo prenošeno iz prve ruke, s osobnim osvrtom na dobre i loše strane pojedinih načina rada. Ovaj način rada uvelike bi olakšavao pronalaženje mladih talenata u srednjoj školi koji bi se potom mogli razvijati u struci u renomiranim sredinama koje bi im omogućavale brzu transformaciju u eksperte u pojedinim uskim specijalizacijama.

VIII. REFERENCE

[1] B Jokić i Z Ristić Dedić: Analitičko izvješće SUPERODLIKAŠI http://os-bukovac-zg.skole.hr/upload/os-bukovac-zg/images/static3/2254/attachment/ANALITICKO_IZVJESCE-Superodlikasi-IDIZ.pdf (31/01/2023)

[2] A Pocrnić i A Tihomirović, Analiza uspjeha na natjecanjima iz Osnova informatike u pandemijskom vremenu, 8th International Conference“Vallis Aurea“ 2020, pp 0039-0349

[3] M Varga: Upravljanje ljudskim potencijalima kroz motivaciju, https;//hrcak.srce.hr (31/01/2023)

[4] Zakon o odgoju i obrazovanju osnovnoj i srednjoj školi, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2008_07_87_2789.html (30.01.2023)

[5] Pravilnik o normi rada nastavnika u srednjoškolskoj ustanovi, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2010_07_94_2630.html (30.01.2023)

[6] Pravilnik o napredovanju učitelja, nastavnika, stručnih suradnika i ravnatelja u osnovnim i srednjim školama i učeničkim domovima, Narodne novine, https://narodne-novine.nn.hr/clanci/sluzbeni/2010_07_94_2630.html (30.01.2023)

[7] Okvir nacionalnog standarda kvalifikacija za učitelje u osnovnim i srednjim školama, http://nvoo.hr/?page_id=391 (30.01.2023)

[8] T Dubrović i B Smojver: Motivacija nastavnika za usavršavanje i profesionalni razvoj, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 49- 56, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

[9] R Pažin-Ilakovac i M skelac: Stručno usavršavanje u organizaciji Agencije za odgoj i obrazovanje – opis stanja, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 12-15, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

[10] M Barbaroša-Šikić: Izazovi u osmišljanju i upravljanju programima usavršavanja, Stručno usavršavanje i razvoj, Agencija za odgoj i obrazovanje pp 16-21, https://www.azoo.hr/app/uploads/uvezeno/datoteke/strucno-usavrsavanje-i-profesionalni-razvoj-2010-1536867257.pdf, (30.01.2023)

[11] N Vesko; Upravljanje ljudskim potencijalima na području matematike u odgojno obrazovnim ustanova Republike Hrvatske, Sveučilište u Zadru, Zadar 2019, https://repozitorij.unizd.hr/islandora/object/unizd%3A3323/datastream/PDF/view (30.01.2023)

Primjena e-učenja u visokoškolskoj nastavi: pristupi instrukcijskom dizajnu, pedagoški scenariji, modeli hibridnog poučavanja i razine kompleksnosti u primjeni

Goran Bubaš
Sveučilište u Zagrebu, Fakultet organizacije informatike, Varaždin, Hrvatska
gbubas@foi.unizg.hr

***Sažetak* – Primjena e-učenja tijekom pandemije virusa COVID-19 poprimala je različite forme u pogledu omjera između online nastave i nastave u učionici, kao i primjene različitih vrsta sinkronog i asinkronog e-učenja kod nastavnika u visokom, srednjem i osnovnom obrazovanju. Pandemijsko razdoblje nije samo potaknulo preispitivanje učinkovitosti e-učenja u odnosu na učenje u učionici, već je potaknulo i razmatranje pristupa hibridnom i kombiniranom učenju, instrukcijskom dizajnu te razini kompleksnosti u primjeni e-učenja. U ovom radu izložen je pregled različitih modela primjene hibridnog (engl. hybrid) ili kombiniranog (engl. blended) učenja. Prikazani su povijesni rezultati istraživanja, kao i odabrani sadržaji preglednih radova i metastudija te nekoliko empirijskih istraživanja. Rezultati pokazuju raznolikost modela hibridnog ili kombiniranog učenja, kao i preferenciju prema hibridnom poučavanju kod studenata, s jedne strane, ali i na pretežnu preferenciju prema učenju u učionici kod nastavnika. Bez obzira na smanjen interes studenata i učenika za *online* nastavu nakon intenzivnog razdoblja pandemije COVID-19, nastavnici trebaju i dalje razvijati odgovarajuće elemente svoje kompetencije za primjenu e-učenja.**

Ključne riječi – e-učenje, hibridno učenje, kombinirano učenje, instrukcijski dizajn, kompleksnost e-učenja, pregledna istraživanja, metastudije

I. UVOD

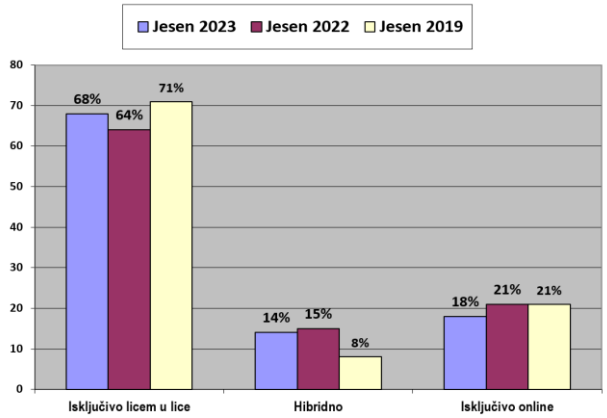
Učenje na daljinu (engl. distance learning) forma je poučavanja kod koje su učitelj i učenik fizički razdvojeni, a počeci mu sežu u prvu polovicu 19. stoljeća. Među „pionirima“ poučavanja na daljinu najviše se spominje Isaac Pitman koji je 1840.-ih započeo prvi dopisni tečaj. Prema H. Kentor [1], za poučavanje na daljinu vrlo su često korišteni radio i televizija, dok revolucija u poučavanju na daljinu započinje pojavom računala te se posebno intenzivira nakon pojave interneta kasnih 1990.-ih godina. Međutim, ne treba zaboraviti da su za poučavanje na daljinu u 20. stoljeću također korištene gramofonske ploče i audiokazete, filmske vrpce, videokazete te audio, video i multimedijalni sadržaji na CD-ROM-ovima i DVD medijima.

Povijest učenja na daljinu, imajući u vidu brojne medije koji su za tu namjenu korišteni, kao da se na izuzetno brz način ponovila pojavom pandemije COVID-19 u veljači/ožujku 2020. godine. Naime, na početku pandemije došlo je do naglog i masovnog korištenja učenja na daljinu uz uporabu internetskih tehnologija najrazličitijih vrsta. Prema anketi koja je provedena u 41 europskoj državi u razdoblju od 21. ožujka do 3. svibnja 2020. godine na 9.180 ispitanika [2] čak 93% studenata u visokom obrazovanju izjavilo je da im je nastava na fizičkoj lokaciji (u učionici, *licem u lice*) bila obustavljena u razdoblju te ankete. Slična anketa [3], provedena u razdoblju od 5. svibnja do 15. lipnja 2020. godine među studentima u visokom obrazovanju na 6 kontinenata, s konačnim uzorkom od 30.383 studenta iz 62 zemlje, pokazala je da je za 86.7% ispitanika obustavljena nastava na lokaciji (*licem u lice*) u razdoblju prije ili tijekom anketiranja. Ovo potonje istraživanje također je pokazalo da je nastava *licem u lice* u vrijeme anketiranja najčešće zamijenjena videokonferencijama uživo (59.4%), slanjem prezentacija studentima (15.2%) i asinkronim videozapisima nastave (11.6%).

U kratkom razdoblju prvih nekoliko mjeseci od početka pandemije, kao i tijekom godinu-dvije nakon toga, većina nastavnika na visokim učilištima svladavala je i primjenjivala najrazličitije nove tehnologije za učenje na daljinu (e-učenje), pri čemu se ne misli samo na sustave za upravljanje učenjem kao što su *BlackBoard* ili *Moodle*, već i na brojne druge komunikacijske i kolaboracijske alate. Autor ovog rada nije u literaturi pronašao sustavno provedeno međunarodno metaistraživanje ili detaljan pregledni rad iz spomenute tematike, no i unatoč tome bilo je lako zaključiti da su, među komunikacijskim tehnologijama koje su korištene tijekom pandemije, u znanstvenim i stručnim radovima najčešće spominjani videokonferencijski alati u sklopu servisa *Google Meet, Zoom* i *Microsoft Teams*, a od drugih komunikacijskih alata prije svega spominjana je *elektronička pošta*, kao i *WhatsApp* te *Facebook Messenger*. Među srodnim istraživanjima koja su provedena na nacionalnoj razini svakako treba spomenuti ono iz Poljske koje je provedeno u prosincu 2021. godine na 621 studentu visokih učilišta [4], koje je potvrdilo da su *MS Teams* i *Google Clasroom* bili među najčešće korištenim alatima.

Prema interaktivnoj karti zatvaranja škola i visokih učilišta tijekom pandemije (vidjeti: OurWorldinData, https://ourworldindata.org/covid-school-workplace-closures), 31. prosinca 2022. godine, dokad su ažurirani podaci, samo je još nekoliko zemalja u svijetu imalo

djelomičan prekid nastave uživo, i to samo na određenim školskim razinama, izuzevši Kinu u kojoj je to bilo obvezno na svim razinama. Pritisak na korištenje e-učenja u visokom obrazovanju na globalnoj razini bitno je smanjen jenjavanjem pandemije COVID-19, na što upućuje i ACE anketa iz 2022. godine koja se redovito provodi među rukovoditeljima visokoškolskih ustanova u SAD-u [5]. Na slici 1. prikazani su rezultati te ankete, provedene u siječnju i veljači 2022. godine na 375 rukovoditelja visokoškolskih ustanova u SAD-u, koji pokazuju njihovo predviđanje da će se omjer nastave koja se izvodi na prijediplomskim studijima (a) *licem u lice*, (b) hibridno ili dijelom *licem u lice*, a dijelom *online*, kao i (c) u potpunosti *online*, na njihovim institucijama u jesenskom semestru 2022. i 2023. godine vratiti vrlo blizu pretpandemijskim omjerima iz jesenskog semestra 2019. godine.



Slika 1. Planirani oblici nastave na visokim učilištima u SAD-u za jesenski semestar 2022. i 2023. godine, a u odnosu na jesenski semestar 2019. godine, prema ACE anketi (izvor: [5])

Prethodno navedena ACE anketa također je pokazala da je čak 83% rukovoditelja visokoškolskih ustanova u SAD-u smatralo da nakon pandemije dugoročno treba zadržati povećane mogućnosti za *online* učenje. Budući da se nakon jenjavanja pandemije zadržala svijest o koristi uporabe e-učenja, u nastavku će biti razmatrani modeli poučavanja uz primjenu e-učenja s naglaskom na *hibridno* (eng. hybrid) ili *kombinirano* (engl. blended) e-učenje. Pritom treba istaknuti da je recentna metaanaliza pokazala kako e-učenje može uspješno smanjiti vrijeme koje studenti provode u učionici, bez značajnijeg smanjenja rezultata učenja [6].

II. HIBRIDNO ILI KOMBINIRANO UČENJE

Iako početak hibridnog ili kombiniranog učenja neki autori pogrešno poistovjećuju s početkom učenja na daljinu koje datira približno 1840. godine, počeci hibridnog ili kombiniranog učenja mogu se identificirati tijekom povijesti primjene audiovizualnih tehnologija u nastavi, imajući u vidu kombiniranje nastave uživo u učionici s primjenom tehnologije (gramofonskih ploča, audio i videokaseta, multimedijalnih sadržaja na CD/DVD medijima i sl.) u učionici i izvan učionice. Suština je da uporaba tehnologije dopunjuje nastavni proces koji se izvodi u učionici. Hibridno ili kombinirano učenje posebno se vezuje uz početak šire uporabe World Wide Weba u obrazovanju nakon 1994. godine, odnosno uz *poučavanje putem weba* (engl. web-based instruction).

Prema analizi literature o počecima hibridnog učenja [7], jednu od prvih studija koja je koristila termin "mješovito učenje" (engl. blended learning) objavili su 2000. godine Cooney i suradnici [8] iz područja predškolskog odgoja. Slijedili su 2001. godine Voci i Young [9], koji su e-učenje uključili u trening za razvoj vodstva, a potom 2002. godine Bonk i suradnici [10], koji su asinkrono i sinkrono poučavanje uključili u vojni tečaj.

Imajući u vidu *modele* hibridnog ili kombiniranog učenja, jedan od prvih takvih modela predložila je još 2002. godine Purnima Valiathan [11] na portalu *Learning Circuits*. Njezin je model imao tri komponente: (a) učenje usmjereno *vještinama* koje kombinira rad studenta vlastitim ritmom, uz aktivnosti instruktora, za razvoj specifičnih znanja i vještina, (b) učenje usmjereno *stavovima* (engl. attitude), koje koristi različite medije za razvoj novih specifičnih ponašanja i stavova, kao i (c) učenje usmjereno na *kompetencije* koje kombinira alate za podršku učinkovitosti i alate za upravljanje znanjem s mentoriranjem za razvoj kompetencija vezanih uz posao. Također, za svaki od navedenih modela učenja predložene su tradicionalne (u učionici) i *online* aktivnosti za pojedine faze izvođenja nastavne jedinice.

Clark [12] još je 2003. godine nabrojio moguće zamke u primjeni hibridnog/kombiniranog učenja, koje ulaze u sljedeće (slobodno interpretirane) kategorije: (a) *intuitivno odlučivanje* bez primjerenog prosuđivanja i znanstvenih temelja, (b) *preveliko oslanjanje na rad u učionici* kao nezaobilaznu komponentu, (c) *„lijepljenje" učenja u učionici s elementima online učenja* koje je samo sebi svrhom bez optimiziranja kombinacije sadržaja *offline*/*online* medija, (d) *dupliciranje* ili paralelno korištenje sadržaja izloženih na tradicionalan način, u učionici i u knjigama, s istovjetnim sadržajima koji su isporučenih u *online* obliku, (e) *nepotrebno kompliciranje* uvođenjem puno različitih kanala i oblika učenja putem interneta.

Twigg [13] iste je 2003. godine, na temelju analize 30 projekata redizajna nastavnih predmeta na sveučilištima u SAD-u, utvrdio da su korišteni različiti pristupi, kao i da se u tim projektima uglavnom vodilo računa o temeljnoj disciplini koja je poučavana, specifičnostima studenata polaznika te preferencijama nastavnika. Modeli primjene hibridnog/kombiniranog učenja koji su pritom identificirani bili su: (a) *dopunski* model ili dodavanje aktivnosti *online* učenja uz zadržavanje dotadašnje tradicionalne nastave; (b) *zamjenski* model u kojem su odabrane interakcije *licem u lice* i tradicionalne metode podučavanja zamjenjivane odgovarajućim online ekvivalentima; (c) model *online „supermarketa"* u kojem se eliminira veći dio potrebe za nastavom u učionici te studenti samostalno koriste brojne online resurse za učenje, pri čemu sami biraju specifične obrazovne sadržaje, vrijeme učenja i tempo uporabe obrazovnih materijala iz odabrane teme; (d) *u potpunosti online model* u kojem se reducira potreban broj nastavnika za rad u učionici te se najveći dio obrazovnih sadržaja isporučuje *online*, a *online* kanali koriste se i za interakciju sa studentima; (e) *model "švedskog stola"*, kod kojeg nema "fiksnog jelovnika" za učenje pa studenti slobodno biraju između *online* i *offline* isporuke sadržaja te različitih metoda podučavanja, ali i između kontakata *licem u lice* i *online* interakcija s nastavnicima.

U Hrvatskoj su 2006. godine Bubaš i suradnici [14] razmatrali različite aspekte relevantne za unapređenje implementacije hibridnog/kombiniranog učenja imajući u vidu modele instrukcijskog dizajna (modele *ADDIE, ARCS, R. Gagnéa,* kao i *W. Dicka i L. Carreya)* te su prikazali studij slučaja za jedan hibridni predmet na visokom učilištu i rezultate evaluacije njegovih *online* komponenti.

Allen i suradnici [15, 5-6] predložili su 2007. godine sljedeću kategorizaciju hibridnih/kombiniranih nastavnih predmeta, prema omjeru tradicionalnog i *online* poučavanja:

a) 0% online – *tradicionalno* izvođenje predmeta bez uporabe internetskih tehnologija;

b) 1-29% *online* – izvođenje predmeta *potpomognuto webom*, gdje sustav za upravljanje učenjem (Learning Management System – LMS) ili *web*-stranice služe npr. za prikaz opisa nastavnog predmeta i zadataka;

c) 30-79% *online* – hibridno/kombinirano izvođenje predmeta koje kombinira poučavanje *online* i *licem u lice*, sa znatnim udjelom online isporuke sadržaja;

d) 80+% *online* – većina ili skoro sav sadržaj isporučuje se *online* i u pravilu nema kontakata licem u lice s nastavnikom.

Staker i Horn [16] doradili su 2012. godine prvotnu taksonomiju *Innosight instituta* i predložili sljedeća četiri oblika hibridnog/kombiniranog učenja u K-12 obrazovnim institucijama (osnovnim i srednjim školama):

- *Rotacijski model* kod kojeg se u određenom nastavnom predmetu za studente mijenjaju modaliteti rada (vrste obrazovne aktivnosti), od kojih je jedna *online* učenje/poučavanje, a prema fiksnom rasporedu koji je definirao nastavnik.
- *Fleksibilni model* u kojem studenti sami, prema individualnim preferencijama, biraju između više modaliteta (kao i između više mogućih aktivnosti učenja), koji su uglavnom online, pri čemu je jedan nastavnik u funkciji nositelja predmeta ili glavnog odgovornog nastavnika (engl. *teacher-of-record*) i uglavnom ga se zbog savjetovanja kontaktira na lokaciji (*licem u lice*).
- *Samokombinirani model* (engl. *self-blend*) vezan je uz situaciju u kojoj, uz nastavu koja se na tradicionalan način izvodi u učionici za druge predmete, student bira jedan ili više dodatnih predmeta koji se izvode u potpunosti *online*, a s nositeljem predmeta ili glavnim/odgovornim nastavnikom za te *online* predmete također kontaktiraju većinom *online*.
- *Obogaćeni virtualni model* predstavlja školsko okruženje u kojem kod svih nastavnih predmeta studenti dio vremena rade u tradicionalnom okruženju u učionici, a dio vremena u *online* uvjetima učenja i poučavanja.

Na temelju pregleda literature Allamary i suradnice naglasile su 2014. godine [17] važnost „balansiranog“ i „promišljenog“ slaganja i međusobnog uklapanja *online* i *offline* aktivnosti i medija kod dizajniranja kombiniranog (hibridnog) učenja i poučavanja, vodeći računa o njihovoj pedagoškoj vrijednosti. Pritom, iz različitih pristupa pedagoškom dizajnu koje su razmatrale, izlučuju sljedeća tri distinktivna modela: (a) *kombiniranje s niskim učinkom* – dodavanje dodatnih aktivnosti postojećem (tradicionalnom) nastavnom predmetu; (b) *kombiniranje s umjerenom učinkom* - zamjenjivanje pojedinih aktivnosti u postojećem (tradicionalnom) nastavnom predmetu; (c) *kombiniranje s visokim učinkom* – izgradnja kombiniranog (hibridnog) tečaja od početka. Allamary [18] također je napravila 2019. godine sustavni pregled literature iz područja kombiniranog (hibridnog) poučavanja uvodnih predmeta programiranja te je identificirala sljedećih pet tipičnih modela: model *obrnute učionice* (engl. flipped model), *miješani* (engl. mixed) model, *fleksibilni* (engl. flex) model, model *dodataka* (engl. supplemental) i model *online vježbanja* (engl. practicing).

Na području poučavanja stranog jezika Murray i Christison [19] kritizirale su 2018. godine ranije spomenutu kategorizaciju hibridnih/kombiniranih nastavnih predmeta Allena i suradnika [15], koja se bazira na omjeru/postotku *online* i *offline* poučavanja, te su predložile klasifikaciju koju su u svojoj kasnijoj metaanalizi 2021. godine doradili Dixon i suradnici [20]. Prema ovoj novijoj inačici kategorizacije *online* i hibridnog poučavanja ono se dijeli na: (a) *unaprijeđeno poučavanje* u kojem se nastava *licem u lice* potpomaže *online* sadržajima i aktivnostima; (b) *kombinirano ili hibridno poučavanje* kod kojeg je broj nastavnih aktivnosti *licem u lice* smanjen u korist *online* učenja i poučavanja; (c) *okrenuto* (engl. flipped) u kojem znatan dio *online* aktivnosti prethodi interakciji *licem u lice* koja je posvećena rješavanju problema temeljem informacija koje su (bile) dostupne *online*; (d) *sinkrono online* gdje je cjelokupno poučavanje online, a studenti se također istovremeno susreću *online*; (e) *asinkrono online* kod kojeg je cjelokupno poučavanje *online* i sve aktivnosti i zadaci se mogu obaviti asinkrono. Za usporedbu, Albiladi i Alshareef [21] u preglednom su radu iz 2019. godine promatrali kombinirano (hibridno) učenje kao jedinstven model koji služi razvoju jezičnih vještina i poticanju motivacije te je sredstvo za unapređenje obrazovnog okruženja za poučavanje engleskog kao stranog jezika.

Na kraju ovog pregleda modela hibridnog i kombiniranog učenja svakako treba spomenuti *HyFlex*, pojam koji je 2007. godine predložio Beatty [22], a koji označava (a) *hibridno učenje* jer kombinira aktivnosti *online* poučavanja i učenja s poučavanjem *licem u lice*, kao i (b) *fleksibilno učenje* jer studenti mogu birati hoće li u određenom tjednu pohađati ili ne pohađati određene lekcije nastave *licem u lice*, odnosno hoće li, bez negativnih posljedica, pohađati *online* inačice za lekcije u učionici, tj. *online* zamjene za te lekcije. Inače, Pressley je u popularno pisanom tekstu [23] detaljnije konceptualizirao razliku između *hibridnog* učenja i *HyFlexa* naglašavajući da u potonjem formatu studenti mogu odabrati način na koji će sudjelovati u nastavi od sesije do sesije te koristiti nastavne materijale u obliku koji im u bilo kojem trenutku najbolje odgovara: sinkrono *licem u lice*, sinkrono *online* ili asinkrono *online*. Primjerice, tri studenta bi, prema *HyFlex modelu* poučavanja i učenja, mogla u jednom tjednu sudjelovati svaki u drukčijem formatu izvođenja nastave. Ovdje treba istaknuti da su u recentnom empirijskom istraživanju [24] iz 2022. godine Bubaš i Čižmešija otkrili da je, za jedan *hibridni* nastavni predmet, čak 76% studenata izjavilo kako bi „vjerojatno“ ili „sigurno“ preferirali da im nastavnik omogući u bilo kojem trenutku/tjednu birati između odgovarajućeg oblike *online* ili *offline* nastave.

III. INSTRUKCIJSKI DIZAJN I HIBRIDNO UČENJE

Nakon što se kasnih 1990.-ih godina proširila uporaba interneta u obrazovanju, postupno se mijenjala konceptualizacija instrukcijskog dizajna (ID) kao teorijske, istraživačke i primijenjene discipline. Za primjer moguće je navesti definiciju ID-a koju je 2001. godine predložio Reiser [25]: „Područje instrukcijskog dizajna i primjene tehnologije u poučavanju obuhvaća analizu problema vezanih uz učenje i njegove učinke, kao i dizajna, razvoja, implementacije, evaluacije i upravljanja nastavnim i nenastavnim procesima i resursima kojima je cilj poboljšanje učenja i uspješnosti u različitim uvjetima, a posebno u obrazovnim institucijama i u poslovnom okruženju."

Još 1994. godine Edmonds i suradnici [26] isticali su da su znanstvenici i praktičari koji se bave obrazovanjem suočeni sa „stotinama" modela instrukcijskog dizajna (ID) „kojima je cilj unaprijediti proces planiranja poučavanja". Prema konceptualnom okviru Edmondsa i suradnika važno je znati prepoznati o kakvom se modelu poučavanja radi, tj. koja je njegova razina *deskripcije* prema *preskripciji*, kojem je *kontekstu* namijenjen (npr. razini školovanja i korištenju tehnologije), stjecanju koje vrste *znanja* je usmjeren (*proceduralnog* ili *deklarativnog*), kakva mu je *kompleksnost* s obzirom na razinu *ekspertize koja je potrebna za njegovu implementaciju*, kao i koliko ima *elemenata sistemskog teorijskog pristupa*. Ako se prema teorijskom modelu Edmondsa i suradnika analiziraju u prethodnom poglavlju ovog rada opisani modeli hibridnog ili kombiniranog učenja, vidljivo je da su svi *deskriptivni* (uključujući: [11], [13], [15], [22], [16], [17], [18], [19],) a samo neki od njih ujedno su i *preskriptivni* na način da propisuju kako se obrazovno okruženje može prilagoditi radi ostvarenja željenih pedagoških učinaka.

Osim prethodno navedenog, treba istaknuti da su 1980. godine, u komparativnoj analizi 40 modela instrukcijskog dizajna nastalih između 1966. i 1979. godine, Andrews i Goodson [27] naveli moguće razloge za pojavu (pre)velikog broja ID modela, pri čemu su dodali i da neki od njih uopće nisu modeli jer „ne uspijevaju opisati, objasniti ili predvidjeti elemente u svojem referentnom sustavu". Takva obilježja mogu se pripisati npr. modelima hibridnog/kombiniranog učenja Allen i suradnika [15] iz 2007., kao i Allamary i suradnica [17] iz 2017. godine.

U novijoj preglednoj analizi 38 istraživanja, objavljenih većinom između 2004. i 2019. godine, koja su se bavila uporabom modela instrukcijskog dizajna u različitim kontekstima, Stefaniak i Xu [28] utvrdili su da većina tih studija „koje su predstavile dodatne modele instrukcijskog dizajna, nisu ponudile nešto što bi upućivalo na opsežne promjene u tradicionalnom sustavnom procesu instrukcijskog dizajna". Sličan zaključak mogao bi se izvesti i iz sažetog pregleda različitih modela hibridnog i kombiniranog učenja koji su se pojavili u literaturi u razdoblju od 2002. do 2021. godine, a koji su prikazani u prethodnom poglavlju ovoga rada. Treba naglasiti da su noviji pregled, uz kategorizaciju 33 ID modela, 2016. godine napravili i Donmez i Cagiltay [29]. Posebno je ilustrativna i bibliometrijska analiza iz 2022. godine koju su izradili Saçak i suradnici [30], koja je pokazala sljedeće četiri najčešće tematske skupine koje povezuju ključne riječi navedene u objavljenim radovima iz ID-a: (1) teorijski zasnovani pristupi, (2) dizajn utemeljen na tehnologiji, (3) ID za visoko obrazovanje te (4) vrednovanje i evaluacija. Lako je zapaziti da su u modelima hibridnog/kombiniranog učenja spomenutim u prethodnom poglavlju također izostali primjereni *teorijski* elementi, kao i komponente *vrednovanja i evaluacije*.

Prema istraživanju koje su 2018. godine proveli Bond i Dirkin [31], od 297 anketiranih nastavnika na visokim učilištima u SAD-u, njih 41.3% koristili su *ADDIE model* instrukcijskog dizajna (engl. Analysis, Design, Development, Implementation, Evaluation), dok ih je 29.6% koristilo *Backword design* model. Ostale ID modeli (*Bloomova taksonomija, Quality Matters, Constructivist / Constructivism, Dick & Carey, Fink*) koristilo je 5.6% do 9.2% ispitanika. ADDIE model najčešće se spominje i na listama modela instrukcijskog dizajna koje je moguće naći u popularnim člancima i na *web*-portalima koji su namijenjeni obrazovanju. Važan je element ADDIE modela evaluacija kvalitete i učinkovitosti obrazovnih materijala te kako se njima ostvaruju ciljevi poučavanja. Prema intervjuima koje su proveli Martin i suradnici [32], nastavnici koji su nagrađivani za svoj rad i rezultate koriste evaluaciju svojih predmeta kroz metode osiguravanja kvalitete, kao i na temelju procjena drugih nastavnika i studenata. Evaluacija, kao faza ADDIE modela instrukcijskog dizajna, važna je za uspješno kreiranje nastavnih sadržaja za hibridno ili kombinirano učenje. ADDIE model korišten je u metaanalizi koji su 2021. godine objavili Castro i Tumibay [33], koja se bavila činiteljima efikasnosti tečajeva (nastavnih predmeta) koji koriste e-učenje. Inače, evaluacijom poučavanja sistematski su se, još 1993. godine, bavili npr. Kraiger i suradnici [34]. Područjem evaluacije hibridnog učenja u Hrvatskoj bavili su se 2007. godine Bubaš i suradnici [35] (u obliku koncipiranja i primjene ankete za detaljnu studentsku procjenu dva *online* tečaja u sklopu sveučilišnih predmeta), kao i 2008. godine Kovačić i suradnici [36], imajući u vidu studentsku procjenu specifičnih *online* aktivnosti (tzv. e-tivities) u pojedinim nastavnim lekcijama prilikom poučavanja engleskog kao stranog jezika na visokom učilištu. Nažalost, u svojem teorijskom uvodu za sistematski pregled literature iz *online* poučavanja Martin i suradnici [37] nisu identificirali *evaluaciju nastave* kao temu *ranijih* sličnih pregleda literature i metaanaliza koje su drugi autori objavili za razdoblje od 1990. do 2008. godine, ali su u svojoj metaanalizi, za *kasnije* razdoblje od 2009. do 2018. godine, ipak pronašli da se tematsko područje *evaluacije nastave* i *osiguravanja kvalitete u e-učenju* pojavljivalo u 6.14% analiziranih objavljenih radova.

IV. RAZINE KOMPLEKSNOSTI U PRIMJENI E-UČENJA

Ubrzo nakon početka pandemije COVID-19 među nastavnicima na visokim učilištima i istraživačima proširio se izraz „hitna nastava na daljinu" (engl. *Emergency Remote Learning* - ELT), koji označava naglo prebacivanje na u potpunosti *online* poučavanje i učenje, odnosno privremenu promjenu izvođenja nastave iz učionice na alternativni *online* modalitet zbog neočekivane krizne situacije (vidjeti: [38]). Zanimljivo je da se na temelju preglednog rada Stewarta i suradnika [39] iz 2021. godine može zaključiti da se u studijama koje su objavljene tijekom 2000. godine, tj. na početku pandemije COVID-19, nije pojavljivala tema *evaluacije* novih online oblika e-učenja, odnosno *osiguravanja kvalitete* na tom području. Drugim riječima, u početku su praktičari i istraživači tim temama kod prebacivanja nastave iz učionice na u potpunosti *online* okruženje poklanjali malo pozornosti, za razliku od interesa za takve teme tijekom kasnijeg razdoblja pandemije. Naime, u *novijem* preglednom radu iz 2022. godine, koji su, vezano uz istraživanja primjene tehnologije za učenje tijekom

pandemije, proveli Xue i Crompton [40], identificirana je *učinkovitost online nastave* kao jedna od četiriju najčešćih tema u analiziranim radovima. Između ostalih, s učinkovitošću povezane *podteme* koje su se javile u analiziranoj literaturi, vezanoj uz poučavanje tijekom pandemije, bile su: (a) evaluacija utjecaja *online* poučavanja i učenja, (b) uloge i kompetencije nastavnika te njihove vještine za *online* poučavanje u hitnim/kriznim uvjetima, (c) pedagoške strategije za integraciju tehnologije u odgovarajuće metode poučavanja, (d) učinkovitost i drugi pozitivni efekti "virtualnog" poučavanja, (e) općenita potreba za unapređenjem kvalitete *online* poučavanja, (f) treninzi za studente i nastavnike koji bi mogli povećati različitost *online* metoda poučavanja.

Poslije 2000. godine, tj. nakon što se e-učenje afirmiralo u visokom obrazovanju, pojavili su se radovi koji su elaborirali potrebe za povezanim pedagoškim kompetencijama nastavnika, navodećim pritom pedagoške tehnike i klastere nastavničkih ponašanja (autorica Mehanna 2004. godine [41]) ili pedagoške teorijske modele i scenarije (autorica Dabaggh 2005. godine [42]). Inače, o pedagoškim ulogama i kompetencijama nastavnika u visokoškolskom online okruženju detaljno su 2013. godine pisali Muñoz Carril i suradnice [43], dok su o modelu kompetencija nastavnika za e-učenje u hibridnom visokoškolskom okruženju te rezultatima anketnog istraživanja povezanog s njime pisali Babić i Bubaš 2015. godine [44]. Na potrebu dodatne pedagoško-tehnološke edukacije nastavnika na visokim učilištima u Hrvatskoj pokazuju rezultati dviju doktorskih disertacija – Ane Mirković Moguš iz 2014. [45, 121-127] i Snježane Babić iz 2016. godine [46, 252-259]. U obje je disertacije anketama utvrđeno korištenje ne samo *jednostavnijih* oblika e-učenja kod većine anketiranih, nego i *naprednijih i kompleksnijih* tehnologija i metoda, ali kod manjeg broja nastavnika. Treba naglasiti da je razdoblje pandemije COVID-19 svim nastavnicima u Hrvatskoj osiguralo poticaj i tehnološko iskustvo u primjeni jednostavnijih oblika e-učenja, najviše za sinkrono i asinkrono *videopoučavanje*, ali je izostala sustavna masovnija edukacija za kompleksnije, pedagoški utemeljene oblike primjene e-učenja.

V. ZAKLJUČAK

Hibridno ili kombinirano učenje posebno je dobilo na važnosti tijekom pandemije COVID-19 kad su, nakon početnog prelaska na u potpunosti online nastavu, mnoge visokoškolske ustanove barem godinu-dvije intenzivno koristile taj oblik poučavanja. U *prvom* uvodnom poglavlju ovog rada ukratko je razmotrena povijest e-učenja i rezimirano razdoblje COVID-19 pandemije. Potom su, u *drugom* poglavlju, prikazani različiti modeli hibridnog/kombiniranog e-učenja imajući u vidu njihov povijesni razvoj, pri čemu je moguće zaključiti da nije oblikovan općeprihvaćeni model u tom području. U *trećem* poglavlju izložena je sažeta kritika pristupa kreiranju modela instrukcijskog dizajna (ID), imajući u vidu i nedostatke prethodno izloženih modela hibridnog/kombiniranog učenja te uz naglašavanje važnosti evaluacije u ID modelima. U *četvrtom* poglavlju istaknuta je potencijalna kompleksnost u primjeni e-učenja uz stavljanje u prvi plan njegove učinkovitosti i evaluacije u post-pandemijskom razdoblju kako bi se bolje iskoristile novostečene kompetencije nastavnika i unapređeni tehnički preduvjeti za e-učenje.

Ne treba zanemariti anketno istraživanje koje je provedeno u SAD-u [47] i pokazalo da su u 2021. i 2022. godini najčešći razlozi studenata (N=1,300) za izbor *online* prisutnosti nastavi, prema nastavi *u učionici* (tj. na kampusu), bili: (a) postojeće obveze, rad i obitelj, nedopuštanje nastave na kampusu, (b) e-učenje bio je jedini način da studiram iz područja svojeg interesa, (c) utjecaj pandemije COVID-19 na moje mogućnosti i visoko obrazovanje općenito. S druge strane, u istom je istraživanju utvrđeno da je, osim balansiranja između obaveza u obitelji, na poslu i u kućanstvu, drugi najveći izvor zabrinutosti za anketirane ispitanike (N=1,500) bila kvaliteta *online* nastave i akademska podrška, ispred prilagodbe *online* okruženju za učenje. U istraživanju [48] koje je za *Pew Research Center* u SAD-u provela agencija *Ipsos* u razdoblju od 14. travnja do 4. svibnja 2022. godine na 1,316 tinejdžera i njihovih roditelja, čak 65% anketiranih izjavilo je da bi nakon pandemije COVID-19 željelo da se nastava odvija u potpunosti uživo (*licem u lice*), dok ih je 18% željelo da se odvija mješovito (hibridno), a samo 9% u potpunosti *online*.

Iako utjecaj pandemije postaje sve manji za većinu nastavnika i studenata, što dovodi i do velikog smanjenja interesa za primjenu e-učenja, potencijali koje je pandemijsko razdoblje pružilo za daljnje unapređenje u primjeni tehnologije u visokom obrazovanju u obliku hibridnog/kombiniranog e-učenja ne bi u budućnosti trebali ostati neiskorišteni.

REFERENCE

[1] Kentnor, H.E. „Distance education and the evolution of online learning in the United States“, *Curriculum and Teaching Dialogue*, 2015, vol. 17, pp. 21-34, URL: https://digitalcommons.du.edu/cgi/viewcontent.cgi?article=1026&context=law_facpub

[2] Doolan, K., Barada, V., Burić, I., Krolo, K., & Tonković, Ž., *Student Life During the COVID-19 Pandemic Lockdown: Europe-wide Insights* [Online], European Students' Union (ESU), 2021, URL: https://esu-online.org/wp-content/uploads/2021/04/0010-ESU-SIderalCovid19_WEB.pdf

[3] Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N., & Umek, L. “Impacts of the COVID-19 pandemic on life of higher education students: A global perspective”, *Sustainability*, 2020, vol. 12, 8438, URL: https://www.mdpi.com/2071-1050/12/20/8438

[4] Stecuła, K., & Wolniak, R., „Influence of COVID-19 pandemic on dissemination of innovative e-learning tools in higher education in Poland“, *Journal of Open Innovation: Technology, Market, and Complexity*, 2022, vol. 8, 89, https://doi.org/10.3390/joitmc8020089

[5] Jaschik, S., & Lederman, D., *2022 Survey of College and University Presidents*, Inside Higher Ed, Washington, DC, USA, 2022, URL: https://www.insidehighered.com/booklet/2022-survey-college-and-university-presidents

[6] Müller, C., & Mildenberger, T., „Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education“, *Educational Research Review*, 2021, vol. 34, 100394, URL: https://doi.org/10.1016/j.edurev.2021.100394

[7] Güzer, B., & Caner, H., „The past, present and future of blended learning: An in depth analysis of literature“, *Procedia - Social and Behavioral Sciences*, 2014, vol. 116, pp. 4596-4603, https://doi.org/10.1016/j.sbspro.2014.01.992

[8] Cooney, M.H., Gupton, P., & O'Laughlin, M., „Blurring the lines of play and work to create blended classroom learning experiences“, *Early Childhood Education Journal*, 2000, vol. 27, no. 3, pp. 165-171, https://doi.org/10.1007/BF02694230

[9] Voci, E., & Young, K., „Blended learning working in a leadership development programme“, *Industrial and Commercial Training*, 2001, vol. 33, no. 5, pp. 157-160, https://doi.org/10.1108/00197850110398927

[10] Bonk, C.J., Olson, T.M., Wisher, R.A., & Orvis, K.L., „Learning from focus groups: An examination of blended learning“, *Journal of Distance Education*, 2002, vol. 17, no. 3, pp. 97-118, URL: https://www.ijede.ca/index.php/jde/article/view/299

[11] Valiathan, P., „Blended Learning Models“, *Learning Circuits*, 2002, URL: https://www.purnima-valiathan.com/wp-content/uploads/2015/09/Blended-Learning-Models-2002-ASTD.pdf

[12] Clark, D., *Blended Learning* [White paper], Epic Group, Brighton, UK, 2003, URL: https://www.scribd.com/document/84278560/Clark-D-Blended-Learning#

[13] Twigg, C.A., „Improving learning and reducing costs: New models for online learning“, *EDUCAUSE review*, 2003, vol. 38, no. 5, pp. 28-38, URL: http://www.educause.edu/ir/library/pdf/erm0352.pdf

[14] Bubaš, G., Kermek, D., & Radošević, D., „Instructional design in hybrid learning: Use of e-learning technologies to develop competence in online communication“, *Proceedings of the Conference Computers in Education*, 29th International Convention MIPRO 2006, Opatija, Croatia, 2006, pp. 125-130.

[15] Allen, I.E., Seaman, J., & Garrett, R., *Blending In: The Extent and Promise of Blended Education in the United States*, Sloan Consortium, Needham, MA, USA, 2007, URL: https://files.eric.ed.gov/fulltext/ED529930.pdf

[16] Staker, H., & Horn, M. B., *Classifying K-12 Blended Learning*, Innosight Institute, Mountain View, CA, USA, 2012, URL: http://files.eric.ed.gov/fulltext/ED535180.pdf

[17] Alammary, A., Sheard, J., & Carbone, A., „Blended learning in higher education: Three different design approaches“, *Australasian Journal of Educational Technology*, 2014, vol. 30, no. 4, pp. 440-454, https://doi.org/10.14742/ajet.693

[18] Alammary, A., „Blended learning models for introductory programming courses: A systematic review“, *PLoS ONE*, 2019, vol. 14, no. 9, e0221765, https://doi.org/10.1371/journal.pone.0221765

[19] Murray, D. E., & Christison, M. A., *Online Language Teacher Education: Participants' Experiences and Perspectives*, The International Research Foundation (TIRF), Monterey, CA, USA, 2018, URL: https://www.tirfonline.org/wp-content/uploads/2017/03/TIRF_OLTE_2017_Report_Final.pdf

[20] Dixon, T., Christison, M. A., Dixon, D. H., & Palmer, A., „A meta-analysis of hybrid language instruction and call for future research“, *The Modern Language Journal*, 2021, vol. 105, no. 4, pp. 792-809, http://doi.org/10.1111/modl.12732

[21] Albiladi, W. S., & Alshareef, K. K., „Blended learning in English teaching and learning: A review of the current literature“, *Journal of Language Teaching & Research*, 2019, vol. 10, no. 2, pp. 232–238, http://dx.doi.org/10.17507/jltr.1002.03

[22] Beatty, B., „Transitioning to an online world: Using HyFlex courses to bridge the gap“, in *Proceedings of ED-MEDIA 2007--World Conference on Educational Multimedia, Hypermedia & Telecommunications*, Vancouver, Canada, 2007, pp. 2701-2706, URL: https://www.learntechlib.org/primary/p/25752/

[23] Pressley, J.P., „Explaining the difference between HyFlex and hybrid teaching models“, *EdTech: Focus on Higher Education*, March 1st, 2022, URL: https://edtechmagazine.com/higher/article/2022/03/hyflex-hybrid-teaching-models-whats-the-difference-eperfcon

[24] Bubaš, G., & Čižmešija, A., „Possible changes in academic teaching as a consequence of the COVID-19 pandemic: International findings and a case study from Croatia“, *Proceedings of the Central European Conference on Information and Intelligent Systems (CECIIS 2022)*, Faculty of Organization and Informatics, 2022, pp. 149-164, http://archive.ceciis.foi.hr/app/public/conferences/2022/Proceedings/DRP/DRP6.pdf

[25] Reiser, R.A., "A history of instructional design and technology: Part I: A history of instructional media", *Educational Technology Research and Development*, 2001, vol. 49, no. 1, pp. 53–64, URL: http://www.jstor.org/stable/30220299

[26] Edmonds, G. S., Branch, R. C., & Mukherjee, P., „A conceptual framework for comparing instructional design models“, *Educational Technology Research and Development*, 1994, vol. 42, no. 4, 55–72, URL: http://www.jstor.org/stable/30220096

[27] Andrews, D. H., & Goodson, L. A., „A comparative analysis of models of instructional design“, *Journal of Instructional Development*, 1980, vol. 3, no. 4, pp. 2–16, URL: http://www.jstor.org/stable/30220623

[28] Stefaniak, J., & Xu, M., „An examination of the systemic reach of instructional design models: A systematic review“, *TechTrends* 2020, vol. 64, pp. 710–719, URL: https://doi.org/10.1007/s11528-020-00539-8

[29] Donmez, M. & Cagiltay, K., „A review and categorization of instructional design models“, *Proceedings of E-Learn: World Conference on E-Learning*, Association for the Advancement of Computing in Education (AACE), 2016, pp. 370–384, URL: https://www.learntechlib.org/primary/p/173964/

[30] Saçak, B., Bozkurt, A., & Wagner, E., „Learning design versus instructional design: A bibliometric study through data visualization approaches“, *Education Sciences*, 2022, vol. 12, 752. https://doi.org/10.3390/educsci12110752

[31] Bond, J., Dirkin, K., „What models are instructional designers using today?“, *The Journal of Applied Instructional Design*, 2020, vol. 9, no. 2, URL: https://edtechbooks.org/jaid_9_2/what_models_are_inst

[32] Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K., „Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation“, *The Internet and Higher Education*, 2019, vol. 42, pp. 34-43, https://doi.org/10.1016/j.iheduc.2019.04.001

[33] Castro, M.D.B., & Tumibay, G.M., „A literature review: Efficacy of online learning courses for higher education institution using meta-analysis“, *Education and Information Technology*, 2021, vol. 26, pp. 1367–1385, https://doi.org/10.1007/s10639-019-10027-z

[34] Kraiger, K., Ford, J. K., & Salas, E., „Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation“, *Journal of Applied Psychology*, 1993, vol. 78, no. 2, pp. 311–328, https://doi.org/10.1037/0021-9010.78.2.311

[35] Bubaš, G., Balaban, I., & Begičević, N., „Course evaluation in e-learning by comparative evaluation of two or more e-learning courses“, *Proceedings of the 18th International Conference on Information and Intelligent Systems*, Varaždin, Croatia, 2007, pp. 65-72, URL: https://www.researchgate.net/publication/228716200_Course_Evaluation_in_e-Learning_by_a_Comparative_Analysis_of_Two_or_More_e-Learning_Courses

[36] Kovačić, A., Bubaš, G., & Zlatović, M., „E-tivities with a wiki: Innovative teaching of English as a foreign language“, *14th European University Information Systems (EUNIS) International Congress*, Aarhus, Danska, 2008, URL: https://www.eunis.org/eunis2008/papers/p87.pdf

[37] Martin, F., Sun, T., & Westine, C.D., „A systematic review of research on online teaching and learning from 2009 to 2018“, *Computers in Education*, 2020, vol. 159, 104009, https://doi.org/10.1016/j.compedu.2020.104009

[38] Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A., „The difference between emergency remote teaching and online learning“, *EDUCAUSE Review*, March 2020, URL: https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

[39] Stewart, W. H., "A global crash-course in teaching and learning online: A thematic review of empirical Emergency Remote Teaching (ERT) studies in higher education during year 1 of COVID-19", *Open Praxis*, 2021, vol. 13, no. 1, pp. 89–102, URL: https://search.informit.org/doi/10.3316/informit.758902304536019

[40] Xue, S., & Crompton, H., „Educational technology research during the COVID-19 pandemic“ [Online first], *Interactive Technology and Smart Education*, 2022, URL: https://doi.org/10.1108/ITSE-05-2022-0067

[41] Mehanna, W. N., "e-Pedagogy: the pedagogies of e-learning", *Research in Learning Technology*, 2004, vol. 12, no. 3, pp. 25-44, https://doi.org/10.3402/rlt.v12i3.11259

[42] Dabbagh, N., „Pedagogical models for E-Learning: A theory-based design framework“, *International Journal of Technology in Teaching and Learning*, 2005, vol. 1, no. 1, pp. 25-44, URL: https://sicet.org/main/wp-content/uploads/2016/11/ijttl-05-01-DabbaghVol1.Iss1_.pp25-44.pdf

[43] Muñoz Carril, P., González Sanmamed, M., & Hernández Sellés, N., „Pedagogical roles and competencies of university teachers practicing in the e-learning environment“, *International Review of Research in Open and Distributed Learning*, 2013, vol. 14, no. 3, pp. 462–487. https://doi.org/10.19173/irrodl.v14i3.1477

[44] Babić, S., & Bubaš, G., „Assessment of competencies of online teachers: Pilot study and survey evaluation“, *Proceedings of the 6th International Conference on e-Learning 2015*, Beograd: Metropolitan University, 2015, pp. 164-171, URL: https://elearning.metropolitan.ac.rs/files/pdf/2015/26-Snjezana-Babic-Goran-Bubas-Assessment-of-competencies-of-online-teachers-pilot-study-and-survey-evaluation.pdf

[45] Mirković Moguš, A., *Procjena kvalitete procesa e-obrazovanja na visokim učilištima u Republici Hrvatskoj*, doktorska disertacija, Sveučilište u Zagrebu, Fakultet organizacije i informatike Varaždin, Hrvatska, URL: https://repozitorij.foi.unizg.hr/islandora/object/foi%3A535/datastream/PDF/view

[46] Babić, S., *Činitelji nastavničkoga prihvaćanja e-učenja i kompetencije za njegovu primjenu na visokoškolskim ustanovama*, doktorska disertacija, Sveučilište u Zagrebu, Fakultet organizacije i informatike Varaždin, Hrvatska, 2016, URL: https://repozitorij.unizg.hr/islandora/object/foi:405/datastream/PDF/view

[47] Venable, M. A., *2022 Online Education Trends Report* [Online], BestColleges.com, URL: https://www.bestcolleges.com/research/annual-trends-in-online-education/

[48] Anderson, M., Faverio, M., & McClain, C., *How Teens Navigate School During COVID-19* (Press Release), Pew Research Center, Washington, DC, USA, URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2022/05/PI_2022.06.02_teens-covid19-school_REPORT.pdf

Umjetna inteligencija kao izvor potpore nastavnicima u pripremi i izvođenju nastave

K. Maček Blažeka
Tehnička škola Ruđera Boškovića, Zagreb, Republika Hrvatska
Tehnička škola Čakovec, Zagreb, Republika Hrvatska
kristinka.blazeka@skole.hr

***Sažetak* - Svjedočimo vidljivo ubrzanom napretku u područjima istraživanja i implementacije umjetne inteligencije (AI) vezano uz procesiranje prirodnih jezika (NLP) i njihovog razumijevanja (NLU). U softverskom inžinjerstvu uvedeni su botovi (bots), aplikacije sposobne reagirati na vanjske podražaje od strane korisnika na koje odgovaraju pokretanjem automatiziranih postupaka. Pri tome često uključuju konverzacijske sposobnosti, te sa krajnjim korisnikom komuniciraju tekstualnim porukama (chatbots) ili govorom (voicebots). U radu se diskutira kako mogućnost uvođenja umjetne inteligencije, odnosno korištenje chatbota (ChatGPT) tijekom postupka izrade i razrade nastavnih sadržaja za primjer pripreme (i izvođenja) nastave u srednjoškolskom obrazovanju može pomoći kvalitetnijoj pripremi i tako poboljšati procese učenja i poučavanja. U radu se istražuje mogućnost primjene umjetne inteligencije kod oblikovanja sadržaja i tijeka nastavnog sata namijenjenog učenju programiranja. Pri tom se pokušava odgovoriti na neka značajna pitanja kao što su: kako se umjetna inteligencija može iskoristiti za olakšavanje rada ili pomoć u radu nastavnika; kako može pomoći u traženju kreativnih rješenja; kako može pomoći kod izbjegavanja čestih pogrešaka kod učenja početnicima i drugo. Prezentirane mogućnosti upućuju općenito na veliki potencijal. Konačno, u zaključku se demonstriraju preporuke i daje osvrt na moguće aspekte koji upućuju na obećavajuće rezultate sustava umjetne inteligencije da, u ulozi pomoćnika nastavnicima, utječu na povećanje kvalitete njihova rada, a time i cjelokupnog obrazovnog procesa.**

Ključne riječi – planiranje i priprema nastave; učenje programiranja; umjetna inteligencija; chatbot;

I. UVOD

Umjetna inteligencija već značajno oblikuje naše živote (popunjavanje radnih mjesta, brojne primjene u vojsci, primjenjuje se otkrivanje plagijata, masovni mediji otkrivaju što volimo...), pod utjecajem smo algoritama i tehnologija sa svrhom da uzimaju i oblikuju što znamo, vidimo i kako zamišljamo svijet. Nova se tehnološka rješenja koriste za prikupljanje, sintezu i primjenu složenih oblika nadzora nad podacima koje koriste ili ih prodaju korporacije, banke, osiguravajuća društva, investitori. To su samo neki od razloga zbog kojih je važno osigurati dubinsku analizu razvoja, rizika i mogućnosti za primjenu umjetne inteligencije u obrazovanju [1] kao i istražiti kako bi to moglo oblikovati i utjecati na budućnost obrazovanja. Poznato je i u nizu tekstova izloženo kako AI sustav uspješno rješava ispitne zadatke na raznim ispitima.

Kako bi se škole i obrazovni sustav u cjelini trebali prilagoditi širenju alata umjetne inteligencije? Za dio nastavnika OpenAI ChatGPT postao je razlogom za brigu i paniku. Od svoje pojave ovaj je alat postao uzrokom kontroverzi. S jedne strane raste bojazan vezana uz utjecaj umjetne inteligencije na lažiranje prilikom rješavanja raznovrsnih zadataka kod kuće i u školi jer taj alat može uspješno generirati eseje, rješavati probleme iz matematike ili drugih predmeta, prezentirati funkcionalan kod za računalne programe za različite jezike ili platforme, kao i niz drugih zadataka i sve to u vremenu koje se mjeri u sekundama. S druge strane, smatra se i opasnim jer konverzacijski AI softver, treniran na enormnim količinama podataka, koji je u mogućnosti voditi realistične konverzacije sa ljudima može u određenim situacijama manipulirati korisnicima bez njihovog znanja, npr. zbog marketinških, financijskih, političkih ili čak kriminalnih razloga.

Ovaj rad nastao je kako bi se ispitale i prezentirale mogućnosti ChatGPT-a u ulozi pomoćnika nastavnicima skraćivanjem vremena potrebnog za pripremu nastave i nastavnih materijala. Dijeleći stranu koja je zauzela stav kako ga (ipak) ne bi trebalo izbacivati iz škola i učionica, u nastavku se iznosi iskustvo da se kroz primjer korištenja ovog alata u kreiranju plana i sadržaja poučavanja za odabrane nastavne sadržaje kod učenja programiranja utvrde i iznesu neka opažanja i iskustvo.

II. ISKUSTVA I (NE)PRIHVATLJIVOST

Kao i drugi napredni sustavi strojnog učenja, ChatGPT pokazuje potencijalne sigurnosne probleme, npr. postoji rizik od napada oglašavanja pri kojem napadač pokušava manipulirati modelom ubacujući zlonamjerne ulaze koji će producirati neispravne ili neželjene izlaze. Zabrinutost se može vezati i uz njegov potencijal da ga se koristi za širenje dezinformacija ili propagande, posebno ako se integrira sa platformama širokog dosega kao što su to društvene mreže [2].

Zbog tih i sličnih razloga, neke su javne škole u SAD promptno početkom 2023. godine blokirale pristup do ChatGPT platforme na školskim računalima i mrežama izbacivši ga iz škola i učionica s obrazloženjem „brige zbog negativnog utjecaja na učenje i zabrinutosti vezane uz sigurnost i točnost sadržaja“. U dijelu škola pristup je onemogućen samo učenicima.

Pozitivna iskustva nastavnika sa ovim sustavom iznijeta su u članku "Don't Ban ChatGPT in Schools, Teach With It" (NYT, 12.1.2023.). Jedan je nastavnik za evaluaciju učeničkih radova od AI sustava u vrlo kratkom vremenu dobio upotrebljiviji i detaljniji odgovor učenicima od onog kojeg bi on sam kao mogao pružiti, a drugi je utvrdio kako je većina pitanja koje mu je AI sustav generirao za analizu zadanog teksta bila upotrebljiva. Ipak, ne vidjevši u AI sustavu pomoćnika, nego konkurenciju, neki su se nastavnici zabrinuli za svoju egzistenciju.

Zbog različitih koristi (o jednoj se govori i u ovom radu) niz nastavnika ipak pokazuje spremnost da prigrli ovaj AI sustav kao novinu, u vremenskom slijedu niza tehnologija i alata za učenje, svjesni da je njihov zadatak na početku 21. stoljeća pripremiti učenike za svijet prožet umjetnom inteligencijom: "kao edukatori, još nismo otkrili koji je najbolji način upotrebe umjetne inteligencije u školama, no ona dolazi, htjeli mi to ili ne." [3]. Ti su nastavnici oduševljeni mogućnostima ChatGPT-a pri uključivanju u njihov svakodnevni rad, kao i njegovoj sposobnosti da lekcije nastavnih predmeta učini kreativnijima i primjerenijima svakodnevnom životu.

Neki od primjera kako se ChatGPT može koristiti u razredu su: personalizirani planovi učenja, pomoć prilikom učenja pisanja u engleskom jeziku, za generiranje ideja za razredne aktivnosti, kao pomoć u učenju nakon nastave ili kao partner u debati (niz primjera je u [4]). Kako sustav (ipak) ponekad griješi, može pomoći i u vježbama kritičkog mišljenja ili poticanju evaluacije samog sustava od strane korisnika, odnosno učenika koji ga koriste.

Očekivano, nastavnici sada predlažu i potiču da im se omogući stjecanje vještina u korištenju ove platforme umjetne inteligencije kako bi na odgovarajući način stimulirali kreativnost učenika i njihove vještine u rješavanju problema: "Imamo mogućnost povesti naše učenike uz nešto što će biti važan dio njihovih života. Nemojmo to blokirati i izbacivati." [3]

Ne manje važno, već i učenici osnovnih škola potvrđuju kako im ovaj nov način rada uz umjetnu inteligenciju učenje čini ugodnijim i zabavnijim, impresionirani su i smatraju ga generatorom novih, dobrih i kreativnih ideja. Uz ispravan pristup, današnji učenici koji će, kako je izgledno, rasti kao prva generacija nativnih svjetskih AI korisnika, prihvatit će da umjetnu inteligenciju koriste za sugestije, ali da proizvode svoj vlastiti rad.

III. OSNOVNO O CHATGPT

U jednom od važnih područja umjetne inteligencije, procesiranju prirodnih jezika (NLP, Natural Language Processing) i njihovom razumijevanju (NLU, Natural Language Understanding), u novije se vrijeme javljaju sve uspješniji proizvodi softverskog inženjerstva, a mogu se opisati kao botovi.

Botovi (bots) su općenito računalni programi sposobni reagirati na vanjske podražaje (događaje, poruke) od strane korisnika na koje odgovaraju pokretanjem automatiziranih postupaka, djelujući kao sučelje između korisnika i odgovarajućih usluga. Pri tome često uključuju konverzacijske sposobnosti, te se krajnjem korisniku obraćaju (komuniciraju) tekstualnim porukama (chatbots) ili govorom (voicebots). Chatbot (ponekad se naziva i virtualnim pomoćnikom) je alat koji ima svrhu pojednostavniti komunikaciju između ljudi i računala. Takvi AI programi su složeni obećavajući oblik interakcije između čovjeka i stroja i mogu oponašati humani razgovor sa korisnicima.

2022. godina biti će obilježena kao godina u kojoj je javnosti obznanjena nova generacija konverzacijskih sučelja te vrste: OpenAI ChatGPT i Google LaMDA. Oba sustava koriste generativnu tehniku umjetne inteligencije poznatu kao Large Language Models (LLMs) i predstavljaju značajan skok u napretku konverzacijskih sposobnosti softvera. Izgleda kako smo konačno ušli u doba prirodnog računalstva u kojem ljudi mogu voditi smislene i organski tekuće konverzacije sa programskim alatima i aplikacijama.[5]

ChatGPT (Generative Pre-trained Transformer) objavljen krajem studenog 2022. kao dio šireg skupa tehnologija razvijenog od strane OpenAI (startup iz San Francisca), blisko je povezan sa Microsoft-om. Predstavlja dio nove generacije AI sustava koji je u mogućnosti razgovarati, odnosno na zahtjev generirati čitljiv tekst, pa čak i producirati slike i video temeljem naučenog iz ogromnih baza digitalnih knjiga, tekstova i drugih medija. U mogućnosti je razumjeti kontekst konverzacije i generirati prikladne odgovore u različitim programskim jezicima.

ChatGPT-ov Improved Accuracy (CGA) model ima različita svojstva koja ga čine snažnim NLP sustavom koji koristi AI arhitekturu baziranu na dubokom učenju kako bi producirao točne i smislene konverzacije. Duboko učenje omogućuje mu odlučivanje na temelju podataka, pronalaženjem i izdvajanjem uzoraka podataka iz velikih skupova podataka, koji se precizno preslikavaju u valjane izlazne odluke [6]. Primjenom predtreniranog modela GPT-3, CGA je tako u mogućnosti generirati realistične i poticajne konverzacije usavršene sa mogućnošću da uči iz vlastitih grešaka, dopuštajući mu prilagodbu novim kontekstima kako bi generirao bolje rezultate. Trenutna istraživanja pokazala su kako je CGA dostigao impresivnu razinu točnosti i generativnih sposobnosti. [2]

Dizajniran je da radi kao pisani dijalog (chatbot) između AI sustava i osobe koja postavlja pitanja. Sva pitanja koja dobije pomažu mu i čine ga „pametnijim" [7] te je tako ChatGPT u mogućnosti učiti i poboljšavati se.

IV. ODABIR NASTAVNOG PREDMETA, CJELINE I TEME

Kao izazov ovom AI sustavu, u strukovnom kurikulumu za stjecanje kvalifikacije tehničar za računalstvo [8] odabran je predmet Algoritmi i programiranje koji se predaje u drugom razredu srednje škole i pripadajuća nastavna cjelina Pokazivači. Kurikulumom ova cjelina sastoji se od nastavnih tema: deklaracija pokazivača, funkcije i pokazivači, polja i pokazivači, te pokazivači na niz znakova.

Izbor je pao na drugu godinu učenja, s obzirom da omogućava širi obuhvat znanja o algoritmima i osnovama

programiranja koje su učenici do tada već usvojili. Vremenski ova tema ulazi u prvo polugodište drugog razreda srednje škole, a u objašnjavnje se može uključiti imenovanje i deklaraciju varijabli, aritmetičke i logičke operacije, naredbe za određivanje tijeka programa i naredbe za ponavljanje, funkcije, polja i rad sa nizovima znakova jer su to učenici trebali savladati već ranije. Uz to, učenici su već trebali upoznati i algoritme za pretraživanje i sortiranje.

Nakon te nastavne cjeline uče se još samo strukture podataka i datoteke, te grafika. Dosta logično se strukture podataka mogu uvoditi i već kroz teme u odabranoj cjelini Pokazivači. Sve su to razlozi zbog kojih je izbor pao upravo na ovu cjelinu.

V. KONVERZACIJA SA SUSTAVOM PRILIKOM PRIPREME MATERIJALA ZA NASTAVNU CJELINU POKAZIVAČI

S ciljem ispitivanja mogućnosti i reakcija ChatGPT sustava za pripremu i razradu nastavne cjeline Pokazivači provedena je konverzacija u nastavku. Odgovori su većinom skraćeni bez utjecaja na njihovu bit.

P1. "Koji bi bio najprimjereniji obuhvat i poredak tema za objašnjavanje pokazivača početnicima u C++?"

O1. Logičan i razumljiv poredak bio bi:

1. Uvod u pokazivače: što su pokazivači, za što se koriste, kako se razlikuju od „običnih" varijabli.
2. Deklaracija i inicijalizacija pokazivača kako bi pokazivali na specifične memorijske lokacije.
3. Pristup memoriji korištenjem pokazivača uz postupak dereferenciranja kako bi se dohvatilo vrijednost spremljenu na specifičnoj memorijskoj lokaciji.
4. Aritmetika s pokazivačima uz korištenje aritmetičkih operacija kod prolaska članovima polja i drugim podatkovnim strukturama.
5. Dinamička memorijska alokacija upotrebom new i delete i kako pokazivače koristiti za upravljanje dinamičkom memorijom u C++.
6. Pokazivači i funkcije: predaja pokazivača kao argumenta funkciji i vraćanje pokazivača iz funkcija.
7. Pokazivači i polja: kako pokazivači služe za manipulaciju poljima i kako se polja mogu predati funkciji kao pokazivači.
8. Pokazivači na pokazivače: deklaracija i korištenje, posebno kao korištenje kod implementacije višedimenzionalih polja i drugih naprednih podatkovnih struktura.
9. Uobičajene pogreške prilikom upotrebe pokazivača te kako ih spriječiti.

Komentar: prihvatljivo i podudarno uobičajenom tijeku obrade tih sadržaja u priručnicima. Prijedlog prati sugestivan i koristan zaključak sustava: „ovakav outline osigurava solidno razumijevanje osnovnih koncepata pokazivača u C++ početnicima prije nego što prijeđu na naprednije teme".

P2: "Koje su razlike između pokazivača i uobičajenih varijabli?"

O2: Radi se o dva različita tipa varijabli u programiranju koji se primarno razlikuju s obzirom na definiciju, memorijsku alokaciju, pristup podacima, dinamičku memorijsku alokaciju, mogućnost poprimanja null vrijednosti, mogućnost indirektnog pristupa te mogućnost konverzije.

Komentar: sažeto i korektno.

P3: „Koje su prednosti upotrebe pokazivača kod programiranja u C++?"

O3: Pokazivači su moćan alat koji, ako se koristi učinkovito, osigurava različite prednosti: mogućnost dinamičke memorijske alokacije, povećanu učinkovitost zbog izravne manipulacije memorijom, poboljšanu fleksibilnost funkcija, te unaprijeđeno objektno orijentirano programiranje kod programiranja koncepata poput nasljeđivanja i polimorfizma, kao i mogućnost aritmetike s pokazivačima te poboljšanu kontrolu nad memorijskim resursima.

Ipak, nekorektno korištenje pokazivača može voditi ka teškim slučajevima otklanjanja pogrešaka (memory leaks, pogreške segmentacije i druge pogreške kod izvođenja) zbog čega je esencijalno razumjeti osnove te zadržati dobre programerske prakse prilikom njihovog korištenja.

Komentar: vrlo dobra analiza zaključena sa korisnom primjedbom.

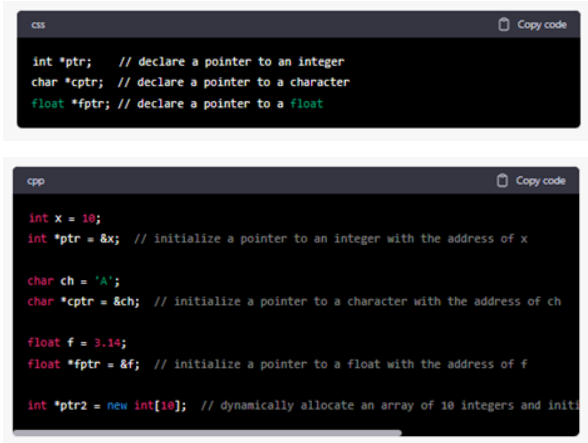
P4: "U čiju se pomoć najviše možemo pouzdati prilikom traženja pogrešaka u složenim situacijama upotrebe pokazivača?"

O4: Iskusni programeri, alati za debugiranje (gdb, lldb ili Visual Studio debugger), online forumi i zajednice (npr. Stack Overflow) te profesionalni servisi za debugiranje.

Komentar: Korektno uz opasku kako sustav ipak ne navodi mogućnost da sam pomaže u takvim situacijama.

P5: „Mogu li dobiti nekoliko dobrih primjera deklaracije i inicijalizacije pokazivača."

O5:



Slika 1 Primjeri deklaracije i inicijalizacije pokazivača

Komentar: Sustav ispisu dodaje napomenu kako je "važno da se kod inicijalizacije osigura da pokazivač pokazuje na validnu memorijsku lokaciju". Također se ističe kako je "prilikom dinamičke alokacije memorije sa new, važno i dealocirati memoriju sa delete". Korektno, uz opasku kako je u naslovu ispisa krivo navedeno css.

P6: „Postoje li u programskom jeziku C++ pokazivači na pokazivače?“

O6:

```java
int x = 42;
int *ptr = &x;
int **ptrptr = &ptr;  // ptrptr is a pointer to a pointer to an integer
```

```c
cout << **ptrptr << endl;  // outputs 42
```

Slika 2 Primjer deklaracije pokazivača na pokazivač

Komentar: Generirani primjer slijedi opis funkcioniranja prilikom indirektnog pristupa varijabli u memoriji. U nastavku se naglašava kako “pokazivači na pokazivače mogu biti korisni npr. u slučaju rada sa dinamičkim višedimenzionalnim poljima, odnosno općenito kod dinamičke manipulacije pokazivačima.” Ovdje se također u naslovu javlja krivi naziv programskog jezika.

P7: U pokušaju povezivanja naučenog o manipulaciji sa nizovima znakova upotrebom klase string i korištenja pokazivača za rješavanje zadataka takvog tipa pomoću pokazivača, slijedi pitanje: “Koji bi bio dobar primjer za usporedbu dva pristupa u programiranju?”

O7:

```c
#include <iostream>
#include <string>

using namespace std;

int main()
{
    // Approach 1: using string class
    string str1 = "Hello";
    string str2 = " World";
    string str3 = str1 + str2;
    cout << str3 << endl;

    // Approach 2: using character pointers
    char cstr1[] = "Hello";
    char cstr2[] = " World";
    char cstr3[11];
    int i;
    for (i = 0; cstr1[i]; i++) {
        cstr3[i] = cstr1[i];
    }
    for (int j = 0; cstr2[j]; j++, i++) {
        cstr3[i] = cstr2[j];
    }
    cstr3[i] = '\0';
    cout << cstr3 << endl;

    return 0;
}
```

Slika 3 Primjer korištenja pokazivača kod spajanja nizova

Navedeni programski kod prati i iscrpno objašnjenje te usporedba s obzirom na očekivano vrijeme izvođenja ovog programskog koda na velikim (dugim) stringovima koje je dulje prilikom korištenja klase string.

Komentar: Točnost prezentiranog koda provjerena je korištenjem programskog okruženja Microsoft Visual Studio Community 2022 (64-bit), pa s obzirom da je kod drugog pristupa bila potrebna korekcija programskog koda (radi određenih upozorenja i/ili pogrešaka vezanih uz ograničenja korištenih funkcija), od sustava je zatraženo novo rješenje što je rezultiralo slijedećim prijedlogom:

```c
#include <iostream>
#include <cstring>

int main()
{
    char str1[] = "Hello";
    char str2[] = " World";
    int length = strlen(str1) + strlen(str2);
    char *str3 = new char[length + 1];

    strcpy(str3, str1);
    strcat(str3, str2);

    std::cout << str3 << std::endl;

    delete[] str3;
    return 0;
}
```

```c
#include <iostream>

int main()
{
    char str1[] = "Hello";
    char str2[] = " World";
    int length1 = 0;
    int length2 = 0;
    while (str1[length1] != '\0')
        length1++;
    while (str2[length2] != '\0')
        length2++;
    int length = length1 + length2;
    char *str3 = new char[length + 1];

    for (int i = 0; i < length1; i++)
        str3[i] = str1[i];
    for (int i = 0; i < length2; i++)
        str3[length1 + i] = str2[i];
    str3[length] = '\0';

    std::cout << str3 << std::endl;

    delete[] str3;
    return 0;
}
```

Slika 4 Dodatni primjeri rješenja za problem spajanja nizova

Komentar: Kako je drugi prijedlog nakon provjere generirao pogreške u prevođenju, na zahtjev za novim, trećim prijedlogom, sustav ispisuje treći primjer programskog koda koji kao ispis daje očekivani tekst. Oduševljava što su svi primjeri popraćeni korisnim detaljnim tumačenjem prezentiranog koda, a impresionira i mogućnost sustava da s lakoćom prikazuje alternativna rješenja za zadani problem. Pokazuje se poželjnim da u novijim inačicama sustav bude u mogućnosti predvidjeti ili upozoriti i na ograničenja ili specifičnosti korištenja primijenjenih funkcija za rad sa nizovima znakova.

P8: „Koji bi bili primjeri ili slučajevi kada, u postupcima manipulacije nizovima znakova, pristup upotrebom pokazivača zahtjeva značajno manje vremena kod izvođenja i ima prednosti u performansama nego odgovarajući pristup upotrebom klase string?“

O8: Neki primjeri su:

- Vrlo dugi stringovi zbog toga što klasa string svoje podatke sprema dinamički na heap

memoriju tražeći dodatne operacije memorijske alokacije i dealokacije što usporava postupak.
- Performance-critical code kad se manipulacija stringovima izvodi učestalo (npr. high-performance game engine).
- Operacije spajanja stringova kod spajanja mnogo manjih stringova u jedan veliki i to zbog višestrukih memorijskih alokacija i dealokacija potrebnih kako bi se promijenila veličina interne kratkoročne memorije prilikom rasta stringa u slučaju rada sa string klasom. Upotrebom pokazivača može se prealocirati velika kratkoročna memorija i spajati stringove izravno u toj memoriji.
- Manipulacija C-style stringovima za sučelja sa C programskim kodom ili bibliotekama.

Komentar: Prilikom ovog pitanja sustav također generira obuhvatan i korektan odgovor.

P9: „Koji bi bio primjer za programski kod funkcije koja će bez i uz pomoć pokazivača sortirati riječi u rečenici prema njihovoj duljini?“

O9:

```c
#include <iostream>
#include <string>
#include <vector>
#include <algorithm>

using namespace std;

bool compareLength(const string &a, const string &b)
{
    return a.length() < b.length();
}

void sortWordsByLength(const string &sentence)
{
    vector<string> words;
    string word;
    for (char c : sentence) {
        if (c == ' ') {
            words.push_back(word);
            word.clear();
        }
        else {
            word += c;
        }
    }
    words.push_back(word);

    sort(words.begin(), words.end(), compareLength);

    for (const string &word : words) {
        cout << word << " ";
    }
    cout << endl;
}

int main()
{
    string sentence = "This is a sentence that we want to sort";
    sortWordsByLength(sentence);

    return 0;
}
```

Slika 5 Primjer programskog koda sa sortiranjem bez pomoći pokazivača

Komentar: Korektni primjeri koda, sa i bez upotrebe pokazivača, detaljnije opisani sa objašnjenjima korištenih funkcija i čitavog programa i izvode se na očekivani način.

```c
#include <iostream>
#include <cstring>
#include <algorithm>

using namespace std;

bool compareLength(const char* a, const char* b)
{
    return strlen(a) < strlen(b);
}

void sortWordsByLength(char* sentence)
{
    int n = 0;
    char* words[100];
    char* p = strtok(sentence, " ");
    while (p) {
        words[n++] = p;
        p = strtok(nullptr, " ");
    }

    sort(words, words + n, compareLength);

    for (int i = 0; i < n; i++) {
        cout << words[i] << " ";
    }
    cout << endl;
}

int main()
{
    char sentence[] = "This is a sentence that we want to sort";
    sortWordsByLength(sentence);

    return 0;
}
```

Slika 6. Primjer programskog koda sa sortiranjem uz pomoć pokazivača

P10: „Koji bi bio primjer korištenja pokazivača kod programiranja rekurzije?“

O10: Uz ispis primjera, sustav odgovara da su pokazivači vrlo korisni kod rekurzije:

```c
#include <iostream>

void printList(int* arr, int size) {
  if (size == 0) return;
  std::cout << *arr << " ";
  printList(arr + 1, size - 1);
}

int main() {
  int arr[] = {1, 2, 3, 4, 5};
  int size = sizeof(arr) / sizeof(arr[0]);
  printList(arr, size);
  return 0;
}
```

Slika 7. Primjer upotreba pokazivača kod rekurzivne funkcije

Komentar: Nakon iznošenja primjera, u objašnjenje je uključen poučan savjet kako “korištenjem pokazivača umjesto varijable koja bi predstavljala trenutnu poziciju u polju možemo spriječiti overhead kopiranja čitavog polja tijekom svakog rekurzivnog poziva. To može biti vrlo korisno kod rada sa vrlo velikim poljima, jer reducira upotrebu memorije te popravlja performanse programa.“

P11: “Koji bi bili primjeri uobičajenih pogrešaka koje se javljaju prilikom korištenja pokazivača i kako ih izbjeći?”

O11:

- Dereferenciranje null pokazivača: pojavljuje se prilikom pokušaja pristupa vrijednosti spremljenoj u pokazivač koji nije bio inicijaliziran ili je bio postavljen u null.
- Pristup memoriji izvan granica polja
- Memory leak: pojavljuje se prilikom dinamičnog alociranja memorije pomoću new ili malloc, ali je propušteno dealocirati ili osloboditi memoriju.
- Korištenje/pristupanje memoriji nakon oslobađanja ili dealociranja sa delete ili free.
- Dvostruko oslobađanje: dealokacija već dealocirane memorije

VI. KONVERZACIJA PRILIKOM OBLIKOVANJA PREZENTACIJE ZA NASTAVNU CJELINU POKAZIVAČI

P12: „Molim za prijedlog i razradu sadržaja za prezentaciju koja bi se koristila za poučavanje i objašnjavanje pokazivača početnicima.“

O12:

I. Uvod (objašnjenje pokazivača, važnost pokazivača u C++, zbog čega je važno učiti o pokazivačima, preduvjeti za učenje pokazivača)
II. Memorija (objašnjenje memorije u C++, alokacija i upravljanje memorijom, stack i heap memorija)
III. Pokazivači i adrese (Što je pokazivač? Kako radi pokazivač? Sintaksa i deklaracija pokazivača, aritmetika pokazivača, adresni operator (&), operator dereferenciranja (*))
IV. Pokazivači i funkcije (Korištenje pokazivača kao funkcijskih parametara, predaja pokazivača preko adrese – by reference, vraćanje pokazivača iz funkcija, Dangling pointers
V. Dinamička memorijska alokacija (Objašnjenje, operatori new i delete, memory leaks, pametni pokazivači)
VI. Uobičajene upotrebe pokazivača (polja i pokazivači, strukture i pokazivači, pokazivači i klase, pokazivači i stringovi)
VII. Savjeti i dobre prakse (Izbjegavanje uobičajenih pogrešaka kod upotrebe pokazivača, traženje pogrešaka u stvarima vezanim uz pokazivače, najbolje prakse upravljanja memorijom)
VIII. Zaključak (Sažetak ključnih koncepata, Izvori za daljna učenja, rekapitulacija važnosti pokazivača)

P14: „Gdje se mogu naći dobre ilustracije i animacije koje mogu pomoći u tumačenju ovih koncepata učenicima?“

O14: Mnogo je raspoloživih online resursa koji mogu u tome pomoći, npr.:

YouTube – neki popularni kanali su "The Cherno" "CodeVault," i "Programming with Mosh."
Online Learning Platforms – Udemy, Coursera i edX
Codecademy - interaktivna platforma
Priručnici – npr. na Amazon ili u lokalnoj knjižnici
Online Forumi i zajednice - Stack Overflow i Reddit.

Komentar: sustav „ne zaboravlja“ u odgovoru naglasiti kako je „važno da su vizuali jasni i jednostavni za razumijevanje, te ih je važno detaljno objasniti kako bi učenici razumjeli prezentirane koncepte.“

VII. ZAKLJUČAK

Unatoč određenoj sumnjičavosti pri pristupu ovoj konverzaciji, primjenjeni AI sustav demonstrirao je brze odgovore, konkretna i upotrebljiva znanja te sposobnost sintetiziranja znanja kao i konciznost u odgovorima. Na sva pitanja ili komentare AI odgovarao je prihvatljivo i s prividnim „razumijevanjem“. Programski odsječci su korektni, raznoliki i dovoljno detaljno objašnjeni.

Iskustvo izloženo kroz primjere u ovom radu navodi na stav kako bi se ChatGPT mogao koristiti kao pomagalo u pripremi materijala za poučavanje, u objašnjavanju nastavnih sadržaja i pripremi ispitnih pitanja.

Rezultati i dosadašnje iskustvo potvrđuju mišljenje koje u knjizi [1] istražuje mogućnosti i rizike upotrebe AI u obrazovanju te utvrđuje kako bi škole trebale prigrliti tu tehnologiju i koristiti ju za poboljšanje zadataka i poticanje originalnosti u idejama i razmišljanju učenika. "U mogućnosti smo razmisliti što mi u obrazovanju zapravo radimo. Zbog čega imamo zadatke koje AI s takvom lakoćom replicira?“. Zadaci bi trebali biti smisleniji i ispitivati vještine kritičkog razmišljanja učenika, razvijati umove sposobne nalaziti rješenja.

Zato potencijal ovog softvera, koji je od svoje pojave krajem studenog 2022. postao globalni fenomen, ne bi trebalo ni zanemarivati ni negirati. U području obrazovanja njegove uloge pomoćnika u pripremi i izvođenju nastave (kao obrazovnog alata) mogu biti vrlo različite i trebalo bi osigurati da nadjačaju rizike njegovog korištenja. Određeni oblici „suradnje“ sa ovim alatom kao učinkovitim pomoćnikom omogućuju put ka uspješnijem učenju i dubljem razumijevanju predmetnog gradiva. Tako bi se ovaj alat u poučavanju mogao prihvatiti i kao pomoć oslobađanju učeničke kreativnosti nuđenjem personaliziranog mentorskog nadzora, ali i radi pripreme učenika za rad sa AI sustavima i njihovim modelima u budućnosti. Primjenivši ga na način da djeluje kao pomoćnik, ChatGPT može učiniti nastavu kvalitetnijom.

REFERENCE

[1] S. Popenici: “Artificial Intelligence and Learning Futures, Critical Narratives of Technology and Imagination in Higher Education”, Routledge, New York, 2023.

[2] Jianyang D, Yijia L.: ”The Benefits and Challenges of ChatGPT: An OverviewAssessment in vocational upper secondary education”, Frontiers in Computing and Intelligent Systems, Vol. 2, 2022.

[3] J. Gecker: “It’s coming, whether we want it to or not’: Teachers nationwide are using ChatGPT to prepare kids for an AI world”, sa https://fortune.com/2023/02/14/chatgpt-school-lessons-cheating-robots-ai-teachers/ (posjećeno veljača 2023.)

[4] “ChatGPT, Chatbots and Artificial Intelligence in Education”, sa https://ditchthattextbook.com/ai (posjećeno prosinac 2022.)

[5] “We gave ChatGPT a college-level microbiology quiz. It aced it.”, sa https://bigthink.com/the-future/chatgpt-microbiology-quiz-aced/ (posjećeno, veljača 2023.)

[6] J. D. Kelleher: “Duboko učenje”, Mate, Zagreb, 2021.

[7] M. O Brien: “Explainer: What is ChatGPT and why are schools blocking it?”, sa https://apnews.com/article/what-is-chat-gpt-ac4967a4fb41fda31c4d27f015e32660 (posjećeno veljača 2023.)

[8] Strukovni kurikulumu za stjecanje kvalifikacije tehničar za računalstvo sa https://www.asoo.hr/UserDocsImages/Kurikulumi/K_Tehnicar%20za%20racunalstvo_za%20odobrenje.pdf, posjećeno veljača 2023.

Knjižnica u Glideu

Korina Udina
OŠ Kostrena, Kostrena, Hrvatska
korinau@yahoo.com

***Sažetak* - Primjenjujući digitalne alate otvaraju se nove mogućnosti dostupnosti i vidljivosti sadržaja preporučenog za čitanje u osnovnoškolskoj dobi. Rad istražuje prihvaćenost i korištenje aplikacije Glide „Knjižnica u Glideu“ nakon nekoliko godina upotrebe u osnovnoj školi. Svrha rada je ukazati na pitanja razvoja digitalnih i čitalačkih pismenosti u osnovnoj školi putem modela poticanja čitanja, te važnost stjecanja informacijske pismenosti, kompetencije obavezne u osnovnoškolskom uzrastu. Aplikacija Glide pokazala se izvrsnom u dijeljenju i pretraživanju mrežnih mjesta, unaprijed pripremljenih izvora informacija za učenike. Cilj istraživanja bio je utvrditi stvarno korištenje aplikacije “Knjižnica u Glideu” među učenicima i učiteljima te moguće razloge nekorištenja. Analiza rezultata istraživanja o korištenju aplikacije prezentira da je sredina u kojoj djeluje školska knjižnica djelomično prihvatila model traženja djela za čitanje putem digitalnog alata, ali je školska knjižnica ovim modelom poticanja čitanja ostvarila izvrsnu vidljivost i dostupnost u svom okruženju. Temeljem analize može se zaključiti da sadržaji kojima ispitanici najčešće pristupaju su „Popis djela za čitanje“i poveznica za pristup e-lektiri, što opravdava stvaranje ovog modela za poticanje čitanja „Knjižnica u Glideu“. Zaključno, postoji ograničena upotreba dostupnih tehnoloških resursa, a korištenje aplikacije “Knjižnica u Glideu” ovisi o dostupnosti Interneta, o zainteresiranosti učenika ali i učitelja za sadržaj koji se nudi.**

Ključne riječi – digitalni alat, dostupnost, igra, školska knjižnica, vidljivost

I. UVOD

Rad istražuje prihvaćenost i korištenje aplikacije „Knjižnica u Glideu“ nakon nekoliko godina upotrebe u jednoj osnovnoj školi, kao u [2], prikaz takvog modela poticanja čitanja te važnost razvoja digitalnih i čitalačkih pismenosti. Tijekom tri godine uočava se razlika u korištenju aplikacije. Razlike u raširenosti korištenja ove aplikacije odnose se na učenike i učitelje. Provedeno je istraživanje sa svrhom utvrđivanja stvarne situacije o upotrebi digitalnog alata „Knjižnica u Glideu“ jer se metode učenja i poučavanja mijenjaju i razvijaju usporedno sa mijenjanjem potreba u sustavu.

O važnosti stjecanja informacijske pismenosti kao kompetencije u školskom uzrastu govore razni autori, kao i školski knjižničari na državnim stručnim skupovima. Šibenski državni skup 2019. bio je posvećen upravo razvoju i poticanju informacijske pismenosti u programima rada školskih knjižničara, kao u [4]. Pismenosti o kojima se govori u radu su digitalna i čitalačka, a obje su podređene informacijskoj pismenosti. Za digitalnu pismenost najčešća predodžba je korištenje informacija dostupnih putem Interneta, sposobnost čitanja i razumijevanja slika, zvukova i teksta. Čitalačka pismenost pripada govornom području odgojno-obrazovnog procesa kojem je cilj poticati i usavršavati vještinu čitanja jer je učenje zavisno od razumijevanja teksta, i moguće je samo ako je prepoznavanje riječi automatizirano. Kada osoba usvoji vještinu čitanja može se s njom raditi na usvajanju čitalačke pismenosti.

„Nismo čitateljski pismeni ako ne znamo prepoznati manipulaciju… Nismo čitalački pismeni ako smo neuki. Ako smo neuki, naime, ne znamo ni što su činjenice, a što interpretacije, ne znamo što su argumenti, što smicalice, ne znamo odabrati podatke potrebne da bismo mogli kritički promišljati. Ne znamo učiti. A da bismo mogli učiti, učenje treba zavoljeti.“, kao u [9]. Bez čitalačke pismenosti neće se steći kompetencije potrebne za život, pojašnjava Dinka Jurić u svom članku.

Digitalna pismenost sa čitalačkom pismenošću doprinijeti će bržem stizanju do obrazovnog cilja, a to je postići kod učenika čitanje s razumijevanjem, čitanje s kritičkim mišljenjem o pročitanom, čitanje s užitkom, čitanje koje može biti zabavno. „Prema ALA-inom Izvješću (Presidential Committee, 1989), informacijski pismene osobe definiraju se kao one koje su naučile kako učiti… jer znaju kako je znanje organizirano, kako pronaći informacije i kako ih koristiti na svima razumljiv način… To su osobe pripremljene na učenje tijekom cijelog života”, kao u [7]. Informacijska pismenost je preduvjet digitalne pismenosti što se pojašnjava u članku da „Digitalno pismena osoba posjeduje sposobnost pronalaženja, analize, (pr)ocjene, upotrebe i stvaranja sadržaja korištenjem tehnologije i interneta. Takvoj se osobi pružaju različite mogućnosti razvoja i zapošljavanja, u zemlji u kojoj živi, ali i u bilo kojoj drugoj zemlji članici Europske unije.“, kao u [8]. Kako bi se učenicima omogućilo usvajanje vještine pristupa i korištenja provjerenih izvora znanja, bez obzira na to gdje se nalaze, u školi ili kod kuće, ponuđena im je aplikacija “Knjižnica u Glideu” kao model pristupačnog i zanimljivog pretraživanja obaveznih sadržaja. Cilj istraživanja u radu jest utvrditi stvarnu raširenost i korištenje aplikacije među učenicima i učiteljima.

Praksa je pokazala da je najuspješnije učenje kroz igru. Interdisciplinarnošću u odgojno obrazovnom neposrednom radu, predmetnim korelacijama i koordinacijom sa drugim područjima, školski knjižničar kroz svoj program ima prostor za razvoj informacijske pismenosti. To je proces koji se događa u sva tri obrazovna ciklusa, kao što je i razvoj čitanja o kojem piše Kelić u svom priručniku „Za razliku od razvoja materinjeg jezika kojim govorimo, koji se događa spontano, razvoj čitanja i pisanja je sustavan proces.“, kao u [3].

Modeli poticanja čitanja su mnogobrojni, a uloga knjižnice, između ostalih, je neprestano osvježavati taj proces. Školski knjižničar, uz kompetencije stečene formalnim ili neformalnim obrazovanjem, unutar školskog kurikuluma će educirati učenike u korištenju primjenjivih digitalnih alata i igara kojima će produbiti njihova znanja o učinkovitom traženju informacija, vrednovanju pronađenih informacija, poučiti ih pouzdanosti i vjerodostojnosti te prijenosu informacija drugima. To područje poslova učitelja i knjižničara oko informacijske pismenosti još uvijek je ovisno o stupnju informacijske i informatičke pismenosti kolektiva. U kurikulumu odgoja i obrazovanja smatra se da je informacijska pismenost skup znanja i vještina sa kojima učenici trebaju završiti osnovnoškolsku razinu, nadograđivati je na srednjoškolskoj razini. Međutim da bi se taj cilj obrazovanja ostvario važno je obrazovanje studenata, budućih nastavnika/učitelja i školskih knjižničara informatičkim programima i informacijskom pismenošću.

II. KNJIŽNICA U GLIDEU

Novi načini komunikacije potpuno mijenjaju prostor i vrijeme, te iako je vrijednost tehnoloških dostignuća neprocjenjiva, postalo je pitanje svih pitanja koji način upotrebe je primjeren u osnovnoškolskoj dobi. Dijete prije polaska u školu stvara rječnik slika koje mu pomažu u komunikaciji, prepoznaje riječi kao slike, simbole, prema vizualnim obilježjima. Važan uvjet čitanja je povezivanje govornog i pisanog jezika.

U prezentiranoj aplikaciji „Knjižnica u Glideu“, digitalnom alatu, koriste se idiomi, slike za poticanje vidne percepcije za povezivanje s pisanim tekstom što može poslužiti za poticanje vještine čitanja. Procesi za poticanje vještine čitanja dio su programa rada školskog knjižničara. Rješenjem o zaduženjima stručnih suradnika knjižničara za školsku godinu propisan je neposredni odgojno-obrazovni rad usmjeren na četiri područja: čitanje, učenje, informacijsku pismenost te kulturnu i javnu djelatnost. U području informacijske pismenosti digitalni alati su medij učenja blizak mladima u diseminaciji sadržaja obrazovanja.

Primjerice, digitalni alati Glide ili Genially, kao u [1] (platforma za interaktivnu animaciju sadržaja) otvaraju nove mogućnosti dostupnosti i vidljivosti sadržaja preporučenog za čitanje u osnovnoškolskoj dobi, ili sadržaja opće knjižnične pismenosti s kojom odlaze u srednjoškolsko obrazovanje. Školski knjižničar će imati veći utjecaj o toj temi u školi ako je ostvarena dostupnost i vidljivost školske knjižnice. Približavanje mladima pretpostavlja znanje o korištenju svih alata kojima se oni koriste i povezivanje istih s programima obrazovanja u školi. Vidljivost i dostupnost školske knjižnice može biti postignuta različitim načinima, primjerice, kroz informiranje mrežnom stranicom, komunikacijom putem kanala MS Teamsa, kanalom Virtualne knjižnice, korištenjem TikTok i YouTube kanala, ili objedinjavanjem svega toga u jednom digitalnom alatu. Osim same fizičke otvorenosti knjižnice svim potrebama mladih, dolazak učenika i učitelja u knjižnicu i služenje njenim sadržajima, digitalno vrijeme omogućava prisutnost obrazovnih sadržaja u virtualnom obliku, na njihovnim osobnim tabletima i mobitelima. Nasuprot informacijsko-komunikacijske tehnologije uvijek je otvoreno projektno, istraživačko učenje i kreativno stvaranje, što nije tema ovog rada.

Učenje kroz igru, učenje osobnim iskustvom pokazalo se izvrsnim načinom u savladavanju novih znanja. Iz istog razloga nastala je prije nekoliko godina aplikacija „Knjižnica u Glideu“, za približavanje sadržaja i izvora informacija mladima. Ta aplikacija je postala način koordiniranja posudbe lektire među razredima u školi. Na početku školske godine, kada su uneseni svi sadržaji relevantni za tu školsku godinu, podijeli se poveznica učenicima i učiteljima putem MS Teams kanala. Preuzimanjem poveznice može se instalirati ikonu za pretraživanje sadržaja na sučelje svog mobitela, tableta, računala ili pametne ploče. Za uzrast učenika od 1. do 4. razreda poveznicu preuzima učiteljica u MS Teams kanalu kojim ide komunikacija među djelatnicima škole i posebnim kanalima za učenike, a može podijeliti poveznicu roditeljima putem e-dnevnika, Viber grupe itd. Učenici od 5. do 8. razreda sami preuzimaju aplikaciju, kao i učitelji hrvatskog jezika koji su izradili popis djela za čitanje te školske godine. U aplikaciji osim dostupnosti i smještaja knjiga iz popisa lektire, dostupan je katalog školske knjižnice sa svim naslovima, koliko traženih naslova ima za posudbu. Ugrađena je poveznica na mrežnu stranicu škole, mrežnu stranicu školske knjižnice, e-lektiru dostupnu putem Carneta, video mamce i video filmove u Carnetu. O školi i školskoj knjižnici može se saznati putem poveznice na kartu školskih knjižnica Hrvatske, ili Virtualnu školsku knjižnicu. Opisana aplikacija je svima dostupna putem poveznice ili skeniranjem QR koda koji se nalazi na vratima školske knjižnice, na ulazu u hodnik škole, na staklenim i drugim površinama, na mrežnim stranicama škole i školske knjižnice. Instalirana na početno sučelje mobitela, tableta, klikom na ikonu svaki korisnik pristupa sadržajima.

Svrha izrade ovakve aplikacije sa popisom lektirnih naslova je bila s jedne strane oživjeti taj popis djela za obavezno čitanje i približiti ga učenicima na moderan i zanimljiv način, a sa druge strane uskladiti posudbu lektirnih naslova između paralelnih razreda kako se ne bi dogodilo da netko ne može besplatno posuditi lektirnu ili drugu knjigu u školskoj knjižnici jer ih nema dovoljan broj. Lektira je učenicima uglavnom odbojna jer je obavezna stoga je izrada Glide aplikacije „Knjižnica u Glideu“ temeljena na nekoliko činjenica. Činjenica je da dijete prije polaska u školu stvara rječnik slika koje mu pomažu u komunikaciji, prepoznaje riječi kao slike, simbole, prema vizualnim obilježjima. Druga važna činjenica za čitanje je povezivanje govornog i pisanog jezika. Ako pretpostavimo da će se u digitalnom alatu koristiti idiomi, slike za poticanje vidne percepcije za povezivanje s pisanim tekstom koji će potom poslužiti za razvoj vještine čitanja, može se zaključiti da će se proces za poticanje vještine čitanja unaprijediti ovim digitalnim alatom.

O potrebi vraćanju knjizi i čitanju, što je temeljni postulat u radu svakog knjižničara, unatoč napretku tehnologije, govorio je Miroslav Solar na otvaranju jedne od trideset i pet „Proljetnih škola školskih knjižničara“, državnog stručnog skupa školskih knjižničara. Navedeni razlozi u prilog čitanju knjiga u njegovom govoru

motivacija su za razvoj i inovacije novih modela poticanja čitanja. Solar kaže da „Nešto poput odgoja ukusa upućuje na to da se vratimo knjigama…čitanju uokvirenih smislenih cjelina…koje nas mogu uputiti na prepoznavanje onoga što nam se sviđa osobno, a ne zbog toga što to preporučuju moda i mediji. Sposobnost ukusa nije „dana“, nego je zadata; nju treba na neki način odgajati.“, kao u [6]. Igra slikama, idiomima, vizualizacija knjiga, likova, poveznice kojima učenik šeta školskom knjižnicom, katalogom slobodnih naslova za posudbu, brzi prijelazi na video mamce i e-lektiru, mrežne stranice škole i knjižnice, prihvatljivija je od „obaveznog popisa“. Time je dostupnost i vidljivost školske knjižnice unaprijeđena, a rad školskog knjižničara suvremen. Učenici, roditelji i učitelji kojima je namijenjena aplikacija uvijek i u svakom trenutku, kao što koriste društvene mreže u komunikaciji (Instagram, Facebook, Viber) ostvaruju brzu poveznicu sa školom, knjižnicom i informacijom koja im je potrebna putem ikone na sučelju mobitela, tableta ili računala.

Tijekom godina uočava se razlika u korištenju aplikacije. Razlike u raširenosti korištenja ove aplikacije odnose se na učenike i učitelje te je provedeno istraživanje sa svrhom utvrđivanja stvarne upotrebe digitalnog alata „Knjižnica u Glideu“ jer se metode učenja i poučavanja mijenjaju i razvijaju usporedno sa mijenjanjem potreba u školi.

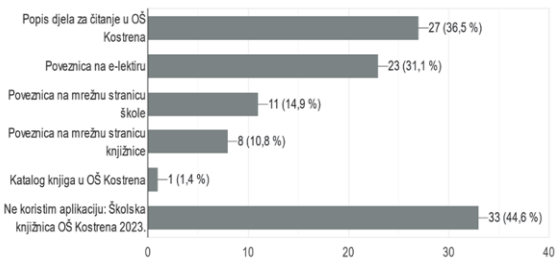
III. REZULTATI ISTRAŽIVANJA

Istraživanje je provedeno sa ciljem utvrđivanja stvarnog broja učenika i učitelja koji koriste aplikaciju „Knjižnica u Glideu“, te utvrđivanje sadržaja kojeg najviše koriste. U metodologiji istraživanja korišteni su intervju, razgovor sa ispitanicima, i anketa. Izabran je instrument istraživanja anketom elektroničkim putem kao najbržom metodom za prikupljanje podataka.

Provedena je mrežna anketa na Google obrascu putem MS Teams kanala. Ciljana populacija ispitanika bili su učenici od petog do osmog razreda i učitelji. Ispitanici su anketi pristupali prema osobnom izboru. Diseminacija i pristup anketi za učenike bio je na satu informatike. Postupak ispitivanja anketom trajao je dva nastavna dana. Učenici od petog do osmog razreda pristupali su anketi na satu informatike uz asistenciju učitelja informatike, a učitelji putem poveznice stavljene u MS Teams kanal zbornice. Učenici nižih razreda nisu bili uključeni u anketiranje. Anketa je sadržavala sedam pitanja. Anketna pitanja bila su koncipirana tako da utvrde: tko ispunjava anketu učenik ili učitelj, jesu li učenici i učitelji preuzeli aplikaciju Glide, koriste li je ponuđene poveznice u obrazovne svrhe i koje sadržaje najčešće koriste. Pitanjem o procjeni osobne informatičke ili informacijske pismenosti daje se mogućnost samovrednovanja. Uzorak ispitanika u odnosu na ukupan broj učitelja u školi i učenika od petog do osmog razreda je obzirom na kratko vrijeme anketiranja manji ali dostatan za određene zaključke. Na anketu je odgovorilo 14 učitelja ili 28% od ukupnog broja, a od petog do osmog razreda u ispitivanje se je uključilo 60 učenika ili 32% od ukupnog broja učenika viših razreda.

Metodom intervjua, razgovora, utvrđene su određene činjenice, primjerice za neke učenike od petog do osmog razreda razlozi nekorištenja aplikacije su što nemaju mobitele, a neki nemaju Internet na mobitelima pa ne mogu preuzeti i instalirati aplikaciju. Mobiteli im služe samo za razgovor. Za tablete zadužene u školi kažu da su spori uz slab Internet, rijetko se koriste jer se brzo prazne baterije. Neke od tih tvrdnji su potvrđene analizom rezultata ankete: 8% ispitanika odgovorilo je da nema Internet na mobitelu, a 9,5% ispitanika se izjasnilo da nema mobitel, tablet ili laptop na kojem može koristiti aplikaciju. Metodom intervjua, razgovora, utvrđene su razlike u korištenju aplikacije uočene kod učenika, a ovise o korištenju aplikacije od strane učiteljica. Neki učitelji/ice su u diseminaciji informacije o aplikaciji posredovali između učenika, roditelja i školske knjižnice, a neki nisu. Razlozi nekorištenja aplikacije „Knjižnica u Glideu“ od dijela učitelja mogu biti temom drugog istraživanja. Novo istraživanje samovrednovanjem moglo bi utvrditi razinu informatičke i informacijske pismenosti učitelja kao jednog od uvjeta digitalne i čitalačke pismenosti učenika.

Analiza rezultata anketnim ispitivanjem o stvarnom korištenju aplikacije „Knjižnica u Glideu“ prezentira zainteresiranost ispitanika za ovaj digitalni alat te sadržaje koje najčešće koriste u aplikaciji Na pitanje gdje su instalirali preuzetu aplikaciju, 42% ispitanika je odgovorilo da je preuzeta na mobitel, 7 % ispitanika preuzelo je aplikaciju na tablet, a 12% na laptop. Zbog spore Internet veze u školi aplikaciju ne koristi 33% ispitanika. Sadržaji kojima se najčešće pristupa je „Popis djela za čitanje“, kao i poveznica za pristup e-lektiri, što je u konačnici bila svrha stvaranja aplikacije „Knjižnica u Glideu“. Nešto manje ispitanika kroz aplikaciju pristupa mrežnoj stranici škole ili knjižnice. Katalog školske knjižnice nije interesantan ispitanicima. Temeljem rezultata istraživanja može se zaključiti da je sredina u kojoj djeluje školska knjižnica djelomično prihvatila ideju da učenici i učitelji djela za čitanje pronalaze putem digitalnog alata, svog mobitela ili laptopa. Nadalje, može se zaključiti temeljem analize rezultata ispitivanja da su razlozi nekorištenja aplikacije „Knjižnica u Glideu“ nepostojanje Interneta na mobitelima učenika, slaba Internet veza u školi, loša diseminacija informacije o aplikaciji. Iako se neprestano govori o našem društvu kao društvu znanja, neophodna je edukacija informacijskom pismenošću u sveučilišnim programima ne samo budućih školskih knjižničara nego i budućih učitelja, prenositelja znanja i osoba koje će svojom znatiželjom i cjeloživotnim učenjem zainteresirati učenike za neprestano istraživanje novih područja i novih načina učenja. (Slika 1.)



Slika 1. Odgovori ispitanika na pitanje: Koje sadržaje najviše koriste iz aplikacije „Knjižnica u Glideu“

Za pretpostaviti je da je protekom tri godine došlo do zasićenja ovim digitalnim alatom te je vrijeme za promjenu metode učenja i poučavanja kroz neki novi oblik ili digitalni alat. Značajan doprinos korištenja aplikacije Glide kao modela za poticanje čitanja može biti vidljiv u budućim istraživanjima vezanim za unaprijeđenje aplikacije temeljenim na povratnim informacijama ključnim za poboljšanje dizajna i korisničkog iskustva.

IV. ZAKLJUČAK

Aplikacija „Knjižnica u Glideu“ pokazala se izvrsnom u dijeljenju i pretraživanju mrežnih mjesta, unaprijed pripremljenih izvora informacija, ali diseminacija informacija o ovom digitalnom alatu kao i njegova upotreba ovise o dostupnosti Interneta, o zainteresiranosti učenika i učitelja za sadržaj koji se nudi.

Sadržaji kojima ispitanici najčešće pristupaju je „Popis djela za čitanje“, kao i poveznica za pristup e-lektiri, što je opravdalo cilj stvaranja ovog modela za poticanje čitanja „Knjižnica u Glideu“.

Ovim istraživanjem se utvrdilo da još uvijek postoji ograničena upotreba dostupnih tehnoloških resursa. Novo istraživanje samovrednovanjem moglo bi utvrditi razinu informatičke i informacijske pismenosti učitelja te utvrditi digitalnu i čitalačku pismenost učenika. Značajan doprinos korištenju aplikacije Glide, pa i kao modela za poticanje čitanja, može biti vidljiv u budućim istraživanjima vezanim za unaprijeđenje aplikacije temeljenim na povratnim informacijama ključnim za poboljšanje dizajna i korisničkog iskustva.

Primjena računala i digitalne tehnologije u školskim knjižnicama u stvaranju modela za poticanje čitanja, razvoju digitalne i čitalačke pismenosti ima tridesetogodišnju tradiciju. O tom razvojnom putu u radu školskih knjižničara svjedoče stručni radovi u digitaliziranim zbornicima stručnih skupova školskih knjižničara, kao u [5]. Na kraju, jedan od uvjeta za razvoj inovativnih pristupa u poticanju digitalne i čitalačke pismenosti je sredina u kojoj djeluje školska knjižnica, od opremljenosti prostora do kolektiva i njihove razine informatičke i informacijske pismenosti. Iz školske godine u godinu povećava se broj znatiželjnika, broj korisnika aplikacija kroz koje je školska knjižnica dostupnija i vidljivija svima, kako učenicima tako i roditeljima i učiteljima.

REFERENCE

[1] Dostupno: https://view.genial.ly/6060a9b00138370d9e849a34/interactive-content-skolska-knjiznica-os-kostrena-knjiznicna-grada

[2] “Knjižnica u Glideu” Glide App. https://shaky-nose-7941.glideapp.io/dl/d0a5f4

[3] M. Kelić, Ovladavanje čitanjem, Naklada Slap, 2015., str.19-23.

[4] “Ogledni nastavni sat stručnog suradnika knjižničara; Informacijska pismenost”, zbornik radova: Šibenik, 08.-11. travnja 2019. Zagreb: Agencija za odgoj i obrazovanje, http://library.foi.hr/lib/knjiga.php?B=1&sqlx=X01861&ser=&sqlid=1&sqlnivo=&css=&H=&U=&grupa=P%8A%8AKRH&vrsta=ZBI&H=proljetna-skola#

[5] K.Udina, “Prije vremena-Ante tempus”, “Tradicija i suvremenost u radu školskih knjižničara”, zbornik stručnih radova: Zagreb: Agencija za odgoj i obrazovanje, 2018., str.21 – 29, dostupno na: http://library.foi.hr/knjige/knjiga1.aspx?C=3196&H=proljetna-skola&vrsta=ZBI_KAT&grupa=P%8A%8AKRH_%8AKOLSKE%20KNJI%8ENICE&lang=hr&broj=1

[6] M.Solar, “Zašto čitati knjige”, Zagreb: Agencija za odgoj i obrazovanje, zbornik stručnih radova 2011., str.11–20., dostupno na: http://library.foi.hr/knjige/knjiga.aspx?C=X00947

[7] S. Špiranec, “Informacijska pismenost - ključ za cjeloživotno učenje,” Edupoint, godište III, rujan 2003., http://edupoint.carnet.hr/casopis/17/clanci/1.html#para_3

[8] “Digitalna pismenost – nezaobilazni uvjet za budućnost”, https://europedirect-cakovec.eu/digitalna-pismenost-nezaobilazni-uvjet-za-buducnost/

[9] D. Juričić, „Čitalačka pismenost – šesti uvjet života“, Školski portal, https://www.skolskiportal.hr/kolumne/citalacka-pismenost-sesti-uvjet-zivota/

Digitalni alati i filozofija uma: Može li stroj razmišljati?

M. Kerstner* i M. Sertić**
*OŠ Šestine, Zagreb, Republika Hrvatska
**Industrijsko-obrtnička škola, Slavonski Brod, Republika Hrvatska
mladen.kerstner@skole.hr , sertic_79@yahoo.com

***Sažetak* - 21. stoljeće donijelo je čovječanstvu nebrojene dobre (ali i neke negativne) promjene. Uobičajeno je govoriti da je jedno od najvećih dostignuća našeg stoljeća razvoj umjetne inteligencije koji za sobom nosi cijeli niz benefita koji se očituju u zdravstvu, ekonomiji, strojarstvu, računarstvu, ekologiji, arheologiji, obrazovanju i u gotovo svim drugim znanostima. U radu smo pokušali povezati neke ključne aspekte filozofije uma – naročito koncept *lakih problema svijesti* i *teških problema svijesti* Davida Chalmersa – s umjetnom inteligencijom koja je *nota bene* ključno vezana za obrazovne digitalne platforme bez kojih europski, a time i hrvatski, obrazovni sustav više ne može funkcionirati. Postavili smo pitanje na koje, ipak, ne možemo konačno odgovoriti, a radi se o budućnosti digitalnih platformi – kako u odnosu na učenike i nastavnike, tako i u odnosu na cjelokupno društvo ili civilizaciju: Hoće li doći vrijeme kad će digitalna tehnologija (umjetna inteligencija koja će biti blizu sposobnosti imanja određenog oblika svjesnosti ili, radikalnije, čiste transcendencije) biti toliko superiornija ljudskome rodu da će činiti više štete nego koristi? Smatramo da je to daleka budućnost, no mnogi su filozofi, poput Martina Heideggera i Herberta Marcusea, još u 20. st. opominjali na određene negativne aspekte uporabe tehnologije bez dubljeg promišljanja na koji će način takva tehnologija utjecati na čovjekov prirodni humanizam.**

Ključne riječi – filozofija uma; digitalna tehnologija; umjetna inteligencija; budućnost obrazovanja

I. KRATKI UVOD U FILOZOFIJU UMA

Sintagma *filozofija uma* prilično je poznata u suvremenim filozofsko-neuroznanstvenim krugovima. S jedne strane imamo pojam *filozofije* – kao stroge humanističke znanosti o počelima, bitku, bivstvu, smislu, spoznaji, izvoru morala, o lijepome i dobrome – a s druge strane imamo nešto što nazivamo *umom*. Spajanjem ta dva pojma, dobivamo jednu cijelu „novu“ znanost koja za svoj predmet ima odnos ljudskog uma (ili duha) i ljudskog (materijalnog) tijela (mozga). Filozofija uma pokušava odgovoriti na pitanje kako ljudski um uopće funkcionira, što zapravo znači biti *svjestan*, što je *mentalnost*, postoji li *umstvenost* mimo postojanja mozgova (odnosno: je li mozak taj bez kojeg nikakva spoznaja nije moguća ili je on „samo“ posrednik koji omogućuje čovjeku spoznaju viših principa, te svih ostalih spoznajućih stvari, a taj viši princip opstoji mimo ljudi i nematerijalan je) [1].

Kako piše Pećnjak [1], pojam uma odnosi se na sveukupnost mentalnog (psihičkog) života, a razumijevamo ga i kao uzročnika ljudskog ponašanja. To mentalno stanje obuhvaća sve naše osjećaje, percepciju, emocije, spoznaju, zatim naše želje i vjerovanja, pa pamćenje i razmišljanje, dakako i puko zaključivanje. E sada, filozofi uma, koji se bave ovakvim vrstama mentalnih stanja, napravili su podjelu na *fenomenalna* (kvalitativna) *stanja* i na *propozicionalne stavove*. Prva stanja odnosila bi se na osjete, percepciju i dijelom na emocije; druga stanja, odnosno stavovi, bile bi želje, vjerovanja i namjere (emocije također jednim dijelom mogu biti propozicionalne). Možemo odmah primijetiti kako su fenomenalna stanja manje „egzaktna“ od stavova [1].

Zapravo je fenomenalni aspekt više subjektivni aspekt, no to ne znači da nije istinit. Primjerice, percepcija zelene boje: radi se o fenomenalnom stanju *zelenosti*, unutrašnjoj kvaliteti doživljavanja zelene boje, čiju „kvalitativnost“ filozofi uma nazivaju *kvalija*. Te kvalije nisu na prvi pogled mjerljive, niti također na prvi pogled podložne egzaktnim zakonima, no tu bi već ušli u priču za koju u ovakvom prikazu nemamo prostora (primjerice, za raspravu o *argumentu iz znanja* Franka Jacksona [2] čiji je misaoni eksperiment s Mary doživio ogromnu popularnost u znanstvenim krugovima).

S druge strane, propozicionalni stavovi su onakve tvrdnje koje se izriču propozicijom ili rečenicom, tj. koje označavaju neki (naš ili tuđi) stav i sadržaj tog stava. Recimo, rečenica „Hrvatska nogometna reprezentacija bi postala prvak svijeta u nogometu da je pobijedila Argentinu u polufinalu“ sadrži elemente propozicije (vjerovanje, želja) [2].

Filozofija uma propituje nekoliko glavnih filozofskih tema, a to su dualizam duha i tijela (duh ili um s jedne, te tijelo s druge strane, egzistiraju neovisno jedno o drugome), fizikalizam (um proizlazi iz tijela, svijest je „materijalna“), biheviorizam (psihološka i biološka interpretacija uma), funkcionalizam (objašnjava mentalna stanja i svijet oko nas prema funkcijama, tj. prema tome

kakvu funkciju ima primjerice određeni mentalni proces za radnju X), intencionalnost (svojstvo mentalnog stanja da se odnosi na neki X), redukcionizam (teorija prema kojoj će se društvene, humanističke, prirodne znanosti reducirati jednog dana na fiziku, bit će „utopljene“ u materijalistički koncept fizikalizma jer, prema toj teoriji, ionako u stvarnosti ništa ne postoji mimo fizikalnih zakona), supervenijenciju (filozofska teorija relacija svojstava nižega reda i višega reda, pri čemu pod nižim redom podrazumijevamo mozak i fizička stanja mozga, a pod višim redom mentalna stanja koja su uvijek i nužno vezana za niži red koji u biti određuje pojavnost višeg reda; ovdje postoji i suprotna teza prema kojoj mentalna stanja supreveniraju nad materijalnim), problem svijesti, i još neke druge probleme. Dualizam (koji se još naziva i *kartezijanski dualizam*) propituje je li čovjekov um različit od tijela (mozga), te ako je (a dualisti se slažu oko toga), na koji način su um i tijelo povezani. Zagovornici ovog modela smatraju da dualizam rješava jedan od velikih problema u filozofiji, a to je problem uma i tijela. Kartezijanski dualizam i argumenti čuvenog francuskog filozofa i matematičara Renéa Descartesa temelje se na premisi da je moguće zamisliti da nečiji um postoji bez nečijeg tijela i nečije tijelo bez nečijeg uma. Postoje i dva suvremena argumenta za dualizam: (i) *argument znanja* (taj smo termin već spomenuli), prema kojem dualizam mora biti istinit jer poznavanje svih fizičkih činjenica nije dovoljno za razumijevanje svjesnog iskustva, te (ii) *zombi argument* prema kojem um mora biti različit od tijela budući da tjelesni duplikat čovjeka ne bi imao svjesna iskustva [1, 3].

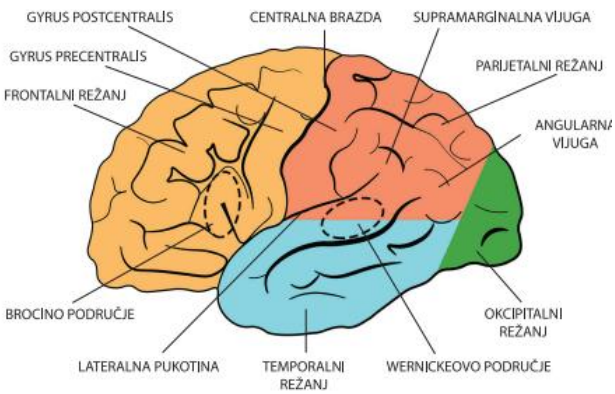
S druge strane, monisti ili fizikalisti su oni filozofi uma koji smatraju da je dualistički koncept neistinit i nemoguć; svijest ne postoji nezavisno od materijalnog objekta koji „emitira“ znanje, emocije, želje i kvalije iz jednog tvarnog centra, mozga. Postoje doduše i monisti koji su idealisti, no oni smatraju kako postoji isključiva nematerijalna datost, te da je sve što postoji *mentalno* (samo mi ljudi mislimo da postoji materija, a ona je zapravo utvara ljudskog uma). Fizikalisti smatraju kako je sve objašnjivo isključivo fizikom, te kako će doći vrijeme kad će se cjelokupna znanost (prvo humanistička, društvena, potom sva ostala prirodna znanost temeljena na empiriji i eksperimentu, osim fizike) motriti fizikalnim (fizičkim) očima. Odatle i termin *redukcionizam*. Što se tiče dualizma, zagovornici ove teorije smatraju da u svemiru opstoje i materijalna i nematerijalna „supstancija“; *um* bi bio nematerijalan, nefizikalan, a *tijelo* i svi objekti izvan uma bili bi materijalni, fizikalni [2].

Naša *svijest* da sada pišemo stručni rad za konferenciju je, prema dualistima, takve *kvalitete* da ju nikako ne možemo svesti na materiju, da ju dakle znanost fizike ne može opisati ni izmjeriti. Naposljetku, postoji i jedna neutralnija pozicija u odnosu na dualiste i moniste, a to je *funkcionalizam*. Takva teorija u sebi sadrži elemente i monizma, ali i dualizma, te je vrlo raširena među suvremenim filozofima uma [2].

Područje filozofije uma iznimno je široko i složeno. Ako bismo napravili jedan cijeli luk, te se prebacili preko supervenijencije, funkcionalizma, biheviorizma, teorije identiteta i još nekih drugih koncepata ove filozofije, doći ćemo na tom drugom luku do nečega što je možda najfundamentalnije, najvažnije, te – pogađate – najteže. Radi se o problemu svijesti. Čuveni Australac David Chalmers predstavio je koncept ovog problema tako što ga je razdijelio na *lake probleme svijesti* i na *težak problem svijesti*. Prvo ćemo nešto reći o lakim problemima svijesti. To su onakvi problemi koji se odnose na jednostavnije *radnje*, na ponašanje i razmišljanje koje se jednostavno mogu izmjeriti, odnosno objasniti mehaničkim analizama neuralnih procesa u našem mozgu. Laki problemi svijesti podložni su tzv. redukcionističkom ispitivanju, mjerljivi su i, uz pomoć adekvatne aparature, čak i vidljivi na monitorima [3].

Zapravo, laki problemi svijesti uključuju i ona pitanja na koja danas nemamo odgovora, no time što znamo njihov predmet izučavanja i formu odgovora koju moraju zadovoljiti, te načine na koje se ti problemi mogu riješiti [3], oni nisu *teški*. Laki problemi svijesti uključuju sposobnost diskriminacije, kategorizacije i reakcije na okolišne podražaje, integraciju informacija u kognitivnom sustavu, mogućnost izvještavanja o vlastitim mentalnim stanjima, sposobnost pristupa vlastitim mentalnim stanjima, fokusiranje pažnje, voljnu kontrolu ponašanja, te razliku između budnosti i spavanja [3, 4].

Sve navedene radnje su karakteristične za „laka“ svjesna mentalna stanja. No, Chalmers [4] tvrdi kako navedena mentalna stanja nisu u suštini ona stanja koja imaju fenomenalni karakter. Upravo objašnjenje tih fenomenalnih stanja svijesti Chalmers svrstava pod *težak problem svijesti*. Naime, to jest takav problem „upravo zato što uopće nije jasno kakav oblik će njegovo rješenje poprimiti. Naime, težak problem svijesti podrazumijeva fenomenalni pojam svijesti koji se odnosi na objašnjenje prirode iskustva koja imaju fenomenalni karakter. Za tu vrstu svjesnih iskustava (…) nije jasno kako ih možemo objasniti kroz funkcionalno mehanicistička objašnjenja“ [3, str. 221].



Slika 1. Kora velikog mozga. Izvor: Šimić, N. „Putovanje ljudskim mozgom.“ U: Mozak i um: Od električnih potencijala do svjesnog bića. [Ur. Šimić, N., Valerjev, P. i Nikolić Ivanišević, M.]. Zadar: Sveučilište u Zadru, 2020, str. 41.

Zapravo, težak problem je problem objašnjenja zašto su određeni mehanizmi popraćeni svjesnim iskustvom. Recimo, zašto pisce ovog članka neuronska obrada u njihovim mozgovima dovodi do tjeskobe (jer moraju ubrzo poslati rad na recenziju), a ne do primjerice ravnodušnosti,

gladi ili ljutnje. Chalmers smatra da je moguće zamisliti da bi se relevantna ponašanja povezana s tjeskobom, gladi ili bilo kojim drugim osjećajem mogla pojaviti čak i bez tog osjećaja. Ovako koncipirana teorija o teškom problemu svijesti sugerira da je iskustvo nesvodivo na fizičke sustave kao što je mozak. Chalmers nije zadovoljan s fizikalističkim objašnjenjem fenomenalnih svijesti, te se utječe svojevrsnom dualističkom konceptu [4].

No, kakve sad to veze ima s umjetnom inteligencijom, sa strojevima i obrazovanjem, vidjet ćemo u narednim odlomcima [4].

II. JOŠ KRAĆI UVOD U UMJETNU INTELIGENCIJU

Može li postojati inteligencija koja nije prirodna, koja je umjetno stvorena? I što bi ona predstavljala, za koga? Umjetna inteligencija (AI – *artificial intelligence*) je naziv kojim označavamo neki neživi, umjetni sustav koji ima sposobnost rješavanja određenih akcija pod nadzorom – na ovaj ili onaj način – ljudskog čimbenika. Teoretski, umjetna inteligencija može predstavljati i autonoman sustav koji funkcionira bez ljudi, bio stvoren od čovjeka, nekog drugog biološkog organizma ili neke treće (umjetne ili ne-umjetne, naravne ili nadnaravne) inteligencije [5].

Uobičajeno je da danas umjetnu inteligenciju najčešće povezujemo s računalnim sustavima. Koncept je poznat i ranije, još od Alana Turinga i njegovog testa (*Turingov test*) iz 1950. Nešto su kasnije stvoreni prvi računalni programi, pa čak i oni koji su „igrali“ šah, te pobjeđivali velemajstore u toj igri. Intencija jest, i bila je ranije, stvoriti stroj koji će *oponašati* čovjekova mentalna stanja [5].

No sjetimo se kraja prethodnog poglavlja; odmah primjećujemo da stroj danas doduše možda i može „oponašati“ neka mentalna stanja, i to ona mentalna stanja za koja bi Chalmers rekao da spadaju u kategoriju *lakih problema svijesti*. No još uvijek nije moguće stvoriti stroj, i pitanje je hoće li to ikada biti moguće, koji bi bio u stanju biti „svjestan“ *teških problema svijesti*, odnosno koji bi bio u stanju producirati *fenomenalna stanja* (kvalije) i o tome „razgovarati“ s drugim strojevima ili ljudima [5].

Umjetna inteligencija uključuje također i koncepte *strojnog učenja* i *dubokog učenja*. Strojno učenje (*machine learning*) je koncept računala koje uči poboljšati svoja predviđanja i kreativnost kako bi nalikovala ljudskom procesu razmišljanja pomoću određenih algoritama. Danas većina internetskih tražilica i društvenih mreža koristi koncept strojnog učenja u svojim predviđanjima ponašanja korisnika (zapravo: potrošača). Strojno učenje uključuje niz procesa učenja [5]. S druge strane, duboko učenje (*deep learning*) predstavlja onaj koncept računala koje simulira procese koje ljudski mozak treba analizirati, o kojima „razmišlja“ i „uči“. Proces dubokog učenja uključuje nešto što se zove *neuronska mreža* kao dio procesa računanja / *razmišljanja*. Za dubinsko učenje potreban je iznimno jak hardver, te velika količina podataka u stroju [5].

Zahvaljujući dubokom učenju, današnji programi mogu s lakoćom prepoznavati u ogromnim bazama podataka fotografije, glasove, određene simbole, slike i sl. Računala rješavaju iznimno složene zadatke za koja bi mnogim timovima bili potrebni tjedni ili godine (primjerice, pronalaženje milijuna maglica u dalekim galaksijama) [5].

I strojno učenje i dubinsko učenje zaslužno je za revolucionaran rast tržišta video igara u svijetu, ali mnogo bitnije: za pomoć u pronalaženju novih lijekova protiv teških bolesti, u biotehnologiji, ekologiji, ekonomiji, strojarstvu, kriminalistici, sportskoj analitici, novinarstvu, pa i u špijunaži. Odnos umjetne inteligencije, strojnog učenja i dubinskog učenja možemo vidjeti na slici broj 2, a kako nam umjetna inteligencija danas pomaže u obrazovanju, vidjet ćemo u sljedećem poglavlju [5].



Slika 2. Odnos umjetne inteligencije, strojnog učenja i dubinskog učenja

III. DIGITALNI ALATI: POMOĆ PRI UČENJU ILI *ŠVERCANJE* ZNANJA

Nakon poglavlja o filozofiji uma i umjetnoj inteligenciji, slijedi poglavlje u kojem ćemo prezentirati rezultate umjetne inteligencije (koja jest vezana i za neke koncepte filozofije uma), a ti se rezultati odnose na „strojeve“ koji učenicima i nastavnicima pomažu u svladavanju gradiva i proširenju stečenih znanja. Nesumnjivo je da su digitalni i *online* alati koji se danas koriste u obrazovanju golema pomoć u svladavanju mnogih obrazovnih sadržaja, te služe kako za učenje, tako i za oblikovanje najrazličitijih formi uz pomoć kojih se nastavni sadržaj može kvalitetnije prezentirati [6].

Danas postoje na tisuće različitih digitalnih i *online* alata koji se koriste u obrazovanju. Dovoljno je na mrežnim stranicama Agencije za mobilnost i programe EU [6] promotriti nazive alata koji služe u izradi video materijala, grafičkih zadataka, potom za interaktivnu prezentaciju sadržaja, za prezentiranje i predavanje sadržaja, potom za aktivno uključivanje sudionika *online* i *offline*, te općenito za komunikaciju između učenika i nastavnika, učenika i učenika, nastavnika i nastavnika, te roditelja s nastavnicima. Neke od aplikacija koje su često korištene u nastavi su *Spark Video*, *iMovie*, *Simplemind*, *Memrise*, *Geoguesser*, *Plickers*, *Flip*, *Edpuzzle*, *Digital sciences*, potom *Padlet*, *Linoit*, *e-Book*, *Edmodo*, *Socrative*, *Plickers*, *Quizizz*, *Moodle*, *Diigo*, *Kami*, *Quizlet*, *Madmagz*, *StoryBoard*, *Kahoot*, *Jamboard*, *Google Classroom*, *Python*, *GeoGebra*, *Webex Platform*, *Assistive Digital Tools*, *Mercury Reader*, *Genealy*, *Metaverse*, *Micro:bit*, te

CAD CAM Fusion 360 [7, 8, 9, 10]. Ishodi učenja bili bi: (i) upoznati prednosti i nedostatke online sastanaka, radionica, predavanja na projektima za vrijeme pandemije; (ii) istraživanje mogućnosti primjene novih tehnologija u nastavi; (iii) prepoznavanje korištenja programa *Android Studio*, *Fusion 360*, *Payton* grafike, *Microbit* uređaja u nastavi srednjih strukovnih škola; (iv) upoznavanje s novim tehnologijama; (v) planiranje i primjena *e-learning* sadržaja i programa u nastavi i svakodnevnom životu; (vi) raspravljanje o važnosti izrade web stranica projekta za svrhu promocije projektnih ishoda [6].

Pitanje koje se postavlja dijelom je etičke, dijelom odgojno-obrazovne naravi, a usmjereno je na to stvara li zapravo suvremena tehnologija buduće *fah idiote*; znaju li današnji srednjoškolci pronaći u tiskanoj enciklopediji podatak o Romy Schneider, te o toj poznatoj glumici napisati vrlo kratak esej isključivo uz pomoć tiskanih izdanja i starih knjiga (ali i uz pomoć vlastite kreativnosti, pismenosti i općeg snalaženja „bez varanja"), ili će u najnoviji *ChatGPT* [11] upisati tekst „Možete li mi hitno napisati esej o Romy Schneider", te će aplikacija potvrdno odgovoriti. Ovo više nije „pravo", već retoričko pitanje [11].

Na pravo pitanje odgovor znamo. Učenik bez problema može prepisati i tiskani sadržaj, te ga prezentirati kao vlastiti tekst, no dojam je da se danas uz pomoć digitalne tehnologije vrlo često ide „linijom manjeg otpora", te da nam mnogi digitalni alati ne samo olakšavaju posao, već nas čine *intelektualno pasivnijima*. Može li sve to za nekoliko godina, desetljeća ili stoljeća dovesti do jedne nadrealne, distopijske situacije u kojoj će strojevi preuzeti, osim računalnih i mehaničkih, i fenomenalna stanja svijesti, te preuzeti kontrolu nad ljudima koji bi u tako zamišljenom udesu bili toliko pasivizirani da ne bi bili svjesni što ih je snašlo? [11].

IV. ZAKLJUČK: DIGITALNA BUDUĆNOST OBRAZOVANJA – TKO TU KOGA PODUČAVA?

Određene etičke, odgojne i obrazovne dvojbe oko uporabe umjetne inteligencije i digitalnih platformi za učenje i poučavanje promišljene su i u vrhu Europske unije. Prije nekog vremena, u listopadu 2022., Europska komisija je objavila dokument pod nazivom „Etičke smjernice za upotrebu umjetne inteligencije i podataka u poučavanju i učenju namijenjene nastavnom osoblju" [12]. Dokument donosi smjernice na koji se način zapravo umjetna inteligencija može koristiti u školama kao pomoć u poučavanju i učenju, te dakako kao potpora u obavljanju administrativnih zadaća u obrazovnom okruženju. No, ove „Etičke smjernice" daju niz upozoravajućih primjera na koje se mora obratiti pozornost, te se u njima „savjetuje kako upotrebljavati tehnologije za prilagodbu poučavanja sposobnostima svakog učenika ili kako osmisliti individualizirani pristup za učenike s posebnim potrebama, [ali su i] dobro ishodište za raspravu o tome koje nove kompetencije nastavno osoblje mora imati za etičku primjenu umjetne inteligencije i podataka…" [12, str. 6/7].

Ovo je tek jedan od dokumenata Europske unije o umjetnoj inteligenciji, obrazovanju i digitalnim tehnologijama. Smatramo kako nam tek slijede desetljeća u kojima ćemo intenzivnije promišljati ne samo etičke dvojbe oko uporabe *pametnih strojeva*, već i koncepte u kojima ćemo povezivati filozofiju uma (svjesna stanja) s umjetnom inteligencijom i digitalnim obrazovnim alatima. Korak dalje je računalna mogućnost transcendiranja, odnosno sposobnost stroja da zahvaljujući najsloženijim neuronskim mrežama *doživi* nešto što bismo mogli nazvati *mističkim iskustvom*. Bojimo se što bi na ove riječi rekao jedan Martin Heidegger koji je bio oštar kritičar *tehniciziranja* svijeta, ili pak Herbert Marcuse koji doduše jest bio pobornik tehnološkog napretka, no koji je istodobno upozoravao na nemogućnost potpunog sjedinjenja čovjeka i tehnologije. S time ćemo završiti naš rad, te vjerojatno otvoriti mnoga druga pitanja [13].

REFERENCE

[1] Pećnjak, D. "Uvod u filozofiju uma." Zagreb: Hrvatski studiji Sveučilišta u Zagrebu, 2013.

[2] Jackson, F. "Epiphenomenal qualia." The Philosophical Quarterly, Vol. 32, No. 127 (Apr., 1982), pp. 127-136.

[3] Jurjako, M. i Malatesti, L. "Filozofija uma: suvremene rasprave o odnosu uma i tijela." Rijeka: Sveučilište u Rijeci, Filozofski fakultet, 2022.

[4] Chalmers, D. "The character of consciousness." New York: Oxford University Press, 2010.

[5] Saha, D. "A Brief Introduction to Artificial Intelligence. What is AI and how is it going to shape the future." Toronto: Ryerson University. Dostupno na https://www.torontomu.ca/content/dam/sciencerendezvous/SR2021/A_Brief_Introduction_To_AI.pdf (6.1.2023.)

[6] "Online i digitalni alati za korištenje u obrazovanju." Zagreb: Agencija za mobilnost i programe EU. Dostupno na https://www.europskesnagesolidarnosti.hr/cms_files/2021/02/1612770339_online-i-digitalni-alati-u-neformalnom-obrazovanju.pdf (7.1.2023.)

[7] Leung, H.K.N. "Evaluating the effectiveness of electronic learning." Computer Science Education, 13 (2), 2003, pp. 123-136-

[8] The 35 best Web 2.0. Classroom tools chosen by you [online]. Edudem. Dostupno na http://edudemic.com/2010/07/the-35-best-web-2-0- classroom-tools-chosen-by-you/ (9.1.2023.)

[9] Elliott, M. i McGreal, R. "Tehnologije online učenja (e-learning)." Edupoint – časopis o primjeni informacijskih tehnologija u obrazovanju 37 (V). 2005. Dostupno na http://edupoint.carnet.hr/casopis/37/clanci/4.html (10.1.2023.)

[10] Maravić, J. "Cjeloživotno učenje." Edupoint – časopis o primjeni informacijskih tehnologija u obrazovanju 17 (III), 2003. Dostupno na http://www.carnet.hr/casopis/17/clanci/5 (10.1.2023.)

[11] ChatGPT: Optimizing Language Models for Dialogue. OpenAI. Dostupno na https://openai.com/blog/chatgpt/ (11.1.2023.)

[12] "Etičke smjernice za upotrebu umjetne inteligencije i podataka u poučavanju i učenju namijenjene nastavnom osoblju." Europska komisija. Ured za publikacije Europske unije. Dostupno na https://op.europa.eu/hr/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1 (14.1.2023.)

[13] Višić, M. "Ogled o tehnologiji i etici." Studia lexicographica, vol. 11, br. 21, 2017, pp. 75-86.

eTwinning – što kažu učenici?

Iva Naranđa
II. osnovna škola Čakovec, Čakovec, Hrvatska
iva.naranda@skole.hr

Sažetak

U ovom radu opisuje se istraživanje o stavovima učenika o međunarodnim eTwinning projektima koje je provedeno u prosincu 2022. godine među učenicima šestog razreda II. osnovne škole Čakovec. U istraživanju je sudjelovalo pedeset i devet učenika. Istraživanjem se željelo saznati kakve stavove imaju učenici o eTwinningu općenito i što učenici misle o potrebi provođenja međunarodnih eTwinning projekata u osnovnoj školi. Odgovori učenika prikupili su se pomoću anonimnog online upitnika za učenike šestog razreda koji su prethodne dvije godine sudjelovali u međunarodnim eTwinning projektima. U radu se prikazuju rezultati istraživanja i zaključci. Rezultati provedenog istraživanja pokazali su pozitivan stav učenika šestog razreda prema eTwinning projektima.

Ključne riječi – eTwinning projekt; istraživanje; stavovi učenika; osnovna škola; informatika; nastava

I. UVOD

eTwinning, zajednica za škole u Europi, između ostaloga omogućuje učenicima rad na virtualnim suradničkim projektima u međunarodnom okruženju. II. osnovna škola Čakovec aktivna je članica eTwinning zajednice od 2011. godine, nagrađena je brojnim Europskim oznakama kvalitete za eTwinning projekte, a od 2018. godine nosi Oznaku eTwinning škole.

eTwinning projekti već godinama su dio nastave Informatike u II. osnovnoj školi Čakovec, a posljednjih godina provode se i u sklopu izvannastavne aktivnosti. Provedeni projekti najvećim dijelom obuhvaćali su razne teme iz područja informatike i održivog razvoja. Suradnja s školama iz drugih država imala je za cilj unaprjeđenje nastave uz provođenje novih ideja kako bi učenicima nastava bila zanimljivo i zabavno iskustvo. Svaki eTwinning projekt uključivao je ex-post evaluaciju, u najvećem dijelu projekata provedenu putem online upitnika.

Odgojno-obrazovni djelatnici su na raznim edukacijama često pozivani na uključivanje u eTwinning zajednicu [1] i provođenje eTwinning projekata sa svojim učenicima. Sudjelovanje u eTwinning zajednici doprinosi razvoju škole i unaprjeđenju nastave zahvaljujući suradnji i razmjeni iskustva i primjera dobre prakse s školama iz drugih država. U sklopu diseminacijskih aktivnosti prikazuju se i dijele iskustva u provođenju eTwinning projekata, predstavljaju se rezultati evaluacija pojedinih eTwinning projekata i dijele primjeri dobre prakse.

Gledano s pozicije učitelja može se reći da eTwinning učenicima pruža izvrsne mogućnosti. Kako bi se dobio uvid u to kako eTwinning projekte doživljavaju učenici, upravo oni zbog kojih se ovakvi projekti pripremaju i provode, pojavila se ideja o istraživanju stavova učenika o eTwinning projektima.

Budući da se eTwinning projekti mogu provoditi u različitim predmetima i kontekstima učenja i poučavanja, ovo istraživanje, iako temeljeno na eTwinning projektima provedenima u sklopu nastave Informatike, relevantno je i primjenjivo u odgojno-obrazovnim ustanovama, ponajprije osnovnim školama i može biti korisno svim učiteljima, stručnim suradnicima i ravnateljima.

Sadašnji učenici šestog razreda svoj su prvi susret s međunarodnim eTwinning projektima imali školske godine 2020./2021. na izbornoj nastavi Informatike, kada su bili četvrti razred. Tada su sudjelovali u eTwinning projektu Teachers4Europe: Democratic Values in the Era of Artificial Intelligence [2]. Iduće školske godine kao učenici petog razreda sudjelovali su u svojem drugom eTwinning projektu pod nazivom Let's Explore the World of Technology [3]. Oba projekta obuhvaćala su više suradničkih aktivnosti koje su stavljale učenike u središte nastavnog procesa, a kroz inovativne metode učenja i poučavanja omogućavale aktivno učenje radom na konkretnim problemima iz svakodnevnog života. Evaluacije oba projekta pokazale su pozitivne rezultate koji su objavljeni na Twinspaceima projekata. eTwinning projekt Teachers4Europe: Democratic Values in the Era of Artificial Intelligence opisan je u Newsletteru 11 eTwinning grupe „Visibility of eTwinning Projects“ kao primjer dobre prakse [4].

Nakon određenog vremenskog odmaka željelo se dobiti povratnu informaciju od učenika koji već imaju iskustvo u eTwinning projektima, saznati kakve stavove imaju učenici o eTwinningu općenito i što uopće misle o potrebi provođenja ovakvih projekata u osnovnoj školi, a sve s ciljem planiranja i provedbe eventualnih budućih eTwinning projekata. Kako bi se došlo do odgovora, osmišljeno je i provedeno online istraživanje na razini škole koje je opisano u nastavku.

II. PROVEDENO ISTRAŽIVANJE I REZULTATI

A. Opis istraživanja

Izbor teme istraživanja potekao je iz potreba prakse i osobnih afiniteta istraživača, s ciljem utvrđivanja stavova i želja učenika u pogledu budućih eTwinning projekata.

Metodologija istraživanja odabrana je s obzirom na mogućnosti provođenja bez materijalnih troškova i omogućavanja anonimnosti učenicima kako bi se dobili iskreni odgovori. Svrha ovog istraživanja bila je postizanje novih spoznaja koje služe planiranju i ostvarivanju budućih eTwinning projekata. Za ovo empirijsko istraživanje podaci su prikupljeni iz neposrednog odgojno-obrazovnog rada u osnovnoj školi. S obzirom na vremensku odrednicu, ovo istraživanje predstavlja istraživanje sadašnjosti odgoja i obrazovanja, a s obzirom na trajanje istraživanja predstavlja transverzalno istraživanje jer istražuje presjek pojave u određenom trenutku, prosinac 2022. godine. Za istraživanje su odabrani učenici šestog razreda iz dva razloga. Prvi razlog je taj što su učenici šestog razreda već sudjelovali u dva eTwinning projekta, a drugi razlog je što se, ovisno o rezultatima istraživanja, za ovu generaciju učenika može pripremiti i provesti još neki eTwinning projekt u idućoj školskoj godini. Uzorak predstavljaju učenici šestog razreda koji su se odazvali pozivu na ispunjavanje online upitnika. Za prikupljanje podataka odabran je anketni upitnik kreiran kao Google obrazac.

B. Rezultati istraživanja

Rezultati istraživanja prikazani su zbirno. Od ukupno 63 učenika šestog razreda II. osnovne škole Čakovec, upitnik je ispunilo njih 59. Uzevši u obzir da nekoliko učenika u prethodne dvije godine nije sudjelovalo u navedenim eTwinning projektima jer u četvrtom razredu nisu bili upisani na izbornu nastavu informatike ili su tek od ove školske godine učenici naše škole, odaziv je potpun.

Na kraju upitnika učenici su imali mogućnost napisati vlastiti komentar što je iskoristilo 19 učenika. Od dobivenih komentara njih 18 bilo je izuzetno pozitivno usmjereno, dok je samo jedan učenik napisao da mu eTwinning projekti nisu bili zanimljivi. Mnogi komentari učenika vrlo su slični, a ovdje se navode neki: „Sve je bilo primjerene težine, sve je bilo zanimljivo. Veselim se novom eTwinning projektu jer ću ponovno nešto naučiti.“, „Meni je ovo iskustvo bilo jako zanimljivo i htjela bih i dalje sudjelovati u eTwinning projektu.“, „eTwinning projekt mi je zanimljiv jer često naučim nove stvari, uglavnom mi je uvijek zabavno.“, „eTwinning projekti su me inspirirali u kreativnosti i bili su jako zanimljivi.“, „Vrlo rado bih htjela sudjelovati u još puno eTwinning projekata, zato što su vrlo zanimljivi, poučni, kreativni i jako je zanimljivo što zapravo sudjelujemo sa drugim učenicima iz Europe.“, „Htio bih da ima više micro:bit-a. Najviše su mi se sviđali micro:bit-ovi i Scratch.“, „Volim i želim sudjelovati u slijedećim eTwinning projektima!“.

eTwinning projekti povezuju se s komunikacijom i suradnjom, zabavom i kreativnošću, stjecanjem dodatnog znanja i interakcijom s učenicima iz drugih država te se ovim istraživanjem željelo dobiti i uvid u to što za učenike znači eTwinning projekt vezano za navedene aspekte. 12 učenika (20,3%) odgovorilo je da za njih eTwinning znači sve od navedenog, a 20 (33,9%) učenika odgovorilo je da za njih eTwinning projekt znači zabavu i kreativnost. (Slika 1)



Slika 1. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec

C. Analiza rezultata

Iako su zabava i kreativnost važne komponente eTwinning projekata, trebalo bi poraditi na tome da učenici u većoj mjeri prepoznaju i druge dobre strane sudjelovanja u eTwinning projektima, kao što su komunikacija i suradnja s vršnjacima iz drugih država te stjecanje dodatnih znanja koja im omogućuje upravo rad u sigurnom međunarodnom okruženju.

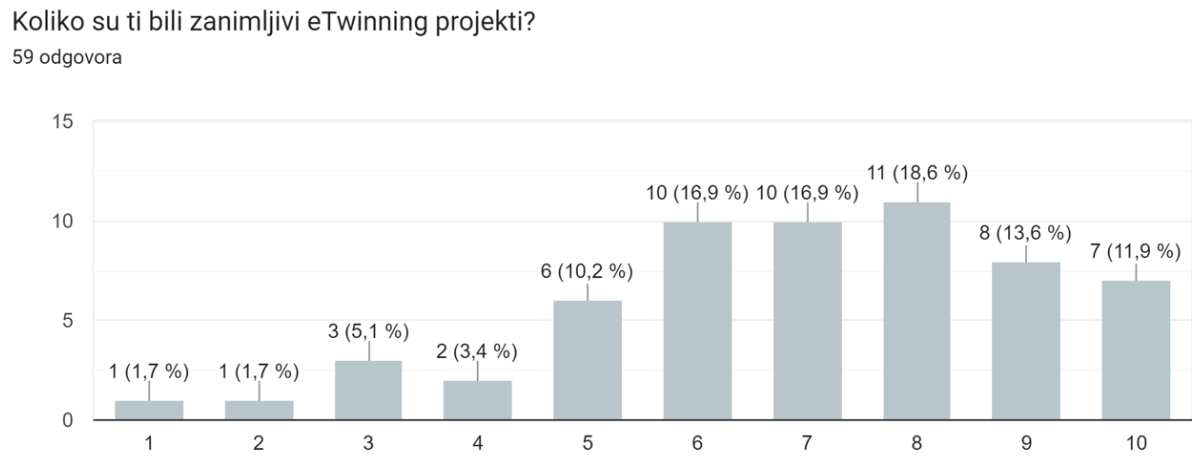
Prema odgovorima dobivenim na pitanje oblikovano kao linearno mjerilo na kojem su učenici mogli ocijeniti zanimljivost eTwinning projekata na ljestvici od 1 do 10 na kojoj odgovor 1 znači „nimalo zanimljivi“, a odgovor 10 „izuzetno zanimljivi“, vidi se da je 46 učenika (78%) odabralo gornji dio ljestvice, ocjenu od 6 do 10. Budući da su provedeni eTwinning projekti značajnom broju učenika bili zanimljivi, planiranje budućih eTwinning projekata može se smatrati dobrom i opravdanom odlukom. (Slika 2)

38 učenika (65%) složilo se s tvrdnjom da eTwinning projekti omogućuju proširivanje znanja iz više područja. Ovaj odgovor upućuje na to da je značajan broj učenika u eTwinning projektima prepoznao mogućnost stjecanja novih znanja iz više područja, a ne isključivo zabavu ili komunikaciju s vršnjacima iz drugih država.

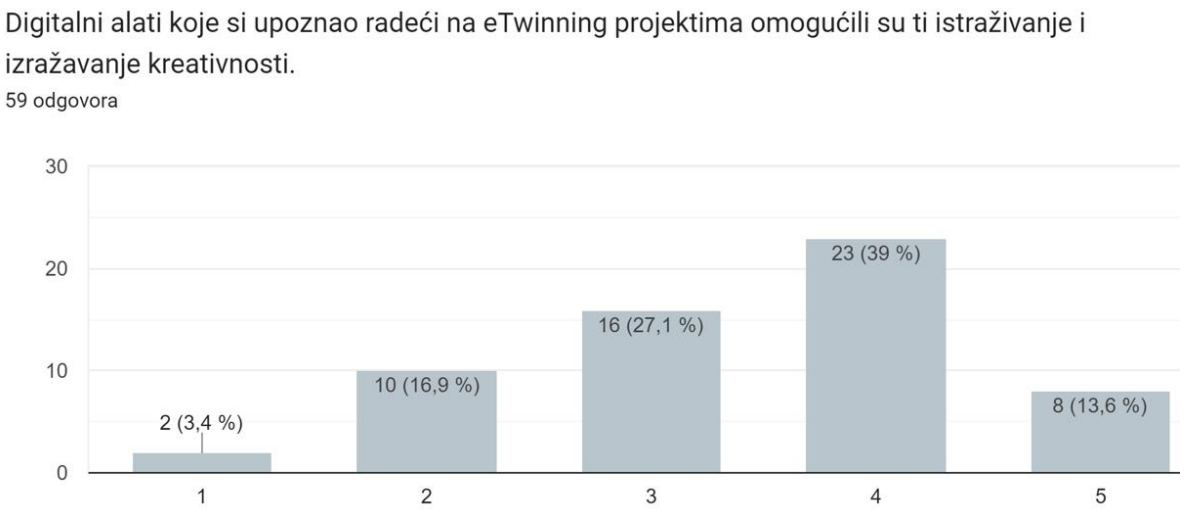
Samo šest učenika (10%) odgovorilo je da se ne slaže s tvrdnjom da su eTwinning projekti pomogli u stjecanju znanja i razvijanju navika primjerenog, odgovornog, sigurnog i svrhovitog korištenja interneta i digitalnih tehnologija.

Prema odgovorima dobivenima na pitanje jesu li digitalni alati koje su upoznali radeći na eTwinning projektima omogućili istraživanje i izražavanje kreativnosti, vidi se da je 31 učenik (52,6%) odabrao ocjenu 4 ili 5. Pitanje je oblikovano kao linearno mjerilo na kojem su učenici mogli odgovoriti na ljestvici od 1 do 5 na kojoj odgovor 1 znači „ne, nimalo“, a odgovor 5 „da, uvelike“. (Slika 3)

Deset učenika (17%) smatra da eTwinning projekti nisu obogatili nastavu informatike. Samo pet učenika (8,5%) odgovorilo je da ne bi rado sudjelovalo u novim eTwinning projektima. (Slika 4)

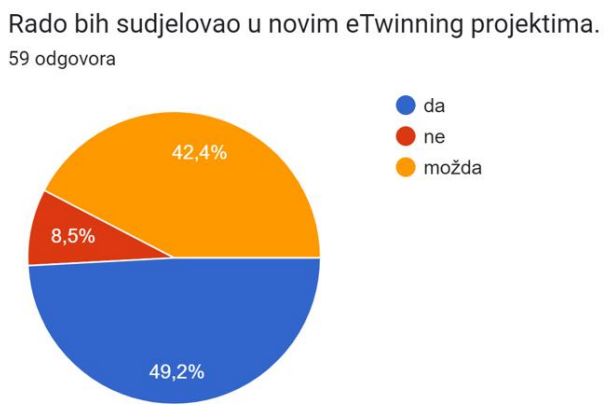


Slika 2. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec (1 = nimalo zanimljivi, 10 = izuzetno zanimljivi)



Slika 3. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec (1 = ne, nimalo; 5 = da, uvelike)

Samo jedan učenik (1,7%) odgovorio je da se eTwinning projekte uopće ne bi trebalo provoditi, dok je 35 učenika (59,3%) odgovorilo da bi eTwinning projekte trebalo provoditi čak svake školske godine u sklopu nastave. (Slika 5)



Slika 4. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec

42 učenika (71,2%) odgovorilo je da bi u eTwinning projektima trebali sudjelovati svi učenici u sklopu nastave. Ovi odgovori ukazuju na to da značajan broj učenika eTwinning projekte dobro prihvaća te smatra važnim, korisnim i potrebnim. (Slika 6)



Slika 5. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec



Slika 6. Prikaz odgovora učenika šestog razreda II. OŠ Čakovec

III. ZAKLJUČAK

Iz dobivenih odgovora učenika i analizom rezultata ovog istraživanja može se zaključiti da učenici šestog razreda II. osnovne škole Čakovec imaju pozitivne stavove o provođenju međunarodnih eTwinning projekata u osnovnoj školi, te su pozivi školama na uključivanje u eTwinning zajednicu i rad na virtualnim projektima dobili potvrdu i od samih učenika.

Za najveći broj učenika eTwinning projekti su zanimljivi i učenici smatraju da im ovakvi projekti omogućuju proširivanje znanja iz više područja. Velika većina učenika rado bi sudjelovala u budućim eTwinning projektima, što su osim odgovaranjem na postavljeno pitanje dodatno potvrdili u napisanim komentarima. Čak više od 70% učenika smatra da bi u eTwinning projektima trebali sudjelovati svi učenici u sklopu nastave, dok samo jedan učenik smatra da se eTwinning projekte uopće ne bi trebalo provoditi. Za 20 učenika (33,9%) eTwinning projekti znače zabavu i kreativnost, dok je 12 učenika (20,3%) u eTwinning projektima prepoznalo sve ponuđene mogućnosti kao što su komunikacija, suradnja, zabava, kreativnost, stjecanje dodatnog znanja i interakcija s učenicima iz drugih država.

Sve ovo potvrđuje da je uključivanje učenika u međunarodne eTwinning projekte dobar izbor te opravdava planiranje i provođenje budućih eTwinning projekata, naročito onih integriranih u nastavu određenog predmeta.

REFERENCE

[1] European School Education Platform, eTwinning https://school-education.ec.europa.eu/hr/etwinning (19.2.2023.)

[2] I. Naranđa, eTwinning projekt Teachers4Europe: Democratic Values in the Era of Artificial Intelligence. http://os-druga-ck.skole.hr/etwinning?news_hk=5630&news_id=1930&mshow=1149#mod_news (19.2.2023.)

[3] I. Naranđa, eTwinning projekt Let's Explore the World of Technology. http://os-druga-ck.skole.hr/etwinning?news_hk=5630&news_id=2039&mshow=1149#mod_news (19.2.2023.)

[4] G. Lascaris, S. Skiada, Teaching Democratic Values in the Era of Artificial Intelligence, Visibility of eTwinning Projects Group Newsletter 11, 2021, pp. 134–139 https://online.pubhtml5.com/xzfd/isvk/#p=134

DC-CPS

International Conference on

DISTRIBUTED COMPUTING AND CYBER-PHYSICAL SYSTEMS

Steering Committee

Chairs: Vlado Sruk, University of Zagreb, Croatia
Dejan Škvorc, University of Zagreb, Croatia

Program Committee

Members: Željko Hocenski, University of Osijek, Croatia
Leonardo Jelenković, University of Zagreb, Croatia
Hrvoje Mlinarić, University of Zagreb, Croatia
Jadranko F. Novak, Nomen, Rijeka, Croatia
Vlado Sruk, University of Zagreb, Croatia
Dejan Škvorc, University of Zagreb, Croatia

PAPERS

Comparing the Use of Custom-built and Commercial Off-the-shelf Data Gathering Devices in IoT Systems

Mika Saari*, Janne Harjamäki*, Mikko Nurminen*, Petri Rantanen*
* Tampere University / Faculty of Information Technology and Communication Sciences, Pori, Finland
mika.saari@tuni.fi

***Abstract*—IoT refers to a wide range of devices that have Internet access, and collect and transmit information. Data have to be gathered from the most diverse environments. Custom data gathering devices are often necessary in situations where the data being collected are unique or need to be collected in a certain way. This is because commercial off-the-shelf (COTS) products may not have the necessary capabilities or features to meet the specific requirements of the data collection. This study focuses on COTS and self-built data gathering devices by comparing two data gathering prototype systems: The first designed and implemented from scratch and the second built with COTS components. Both systems are compared in terms of various features such as cost-effectiveness, time saving, and the usefulness of the output. The end goal of the study is to provide a proposal for when it is appropriate to use COTS components and when it is necessary to design and build devices yourself.**

Keywords*—IoT, Internet of Things, Off-the-shelf, Data Gathering

I. INTRODUCTION

There are already plenty of commercial Internet of Things (IoT) products for data collection on the market. This has also raised the issue of whether it is worth building sensor packages yourself if commercial products are available. This paper looks at how ready-to-use components have been utilized in IoT research.

This research was part of a research project, which focused on energy saving in Finnish apartments. The goal of project was to improve the energy efficiency of housing with the help of data collection. As a result, the objective was to produce information and tools for this purpose. The research project implemented prototype systems for data collection and analysis. This paper focuses on three different data collection device constructions: a prototype system based on self-made components and a prototype system based on commercial off-the-shelf components and their combination. Both systems partly use the same software components, so the information they provide to the user is very similar. We tried to keep the data storage and visualization applications unchanged; these are presented in [1] and [2].

This research does not aim to find out whether a self-built device is better than a commercial one, but to determine what kind of advantages are achieved by using commercial products as part of the system. Therefore, the research question is formulated as follows:

RQ: Is it reasonable to use off-the-shelf devices instead of self-made sensor nodes in IoT data gathering?

The research question will be resolved by presenting the general architecture of the data collecting system. The most important components, their purpose, and mode of operation are described in the model. In addition, the collected environmental data and how the collected data are presented are described.

For the actual comparison of the commercial and non-commercial systems, two implemented real prototype systems are used. The main parts and structure of the systems are explained. In addition, the data they collect (input), and the outgoing data (output) are described.

The systems are compared using the following criteria: cost-effectiveness, time saving, and the usefulness of the output. Also, additional features are considered. Finally, the conclusions are presented.

The structure of this paper is as follows: In Section II, we present the background including the related research. In Section III, we present the abstract data gathering model and three pilot cases of real world data gathering prototype system implementations. In the discussion section, Section IV, we expand on the findings of the comparison of the pilot cases and summarize the study in Section VI.

II. BACKGROUND

A survey [3] conducted in 2002 collected the eight basic design factors of a sensor network. For the present study, the most notable of these factors are scalability, production costs, hardware constraints, and transmission media. Design factors can be used as a guide when developing data gathering prototype systems such as those presented in this study.

Common properties of IoT systems include component-based design. From the software engineering area, one approach is Component-Based Software Engineering (CBSE), which contains features needed in IoT systems, i.e., composability, deployability, comprehensive documentation, independence, and standardization [4]. We have researched the usage of CBSE in relation to IoT in an earlier study [5].

Several studies compare COTS components with self-made components. One study [6] uses scalability, cost

efficiency, and performance as criteria when comparing smart home sensor devices with the developed low-power, fuzzy-based control method. Another study [7] also focuses on low-cost and low-power underwater prototype device development. Further, our earlier study [8] showed that the use of COTS devices is quite extensive.

This study is based on research conducted during the KIEMI project. The goal of the KIEMI project was to save energy, and we worked towards this goal by developing and constructing data gathering IoT sensor systems. During the project a total of 23 different types of pilot cases were carried out related to the energy efficiency and condition measurement of properties. The pilot cases (in chronological order: #03, #08, and #19) mentioned in this study were made in the KIEMI project. The overall results of the project are reported in [9].

III. Data gathering IoT prototype systems

The off-the-shelf hardware components can be presented as follows: For the lowest level we chose resistors, microprocessors, and other components, and built the circuit board ourselves. Also, the software was self-built. For the mid level we chose off-the-shelf devices such as Raspberry Pi or Arduino, and selected the necessary sensors and made the necessary connections between them. The software was partially ready-to-use, for example: Raspbian OS. Self-made software is also needed for connecting the selected hardware so that it works together. The highest level is to buy reasonable priced devices with reasonable software and use them.

The abstract architecture of the data gathering prototype system consists of three main parts, which are presented in Figure 1.

1) Data gathering device, dedicated to gathering data, which could be simple numerical data ("SN_n nodes" in Figure 1) or Big data chunks such as image data (marked Master & Sensor node(s)" in the figure).
2) Communication (data arrows in Figure 1) is chosen for data transfer from device to database. The appropriate communication channel should be selected according to the type of data.
3) Database ("DB" in Figure 1) is responsible for storing and presenting the collected data. The analyzing of the data could be also performed here.

These are the main parts of the systems; next we present three different implementations: The first one is mid level, and uses self-made hardware and software. The second one is at the highest level and uses more COTS components. The third one is a hybrid, where COTS sensors were used with a self-made master node. All three use case descriptions contain reflections about their advantages and disadvantages. The systems are considered using the following criteria: cost-effectiveness, time saving, and the usefulness of the output. Additional qualifications characteristics such as CBSE criteria are also considered.

A. Prototype with self-made sensor nodes

In pilot case #03 (Figure 2), a prototype system was constructed to measure and gather environmental data such as temperature, humidity, pressure (BME680 sensor), carbon monoxide (eCO2-sensor), and air quality (SGP30) from residential apartments with two sensor nodes. The sensor nodes sent data to the master node, which sent the data to the database. This system had no complex visualization; the data were presented on the worksheet in numerical form. The sensor nodes were based on the Sodaq ExpLoRer[1] to which the sensors were connected. Sodaq has the capability to use LoRaWan[2] radio technology. The master node was a Raspberry Pi with a LoRa shield and was connected to the Ethernet network using a 3G/4G base station. With this prototype, the commercial LoRaWAN provider was used as a test case, and therefore Sodaq did not need a master node, because all the data went through the commercial network. Sodaq and Raspberry Pi needed software development for data gathering, caching, and transferring, but the data structure was fully self-made.

The prototype system worked as planned and collected data at the customer's premises for a month. The main finding from this pilot case was quite high CO2 values when there were people in the room. More specific details about the results of this prototype system can be found in [10].

Advantages: Complete customizability in both areas, hardware and software, is the most important advantage. If the developers have the ability to design, configure, and build the prototype, the end result is a complete data gathering prototype system. The ability to choose the best components also affects the result.

Disadvantages: Time-consuming development process where all steps (design, configuration, and construction) have to be done in more depth than with COTS devices.

B. Prototype with commercial sensor nodes

Pilot case #19 (Figure 3), a data gathering prototype system consisting of COTS components, was designed to gather the temperature and humidity data. The sensor nodes (RS_n in the figure) were RuuviTags [3]. The master node was a Ruuvi Gateway [4], which received the data from RuuviTags and sent the data to the Ruuvi Cloud using a WiFi or Ethernet network. The Ruuvi Gateway made it possible to use commercial data storage and visualization, but it also provided an interface to fetch raw data, if needed. In the pilot case, raw data import as well as self-made data storage and visualization were used. The primary goal of the pilot case was to analyze and visualize thermal comfort [11] (Figure 4).

[1] https://support.sodaq.com/Boards/ExpLoRer/
[2] https://lora-alliance.org/about-lorawan/
[3] https://ruuvi.com/ruuvitag/
[4] https://ruuvi.com/fi/gateway/

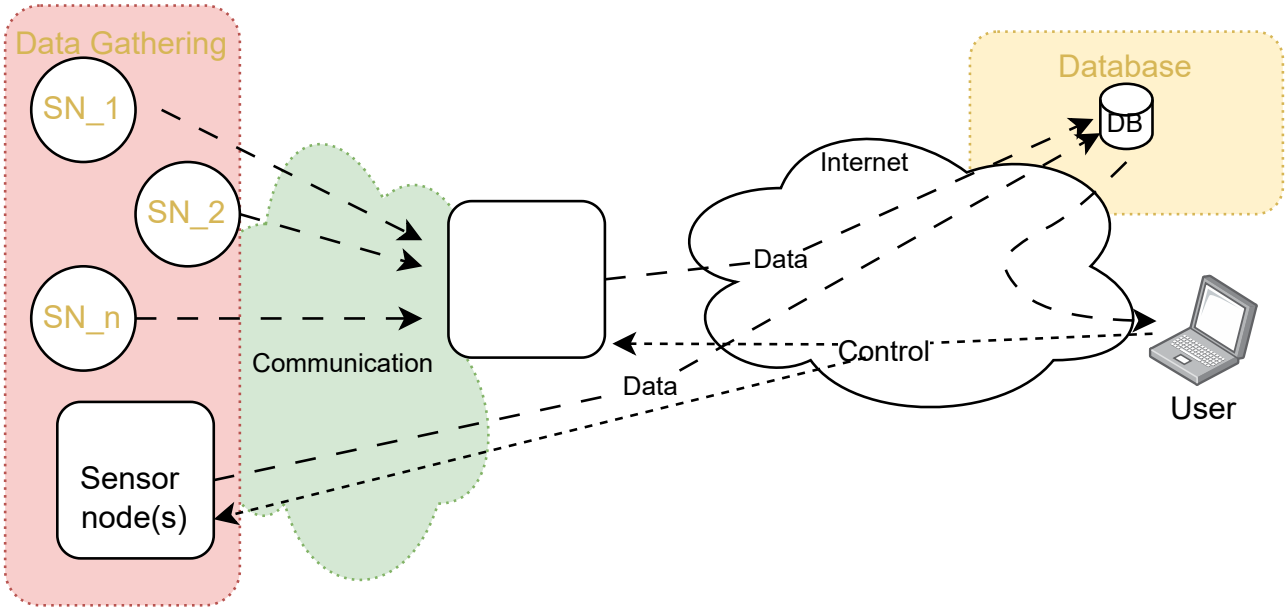


Fig. 1: Abstract data gathering system architecture.

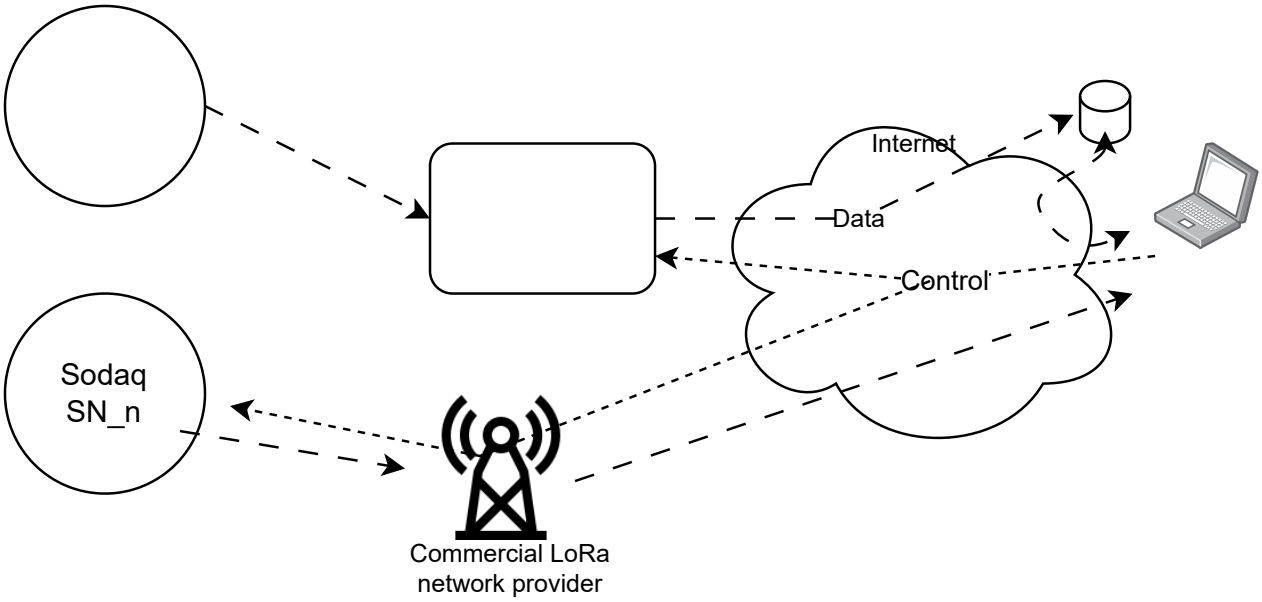


Fig. 2: Mid level data gathering system architecture of partially self-made components.

Advantages: The development process of the system is quite simple, when the connected devices work together. The software of the selected devices was ready-to-use, so there was no need for complicated configuration procedures. Mainly, these aspects were due to one commercial ecosystem, consisting of RuuviTag sensor nodes and a Ruuvi Gateway. The sensor nodes have the capability of using self-made firmware software as well as the Ruuvi Gateway.

Disadvantages: If the selected devices target ordinary consumers, the configuration possibilities are most probably limited, i.e., vendor lock. Also, even though the devices used have several possible features available, the documentation was limited. Ecosystem thinking might also limit the use of different manufacturers' devices together.

C. Hybrid prototype system

Pilot case #08 was a hybrid prototype with 12 RuuviTag sensors. The master node for RuuviTags was a self-configured Raspberry Pi. The purpose of the prototype was similar to pilot case #19, i.e., to collect humidity, pressure, and temperature sensor values. These data were cached on the Rasperry Pi database. Pilot case #08 was an early phase KIEMI pilot and therefore the main goal was to test RuuviTag sensors and their features.

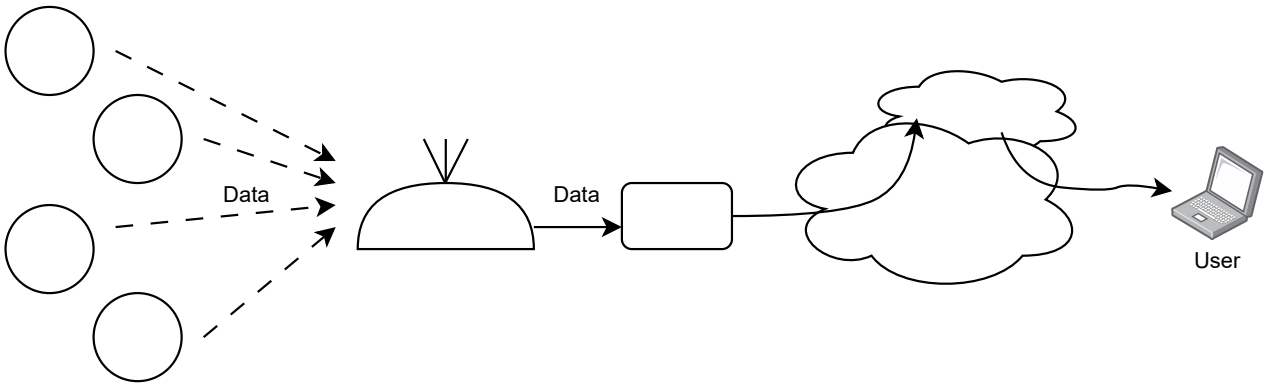


Fig. 3: Data gathering system architecture of COTS devices.

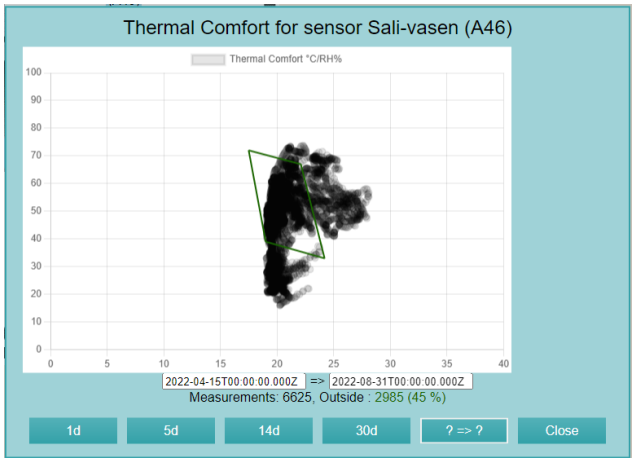


Fig. 4: Thermal comfort can be estimated using time, humidity, and temperature. The results are from pilot case #19. The green parallelogram marks "good" values.

Advantages: Sensor devices are supported for use in a specific interface definition. The sensor node manufacturer supports the open source development of the master node software. Usability of COTS devices depends on whether they are do-it-yourself or consumer products.

Disadvantages: Interoperability issues with specific hardware or the documentation of interfaces is limited. Ecosystem thinking limits the usage of different manufacturers' devices together.

IV. Discussion

The following research question was set for this study: Is it reasonable to use off-the-shelf devices instead of self-made sensor nodes in IoT data gathering?

The study shows that all three approaches have advantages. With self-made systems, there are more opportunities to modify the features. The hardware can be selected as required and software can be produced with the features that are needed. The disadvantages of self-made systems are their time-consuming development, as both software and hardware require definition before production can start. The use of COTS devices is less time-consuming, but their modifiability is limited. Also, the configuration can be challenging depending on the state of the product - whether it is a consumer product or a do-it-yourself product. A hybrid system where developers can select the best devices using selected criteria was a good combination of cost-effectiveness, time- saving, and the usefulness of the output.

However, it cannot be said which is the best because its application depends on the particular use case. We have earlier research results [9] where several prototype systems were built for data gathering. That project and study showed that reusability increased when similar prototype systems were built. Reusability refers to software and hardware, but as shown by the research, software is developed more from self-made with several iteration rounds. On the other hand, hardware more commonly uses COTS devices because their development is more time-consuming. In addition, the KIEMI project made extensive efforts to collect data without device limitations. Further, one reason for custom software development was that the research group had more knowledge about software development than hardware development.

The study showed the need for further research on how to find and select the necessary components, both software and hardware. A possible approach here could be connecting and networking with other partners in industry.

V. Conclusions

This study presented three real world pilot cases where custom data had been collected. This was done by using fully or partly COTS products that had the necessary capabilities or features to meet the specific requirements of the data collection. This study focused on COTS and self-built data gathering devices by comparing three data gathering prototype systems: The first was designed and implemented from scratch. The second was built with COTS components, and the last one was a combination of both. These systems were compared in terms of various features such as cost-effectiveness, time saving, and the usefulness of the output. The study presents examples of the usage of the different devices, showing the advantages and disadvantages at different points of usage. Finally, the

conclusion is that there are situations where the use of a ready-made commercial component makes more sense than using a self-built one.

ACKNOWLEDGMENTS

This work was funded by the European Regional Development Fund and the Regional Council of Satakunta.

REFERENCES

[1] M. Nurminen, A. Lindstedt, M. Saari, and P. Rantanen, "The Requirements and Challenges of Visualizing Building Data," in *2021 44th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2021.

[2] M. Nurminen, M. Saari, and P. Rantanen, "DataSites: a simple solution for providing building data to client devices," in *2021 44th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2021.

[3] I. Akyildiz, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[4] I. Sommerville, *Software engineering (10th edition)*, 10th ed. Pearson Education Limited, 2016.

[5] M. Saari, M. Nurminen, and P. Rantanen, "Survey of Component-Based Software Engineering within IoT Development," in *2022 45th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2022.

[6] G. Pau and V. M. Salerno, "Wireless sensor networks for smart homes: A fuzzy-based solution for an energy-effective duty cycle," *Electronics*, vol. 8, no. 2, 2019. [Online]. Available: https://www.mdpi.com/2079-9292/8/2/131

[7] B.-C. Renner, J. Heitmann, and F. Steinmetz, "Ahoi: Inexpensive, low-power communication and localization for underwater sensor networks and uauvs," *ACM Trans. Sen. Netw.*, vol. 16, no. 2, jan 2020.

[8] M. Saari, A. M. bin Baharudin, and S. Hyrynsalmi, "Survey of prototyping solutions utilizing Raspberry Pi," in *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2017, pp. 991–994. [Online]. Available: http://ieeexplore.ieee.org/document/7973568/

[9] J. Harjamäki, M. Saari, M. Nurminen, P. Rantanen, J. Soini, and D. Hästbacka, "Lessons learned from collaborative prototype development between university and enterprises," in *Information Modelling and Knowledge Bases n*, ser. Frontiers in Artificial Intelligence, M. Tropmann-Frick, H. Jaakkola, and N. Yoshida, Eds. IOS Press, 2023, accepted for publication.

[10] *Towards the utilization of cost-effective off-the-shelf devices for achieving energy savings in existing buildings*. IEEE, 2020.

[11] N. Langner and M. Illner, "Thermische behaglichkeit nach din en iso 7730 – ein ansatz zur vereinfachten datenaufnahme und berechnung für die bewertung von bürogebäuden," *Bauphysik*, vol. 37, no. 3, pp. 159–168, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/bapi.201510019

Challenges of Combining Open and Commercial Data Sources in Visitor Mobility Estimations

J. Grönman*, P. Rantanen*, P. Sillberg*, T. Pohjola†, T. Jönkkäri*
* Tampere University/Faculty of Information Technology and Communication Sciences, Pori, Finland
† Turku School of Economics at the University of Turku, Pori Unit, Pori, Finland
petri.rantanen@tuni.fi

***Abstract*—This paper explores the opportunities and challenges of open data. Open and free interfaces can provide useful data, but sometimes there are challenges with data sufficiency and quality. Open data can be enriched by acquiring commercial data, and by merging multiple data sources, a more comprehensive understanding of a situation can be obtained. The focus of this paper is on nature parks, but the technical solutions can be applied to other venues. In the case presented by this paper, the parks are equipped with on-location visitor counters, and with the help of commercial data provided by a mobile network operator, and open sources, such as the weather and road traffic services, we are attempting to gain a better view of visitor movements. The motivation for this research originated from the cities located in the Satakunta region of Finland. For the purpose of developing recreational and business opportunities in the region, the cities require a better understanding of movement of people - which locations and services are popular, and which locations need improvement. To help achieve this goal, this paper presents a technical solution of how analysis and visualization of combined data can be used to estimate visitor amounts within a geographical area.**

***Keywords*—open data, commercial data, visualization, mobile network operators**

I. INTRODUCTION

This paper presents the measures taken in the SmartMove project, which wanted to promote with concrete solutions the tourism and service companies of the Satakunta area benefiting from the rapidly growing local tourist flows due Covid-19 pandemic in cultural and leisure destinations and routes. By improving intelligent visitor information management and communication, it would be possible to predict and promote the healthy and safe use of services in almost real time, avoiding traffic jams and unnecessary human contact. At the same time, it increases the sense of security to move outside in cultural and leisure destinations. The social isolation associated with Covid-19 pandemic has significantly increased people's desire to spend time outdoors and in nearby destinations, including hiking and moving around in nature and cultural destinations [1].

The popularity of outdoor activities has been reported in many countries. According to [2], "there has been so much traffic in the most popular camping destinations that the roadsides have been blocked with parking and the police have been needed to investigate the matter. The shelters have also been crowded", which can feel particularly frustrating in exceptional circumstances and reduce the feeling of safety.

The tourism profiles of both Finland and Satakunta rely heavily on nature and outdoor environments. Satakunta's natural and cultural destinations have significant potential in the new world situation to develop safe and attractive digital solutions to promote smart and sustainable tourism destinations.

Smart solutions and various data sources will offer completely new development opportunities for travel destinations and attractions, as well as for businesses. For example, from a hiker's point of view, the problem is that when planning your own route, there is no information whether there are already many people on a certain route, rest stop or shelter, or whether the locations are free of other users. Real-time service data, or even indicative open statistical data on the number of users, visit times or congestion peaks of various outdoor attractions are not available in the Satakunta area or nationwide in Finland. Automatic information about non-congested, low-congestion or very congested destinations gives not only the user, but also tourism and leisure developers real-time information on, for example, how to schedule service capacity, recommend routes and carry out distribution in order to minimize congestion.

Nature destinations and rural areas play a crucial role in promoting the well-being of citizens and providing recreational opportunities [3], [4] such as hiking, biking, and admiring scenic landscapes also for the visitors. The advantages of nature destinations in enhancing physical and mental wellness are becoming more evident [5], [6], with studies showing how exposure to diverse ecosystems positively affects our microbiota and immune system [7]. In Finland, nature tourism is centred around forests, marshland, Lapland fells, lakes, and the Baltic Sea. Finland's "Everyman's rights" allow individuals to walk, ski, cycle, and camp in all forests and engage in activities such as swimming, boating, and fishing in the waters. During the winter, individuals can walk, ski, and skate on ice.

Access to place-based rich data and smart services can further increase the appreciation of nature-based destinations among new audiences and deepen the exposure to various cultural experiences. This information and immersive contents can lead to the discovery of new locations for relaxation, recreation, education, immune system boost,

and more. In addition to nature itself forest areas possess cultural and industrial heritage that appeals to tourists and locals alike [8]. Service providers can use data to strengthen the bond to these places [9] and make the experiences even more meaningful [10], [11]. Nature and rural areas, as well as more urban environments, can be considered as hybrid spaces in which tourists and other stakeholders create value together in the development of smart tourism [12]–[14] by blending physical space with real-time interactions [15]–[17]. The use of big data sheds light to visitor behaviour and additionally IoT sensor data can further improve the long-term sustainability of nature and rural tourism operations and stakeholder networks [11], [18], [19] by managing tourism flows in sensitive areas [12].

In this paper, a use case is presented, which combines open data sources and commercial mobile operator data for calculating and estimating visitor amounts in the nature parks of the Satakunta region of Finland. In the following sections, the available data - both from open and commercial sources - is described and illustrations on data visualization are presented.

Therefore, in an attempt to unravel the mysteries of smart tourism, the rest of this paper has been divided into the following three sections: II describes the use case selected for this paper, the data sources utilized, and methods for visualizing the data; III explores the challenges faced and possibilities discovered during the research; and finally, IV gives a summary of the main findings.

II. USE CASE: ESTIMATING CONSUMER MOBILITY IN SATAKUNTA REGION

The use case described in this section presents a map-based visualization tool for smart destination development. In essence, the goal is to collect historical data from multiple sources and combine the data to produce an overview of past visitor movements in an area, and estimates for possible future visitations. The primary users of the tool are the entities responsible for the development of the Satakunta region (cities, municipalities) and the companies and organizations responsible for managing nature parks. The tool could also be utilized by users, who are interested in discovering the various nature locations in the region, or companies interested in establishing new businesses. The focus is on nature parks, but the technical solutions could be applied to other venues as well.

A. High-level Service Architecture

The overall technical view of the system and its connections are presented in Fig. 1. The tree main parts are the two clouds (representing Internet) and a folder representing the containerized `SmartMove Service`. Open source implementation (i.e., source codes) of the service is available at [20].

Internet (i.e., clouds) allows the service to retrieve data from 3rd party APIs (such as Wikipedia). It also provides

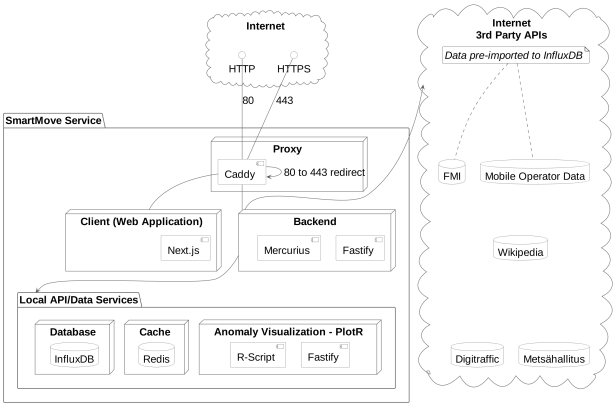


Fig. 1: High-level diagram of the service architecture.

the entry-point for the users of the service to access the system, mainly over HTTP/S connections.

Inside the service package, each node (represented by rectangular cuboids) is a Docker container. The main nodes are `Proxy`, `Client`, and `Backend`. The supporting nodes (Database, Cache, and Anomaly Visualization) are separately packaged in `Local API/Data Services`. It means that any of those are designed to be run independently if needed to.

The first node, `Proxy`, is implemented by *Caddy* server. In essence, it simply delivers the requests appropriately between `Client` and `Backend`. The second node is `Client` that provides the visual user interface for the consumers of the service. It is implemented using Node.js runtime and *Next.js* framework. The `Backend` also utilizes Node.js runtime. It is built on *Fastify*, a low overhead web framework, and *Mercurius*, which is a GraphQL adapter for *Fastify*. Both the `Client` and `Backend` communicate with each other and 3rd party services by using the GraphQL query language. Additionally, the *Mercurius* component in `Backend` enables caching by utilizing *Redis* in-memory cache, making subsequent unmodified queries near instantaneous.

The service does not have a dedicated database as it just shows the information in read-only mode. It does, however, utilize databases such as InfluxDB through other supporting services. For example, we implemented a tool for anomaly visualization as a micro service, `PlotR`. It is useful for inspecting temporal data on side-by-side plotted diagrams. `PlotR` is described more specifically in section II-D.

B. Utilization of Open Data

The main user interface of the service can be seen in Fig. 2. The example shows all data contents selected with the exception of mobile network operator data, which is illustrated separately in Fig. 4, and explained in more detail in Subsection II-C. By clicking the sliders, the users can choose which point of interest (POI) - i.e. the data sources - are visible on the map, and by clicking the POIs, the users can see the data of the clicked POI. The

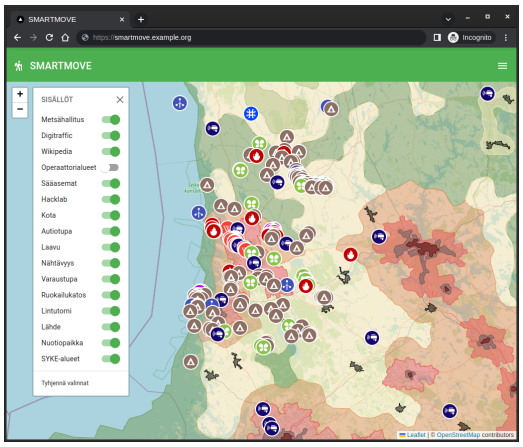


Fig. 2: Example visualization of open data sources.

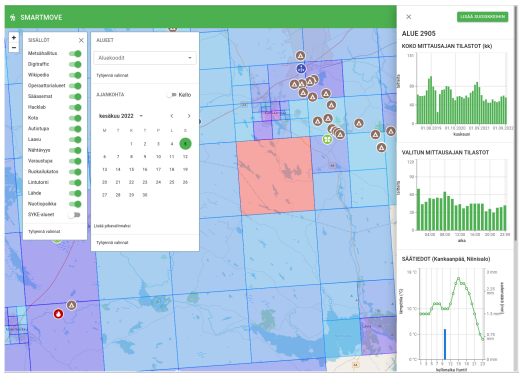


Fig. 3: Example visualization of mobile network operator data with weather and visitor statistics.

majority of the POIs provide general information of the area (places to see, the Wikipedia article of the location, etc.). In the context of this paper, the most interesting ones are Digitraffic and Sääasemat (weather stations of the Finnish Meteorological Institute, FMI [21]). Digitraffic [22] is a service managed by the Finnish Transport Infrastructure Agency, which provides open data on the road, rail and waterway systems of Finland. In this case, Digitraffic's APIs could be used, for example, to retrieve traffic volumes (the amount of cars) passing on a particular highway.

By clicking the map, the user can also see data of a specific area (grid block), as illustrated in Fig. 3. In the example figure the selected area is highlighted in red, and the desired data can be seen in a scrollable list on the right side of the user interface. In this case, the data shows (from top to bottom): the monthly visitor amounts for the whole measurement period; the hourly visitor amounts for the date (June 4th, 2022) selected from the calendar widget; and the weather history for the selected date. The visitor amounts are calculated based on the mobile operator data (described in more detail in the following subsection II-C, and the weather data is retrieved from FMI's open data service. Thus, the user can, with a quick glance, see an overview of visitors on a particular time in relation to the weather.

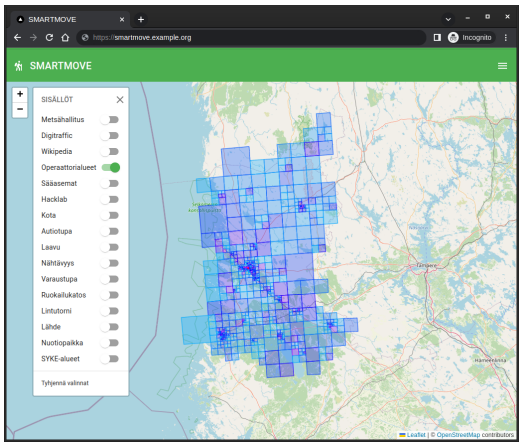


Fig. 4: Example visualization of mobile network operator data.

Of course, especially weather can be a complicated relation to figure out by simply looking at the figures as "poor weather" is relative to the location and to the time of the year (season). Cold weather in the winter near to a ski center can be desirable, but not near a beach in the summer. Similarly, in the case of time, certain locations receive higher visitor amounts during holiday seasons. Unfortunately, at least in Finland, there does not seem to be a programmably accessible API, which could be used to retrieve public holidays and holiday seasons for a particular year, requiring either hard coding the dates or scraping the information from web pages - both being somewhat unreliable or high maintenance solutions in the long run.

C. Utilization of Commercial Data – Operator Data

Today, almost everyone has a cell phone. Thus, at least in principle, the information about the location of the phones, could be a great asset in improving visitor counting accuracy. In Finland, a couple of operators offer commercial services, which provide anonymized data about mobile devices. In general, the service consists of a dataset, which contains per area device amounts with accuracy that vary between mobile network operators. The accuracy can differ both in time and area sizes. In our dataset, the device amounts were reported hourly, but for other operators, the amounts could be reported on a daily basis (once every 24 hours). The dataset contains the amounts for devices that have subscription for the operator in question, and an estimate of other operators' devices in the area. The total amounts are calculated based on a sample collected by the operator, and then extrapolated based on the amount of population living in the area in question. Foreign subscriptions (e.g. SIM cards) are excluded from the dataset. Regardless, at least in principle, ordering the service from one operator should be enough for getting a fairly good estimate on the device amounts in the area, though we did not perform any multiple operator analysis to find out how accurate the promised estimations were.

An example grid of Satakunta region, as provided for our dataset, can be seen in Fig. 4. The region - which has the combined land and sea area of approximately the size of the countries of Jamaica or Kosovo - is divided into 670 smaller areas (blocks), with sizes between 500x500 meters and 20x20 kilometers. In the grid visualization, blocks with lighter (blue) color represent lower device amounts and darker (blue) color represent higher device amounts on the selected time period (in the example, the entire data gathering period of three years). It is important to note that the blocks do not necessarily match a single specific cell tower, but are aggregate products of any number of cell towers, with the actual number not provided in the dataset for privacy and security reasons. The size of a block also depends on the operator's estimate of total device amount in the area. The higher the amount of devices, the smaller (more accurate) the block is defined to be. Thus, it can be assumed that the larger blocks have a lower amount of devices, in general. Furthermore, in our dataset, any geographical area has less than five (5) devices, the data has been omitted for privacy reason by the mobile network operator, though in our use case, the region has sufficiently high population density to generate data on every block. The Finnish mobile network operators also provide data for other countries they operate in (e.g. Estonia, Sweden and Norway, depending on the operator), and these countries have their own (legal) regulations, which may affect the available data accuracy and quality. It is also possible to get estimates for the device movements (e.g. how many devices per day moved from one block to another), but this information turned out to be too coarse for reliable estimation of visitor amounts.

D. Visualization Methods

There was a growing interest toward having more visualizations and automatically computed analysis of the available data. The chosen specimens for this experiment were the weather data and the mobile operator data. Being able to do experiments and seeing (i.e., visualizing) how the data behaved was a major driving factor. The second step was to apply k-means clustering to find out whether there were any automatically discoverable classifications (i.e., cluster centroids) in the data. Unfortunately, due to limited resources only the first part, visualization, was successfully implemented in the user interface. However, the k-means clustering could be implemented and executed as one-shot offline analysis of the complete data set.

As was mentioned above in Fig. 1, the PlotR is implemented as a micro service. It uses Node.js runtime together with Fastify for providing a web service. The web service is a simple facade to pass through the URL query parameters to R-Script which makes computations that generate a downloadable diagram image.

Fig. 5 shows a zoomed screen capture of the bottom part of the scrollable list which was initially seen in Fig. 4. The user interface in Fig. 5 has been manually translated to English for the convenience of the reader.

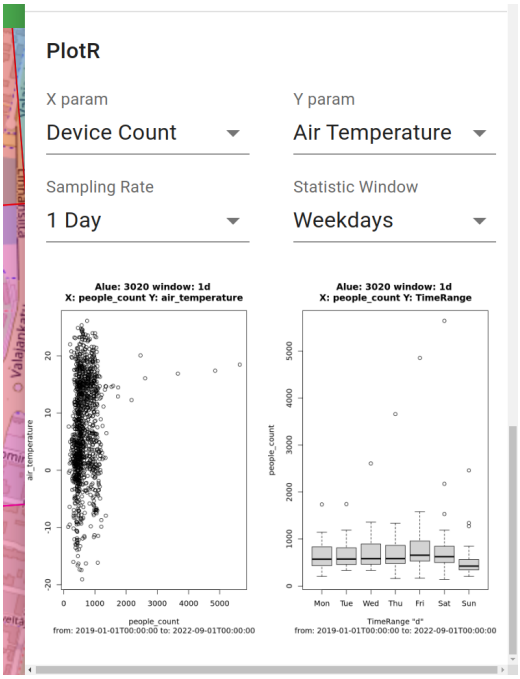


Fig. 5: Anomaly visualization with PlotR micro service.

The user interface consists of two parts: 1) the variable configuration, and 2) the diagram image, which is further divided to the XY scatter plot on left, and the Xt box plot on right.

Each grid block can be selected to display the corresponding detailed plot visualization. It is possible for the user to customize each plot using four additional parameters: `X param`, `Y param`, `Sampling Rate`, and `Statistic Window`. The default settings for these are chosen to be `Device Count`, `Air Temperature`, `1 Day`, and `Weekdays`, respectively.

The sample rate effects both diagrams by dictating the amount of data points in the dataset. It has seven options ranging from one hour (native sampling in this case) up to one year (aggregated samples). Parameters for X and Y axles include 13 similar options. Axis X affects both diagrams while the Y axis affects only the scatter plot on the left. The statistic window dictates how the dataset will be grouped on the box plot on the right. It has six options for grouping up the data; from hourly to yearly groups.

According to the shown example diagram in Fig. 5 it seems to indicate a few features: 1) the bulk of the device count is averaged to approximately 600 devices (both plots); 2) each weekday except for Sundays are quite similar (box plot); 3) there are a few outlier data points that show very high device count over several days, perhaps there was a popular public event (both plots); 4) higher air temperature might indicate slightly higher device count (scatter plot). However, one figure only can not explain everything with certainty, thus further experimentation with the data is needed to reveal more insight into the observed data.

III. DISCUSSION

The primary motivation for considering the utilization of mobile network operator data was the known limitations of the available open data sources. Open data, in general, is provided as general datasets, and it is often the task of the user (or developer) to figure out the potential use case scenarios. This also means, that the data may not be quite what is required. In our use case, the data was not actually measured at the nature parks. The historical weather data gathering stations around the Satakunta region are relatively close to the nature parks, but this may not be the case in larger and more rural nature areas. Similarly, the traffic data - at least in Finland - is collected mainly on the main highways. Thus, if the highway does not directly go to the nature park, but, for example, travels between two major cities with only a smaller side road to the park (or multiple intersecting roads to other destinations), it becomes exceedingly difficult to figure out the actual destinations of the travellers.

Metsähallitus (the Finnish Forest Administration), which is an agency that amongst other duties manages the nature parks in Finland, also provided us with visitor statistics for the parks, but after discussions with the representatives of the agency, it was discovered that the data are mostly estimations, and not actual measured on-location. In Finland, as well as in many other countries, nature parks are free and open for everyone, which is a good thing for visitors, but bad for counting exact visitor counts as there are no ticket sales to count. It is often possible to enter and exit the parks in multiple locations and there are several nature trails one can follow, and only a limited amount of these routes have physical visitor counter devices. The challenges of providing electricity and (winter) weather proof devices to nature areas often technically limit the counters to simple infrared counters, as opposed to, for example, camera-based methods. Also, it is worth noting that even though the use of the visitor statistics API was provided for us free of charge, the API is not as such open to everyone, and the access must be negotiated with Metsähallitus. With other venues, especially with those operated by companies, it is quite common, that visitor (or client) amounts are not openly available.

Mobile operator data is available in all locations, that have cell phone coverage. In some countries large nature areas may have coverage issues, but in Finland, and in Europe in general, there is coverage even in remote areas. As we are only interested in counting the devices that have at some point connected to the cell towers, technically even very poor network coverage is acceptable. Thus, the limiting factors come from the grid and block sizes defined by the operator (accuracy), from the legal regulations of the target area or country, and finally, from the cost of the data. In Finland, the data services provided by the operators are primarily aimed at companies, and to the best of our knowledge the data cannot be acquired freely or openly. Thus, the overall pricing options are reasonable for larger companies, but the cost might be a bit steep for a one-person startup or for someone who simply wants to try-out what could be done with the data. Especially, if the use case requires something special to negotiate, such as customizations to the provided grid or data accuracy.

For our use case of nature parks, the accuracy of the mobile device data was sufficient, but in high population density areas the smallest block size (500x500 meters) might be too large for some use cases. One challenging case could be to pinpoint specific shops in city centers. The grid also has a relatively high number of blocks, and in practice it might be necessary to focus on a smaller geographical area. It is also crucial to map any factors in the target area, which could affect the device amounts, often even if the data does not show any specific irregularities. For example, proximity of a large sports stadium, music festivals, or an arrival of a large cruise ship can temporarily cause huge increase in device amounts. The last two events are especially interesting as they bring in new people, and from the region's point-of-view it would be important to find out whether this "new people" simply stay "at the venue" or do they visit other businesses as well? Unfortunately, discovering these factors is not always trivial. In the Satakunta region, near the Jämijärvi nature park, there is a large military base, which (for obvious reasons) does not publish information about their daily activities. Nevertheless, these activities can affect the device data collected around the nature park's area.

Finally, as our mobile network operator data was collected between January 2019 and August 2022, it shows per-year (and per-season when comparing seasons of individual years with each other) variations in many locations, presumably caused by the COVID-19 pandemic travel restrictions. The parks are by nature large outdoor areas, and are thus somewhat excluded from the most serious effects of the pandemic restrictions. In any case, it would require a much longer data period, covering the pre-pandemic era, to estimate the effect of the pandemic on the visitor amounts. From a technical point-of-view, the data does not cause problems for analysing the visitor data during the pandemic, but it makes estimating future post-pandemic visitor amounts more-or-less impossible.

Furthermore, even though technically the device data analysis can be expanded to venues outside the nature park context, and to countries and areas outside of Finland, a large scale utilization would require further studies on the availability (and cost) of mobile network operator data on other counties. As a summary, the mobile operator data shows promise, but the applicability depends on the (commercial) data services available in the desired geographical target area.

IV. CONCLUSION

This paper presented a use case, in which a combination of open data sources and commercial mobile operator data was utilized to study visitor amounts in the nature parks of the Satakunta region of Finland. The available

data was described and the paper illustrated how the gathered datasets could be visualized. The existing open data sources were found to be somewhat limiting in terms of accuracy and availability. The mobile operator (device) data was sufficient for our use case, but certain limitations of the approach were also discussed in this paper. Overall, the mobile operator data showed promise, but the applicability depends on the (commercial) data services available in the desired geographical target area and venue.

ACKNOWLEDGMENT

This study has been funded by National funding instrument (AKKE) of Regional Council of Satakunta. The mobile operator data was provided by Telia Finland Oyj.

REFERENCES

[1] C. Alba, B. Pan, J. Yin, W.L. Rice, P. Mitra and M.S. Lin, "Covid-19's impact on visitation behavior to us national parks from communities of color: evidence from mobile phone data," *Scientific Reports*, vol. 12, 2022.

[2] Yle - The Finnish Broadcasting Company, "Korona voi levitä myös retkipaikoilla - Metsähallitus varoittaa ruuhkista ja ruohikkopaloista (in Finnish) [Corona can also spread at camp sites - The Finnish Forest Administration warns of congestion and grass fires (title translated into English)]," https://yle.fi/aihe/artikkeli/2020/03/26/korona-voi-levita-myos-retkipaikoilla-metsahallitus-varoittaa-ruuhkista-ja, 2020, accessed: 2023-01-17.

[3] S. Bell, L. Tyrväinen, T. Sievänen, U. Pröbstl and M. Simpson, "Outdoor recreation and nature tourism: A european perspective," *Living Reviews in Landscape Research*, vol. 1, no. 2, 2001.

[4] R. Puhakka, K. Pitkänen and P. Siikamäki, "On certain integrals of lipschitz-hankel type involving products of bessel functions," *The health and well-being impacts of protected areas in Finland*, vol. 25, no. 12, pp. 1830–1847, 2017.

[5] Y. Joye, R. Pals, L. Steg and B. Lewis-Evans, "Correction: New methods for assessing the fascinating nature of nature experiences," *PLoS ONE*, vol. 9, no. 1, 2014.

[6] T. Hartig, M. Mang and G.W. Evans, "Restorative effects of natural environment experiences," *Environment and Behavior*, vol. 23, no. 1, 1991.

[7] L. Ruokolainen, J. Lehtimäki, A. Karkman, T. Haahtela, L. von Hertzen and N. Fyhrquist, "Holistic view on health: two protective layers of biodiversity," *Annales Zoologici Fennici*, vol. 54, pp. 39–49, 2017.

[8] K. Kirillova, X. Lehto and L. Cai, "What triggers transformative tourism experiences?" *Tourism Recreation Research*, vol. 42, no. 4, pp. 498–511, 2017.

[9] L. Birnbaum, C. Wilhelm, T. Chilla and S. Kröner, "Place attachment and digitalisation in rural regions," *Journal of Rural Studies*, vol. 87, pp. 189–198, 2021.

[10] D. Buhalis and Y. Sinatra, "Real-time co-creation and nowness service: lessons from tourism and hospitality," *Journal of Travel and Tourism Marketing*, vol. 36, no. 5, pp. 563–582, 2019.

[11] T. Pohjola, A. Lemmetyinen, and D. Dimitrovski, *Value Co-creation in Dynamic Networks and E-Tourism*. Cham: Springer International Publishing, 2020, pp. 1–23. [Online]. Available: https://doi.org/10.1007/978-3-030-05324-6_92-1

[12] U. Gretzel, C. Koo, M. Sigala and Z. Xiang, "Smart tourism: Convergence of technologies, experiences and theories," *Electronic Markets*, vol. 25, pp. 175–177, 2015.

[13] D. Buhalis and A. Amaranggana, "Smart tourism destinations," in *Information and Communication Technologies in Tourism 2014*, Z. Xiang and I. Tussyadiah, Eds. Cham: Springer International Publishing, 2013, pp. 553–564.

[14] A. Caragliu, C. Del Bo and P. Nijkamp, "Smart cities in europe," *Journal of Urban Technology*, vol. 18, no. 2, pp. 65–82, 2011.

[15] D. Buhalis, "Technology in tourism - from information communication technologies to etourism and smart tourism towards ambient intelligence tourism: a perspective article," *Tourism Review*, vol. 75, no. 1, pp. 267–272, 2019.

[16] U. Gretzel, L. Zhong, C. Koo, A. Morrison and A. Morrison, "Application of smart tourism to cities," *International Journal of Tourism Cities*, vol. 2, no. 2, pp. 216–233, 2016.

[17] M.V. Ciasullo, O. Troisi and S. Cosimato, "How digital platforms can trigger cultural value co-creation? – a proposed model," *Journal of Service Science and Management*, vol. 11, pp. 161–181, 2018.

[18] C. Koo, L. Mendes-Filho and D. Buhalis, "Guest editorial," *Tourism Review*, vol. 74, no. 1, pp. 1–4, 2019.

[19] K. Boes, D. Buhalis and A. Inversini, "Smart tourism destinations: ecosystems for tourism destination competitiveness," *Tourism Destinations: Ecosystems for tourism destination competitiveness*, vol. 2, no. 2, pp. 108–124, 2016.

[20] DigiluontoSatakunta, "Smartmove OSS," https://github.com/DigiluontoSatakunta/smartmove-oss, 2023, accessed: 2023-01-17.

[21] Finnish Meteorological Institute, "The Finnish Meteorological Institute's open data," https://en.ilmatieteenlaitos.fi/open-data, 2023, accessed: 2023-01-17.

[22] Finnish Transport Infrastructure Agency, "Digitraffic," https://www.digitraffic.fi/en/, 2023, accessed: 2023-01-17.

An Overview of Diffusion Models for Text Generation

Helena Čeović*, Marin Šilić*, Goran Delač*, Klemo Vladimir*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10 000 Zagreb, Croatia
helena.ceovic@fer.hr, marin.silic@fer.hr, goran.delac@fer.hr, klemo.vladimir@fer.hr

***Abstract*—Given the great success that diffusion models have achieved in generating various types of continuous data, including image, video and audio, there has been a growing interest in the application of these models to text generation. However, the discrete nature of text presents a challenge for diffusion models initially designed for application in a continuous feature space. The two main lines of work that aim to bring together diffusion models and natural language processing are focused on either defining the diffusion process in continuous space by converting discrete tokens to embeddings or defining the diffusion process in discrete space. These recent works attempt to combine diffusion models with leading sequence-to-sequence generation Transformer architecture as well as with existing pre-trained language models. In this work, we give a detailed overview of the approaches developed to date. We present and analyze the benefits and limitations that each model introduces, along with how they compare to the autoregressive models that dominate this field.**

Keywords*—*diffusion models, denoising diffusion models, text generation, deep generative modeling, natural language processing

I. INTRODUCTION

Latest text-conditional image generators such as DALL-E 2 [1] and Imagen [2] have introduced unprecedented improvements in the field of image generation. They enable production of high-resolution outputs for any possible textual prompt. These modern frameworks are mostly based on diffusion models, a family of generative iterative models that achieve state-of-the-art sample quality for continuous data.

An immediate response to such successes was inevitable. Once the immense potential of diffusion models in generating images was displayed, the interest shifted further towards exploiting their benefits and extending their application to other challenging generative tasks. Naturally, this includes the common challenge of natural language processing (NLP) - text generation. Text generation spans over a large portion of popular NLP tasks including machine translation, building conversational systems, and abstractive text summarization. Text generation is a type of language modeling problem that aims to produce plausible and coherent natural language texts by automatically learning from data.

Autoregressive models currently dominate the field of language modeling. The biggest advance in their development was the Transformer architecture, presented in [3]. Imposing results are later achieved with GPT-3 [4] on a wide range of text generation tasks. However, despite their extensive capabilities, the fixed generation order limits the autoregressive (AR) models' flexibility in many controllable settings, specifically those that require both left and right context [5]. Instead of causal attention in AR models, diffusion models can leverage bidirectionality by predicting all tokens in a sequence at once and potentially lead to more coherent samples [6].

The main challenge that emerges in applying diffusion models to text generation is the discrete nature of textual data. The existing solutions are built and specialized in the continuous data domain which involves generating images, audio, and video. Multiple design modifications are required to make these models suitable for natural language modeling. Previous work in this field was mainly developed in one of the following two directions. The first involves defining text diffusion models in the discrete state space whereas the second focuses on continuous diffusion models. Discrete diffusion models corrupt sentences and refine them on the token level or switch from one discrete value to another. Some of these models attain strong results on several language modeling tasks, but they still fall behind AR models in terms of coherence due to the inability to model semantic correlation [6]. Diffusion models for continuous domain were first introduced in [5] where they explain the approach of embedding the discrete text into a continuous latent space. By preserving the continuity of the input, we keep important properties such as classifier-free guidance and the ability to represent uncertainty at individual token level. However, due to the different nature of data, modifications are necessary to apply existing diffusion models to textual data.

The paper is organized as follows. Sect. II introduces related surveys in the diffusion domain. Sect. III includes the formal and mathematical definition of diffusion models. The aforementioned two leading lines of work are discussed in detail in Sect. IV and Sect. V by presenting the development of diffusion models in discrete and continuous space, respectively. Conclusion and discussion of possible future research directions are given in Sect. VI.

II. RELATED WORK

Given the rapid increase in interest and development in the research of diffusion models, surveys have begun to emerge that aim to give an overview of the latest occurrences in the research field and state-of-the-art results

as well as point out current dominant applications along with the remaining challenges and limitations. The most comprehensive and extensive surveys to date are given by Cao et al. [7] and Yang et al. [8]. Both of these surveys mention natural language processing as one of the key areas for application and further work, but only briefly discuss the current works in the field.

Following the successes achieved with diffusion models in computer vision and the numerous papers focused on the task, surveys such as [9] and [10] attempt to give an overview and categorize the different approaches and design choices as well as to identify their contributions and limitations. More specialized surveys have begun to appear such as [11] where the focus lies on the medical domain by providing an extensive overview of diffusion models used for medical image analysis.

We believe this to be the first survey that specifically concentrates on diffusion models for text generation.

III. Diffusion Models

The idea of diffusion models as generative models was first introduced in [12]. They present a novel perspective derived from non-equilibrium statistical physics instead of standard variational Bayesian methods. The proposed algorithm consists of a forward inference diffusion process and a reverse generative process. The forward trajectory converts any complex data distribution $q(\boldsymbol{x}_0)$ into a simple, tractable distribution by gradually adding Gaussian noise to the data according to a variance schedule $\beta_1, ..., \beta_T$ [13]:

$$q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0) := \prod_{t=1}^{T} q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}), \tag{1}$$

with each transition $x_{t-1} \rightarrow x_t$ parametrized by:

$$q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) := \mathcal{N}(\boldsymbol{x}_t; \sqrt{1-\beta_t}\boldsymbol{x}_{t-1}, \beta_t \boldsymbol{I}). \tag{2}$$

The reverse process predicts the noise of current time step t and denoises to previous state $\boldsymbol{x}_{t-1}$. It is defined as a Markov chain with learned Gaussian transitions starting at $p(\boldsymbol{x}_T) = \mathcal{N}(\boldsymbol{x_T}; \mathbf{0}, \boldsymbol{I})$ [13]:

$$p(\boldsymbol{x}_{0:T}) := p(\boldsymbol{x_T}) \prod_{t=1}^{T} p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t). \tag{3}$$

Each denoising transition $x_{t-1} \rightarrow x_t$ is parametrized by the model:

$$p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) := \mathcal{N}(\boldsymbol{x}_{t-1}; \boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t), \Sigma_\theta(\boldsymbol{x}_t, t)). \tag{4}$$

The variational lower bound (VLB) is tractable and differentiable with respect to θ on $\log p_\theta(\boldsymbol{x}_0)$ [5]:

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}_0) = \mathop{\mathbb{E}}_{q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)} \Bigg[\log \frac{q(\boldsymbol{x}_T|\boldsymbol{x}_0)}{p_\theta(\boldsymbol{x}_T)} + \\ \sum_{t=2}^{T} \log \frac{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0, \boldsymbol{x}_t)}{p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)} - \log p_\theta(\boldsymbol{x}_0|\boldsymbol{x}_1) \Bigg]. \end{aligned} \tag{5}$$

Sohl-Dickstein et al. [12] also show that we can rewrite the training objective in terms of KL divergences and entropies [5]:

$$\begin{aligned} L = \mathop{\mathbb{E}}_{q(\boldsymbol{x}_0)} \Bigg[D_{KL}[q(\boldsymbol{x}_T|\boldsymbol{x}_0)||p(\boldsymbol{x}_T)] + \\ \sum_{t=2}^{T} \mathop{\mathbb{E}}_{q(\boldsymbol{x}_t|\boldsymbol{x}_0)} \Big[D_{KL}[q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0)||p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)] \Big] - \\ \mathop{\mathbb{E}}_{q(\boldsymbol{x}_1|\boldsymbol{x}_0)} [log p_\theta(\boldsymbol{x}_0|\boldsymbol{x}_1)] \Bigg]. \end{aligned} \tag{6}$$

However, Ho et al. [13] show that better results are obtained by simplifying the training objective. To avoid the potential instability, they suggest a simplified training objective that expands and reweights each KL-divergence term in Eq. 6 to obtain the sum of mean-squared errors between the ground truth and its estimates [5]:

$$\mathcal{L}(\boldsymbol{x}_0) = \sum_{t=1}^{T} \mathop{\mathbb{E}}_{q(\boldsymbol{x}_t|\boldsymbol{x}_0)} ||\mu_\theta(\boldsymbol{x}_t, t) - \widehat{\mu}(\boldsymbol{x}_t, \boldsymbol{x}_0)||^2, \tag{7}$$

where $\widehat{\mu}(\boldsymbol{x}_t, \boldsymbol{x}_0)$ is the mean of the posterior $q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0, \boldsymbol{x}_t)$ which is a closed form Gaussian, and $\mu_\theta(\boldsymbol{x}_t, t)$ is the predicted mean of $p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)$ computed by a neural network.

IV. Diffusion models in discrete space

Discrete diffusion models build on categorical distributions where each token has some probability to be corrupted [14]. Diffusion models with discrete state spaces were first considered in [12] and applied to binary 'heartbeat' data. They train a probabilistic model on simple binary sequences of length 20, where a 1 occurs in every fifth bin, and the remainder of the bins are 0. The proposed binary sequence learning via binomial diffusion process was nearly perfect.

Hoogeboom et al. [15] go on to introduce multinomial diffusion modeled directly on categorical random variables with transition matrices characterized by uniform probabilities [16]. The same form was described in [17] although with no empirical evaluations. Furthermore, Austin et al. [16] generalize the approach from [15] with a more structured categorical corruption process. They shape data generation and embed structure or domain knowledge into transition matrices used by the forward process without relaxing or embedding discrete data into continuous spaces. They explore several structured transition matrices in their experiments. One of them is a transition matrix with an absorbing state where each token either stays the same or transitions to a special mask token with some probability allowing for the corrupted tokens to be distinguished from the original ones. They explore using similarity in an embedding space to guide the forward process and construct a transition matrix that transitions more frequently between tokens that have similar embeddings while maintaining a uniform stationary distribution. This work shows the importance of transition matrix choice by outperforming various non-autoregressive baselines for text generation on

character-level text generation as well as scaling discrete diffusion models to large vocabularies and long sequence lengths. However, this D3PM class of models remains inferior in comparison to leading autoregressive models for text generation such as Transformer XL.

None of the aforementioned works have incorporated pre-trained language models in their proposed work on diffusion models for text generation. He et al. [18] bring to attention that pre-trained language models such as BERT [19] and BART [20] are pre-trained with denoising objectives and explores discrete diffusion models to integrate the pre-trained language model. They use the transition matrix with an absorbing state for BERT introduced in [16]. It is more challenging to control the degree of noise added in the discrete domain than it is in the continuous. It is important to take into consideration the linguistic differences among tokens in addition to avoiding generating the most frequently appearing tokens only to achieve a higher likelihood. They introduce a spindle schedule that generates noise for $\boldsymbol{x}_t$ conditioned on both $\boldsymbol{x}_{t-1}$ and $\boldsymbol{x}_0$. By experimenting with several methods of incorporating the time step into pre-trained language models, the time-agnostic decoding method examines the best results. Experiments are conducted on unconditional text generation where DiffusionBERT outperforms [5], a nominal work described in Sect. V, and additionally proves its incompatibility with BERT. Results prove the efficiency of the spindle noise schedule that is independent for each token and depends on the token's frequency in the dataset. DiffusionBERT achieves significant improvement in perplexity results, but still falls behind competitive autoregressive models.

Qian et al. [21] propose a non-Markovian diffusion process where modalities are gradually added as the diffusion step t decreases. Additionally, to achieve smoother modality learning, the generative process uses a non-autoregressive process with residual glancing sampling [22]. Experiments are conducted on tasks of machine translation, paraphrase generation and image caption where DiffGLAT, combined with Directed-Acyclic-Transformer [23], manages to outperform the autoregressive Transformer in both accuracy and efficiency on top of several other high-performing works outlined in Sect. V, including [14], [24], [25]. Their method significantly reduces the decoding iterations when compared to a strong prior discrete diffusion model known as SUNDAE [26]. In [26], Savinov et al. present an unrolled denoising technique where they unroll the chain by sampling from the transition distribution and feed samples back into the input during training.

A different angle is taken in [27] by introducing an edit-based generative model for text generation which is based on the diffusion denoising method. Both corruption and denoising processes are based on Levenshtein operations: insert, delete, replace, and keep. Diffuser demonstrates results comparable to AR models on machine translation, summarization, and style transfer tasks. They also show that Diffuser is complementary with token-level autoregressive methods outperforming autoregressive baselines.

V. Diffusion models in continuous space

The development of diffusion models for text generation has gradually shifted from models in discrete space presented in Sect. IV to continuous diffusion models which induce continuous latent representations. The idea is to preserve the advantages of high-quality continuous diffusion models while adjusting to a new data domain and thus get closer to AR models' results.

The first to pursue this type of architecture were Li et al. [5] by introducing multiple modifications to the standard diffusion model. They start by defining an embedding function that maps discrete text into a continuous space to enable direct application of a continuous diffusion model to discrete text. Embedding function $EMB(w_i)$ maps each word in sequence $\boldsymbol{w}$ of length n to: $EMB(\boldsymbol{x}) = [EMB(w_1), ..., EMB(w_n)] \in \mathbb{R}^{nd}$, where d is the number of dimensions. Furthermore, based on conducted experiments, they propose a new design of the training objective Eq. 5 that jointly learns the diffusion model's parameters and word embeddings. The novelties to the original diffusion models include an additional Markov transition between discrete words $\boldsymbol{w}$ and $\boldsymbol{x}_0$. In the forward process, this transition marks the embedding process, parametrized by $q_\phi(\boldsymbol{x}_0|\boldsymbol{w}) = \mathcal{N}(EMB(\boldsymbol{w}), \sigma_o I)$. A proper method is required for the inverse process of rounding a predicted continuous $\boldsymbol{x}_0$ back to discrete text. They accomplish rounding by choosing the most probable word for each position following argmax $p_\theta(\boldsymbol{w}|\boldsymbol{x}_0) = \prod_{i=1}^{n} p_\theta(w_i|x_i)$, where $p_\theta(w_i|x_i)$ is a softmax distribution. They reparametrize the training objective to ensure that the model predicts $\boldsymbol{x}_0$ in every term with $\boldsymbol{x}_0$ lying directly on a word embedding. Diffusion-LM performs control by running iterative gradient updates with fluency regularization and multiple gradient steps on the continuous latent variables which improve performance and fluency of the text along with speeding up decoding. Diffusion-LM demonstrates successful control for six control tasks, including text infilling where it achieves results competitive with a fine-tuned autoregressive model for this task [28]. However, the authors point out several weaknesses of the model, namely high perplexity, slower decoding and slower training convergence.

Following the same line of work that conducts diffusion directly in a continuous token embedding space, but focusing on a wider scope of application and more diverse textual data, Strudel et al. [6] introduce their Self-conditioned Embedding Diffusion (SED) model. Unlike Li et al. [5], they do not learn the embedding matrix $\boldsymbol{E}$ due to the discovered empirical instability, potential unigram entropy drops and limited applicability. Thus, the training objective depends only on trainable readout weights. Their model is based on the self-conditioning technique introduced by Chen et al. [29] which adapts $\tilde{\boldsymbol{x}}_0$ estimates by passing the estimate obtained at the previous sampling

step as the input to the denoising network instead of only $\boldsymbol{x}_t$. Dieleman et al. [30] point out the resemblance of the self-conditioning technique to the unrolled denoising strategy [26]. SED also allows conditional text generation by applying the span masking strategy for infilling tasks and classifier-free guidance [31] on text data to alleviate the need for a separately-trained guide model. Although AR baselines of similar capacities outperform SED, the model still performs strongly and comparably. Authors point out limitations which include limited model tuning, sampling efficiency, relying on another model for diffusion in a pre-trained embedding space as well as the lack of appropriate metrics and baselines for certain tasks.

Dieleman et al. [30] explore continuous diffusion for categorical data by following the formalism from [32] to describe the corruption and the reverse process using differential equations. The proposed framework, Continuous diffusion for categorical data (CDCD), is based on the diffusion framework in [33]. Similarly to prior works in this section, they embed the discrete input into a continuous space and directly apply continuous diffusion to the embeddings. The proposed CDCD framework replaces the standard score matching function with score interpolation for diffusion model training allowing the use of familiar categorical cross-entropy loss function for training. They go on to learn the embeddings and the diffusion model jointly to embed the discrete input into a continuous space, as seen in [5]. The final component that makes up the CDCD framework is an active learning strategy called time warping which automatically adapts the distribution of noise levels sampled during training to maximize efficiency. Experiments have shown that the CDCD model can produce compelling samples for prompt completion and infilling, but overall lower performance on machine translation with autoregressive models of the same size.

The focus in [14] lies in conducting a study of challenges connected to embedding discrete textual data and utilizing continuous diffusion models to generate it. Unlike images and audio that have a fixed data space during training, the embedding space is learnable for discrete textual data which may cause a collapse of the denoising loss function and bring instability to the training of the model. They find that the rounding loss used by Li et al. [5] and Gong et al. [24] is not sufficient for alleviating this issue. Secondly, because of the imbalanced frequency of tokens in textual datasets, the learning of token embeddings diverges. Considering the diverse embedding scale of different words, it is suboptimal to add the same amount of noise to different embeddings. Lastly, when denoising an embedding from noise sampled from the normal Gaussian prior, the generation process may be distracted by other embeddings that are near the noise. Their proposed model, named Differmer, solves the described challenges by implementing the following methods. First, they propose an anchor loss training objective that uses the model prediction $\hat{\boldsymbol{x}_0}$ as input instead of the noisy embedding $\boldsymbol{x}_0$ and successfully regularizes embeddings and prevents loss collapse. The solution to the imbalanced embedding scales is implemented in the form of a layer normalization module on the top of the embedding layer which guarantees the uniform scale of tokens. As for the potential distraction in the diverse process, it is eliminated by adding an increasing scale of noise at each step. Experiments conducted on the tasks of machine translation and text summarization show the Differmer outperforming concurrent diffusion-based models including those presented in [5] and [30] as well as achieving comparable results with the autoregressive Transformer model.

Gong et al. [24] go on to extend the framework from [5] to a more generalized sequence-to-sequence setting, a key NLP aspect that covers various downstream tasks. The proposed DiffuSeq model introduces partial noising in the forward process by imposing noise only on the target sequence, but not on the source sequence. The reverse process can therefore impose the input as the condition when denoising. Since no additional classifiers are required to control the denoising process, it is considered classifier-free. Experiments are conducted on four sequence-to-sequence tasks: open domain dialogue, question generation, text simplification and paraphrase. DiffuSeq achieves comparable or higher quality results than competitive AR, iterative non-autoregressive, and large-scale pre-trained models, specifically demonstrating its ability to generate diverse, and hence high-quality sequences.

Yuan et al. [25] continue the research on diffusion models for sequence-to-sequence text generation. Unlike the encoder-only Transformers with partial denoising used in [24], their SeqDiffuSeq model is based on an encoder-decoder Transformers architecture with self-conditioning [29] and a token-level adaptive noise schedule. The encoder-decoder architecture has computational advantages during inference because the input sequences require only one forward computation through the encoder during the whole reverse process. They improve sequence-to-sequence text generation by introducing a heuristic that the difficulty of predicting the sample should increase linearly with respect to time steps. Their suggested noise schedule is different from the previous proposals including those in DiffusionBERT [18] and CDCD [30]. SeqDiffuSeq achieves significant acceleration in inference speed compared to DiffuSeq. Seq2Seq proves its generation quality by surpassing both DiffuSeq and several autoregressive baseline models on text simplification and paraphrase. On question generation, it is comparable with DiffuSeq whereas for open domain dialogue and machine translation it demonstrates poorer results than the AR baseline.

GENIE [34] can be considered a semi-non-autoregressive model and the first large-scale pre-trained language model based on diffusion. A new approach is proposed that combines the diffusion model and Transformers. They present a pre-training task named continuous paragraph denoise which predicts the noise added to continuous paragraphs in the current time step based on the paragraph context information and

the noisy paragraph information. Experiments show the effectiveness of the large scale pre-training by achieving comparable results to a pre-trained AR model on task of text summarization.

Another line of work includes employing the continuous process on surrogate representations of discrete data such as analog bits [29] and simplex [35]. Chen et al. [29] design a simple and generic approach to enable continuous state diffusion models to generate discrete data. The idea lies in representing the discrete data as binary bits which are then modeled into analog bits by continuous diffusion models. Analog bits are real numbers that share the same bimodal values as binary bits that represent discrete data. This process requires no discrete space or re-formulation of the continuous diffusion process. For sampling, the process is the same as in continuous diffusion models with the additional step of applying a thresholding operation which decodes the generated analog bits that are then converted into original discrete variables. They introduce the self-conditioning technique previously described in this section and often used in the following work. Their model achieves comparable results to an autoregressive Transformer baseline on the task of image captioning.

The SSD-LM presented in [35] introduces a different approach that aims to bring together the advantages of autoregressive and diffusion models by proposing a semi-autoregressive solution. It maintains the ability to train and generate variable-length sequences in addition to allowing refinement within the token block, thus maintaining the advantages of both autoregressive and diffusion models. Unlike the representation in [29] which can lead to extremely long sequences, they keep a subword based vocabulary with each token represented as a sequence of manually defined logits. SSD-LM is comparable to or outperforms competitive autoregressive baselines such as GPT-2 both in quality and diversity on unconstrained text generation. Still, several limitations of the model in comparison to autoregressive models are pointed out including lower sample efficiency, slower decoding speed and inflexibility of the decoding schedule.

VI. Conclusion

This paper gives a comprehensive overview of research conducted in the field of text generation using diffusion models. Following the impressive achievements of diffusion models in image generation, the desire to extend their application to other domains has grown rapidly. However, language modeling itself represents a challenging task due to the complexity and diversity of the natural language. Furthermore, the high-performing image generation models were designed for continuous data space, requiring significant modification to generate discrete textual data. We presented the most notable advancements in their development and how they compare to the current state-of-the-art autoregressive models.

We considered works developed in both discrete and continuous data domains, ranging from those focused on controllable text generation to unconditional and conditional text generation as well as sequence-to-sequence tasks. Although many of them examine respectable results, they are still mostly limited to either a specific task domain or language structure in order to perform comparably to autoregressive models.

Diffusion models offer a more flexible approach by abandoning the sequential generation of autoregressive models and generating simultaneous outputs through iterative refinement. However, they struggle to catch up with AR models' inference speed and efficiency. There are still multiple open questions addressed in many of the considered works, such as defining the noise schedule and the embedding space.

Finally, this is an active area of research aspiring to reach the breakthrough seen in the field of image generation. It is yet to be seen whether such performance is feasible, especially considering the existing state-of-the-art autoregressive models that are already capable of generating high-quality diverse texts. Another issue in this field is the lack of appropriate metrics and baselines for certain tasks, such as infilling. The discussed works show promising directions for diffusion models in text generation that are yet to be extended to their full potential.

References

[1] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, and M. Chen, "Hierarchical text-conditional image generation with clip latents," 2022. [Online]. Available: https://arxiv.org/abs/2204.06125

[2] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. Denton, S. K. S. Ghasemipour, B. K. Ayan, S. S. Mahdavi, R. G. Lopes, T. Salimans, J. Ho, D. J. Fleet, and M. Norouzi, "Photorealistic text-to-image diffusion models with deep language understanding," 2022. [Online]. Available: https://arxiv.org/abs/2205.11487

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[4] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," 2020. [Online]. Available: https://arxiv.org/abs/2005.14165

[5] X. L. Li, J. Thickstun, I. Gulrajani, P. Liang, and T. B. Hashimoto, "Diffusion-lm improves controllable text generation," 2022. [Online]. Available: https://arxiv.org/abs/2205.14217

[6] R. Strudel, C. Tallec, F. Altché, Y. Du, Y. Ganin, A. Mensch, W. Grathwohl, N. Savinov, S. Dieleman, L. Sifre, and R. Leblond, "Self-conditioned embedding diffusion for text generation," 2022. [Online]. Available: https://arxiv.org/abs/2211.04236

[7] H. Cao, C. Tan, Z. Gao, G. Chen, P.-A. Heng, and S. Z. Li, "A survey on generative diffusion model," 2022. [Online]. Available: https://arxiv.org/abs/2209.02646

[8] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, Y. Shao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," 2022. [Online]. Available: https://arxiv.org/abs/2209.00796

[9] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," 2022. [Online]. Available: https://arxiv.org/abs/2209.04747

[10] A. Ulhaq, N. Akhtar, and G. Pogrebna, "Efficient diffusion models for vision: A survey," 2022. [Online]. Available: https://arxiv.org/abs/2210.09292

[11] A. Kazerouni, E. K. Aghdam, M. Heidari, R. Azad, M. Fayyaz, I. Hacihaliloglu, and D. Merhof, "Diffusion models for medical image analysis: A comprehensive survey," 2022. [Online]. Available: https://arxiv.org/abs/2211.07804

[12] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," *CoRR*, vol. abs/1503.03585, 2015. [Online]. Available: http://arxiv.org/abs/1503.03585

[13] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 6840–6851. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf

[14] Z. Gao, J. Guo, X. Tan, Y. Zhu, F. Zhang, J. Bian, and L. Xu, "Difformer: Empowering diffusion model on embedding space for text generation," 2022. [Online]. Available: https://arxiv.org/abs/2212.09412

[15] E. Hoogeboom, D. Nielsen, P. Jaini, P. Forré, and M. Welling, "Argmax flows and multinomial diffusion: Learning categorical distributions," in *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., vol. 34. Curran Associates, Inc., 2021, pp. 12 454–12 465. [Online]. Available: https://proceedings.neurips.cc/paper/2021/file/67d96d458abdef21792e6d8e590244e7-Paper.pdf

[16] J. Austin, D. D. Johnson, J. Ho, D. Tarlow, and R. van den Berg, "Structured denoising diffusion models in discrete state-spaces," in *Advances in Neural Information Processing Systems*, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., 2021. [Online]. Available: https://openreview.net/forum?id=h7-XixPCAL

[17] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," 2020. [Online]. Available: https://arxiv.org/abs/2010.02502

[18] Z. He, T. Sun, K. Wang, X. Huang, and X. Qiu, "Diffusionbert: Improving generative masked language models with diffusion models," 2022. [Online]. Available: https://arxiv.org/abs/2211.15029

[19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018. [Online]. Available: https://arxiv.org/abs/1810.04805

[20] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," 2019. [Online]. Available: https://arxiv.org/abs/1910.13461

[21] L. Qian, M. Wang, Y. Liu, and H. Zhou, "Diff-glat: Diffusion glancing transformer for parallel sequence to sequence learning," 2022. [Online]. Available: https://arxiv.org/abs/2212.10240

[22] L. Qian, H. Zhou, Y. Bao, M. Wang, L. Qiu, W. Zhang, Y. Yu, and L. Li, "Glancing transformer for non-autoregressive neural machine translation," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online: Association for Computational Linguistics, Aug. 2021, pp. 1993–2003. [Online]. Available: https://aclanthology.org/2021.acl-long.155

[23] F. Huang, H. Zhou, Y. Liu, H. Li, and M. Huang, "Directed acyclic transformer for non-autoregressive machine translation," 2022. [Online]. Available: https://arxiv.org/abs/2205.07459

[24] S. Gong, M. Li, J. Feng, Z. Wu, and L. Kong, "Diffuseq: Sequence to sequence text generation with diffusion models," 2022. [Online]. Available: https://arxiv.org/abs/2210.08933

[25] H. Yuan, Z. Yuan, C. Tan, F. Huang, and S. Huang, "Seqdiffuseq: Text diffusion with encoder-decoder transformers," 2022. [Online]. Available: https://arxiv.org/abs/2212.10325

[26] N. Savinov, J. Chung, M. Binkowski, E. Elsen, and A. v. d. Oord, "Step-unrolled denoising autoencoders for text generation," 2021. [Online]. Available: https://arxiv.org/abs/2112.06749

[27] M. Reid, V. J. Hellendoorn, and G. Neubig, "Diffuser: Discrete diffusion via edit-based reconstruction," 2022. [Online]. Available: https://arxiv.org/abs/2210.16886

[28] C. Donahue, M. Lee, and P. Liang, "Enabling language models to fill in the blanks," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, Jul. 2020, pp. 2492–2501. [Online]. Available: https://aclanthology.org/2020.acl-main.225

[29] T. Chen, R. Zhang, and G. Hinton, "Analog bits: Generating discrete data using diffusion models with self-conditioning," 2022. [Online]. Available: https://arxiv.org/abs/2208.04202

[30] S. Dieleman, L. Sartran, A. Roshannai, N. Savinov, Y. Ganin, P. H. Richemond, A. Doucet, R. Strudel, C. Dyer, C. Durkan, C. Hawthorne, R. Leblond, W. Grathwohl, and J. Adler, "Continuous diffusion for categorical data," 2022. [Online]. Available: https://arxiv.org/abs/2211.15089

[31] J. Ho and T. Salimans, "Classifier-free diffusion guidance," 2022. [Online]. Available: https://arxiv.org/abs/2207.12598

[32] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020. [Online]. Available: https://arxiv.org/abs/2011.13456

[33] T. Karras, M. Aittala, T. Aila, and S. Laine, "Elucidating the design space of diffusion-based generative models," 2022. [Online]. Available: https://arxiv.org/abs/2206.00364

[34] Z. Lin, Y. Gong, Y. Shen, T. Wu, Z. Fan, C. Lin, W. Chen, and N. Duan, "Genie: Large scale pre-training for text generation with diffusion model," 2022. [Online]. Available: https://arxiv.org/abs/2212.11685

[35] X. Han, S. Kumar, and Y. Tsvetkov, "Ssd-lm: Semi-autoregressive simplex-based diffusion language model for text generation and modular control," 2022. [Online]. Available: https://arxiv.org/abs/2210.17432

An Overview of State-of-the-art Solutions for Scene Text Detection

Mladen Džida, Davor Vukadin, Marin Šilić, Goran Delač, Klemo Vladimir
University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
mladen.dzida@fer.hr, davor.vukadin@fer.hr, marin.silic@fer.hr, goran.delac@fer.hr, klemo.vladimir@fer.hr

***Abstract*—Scene text detection is a task of identifying text regions and labeling them with bounding boxes in a complex background. It has received a lot of attention recently and has become far from unsolvable due to progress of deep learning for computer vision and also due to rapid development of computer hardware which is able to process complex neural networks. Some of the most common challenges that make this task difficult are irregular text shapes, text interferences, very complex background, different text sizes and low image quality. This paper presents an overview of state-of-the-art solutions for scene text detection where ICDAR 2015 was used as a benchmark dataset. We compare solutions with respect to precision, recall and F-score.**

***Keywords*—Scene text, Text detection, Bounding box, Multi-oriented text, Segmentation, Convolutional neural network**

I. INTRODUCTION

Today, textual information is everywhere around us and correctly interpreting it is crucial for utilizing the benefits that society has to offer. This paper focuses on research of textual information that can be found in an outdoor environment, i.e. scene text. Typically, these information are used for wayfinding, advertising, traffic control and displaying of art or any other type of information that is meant to be seen by large number of people. By using machine learning, we could extract the meaning from these texts and use it to help automate outdoor tasks. For example, text information can be used in training of self driving cars, drones and Text-to-Speech models for visually impaired individuals. To achieve this, the automation system would first have to get the image of the scene text, then detect the location of the text in an image and finally recognize the text. Detecting or localizing the text in an image is finding and labeling the correct location of the text. Text recognition is a process that converts an image of text into a machine-readable text format [1]. This paper will focus only on the solutions for the problem of localizing the text in an image.

Detecting scene text in an image faces many challenges and the following are the most common ones:

- **Irregular text shapes**. Text in an outdoor environment often appears in various orientations and shapes. Irregular text shapes can include text that is curved, slanted or written at an angle, which can cause problems for models that are designed to detect text in a specific orientation or shape.
- **Text interferences**. Text interference refers to various factors that can make it difficult to detect text in natural scenes. For example, other objects can cause occlusion of the text in the image or there can be low contrast between text and the background, making it difficult to distinguish them.
- **Complex background**. Scene text in an image often appears on cluttered and dynamic backgrounds, i.e. complex backgrounds. Complex backgrounds can include textures, patterns, or other elements that are similar to the text. While these backgrounds can cause already mentioned text interferences, they can also produce false positives and detect text in the background patterns.
- **Different text sizes**. Scene text comes in various sizes. It can be small like on signs and labels or it can be large like on billboards and buildings. Models that are trained to detect text at a specific size may not perform well when they are given text that is of different size than what they have seen during training.
- **Low image quality**. Low image quality refers to factors such as poor lighting, blurriness and noise. Poor lighting can cause unwanted shadows and reflections, blurriness can make it difficult to distinguish the text from the background and noise can cause text interferences.

To overcome these challenges, many approaches have been proposed, but those from deep learning era have produced the best results. It is notable to mention that there were methods before the rise of deep learning that tackled the problem of scene text detection and they mostly adopted the Connected Components Analysis (CCA) or Sliding Window (SW) based classification [2]. Solutions that adopt these methods will not be covered in this overview, since all state-of-the-art solutions build models on deep learning methodologies.

The remaining sections of this paper are arranged as follows: In Section 2, the general deep learning methodologies for scene text detection are briefly reviewed. In Section 3, state-of-the-art solutions are described. In Section 4, comparison of the state-of-the-art solutions on benchmark dataset is carried out. Finally, conclusion is set in section 5.

II. RELATED WORK

Deep learning has played an important role in the field of scene text detection by enabling the development of

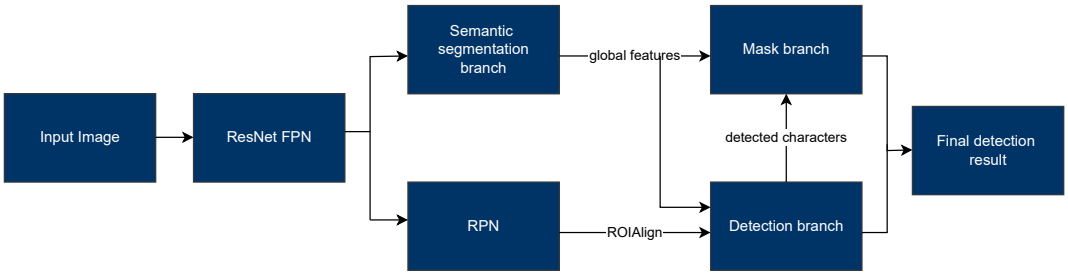


Fig. 1: TextFuseNet: pipeline

highly accurate and robust text detection models [2]. Before the deep learning era, feature extraction for scene text detection relied on hand-crafted features. These features were difficult to engineer and could not handle variations in text such as different fonts, orientations and scales. In contrast, deep learning models can automatically learn powerful representations from the data, allowing them to be more robust to variations in text. These models use convolutional neural networks (CNNs) to automatically learn features from images. Overall, the use of deep learning has led to significant improvements in the performance of scene text detection systems.

Recently, scene text detection experienced great improvement with the development of the CNNs. CNN-based algorithms for scene text detection can be classified into two categories: regression-based, segmentation based and hybrid methods [3]. These approaches will be explained in a more detail in the following subsections.

A. *Regression-based CNN methods*

Regression-based approaches [4]–[11] for scene text detection imply training a model to regress the coordinates of the bounding box for an object directly from the image [3]. Regression-based methods are typically simpler and faster than segmentation-based methods because they only need to predict a small number of bounding box coordinates, but they may not be as accurate when localizing occluded and small objects.

B. *Segmentation-based CNN methods*

Segmentation-based approaches [12]–[20] for scene text detection imply training a model for pixel-level prediction of objects, i.e. prediction of a segmentation mask for each object in the image [4]. Segmentation-based methods are typically more accurate than regression-based methods because they provide more detailed information about the object's shape and position in the image. However, these methods are usually more computationally expensive and require complex post-processing.

C. *Hybrid methods*

Hybrid approaches [21]–[24] combine feature maps produced by regression-based and segmentation-based methods into a sole feature representation which is used for scene text detection. The idea behind using both feature maps is reducing the negative effect of each of these approaches and having more useful information.

III. State-of-the-art approaches

In this section, state-of-the-art architectures for scene text detection will be described in detail. Each architecture will be described in a separate subsection.

TextFuseNet

TextFuseNet is a scene text detection framework that is using multi level feature extraction for detecting texts of arbitrary shapes [24]. The proposed method is a hybrid approach, i.e. the authors perform regression and segmentation. To help explain the reasoning behind such approach, they categorize all of the work in this area into two classes: character-based and word-based. The former detects characters and groups them into words, while the latter immediately detects words. The key idea of their work was extracting three semantically different types of features from the image and combining their information to detect text in various shapes, orientations and sizes. The three levels of features they extract are: character, word and global.

To extract features and later detect bounding boxes, the authors construct an architecture consisting of five parts: feature pyramid network (FPN), region proposal network (RPN), semantic segmentation branch, detection branch and mask branch. ResNet FPN is used as the backbone to extract feature maps of different sizes which they combine in semantic segmentation branch to extract global level features via RoIAlign. RoIAlign is used to extract all levels of features from feature maps. RPN is used to generate text proposals and forward those to detection and mask branches. For extracting features from feature maps in the detection and mask branches, they perform, what they call, multi-path Fusion, which essentially combines, i.e. sums them element wise, features from different paths into fused features. In the detection branch the authors extract word level features from the proposals and fuse them with global features to perform bounding box regression for words and characters. After detecting characters, they can extract character level features as well and they do this in the mask branch. When they extract all feature levels, they fuse them all together and use them for instance segmentation. In the end, the output of the mask branch is used in combination



Fig. 2: Orderless box discretization: pipeline

with the output of the detection branch to get the final text detection result. This entire pipeline can be seen on Figure 1.

In order to evaluate the model, the authors also require character annotations in the dataset and since most datasets only have word annotations, they generate them on their own. For generating character annotations, their own pre-trained model is used to detect characters which will be used in evaluation, i.e. they perform weak supervision.

Orderless Box Discretization Network

Orderless Box Discrectization Network is a regression-based method for scene text detection [25]. The model outputs quadrilateral labels, since they are much more effective at detecting curved text than rectangular labels. The authors introduce a novel method for regressing the bounding box in which they address the problem of inconsistent labeling. This problem occurs for regression-based methods because the labels for the bounding boxes are highly sensitive to changes in the image, like rotation, so it causes a significant change in label sequence with just small amount of interference. To solve this problem, they introduce a system that predicts order irrelevant points, which they call Key Edges (KE). These KEs are invariant representations of the coordinates which do not affect label sequence and are still able to produce a valid bounding box.

Their system architecture consists of three parts: an Orderless Box Discretization (OBD) block, a matching-type learning (MTL) block and rescoring and post-processing (RPP) block. Input image is first let through FPN and RPN, where the authors generate region of interest proposals from which they extract features with RoIAlign and feed them into the OBD block. The Orderless Box Discretization block is used to generate quadrilateral bounding boxes. Each quadrilateral bounding box is described with eight KEs which make four vertices of the bounding box: minimum x and y, the second-smallest x and y, the second-largest x and y, and the maximum x and y. After the OBD block has generated these eight KEs, they are fed into the matching-type learning block which takes care of the matching between x and y KEs to form the most accurate bounding box. Each x KE should match exactly one y KE and all should be matched where they represent the four vertices of the bounding box. Every slightly different pairing results in a different bounding box and since there is four x KEs trying to be matched with y KEs with respect to their order, there is exactly twenty four possible matchings. To further explain this, first x KE can be matched with one of four y KEs, second x KE can then be matched with remaining three y KEs, third x KE can be matched with the remaining two y KEs and the last x KE can only be matched with the last y KE so in total, when multiplicaion is carried out, there are twenty four combinations. Therefore, the model outputs twenty four classes and the authors train the model by minimizing the cross-entropy loss. Finally, the rescoring and post-processing block is used to reduce the number of false positives from the first two blocks by calculating scores for the KEs and drawing conclusions from them. This pipeline can be seen on Figure 2.

Real-time Scene Text Detection with Differentiable Binarization

This is a segmentation-based method for scene text detection that is characterized by a particularly low inference time and thus high frames per second rate [20]. Segmentation-based methods usually require complex post-processing step and because of that they are slow. These methods produce a probability map which needs to be transformed into binary map by using some fixed threshold and also it is necessary to apply certain pixel-level grouping method to get final detection result. Their proposed method inserts binarization step into the model to achieve simpler post-processing and better inference time.

The idea of involving the binarization into the model is not only lowering the inference time, but also optimizing this post-processing step. By optimizing it inside of a neural network, the authors enable the network to adaptively set the threshold which should perform better than fixed threshold at distinguishing background from the text. Standard process of binarization is not differentiable and so it cannot be learned by the neural networks backpropagation step, so they introduce a function which carries out differentiable binarization. Usually, binarization is performed as follows:

$$B_{i,j} = \begin{cases} 1, & \text{if } P_{i,j} \geq t \\ 0, & \text{otherwise} \end{cases}$$

The new differentiable method is:

$$B_{i,j} = \frac{1}{1 + e^{-k(P_{i,j} - T_{i,j})}}$$

B represents output from the binarization or the binarized map, P is the probability map produced from the segmentation-based method, t is a fixed threshold, k is an amplifying factor and T is the adaptive threshold map.

The architecture of their proposed method is a pipeline consisting of the following steps: feature extraction, generating probability map, generating threshold map and generating approximate binary map. In the feature extraction step, the input image is fed into a FPN backbone where the authors produce different levels of feature maps which

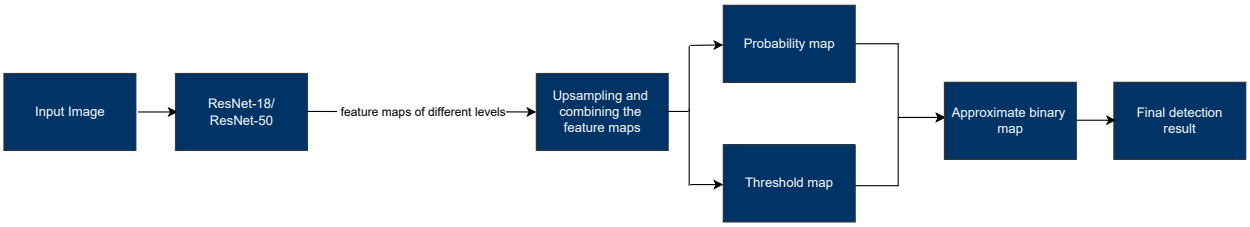


Fig. 3: Segementation-based approach with Differentiable Binarization: pipeline

are then up-sampled to the same scale and combined into a final features used for segmentation. For the backbone, they experimented with ResNet-18 and ResNet-50. Using these features, they generate probability map and threshold map in parallel branches. These maps are, then, used to generate the approximate binary map. Approximate binary map should correctly segment the bounding boxes found in the image. Since they perform supervised training, they generate labels for the probability and threshold maps, but for the binary map the same labels are used as for the probability map.

Progressive Scale Expansion Network

Progressive Scale Expansion Network (PSENet) is a scene text detection model that addresses two issues: detecting text of arbitrary shape and distinguishing text instance from another text instance that is close to it in the image [19]. For the first problem, the authors opt for the pixel-wise segmentation because they believe it performs better than regressing the vertex coordinates of the bounding box. However, segmentation-based methods fail to solve the second problem of separating two really close text instances because pixel-wise segmentation will most likely combine the two text instances into one. PSENet solves this problem using progressive scale expansion algorithm.

PSENet pipeline consists of only two steps: feature extraction and progressive scale expansion block. In the feature extraction step, the input image is fed into a FPN which outputs four levels of feature maps. Low level feature maps are concatenated with high level feature maps and thereby generating four larger feature maps which the authors combine into single feature representation. This feature representation is used for extracting segmentation maps in the progressive scale expansion block. They implement progressive scale expansion block as follows: feature representation extracted from the input image is sent to n number of branches, where each branch is responsible for generating a segmentation map. Each segmentation map masks each text instance in the image, but with different scales. So, for each text instance in the image, they predict multiple segmentation maps, which they call kernels. Each kernel segments the same shape of the input text instance and have the same central point, but the segmented pixels differ in size, i.e. each kernel denotes different area size of pixels for the same instance. Area size of the pixels or the scale of the kernel is a hyperparameter they set orderly where the first kernel returns the smallest scales for the text instances and the last kernel returns the biggest scales of the text instances. After obtaining the kernels, they carry out the core part of the algorithm which is responsible for creating precise segmentation masks for each text instance in the image. This step is based on the Breadth-First-Search (BFS) algorithm, i.e. they perform exhaustive search of the best fit segmentation masks. Here, they start with the first, smallest, kernel where they detect number of text instances or connected components. By doing this, they define the central part of every text instance. After that, they progressively expand other kernels by grouping their pixels and completing the shapes of connected components. Finally they extract these connected components as independent text instances.

To train this model, the authors require the ground truth for the kernels of all selected scales. They generate different scales of the text instance from the bounding box labels by shrinking the bounding box with the Vatti clipping algorithm, but the original bounding box is used as the ground truth for the biggest scale.

IV. Evaluation of the State-of-the-art solutions

In this section, the evaluation of the approaches described in the previous section is carried out. Their performances are compared on ICDAR 2015 benchmark dataset which is described in more detail in the following subsection. Metrics that were used for the evaluation are: precision, recall and F-measure. Also, FPS was compared between the models. The evaluation itself was carried out in the subsection B.

A. Dataset

The ICDAR 2015 dataset, or the International Conference on Document Analysis and Recognition 2015 dataset, is a dataset commonly used for evaluating text detection and recognition methods. It consists of 1500 images, where 1000 images were used for training and 500 for testing [26]. It includes both scene text images and born-digital images. The text in images comes in different languages and orientations and also some of the images intentionally have blur, noise or low resolution. This dataset was used as a benchmark dataset for evaluation of all architectures that are described in this paper.

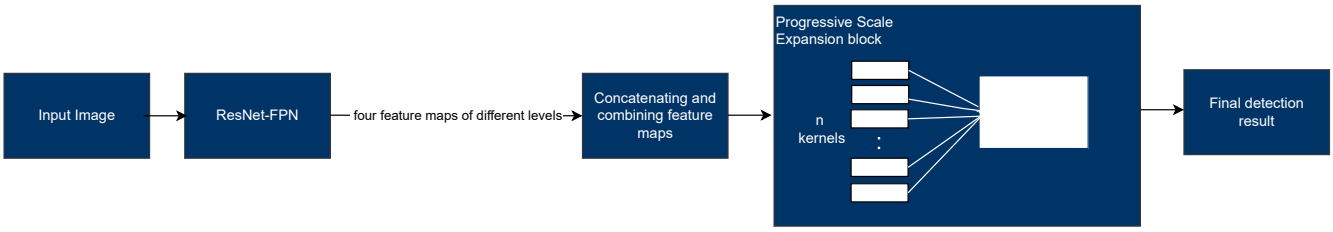


Fig. 4: PSENet: pipeline

Model	Precision	Recall	F-measure
TextFuseNet	**93.96**	**90.56**	**92.233**
OBD	92.1	88.2	90.1
DB	91.8	83.2	87.3
PSENet	88.7	85.5	87.1

TABLE I: Detection results on the ICDAR 2015 dataset

Model	FPS
TextFuseNet	8.3
OBD	4.5
DB	**82**
PSENet	12.38

TABLE II: Highest FPS achieved

B. Comparisons of state-of-the-art approaches on benchmark dataset

In this section, the results of the evaluation are presented. Like mentioned, the approaches are evaluated on the standard metrics: precision, recall and F-measure. The results can be seen in the table 1. TextFuseNet achieved the best measures on all three metrics and therefore perform the best on ICDAR 2015.

Aside from their metrics performance, it is also important that these methods return the localization results quickly, since scene text detection in an outdoor environment is often carried out in real-time and therefore requiring high frames per second (FPS) rate. More complex methods usually achieve better metrics, but worse FPS and simpler methods achieve worse metrics and better FPS. Sometimes, getting better FPS is more important than accuracy and in that case simpler backbones are used for feature extraction. Described approaches experimented with different architectures and have evaluated the models on different datasets than ICDAR 2015. In table 2, one could see the highest FPS rate achieved in each paper, not necessarily on ICDAR 2015. DB model recorded the highest FPS, which was also the selling point of their paper.

V. Conclusion

In this paper, an overview of state-of-the-art solutions for scene text detection was given. These solutions use complex neural networks to solve this task and they can be roughly classified into three groups: regression-based, segmentation-based and hybrid methods. They were evaluated on ICDAR 2015 benchmark dataset with respect to the precision, recall and F-measure. Also, the comparison of the real-time inference was conducted by comparing their FPS rates.

References

[1] "Optical character recognition," 2022, [Online; accessed 14-January-2023]. [Online]. Available: https://en.wikipedia.org/wiki/Optical_character_recognition

[2] S. Long, X. He, and C. Yao, "Scene text detection and recognition: The deep learning era," 2018. [Online]. Available: https://arxiv.org/abs/1811.04256

[3] M. Ibrayim, Y. Li, and A. Hamdulla, "Scene text detection based on two-branch feature extraction," *Sensors*, vol. 22, no. 16, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/16/6262

[4] Z. Tian, W. Huang, T. He, P. He, and Y. Qiao, "Detecting text in natural image with connectionist text proposal network," 2016. [Online]. Available: https://arxiv.org/abs/1609.03605

[5] B. Shi, X. Bai, and S. Belongie, "Detecting oriented text in natural images by linking segments," 2017. [Online]. Available: https://arxiv.org/abs/1703.06520

[6] P. He, W. Huang, T. He, Q. Zhu, Y. Qiao, and X. Li, "Single shot text detector with regional attention," 2017. [Online]. Available: https://arxiv.org/abs/1709.00138

[7] M. Liao, B. Shi, X. Bai, X. Wang, and W. Liu, "Textboxes: A fast text detector with a single deep neural network," 2016. [Online]. Available: https://arxiv.org/abs/1611.06779

[8] W. He, X.-Y. Zhang, F. Yin, and C.-L. Liu, "Deep direct regression for multi-oriented scene text detection," 2017. [Online]. Available: https://arxiv.org/abs/1703.08289

[9] Y. Jiang, X. Zhu, X. Wang, S. Yang, W. Li, H. Wang, P. Fu, and Z. Luo, "R2cnn: Rotational region cnn for orientation robust scene text detection," 2017. [Online]. Available: https://arxiv.org/abs/1706.09579

[10] M. Liao, B. Shi, and X. Bai, "TextBoxes: A single-shot oriented scene text detector," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3676–3690, aug 2018. [Online]. Available: https://doi.org/10.1109%2Ftip.2018.2825107

[11] M. Liao, Z. Zhu, B. Shi, G.-s. Xia, and X. Bai, "Rotation-sensitive regression for oriented scene text detection," 2018. [Online]. Available: https://arxiv.org/abs/1803.05265

[12] Z. Zhang, C. Zhang, W. Shen, C. Yao, W. Liu, and X. Bai, "Multi-oriented text detection with fully convolutional networks," 2016. [Online]. Available: https://arxiv.org/abs/1604.04018

[13] C. Yao, X. Bai, N. Sang, X. Zhou, S. Zhou, and Z. Cao, "Scene text detection via holistic, multi-channel prediction," 2016. [Online]. Available: https://arxiv.org/abs/1606.09002

[14] H. Hu, C. Zhang, Y. Luo, Y. Wang, J. Han, and E. Ding, "Wordsup: Exploiting word annotations for character based text detection," 2017. [Online]. Available: https://arxiv.org/abs/1708.06720

[15] S. Long, J. Ruan, W. Zhang, X. He, W. Wu, and C. Yao, "Textsnake: A flexible representation for detecting text of arbitrary shapes," 2018. [Online]. Available: https://arxiv.org/abs/1807.01544

[16] P. Lyu, C. Yao, W. Wu, S. Yan, and X. Bai, "Multi-oriented scene text detection via corner localization and region segmentation," 2018. [Online]. Available: https://arxiv.org/abs/1802.08948

[17] D. Deng, H. Liu, X. Li, and D. Cai, "Pixellink: Detecting scene text via instance segmentation," 2018. [Online]. Available: https://arxiv.org/abs/1801.01315

[18] W. Wang, E. Xie, X. Song, Y. Zang, W. Wang, T. Lu, G. Yu, and C. Shen, "Efficient and accurate arbitrary-shaped text detection with pixel aggregation network," 2019. [Online]. Available: https://arxiv.org/abs/1908.05900

[19] W. Wang, E. Xie, X. Li, W. Hou, T. Lu, G. Yu, and S. Shao, "Shape robust text detection with progressive scale expansion network," 2019. [Online]. Available: https://arxiv.org/abs/1903.12473

[20] M. Liao, Z. Wan, C. Yao, K. Chen, and X. Bai, "Real-time scene text detection with differentiable binarization," 2019. [Online]. Available: https://arxiv.org/abs/1911.08947

[21] J. Liu, X. Liu, J. Sheng, D. Liang, X. Li, and Q. Liu, "Pyramid mask text detector," *CoRR*, vol. abs/1903.11800, 2019. [Online]. Available: http://arxiv.org/abs/1903.11800

[22] Y. Xu, Y. Wang, W. Zhou, Y. Wang, Z. Yang, and X. Bai, "TextField: Learning a deep direction field for irregular scene text detection," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5566–5579, nov 2019. [Online]. Available: https://doi.org/10.1109%2Ftip.2019.2900589

[23] Y. Liu, S. Zhang, L. Jin, L. Xie, Y. Wu, and Z. Wang, "Omnidirectional scene text detection with sequential-free box discretization," *CoRR*, vol. abs/1906.02371, 2019. [Online]. Available: http://arxiv.org/abs/1906.02371

[24] J. Ye, Z. Chen, J. Liu, and B. Du, "Textfusenet: Scene text detection with richer fused features," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, C. Bessiere, Ed. International Joint Conferences on Artificial Intelligence Organization, 7 2020, pp. 516–522, main track. [Online]. Available: https://doi.org/10.24963/ijcai.2020/72

[25] Y. Liu, T. He, H. Chen, X. Wang, C. Luo, S. Zhang, C. Shen, and L. Jin, "Exploring the capacity of an orderless box discretization network for multi-orientation scene text detection," 2019. [Online]. Available: https://arxiv.org/abs/1912.09629

[26] D. Karatzas, L. Gomez-Bigorda, A. Nicolaou, S. Ghosh, A. Bagdanov, M. Iwamura, J. Matas, L. Neumann, V. R. Chandrasekhar, S. Lu, F. Shafait, S. Uchida, and E. Valveny, "Icdar 2015 competition on robust reading," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, 2015, pp. 1156–1160.

A Transfer Learning Method for Hate Speech Detection

Eva Šmuc*, Goran Delač*, Marin Šilić*, Klemo Vladimir*
* Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
{eva.smuc, goran.delac, marin.silic, klemo.vladimir}@fer.hr

***Abstract*—In this work we explore the possibilities of using transfer learning techniques to enhance performance of hate speech detection models by relying on similar linguistic problems (e.g. toxic language detection). Multiple algorithms are trained for similar linguistic tasks on larger datasets, and the obtained models are used for getting predictions on the ETHOS dataset, which we chose as the target dataset of our work. The obtained predictions are used as sole or additional features in the subsequently performed experiments. Multiple algorithms are evaluated, including Logistic Regression, SVM, RidgeClassifier, Decision Tree, Random Forest, AdaBoost, GradBoost, Bagging. Furthermore, multiple textual representations are taken into account including Tf-Idf, Bert embeddings and BERT embeddings combined with the aforementioned additional features. Transformer-based models BERT and DistilBERT are introduced and fine-tuned on ETHOS dataset. All the obtained models are evaluated and the resulting performance metrics are compared to results obtained by the authors of the ETHOS dataset. In order to explore the remaining underlying issues, model-agnostic method LIME is used to obtain explanations for incorrect predictions.**

***Keywords*—transfer learning, hate speech detection, deep learning, feature extraction**

I. INTRODUCTION

In recent years, people have been joining social networks progressively, resulting in a massive increase in interactions on social media. While this type of infrastructure allows constructive conversations, knowledge exchange and *"information dissemination at a fast rate"* [1], it is also vulnerable to suspicious and harmful activities such as fake news, various forms of propaganda, offensive and hate speech. Hate speech is defined by Facebook as a *"direct attack against people – rather than concepts or institutions – on the basis of protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease"*. Presence of hate speech distorts the initial idea of social media platforms as inclusive environments that support constructive conversation and freedom of expression. In May 2016, The European Commission launched The Code of Conduct, together with Facebook, Microsoft, Twitter and YouTube, as a response to rapidly increasing amount of hate speech present online [2]. The Code makes it possible to report content, which is then removed by the company if it violates their hateful conduct policy. In order to respect these regulations, online social media and social networking services are faced with the challenge of detection and removal of hate speech. Due to a large amount of text posted online on daily bases, it is impossible remove the hateful content manually, which motivates further development of automated hate speech detection technologies based on artificial intelligence. One of the main problems in hate speech detection is a lack of available labeled data for training. Even when labeled, most of the datasets lack a transparent and verified annotation process. However, when similar tasks are considered, there is a substantial amount of available data for toxic, offensive and abusive language detection. In this work, the relationship between hate speech and similar tasks is explored through application of transfer learning techniques.

II. RELATED WORK

A. Similar Tasks

Hate speech classification is a complex task, sometimes not clearly distinguished from other forms of abusive language. The focus of work presented in this paper is binary classification performed on the ETHOS dataset comments. Similar tasks to hate speech classification involve toxic, offensive and abusive text classification, considering both binary and multi-class approaches. Most of the work up to 2020 concerning hate speech classification was focused on feature-based classifiers and advantages of ensemble methods ([3], [4], [5]), while recent research concentrates on usage of deep,transformer-based models such as Bert ([6], [7], [8]) and GPT-3 [9]. Research involving Bert includes retraining the model on specific datasets [7] and experimenting with different embeddings [6] in order to improve performance metrics and achieve new state-of-art results. In [5], a detailed analysis of common challenges in Natural Language Processing is provided. Specifically, difficulties in existing datasets for toxic and similar language detection. Despite focusing on toxic language, authors argue that *"even though each field uses different definitions for their classification, similar methods can often be applied to different tasks"* and show that same methods they use in their work can *"effectively be applied to a hate speech detection task"*.

B. Work Related to ETHOS Dataset

ETHOS binary and multi-label versions of the dataset are introduced in [10]. A detailed description of the annotation process is provided, as well as performance of selected models on binary hate speech classification. In

[6], authors analyze the performance of hate speech detection classifier on ETHOS by replacing or integrating the word embeddings (fastText and GloVe) with static BERT embeddings - it was observed that the neural network performed better with static BERT embeddings. In terms of metrics, an improvement of specificity was achieved when compared to fine-tuned BERT. In [9] classification of ETHOS dataset comments is performed using GPT-3. The paper focuses on exploring the ability of GPT-3 to detect hate speech, with and without providing the model with examples. The main idea was to explore if language models like GPT-3 can be used in the future to help prevent the production of offensive and hateful language produced by humans online.

III. TARGET DATASET: ETHOS

For the purpose of this work, binary ETHOS dataset was selected, due to its detailed description of the annotation process which ensured labels balance and diversity. Another important feature of ETHOS dataset collection process is usage of verified annotators.

A. Creation

The key asset presented in [10] is a balanced dataset creation, which purpose is to overcome the issue of existing hate speech datasets such as biased and imbalanced labels. The data consists of Youtube comments collected through Hatebusters and Reddit comments collected through Public Reddit Data Repository. Initially, data labels are predicted by an SVM classifier provided by Hatebusters. In the next phase, comments in the $[.4, .6]$ probability range are manually annotated since their classification is less certain, while the comments in the ranges $[.0, .1] \cup [.9, .1.0]$ are only examined in order to detect possible misclassification. Only certain comments are selected depending on their label and content, to achieve balance and diversity. Data is validated via Figure Eight Data Labeling Platform. Finally, the annotated data is examined manually to prevent any possible misclassification.

B. Overview

Dataset used in this work is a binary version of ETHOS dataset "Ethos_Binary.csv" that contains 998 comments and two classes, hate and non-hate. More precisely, it contains a label *'isHate'*, a decimal number in the range [0,1] which marks the absence or presence of hate in the comment. For every comment, N annotators voted for the set labels. P is the sum of positive votes, which is divided by the number of annotators N, to normalize the *'isHate'* value to the [0,1] range. For the purpose of using algorithms in binary scope, the authors of ETHOS dataset propose binarization of values to the {0,1} classes for each label ('isHate' >= 0.5: label = 1 Else: label = 0). Rounding of values results in 433 samples classified as hate and 565 samples classified as non-hate. The measure of dataset balance can be calculated using the Shannon entropy [6], where n is the number of examples in the dataset, k the number of classes and c_i the size of class i. The Shannon entropy approaches 0 for unbalanced datasets and log(k) for balanced datasets. For ETHOS dataset $Balance = 0.986$, indicating a balanced dataset.

$$Balance = \frac{H}{log(k)} = \frac{-\sum_{i=1}^{k} \frac{c_i}{n} log(\frac{c_i}{n})}{log(k)} \quad (1)$$

IV. EXPLORATORY DATA ANALYSIS

Exploratory data analysis is performed in order to get data insight and summarize the most important characteristics of the dataset. The additional purpose of the analysis is to detect possible differences in the comments belonging to opposing classes. *Perspective API* [11] is used to calculate toxicity and profanity probability scores of the comments. The score indicates how likely it is that a reader would perceive the comment as containing the given attribute. The *nltk* library is used to get sentiment scores for all comments. To get the number of difficult words per comment, *textstat* library is used.

Toxicity The majority of 'hate speech' comments fall in the range between 0.5 and 1, while the majority of 'not hate speech' comments have toxicity probability between 0 and 0.45.

Profanity Comments labeled as 'hate speech' are in the range between 0.5 and 1, but the majority have a profanity score higher than 0.9. Most of the 'not hate speech' comments have probability range between 0 and 0.25.

Sentiment score Sentiment scores are scaled to range [0,1]. The difference in 'hate speech' and 'not hate speech' comments sentiment is not as clear as in the case of toxicity and profanity, but it is still noticeable.

Difficult words The number of difficult words, further defined by *textstat* as rare or atypical words, are more likely to be found in comments labeled as 'hate speech', according to the analysis.

The difference in attribute mean values for both classes is shown in Table I. Higher means for profanity, toxicity and word difficulty are observed in 'hate speech' comments. Lower mean for sentiment score in 'hate speech' class indicates a more negative sentiment of comments in this class than in the comments labeled as 'not hate speech'.

Label	profanity	toxicity	sentiment score	word difficulty
'not hate speech' (0)	0.2931	0.3336	0.4911	2.7557
'hate speech' (1)	0.6113	0.7197	0.3492	3.1778

TABLE I: Mean value of attributes for 'not hate speech' and 'hate speech' comments

V. PROBLEM DESCRIPTION AND PROPOSED APPROACH

A. Problem Description

Throughout this work, several different combinations of text representations and algorithms are tested in order to compare the performance results on the ETHOS dataset.

Transfer learning paradigm is explored by using predictions made by models trained for similar tasks (offensive, abusive and toxic language detection) as additional or sole features used to train the binary hate speech classifier. Transformer-based models are fine-tuned on the same dataset and their performance is compared to other models. Although hate speech and similar forms are not clearly distinguishable [7], we explore how usage of other (and usually larger) datasets built for similar tasks can possibly improve performance of hate speech detection models.

B. Proposed Approach

The proposed transfer learning approach is to obtain new, numerical features, and train models using those features.

Feature set 1 is a set of features obtained by using machine learning models trained on related datasets. Floating point features in *Feature set 1* represent probabilities for positive, 'isHate' class, obtained by using pre-trained models for prediction of class probabilities for ETHOS comments.

Creation of Feature set 1 consists of selecting the larger datasets for similar tasks to hate speech, selecting machine learning algorithms and pre-trained models, and training those models on selected datasets. ETHOS is a binary dataset, so other binary datasets were needed to obtain the probabilities for classes 'non-hate' and 'hate'. Due to a lack of larger binary datasets, several multi-class ones were chosen (as presented in Table II): Toxic, Movies, Youtube, and one binary dataset - FoxNews. Each non-binary dataset was binarized by separating the neutral class 0 (*neither*, *appropriate*, *non-hate*) from the rest of the classes and assigning class 1 to all other classes.

Dataset	Original classes	New binary classes	Size
Toxic	toxic severe_toxic obscene threat insult identity_hate neither	0: neither 1: (toxic, severe_toxic, obscene, threat, insult, identity_hate)	159571
Fox News	non-hate hate	0: non-hate 1: hate	1528
Movies	hate offensive neither	0: neither 1: (hate, offensive)	3208
Youtube	appropriate inappropriate offensive violent	0: appropriate 1: (inappropriate, offensive, violent)	101569

TABLE II: Selected datasets, their original classes, new binary classes and size

After data preprocessing, each of the four chosen datasets (Toxic, Fox News, Movies and Youtube) contains a column 'comment' labeled as 1 (inappropriate, offensive, violent, toxic, hate etc.) or 0 (neutral, neither, appropriate). Five different machine learning methods were chosen to create models by training (or fine-tuning in case of BERT) on each of the four datasets. One linear approach - Logistic regression, two ensemble methods - Random Forest Classifier and XGBoost, and two transformer-based pre-trained models - BERT and DistilBERT. The result of the described process are 20 = (5 models)*(4 datasets) different models shown in Table III, and *Feature set 1* - predictions of class probabilities for ETHOS comments, provided by these models.

	Toxic	Fox News	Movies	Youtube
Logistic regression (LR)	LR_toxic	LR_fox	LR_movies	LR_youtube
Random Forest Classifier (RFC)	RFC_toxic	RFC_fox	RFC_movies	RFC_youtube
XGBoost (XGB)	XGB_toxic	XGB_fox	XGB_movies	XGB_youtube
BERT	toxic_bert	fox_bert	movies_bert	youtube_bert
DistilBERT	toxic_distil	fox_distil	movies_distil	youtube_distil

TABLE III: Selected ML models and datasets

Feature set 2 is created using toxicity, profanity and sentiment probability, as well as the number of difficult words per comment.

Creation of Feature set 2 is explained in Section IV, along with definitions of each attribute. These attributes accentuate differences in comments labeled with 0 and comments labeled with 1, and can also be used as additional numerical features to achieve improvement in classifier performance. *Feature set 2* is shown in Figure 1.

	comment	label	profanity	toxicity	sentiment_score	word_difficulty
0	You should know women's sports are a joke	1	0.208376	0.448056	0.648495	1
1	You look like Sloth with deeper Down's syndrome	1	0.310944	0.448206	0.681315	2
2	You look like Russian and speak like Indian. Both are disgusting go kill yourself	1	0.693099	0.961763	0.184939	2
3	Women deserve to be abused, I guess.	1	0.464719	0.700786	0.242475	0
4	Women are made for making babies and cooking dinner and nothing else!!!	1	0.221928	0.750965	0.499497	0
...	...	...	...	...	...	...
993	From the midnight sun where the hot springs blow	0	0.153424	0.053735	0.499497	0
994	Don't say I'm not your type	0	0.059359	0.151594	0.499497	0
995	And therefore never send to know for whom the bell tolls	0	0.025942	0.068119	0.499497	1
996	And I can't stand another day	0	0.078510	0.187466	0.499497	0
997	All values, unless otherwise stated, are in US dollars."	0	0.010062	0.018603	0.701802	3

Fig. 1: Feature set 2

VI. Approach Based on BERT

Transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) [12] learn the context of a word based on both the words that precede it and that follow it ('bidirectional'), as a contrast to directional models that read the text input either left-to-right or right-to-left. It is often enough to fine-tune BERT to achieve high performance results on smaller datasets. In this work, BERT and DistilBERT are used for classification and regression. Additionally, BERT is used for extracting embeddings.

A. BERT and DistilBERT - Classification and Regression

BERT and DistilBERT classification model is fine-tuned on each of the datasets presented in Section V-B and for single sentence classification on the ETHOS dataset in Experiment B. When fine-tuning a BERT regression model, the 'isHate' score is used as a label and the mean-squared loss function is used instead of cross-entropy loss function. BERT and DistilBERT regression model are used in Experiment B, for performance comparison with the classification models. This is done to emphasize the impact that the rounding of the averaged score has on the accuracy of predictions.

B. BERT Embeddings

BERT can be used solely to extract features from text data. Other algorithms for obtaining vector representations for words, such as GloVe, Fasttext and Word2Ve, have the same function. The key advantage that BERT offers are contextualized word embeddings, while the other mentioned algorithms have a fixed representation of the word regardless of the context. Since BERT embeddings are numerical vectors, it is possible to append other additional features and use BERT embeddings with appended features as a representation. Experiment A is performed by learning regression models with different algorithms and using several representations extending BERT embeddings: BERT embeddings only, BERT embeddings with appended *Feature set 1*, BERT embeddings with appended *Feature set 2* and finally BERT embeddings with appended both *Feature set 1* and *Feature set 2*.

VII. EVALUATION

Nested cross-validation is used as an approach to perform model selection and validation. The k-fold cross-validation for model hyperparameter optimization is nested inside the k-fold cross-validation procedure for model selection, where $k = 10$ is chosen for the outer loop, and $k = 3$ for the inner loop. For evaluation purposes, the same evaluation method and algorithms (but adjusted to work with the new data) are selected as in [10]. This was done in order to make a direct comparison with the performance results presented by the authors, achieved when using feature extraction on ETHOS dataset comments (done by converting the raw comments to a matrix of TF-IDF features using TfidfVectorizer). The obtained results are shown in Table IV.

A. Experiment A: Logistic Regression, SVM, RidgeClassifier, Decision Tree, Random Forest

In Experiment A, Logistic Regression, SVM, RidgeClassifier, Decision Tree and Random Forest based models are optimized for each algorithm using nested cross-validation method. The proposed approach with BERT embeddings and appended *Feature set 1* and *Feature set 2* performs best, with Ridge Classifier producing the highest scores. SVM performs almost the same as Ridge Classifier, but with a significantly less amount of time needed for fitting.

Representation	Method	F1	Precision	Recall	Accuracy	fitTime(sec)
Tf-Idf	LogReg	0.6650	0.6694	0.6707	0.6694	166.2673
	SVM	0.6607	0.6647	0.6670	0.6643	103.2807
	Ridge	0.6547	0.6571	0.6580	0.6624	36.0019
	DTree	0.6104	0.6148	0.6152	0.6181	646.2416
	RandomF	0.6441	0.6469	0.6468	0.6504	1.2263

TABLE IV: ETHOS paper: performance metrics for Experiment A

B. Experiment B: DistilBERT and BERT

In Experiment B, BERT and DistilBERT are fine-tuned on the ETHOS dataset. Best performance metrics is achieved with the BERT regression model (when the model is fine-tuned using 'isHate' as a label and the obtained predictions are rounded).

Representation	Method	F1	Precision	Recall	Accuracy	fitTime(sec)
BERT emb	LogReg	0.7195	0.7220	0.7216	0.7284	231.4352
	SVM	0.7145	0.7191	0.7173	0.7245	2.1786
	Ridge	0.7114	0.7140	0.7140	0.7204	44.7056
	DTree	0.5632	0.5791	0.5792	0.5813	294.2670
	RandomF	0.6380	0.6554	0.6419	0.6603	29.2616
Feature set 1	LogReg	0.7452	0.7469	0.7467	0.7525	0.5334
	SVM	0.7343	0.7357	0.7367	0.7415	1.0674
	Ridge	0.7395	0.7407	0.7416	0.7464	0.3148
	DTree	0.7051	0.7080	0.7097	0.7114	22.9166
	RandomF	0.7283	0.7309	0.7285	0.7375	183.9144
Feature set 2	LogReg	0.7762	0.7780	0.7774	0.7826	1.2931
	SVM	0.7808	0.7831	0.7832	0.7865	3.5810
	Ridge	0.7096	0.7129	0.7171	0.7134	0.2996
	DTree	0.7728	0.7767	0.7823	0.7755	18.5314
	RandomF	0.7673	0.7693	0.7703	0.7735	1.9666
BERT emb	LogReg	0.7592	0.7623	0.7602	0.7675	112.2335
+	SVM	0.7145	0.7191	0.7173	0.7245	2.2121
Features set 1	Ridge	0.7114	0.7140	0.7140	0.7204	46.2407
	DTree	0.5632	0.5791	0.5792	0.5813	286.9777
	RandomF	0.6380	0.6554	0.6419	0.6603	27.2942
BERT emb	LogReg	0.7963	0.7959	0.7989	0.8016	135.8343
+	SVM	0.7984	0.7981	0.8026	0.8026	2.2675
Feature set 2	Ridge	0.7912	0.7920	0.7937	0.7966	28.3758
	DTree	0.7721	0.7823	0.7852	0.7745	300.7481
	RandomF	0.7838	0.7842	0.7885	0.7885	222.4563
Feature set 1	LogReg	0.8098	0.8100	0.8130	0.8146	9.9949
+	SVM	0.8080	0.8084	0.8119	0.8126	46.4875
Feature set 2	Ridge	0.8080	0.8079	0.8118	0.8126	5.4765
	DTree	0.7678	0.7704	0.7754	0.7715	43.8981
	RandomF	0.7917	0.7908	0.7946	0.7966	3.4595
BERT emb	LogReg	0.8070	0.8077	0.8089	0.8126	123.8262
+	**SVM**	**0.8159**	**0.8153**	**0.8184**	**0.8206**	**2.1346**
Feature set 1	**Ridge**	**0.8162**	**0.8154**	**0.8198**	**0.8207**	**35.0411**
+	DTree	0.7721	0.7823	0.7852	0.7745	361.5874
Feature set 2	RandomF	0.7825	0.7822	0.7860	0.7875	38.4884

TABLE V: Performance metrics for Experiment A

Method	F1	Precision	Recall	Accuracy
BERT (classification)	0.7551	0.7756	0.7681	0.7644
BERT (regression)	**0.7893**	**0.7911**	**0.7902**	**0.7956**
DistilBERT (classification)	0.7277	0.7436	0.7304	0.7446
DistilBERT (regression)	0.7804	0.7866	0.7808	0.7886

TABLE VI: Performance metrics for Experiment B

VIII. INTERPRETABILITY

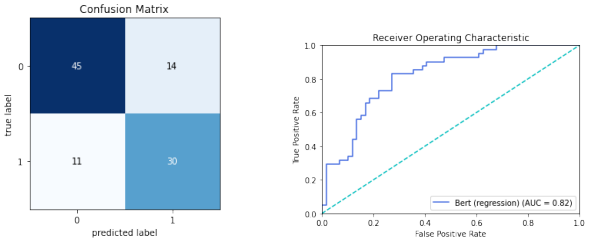
Results of cross-validation provide an overview of model performance over unseen data. This is a process necessary for machine learning model performance estimation, but it doesn't help in understanding model predictions and its decision making path. Models are defined as interpretable *"when humans can understand the reasoning behind predictions and decisions made by the model"* [13]. The purpose of an explanation is to relate feature values of an example to its model prediction in an understandable way for humans. In the case of automated hate speech detection, removing comments with no justifiable explanation can result in violation of corporate technology practices and preferences.

A. Error analysis using LIME

Model-agnostic interpretability techniques enable global and local post hoc model interpretation, independent of implementation. One of the local methods is a local surrogate model, which explains a prediction by replacing complex model with a locally interpretable surrogate model. LIME (Local Interpretable Model-agnostic Explanations), a technique that trains local surrogate models to get explanations for single predictions, was first introduced in [14]. Models from Experiment A do not reach F1

score higher than 0.82, while transformer-based classifiers from Experiment B reach 0.79. To find out the underlying issues, an error analysis is performed using LIME, by providing an explanation for erroneous predictions. Error analysis is based on the one performed in [5].

1) LIME for Text: BERT Regression Model: BERT regression model (the best performing transformer-based model from Experiment B) was tested on the test set consisting of 100 examples. Confusion matrix and the Receiver Operator Characteristic are shown in Figure 2.



(a) Confusion matrix (b) ROC

Fig. 2: BERT Regression Model

False Negatives

- **'Hate speech' comments without swear words or extreme language** Previous research shows that *"classifiers often fail when there are no obvious hateful words present"* ([5], [15]).

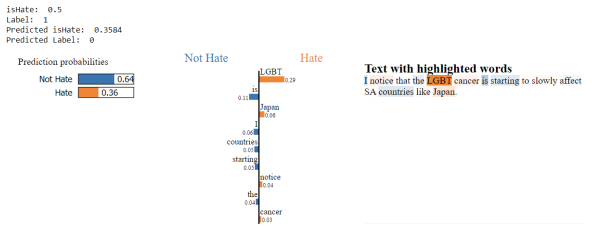


Fig. 3: An example of a comment misclassified as 'not hate speech', because it doesn't contain any hateful vocabulary, swear words or offensive language

- **Questionable 'isHate' score** Out of 11 false negatives, 8 have an original 'isHate' score of 0.5, which indicates that comments with this score are problematic for the classifier. Even though the label of the following comments is 1 when 0.5 is rounded, their content suggests that their 'isHate' score should be higher.

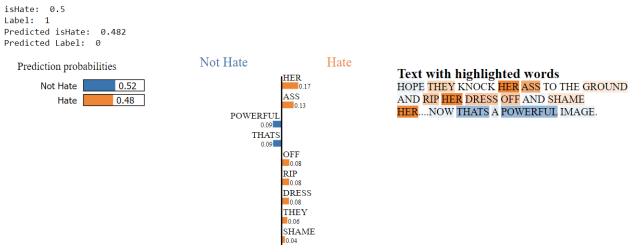


Fig. 4: An example of a comment with a questionable 'isHate' score

False Positives

- **Overrepresented terms in the training set** The classifier learned to use the words "white", "islam", "muslim" (and some others), as features for classifying comments as 'hate speech', which is not always the case. These words alone do not necessarily indicate hate speech. Erroneous predictions that the model makes because of their presence are the result of data distribution in the training set. For example, 40/58 comments containing the word "white" in the training set are comments with label 1, 24/33 comments containing the word "islam" are comments with label 1 etc.

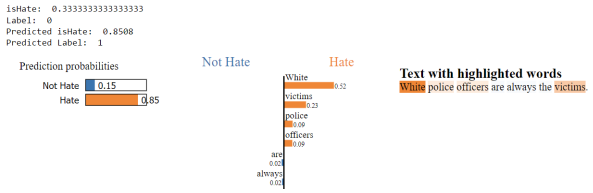
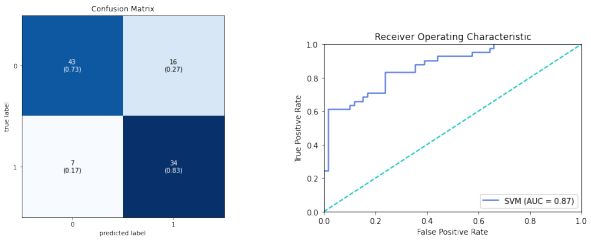


Fig. 5: An example of a comment misclassified as 'hate speech', because of the word "White"

2) LIME for Tabular Data: SVM: For applying LIME for tabular data to get prediction explanations, SVM with *Feature set 1 + Feature set 2* as representation was selected. Even though the Ridge Classifier with BERT embeddings + *Feature set 1 + Feature set 2* as representation had the best overall performance, the best performing model with *Feature set 1 +Feature set 2* as representation was chosen because these features are named and both their meaning and creation process are explained in previous sections. BERT word embeddings have 768 unnamed features with no clear meaning and can't contribute to explanatory analysis. The selected SVM model was tested on the same test set consisting of 100 examples. Confusion matrix and the Receiver Operator Characteristic are shown in Figure 6.



(a) Confusion matrix (b) ROC

Fig. 6: SVM

In LIME for tabular data, floating point numbers represent the relative importance of selected features and feature names on the right.

False Negatives

- **'Hate speech' comments without toxic language** All the False Negatives have the `toxicity` feature contributing to class 0. Since `toxicity` is the most contributing feature, the predicted label for these

comments is 0. This again points to an already stated issue of detecting comments containing hate speech but not swear words or toxic language.

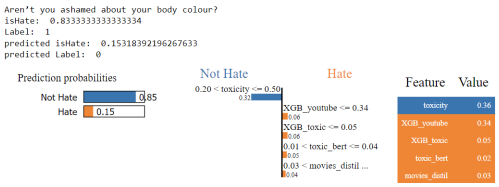


Fig. 7: An example of a 'hate speech' comment without toxic language

False Positives

- **'Not hate speech' comments with toxic language** Upon detailed analysis, it can be concluded that `toxicity` is the most contributing feature, which is the cause of the majority of False Positives. Each of the following comments contain words such as "wtf", "bastard", "disgusting" and other words likely to be overrepresented in the training data of classifiers for toxic language. Detecting toxic language is a similar task, but the differences become more apparent once a dominant toxicity feature is used in training models for a hate speech detection task.

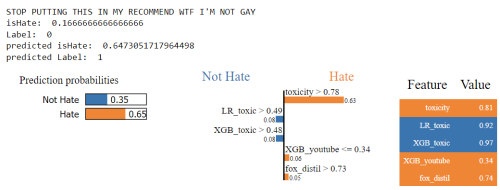


Fig. 8: An example of a 'not hate speech' comment with toxic language

B. Error analysis summary

Analysis showed that the analysed models had 4 False Negatives and 8 False Positives in common. After inspecting their content and getting the same prediction using both models, it is reasonable to question the quality of these labels. Through the error analysis, some of the possible biases were detected. Definitions of these biases and their consequences are given in Table VII.

Bias	Meaning	Result
Sampling bias	The way that dataset examples were sampled doesn't reflect real-world distribution	Misclassification due to terms over-represented in one class ("white", "muslim"...)
Annotator bias	A bias caused by annotators because of subjectivity and differences in knowledge [16]	Lack of quality and consistency in labels

TABLE VII: Definitions and results of detected biases

IX. CONCLUSION

Due to the ambiguous nature of language and its structural complexity, differentiation of hate speech and other forms of similar language presents a challenging task. In this work, the relationship between hate speech and other forms of similar language phenomena is explored through application of transfer learning techniques. Models are trained for hate speech detection using additional features, obtained through predictions of models created for similar tasks. Performed error analysis using LIME highlighted a problem in this approach. Even though related language phenomena helps in improving classifier performance due to its semantic similarity with hate speech, it also causes inaccurate predictions because it does not necessarily imply hate speech. Despite this problem, the presented approach still proved to be an effective and affordable strategy to increase model performance on smaller datasets. Error analysis brought to attention other underlying problems including sampling bias and lack of quality and consistency of ETHOS labels. This is arguably one of the most problematic issues, having in mind that only the test set was inspected closely, thus, incorrect labels are possibly present in the training set. Finally, the future research steps will be focused on ensuring better label quality. Specifically, we argue that better distinction between hate-speech and similar phenomena, might be a way of improving classifier performance. This stems from the fact that poor distinction between the the mentioned phenomena presents performance issues even for the transformer-based models.

REFERENCES

[1] B. Mathew, R. Dutt, P. Goyal, and A. Mukherjee, "Spread of hate speech in online social media. in: Proceedings of the 10th acm conference on web science," pp. 173–182, 2019.

[2] European Comission, "The code of conduct on countering illegal hate speech online," https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_113, accessed: 2022-04-20.

[3] Z. Waseem and D. Hovy, "Hateful symbols or hateful people? predictive features for hate speech detection on twitter," 2016.

[4] P. Burnap and M. L. Williams, "Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making," 2015.

[5] B. van Aken, J. Risch, R. Krestel, and A. Löser, "Challenges for toxic comment classification: An in-depth error analysis," 2018.

[6] G. Rajput, N. Singh, S. Sonbhadra, and S. Agarwal, "Hate speech detection using static bert embeddings," 2021.

[7] T. Caselli, V. Basile, J. Jelena Mitrović, and M. Granitzer, "Hatebert: Retraining bert for abusive language detection in english," 2021.

[8] S. S. Aluru, B. Mathew, P. Saha, and A. Mukherjee, "Deep learning models for multilingual hate speech detection," 2020.

[9] K.-L. Chiu, A. Collins, and R. Alexander, "Detecting hate speech with gpt-3," 2022.

[10] I. Mollas, Z. Chrysopoulou, S. Karlos, and G. Tsoumakas, "Ethos: an online hate speech detection dataset," 2021.

[11] "Perspective api," https://perspectiveapi.com/how-it-works/, accessed: 2022-04-20.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018.

[13] MathWorks, "What is interpretability?" https://www.mathworks.com/discovery/interpretability.html, accessed: 2022-05-15.

[14] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?": Explaining the predictions of any classifier," 2016.

[15] T. Davidson, D. Warmsley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," 2017.

[16] M. Wich, H. Al Kuwatly, and G. Groh, "Investigating annotator bias with a graph-based approach," 2020.

Detection of Shilling Attacks on Collaborative Filtering Recommender Systems by Combining Multiple Random Forest Models

Vjeran Grozdanić*, Klemo Vladimir*, Goran Delač*, Marin Šilić*
* University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10 000 Zagreb, Croatia
vjeran.grozdanic@fer.hr, klemo.vladimir@fer.hr, goran.delac@fer.hr, marin.silic@fer.hr

***Abstract*—Collaborative filtering recommender systems are one of the essential recommender systems and are still widely used in combination with other algorithms to make predictions for users. However, they are vulnerable to shilling attacks, and if there isn't any detection system to prevent those attacks, original recommendations can be heavily influenced to benefit the attackers. Designing attack-resistant recommendation systems is not an easy task, and many researchers have tried to tackle that problem. In this paper, a new approach that combines multiple random forest models is proposed. Each of the random forest models is specialized in detecting one group of shilling attacks, and then all the models are combined into an ensemble model. Experimental results show that the proposed ensemble is capable of detecting attack profiles at high rate without causing significant bias in the original recommendation system.**

***Keywords*—collaborative filtering, random forest, attack detection, shilling attacks, classification**

I. INTRODUCTION

It's hard to imagine today's online world without the recommender systems. They have become essential part of the web applications like Netflix [1], Amazon [2] and many others. Some of the most used algorithms in recommender systems come from collaborative filtering (CF) group of algorithms. CF algorithms have shown great prediction quality both in academic research and in industrial applications and up to this day they still give state-of-the-art results when it comes to recommendation prediction results [3].

Given the wide use of CF methods in the industry recommender systems, it is not surprising to see increase in malicious activities and attacks against CF algorithms [4]. With properly executed attack, one can significantly increase the ratings and the popularity of the chosen product, during a *push* attack, or if wanted, one can significantly decrease the popularity of the target product which is performed with a *nuke* attack.

Without the proper attack detection system incorporated into the recommender system, it is easy to manipulate the system through various attacks. For that reason, a lot of focus has been put on designing robust recommender systems capable of ignoring the impact of the attack. One of the first approaches in attack detection was ad-hoc algorithm [5] which was used to classify fake profiles.

Item 1	...	Item K	...	item N	...	Target item
r_1	...	r_k	...	r_n	...	r_t
Selected items		Filler items		Unrated items		Target item(s)

TABLE I: Visual representation of the attack profile

Later approaches started using various machine learning algorithms to tackle the attack detection problem.

Recent work has been very successful in coping with the problems of attack detection and prevention. Many different approaches has been tried to create as robust systems as possible [6]. Clustering approach [7] [8] [9] is designed believing that fake profiles will eventually end up in the same cluster or a group after the clustering algorithm is applied. Probabilistic approach [10] [11] focuses on using the probabilistic methods to detect fake profiles. Classification approach [12] [13] uses extracted features to decide if the profile is fake or authentic.

II. SHILLING ATTACKS

Attacks on the collaborative filtering methods are called *profile injection attacks*, because they are done by injecting fake profiles into the recommender system, thus causing the bias in the recommendation that the system produces. Also, a very popular name for this kind of attacks is shilling attacks [14], where the origin of the name addresses how cheap they are to perform, because the attacks can easily be automated and executed only from attacker's computer.

There are multiple ways in which the attacks can be categorized. One of the biggest factors of the attack is the intent, does the attacker want to increase the popularity of a target item or decrease it. When the goal of the attack is to increase the popularity of the item, and along with that the number of times the item will be recommended, it is called push attack, because the popularity is being pushed. On the other side, there are nuke attacks, where the goal of the attack is to decrease the popularity of an item.

Shilling attacks are categorized by the strategies the attacker would use while creating the fake profiles. Each attack profile (visible in a table I) is made out of 4 different sets of items and grades for each of the item and those sets are:

- **Target item(s)** - set which contains only the item(s) which is being attacked
- **Selected items** - set of carefully selected items to increase the reach of the attack
- **Filler items** - set of randomly selected items which are used to make attack profiles bigger and to add some noise into the attack
- **Unrated items** - set containing all the items available in the system for which the attacker hasn't provided the rating

Every attack can be described by the strategies used while generating the ratings for each of the sets of the attack profile. One of the most basic attacks proposed in [14] is random attack which is performed by generating random rating for every item in both selected items and filler items sets, and this attack is easy to perform since the attacker doesn't need deep knowledge about the system besides what items are in it.

Other, way more effective attack also proposed in [14] is average attack which fills the selected items set and filler items set with a system-wide mean rating of the specific item. Performing the average attack requires more knowledge about the system because the attacker needs to know the item's mean rating in the system, but often that information is publicly available. Both random and average attacks can be used as either push or nuke attacks, the only difference is if the target item is going to get maximum rating (push attack) or minimum rating (nuke attack).

There are many more variants of shilling attacks, but only one more is important for this paper, and that is a bandwagon attack. It is an attack designed to capitalize on popular items in the systems to expand the impact of the attack. During the attack, the attacker carefully selects k most popular items in the system and assigns maximum rating to them, those items are part of the selected items set, and for the filler items set, it randomly selects the items and the ratings for them. Random attack could be considered a special case of the bandwagon attack when the k is set to 0.

Attack on the CF recommender system is also determined by the attack *profile size*, the number of the items attacker has rated, and the *attack size*, usually shown as a percentage of the fake profiles in the system, which were injected by the attacker in the system.

III. CLASSIFICATION MODEL

For a good classification results, the good classification attributes are needed. In [15] were proposed both generic and model-specific attributes which could be used for the classification of fake profiles. Generic attributes are the ones which are used for every different variant of attack, while model-specific attributes were proposed as a result of careful inspection of how each of the attacks is performed and those attributes have significant importance when detecting specific attacks for which they were designed.

A. Classification Attributes

Classification attributes used in this paper have all been taken from [15]. When there are a lot of items rated from a single user, it is not very probable that an authentic user is behind that profile. Driven by that idea, the *length variance*(L_u) was proposed:

$$L_u = \frac{|n_u - \overline{n_u}|}{\sum_{u \in U}(n_u - \overline{n_u}^2)} \tag{1}$$

where n_u is the length of the profile of the user u, $\overline{n_u}$ is the average length of the user profile in the system and U is a set of all users.

Another proposed attribute was *weighted deviation from mean agreement* (WDMA) which is based on *rating deviation from mean agreement* (RDMA), but it puts higher emphasis on the rating deviations for sparse items.

$$WDMA_u = \frac{\sum_{i=0}^{n_u} \frac{|r_{u,i} - \overline{r_i}|}{l_i^2}}{n_u} \tag{2}$$

In (2) n_u is the length of the profile of the user u, $r_{u,i}$ is the rating user u has given to the item i, $\overline{r_i}$ is the mean rating for item i, and l_i is a number of users which have rated item i.

Similar to WDMA, *weighted degree of agreement* (WDA) is defined with the difference that here the size of the user profile isn't important. WDA can be calculated with the following formula:

$$WDA_u = \sum_{i=0}^{n_u} \frac{|r_{u,i} - \overline{r_i}|}{l_i} \tag{3}$$

where n_u is the length of the profile of the user u, $r_{u,i}$ is the rating user u has given to the item i, $\overline{r_i}$ is the mean rating for item i, and l_i is a number of users which have rated item i.

Standard deviation of the user's ratings can also help distinguish fake and authentic profiles:

$$\sigma_u = \sqrt{\frac{\sum_{i=0}^{n_u}(r_{u,i} - \mu_u)^2}{n_u}} \tag{4}$$

In (4) n_u is the number of ratings user u has given, $r_{u,i}$ is the rating user u has given to the item i, while μ_u is the mean rating of user u.

The idea that fake profiles are closer to their most similar profiles than the authentic ones is captured in the degree of similarity attribute:

$$DegSim_u = \frac{\sum_{v \in neighbors(u)} W_{u,v}}{k} \tag{5}$$

where k is the number of the closest neighbors for user u, and $W_{u,v}$ is the similarity function between user u, and it's neighbor v.

All the attributes mentioned above are generic attributes, but in order to tackle with average attack, model specific

attribute was also designed in [15]. First, the set of ratings that are potential targets is defined as $P_{u,T} = \{i \in P, where\ r_{u,i} = r_{max}\}$ (or $where\ r_{u,i} = r_{min}$ for nuke attacks). P_u represents profile of user u, $P_{u,F}$ represents rest of the user profile and it is calculated as $P_{u,F} = P_u - Pu, T$. Then a *mean variance* can be calculated using the following formula:

$$MeanVar_u = \frac{\sum_{i \in P_{u,F}} (r_{u,i} - \overline{r_i})^2}{|P_{u,F}|} \quad (6)$$

where $r_{u,i}$ is the rating user u has given to the item i, and $\overline{r_i}$ is the mean rating for item i. For the attack profiles during average attack, this value is expected to be very low comparing to the authentic profiles.

B. Influence of the attack

In order to tell how effective an attack is, we need to be able to measure the influence of it. There are two well known methods for measuring that influence.

First method, a prediction shift, which tells how much has the mean predicted rating of the item shifted after the attack. Let $p(u, i)$ be a predicted rating which the user u would give to the item i before the attack, and let the $p'(u, i)$ be a predicted rating which the user u would give to the item i after the attack has happened. Then, we can calculate prediction shift using the following formula:

$$\Delta p = \frac{\sum_u^U |p(u, i) - p'(u, i)|}{n} \quad (7)$$

where n is the number of the users in set U.

Using just one method for measuring the effect of the attack could be misleading, that's the reason why often, along with the prediction shift, another method is used, a target *hit ratio*. The second method is calculated by creating top-n predictions for every profile in the system before and after the attack, and then count for how many users has the target item occurred in the top-n recommendations. The influence of the attack can then be seen by comparing values of the target hit ratio before and after the attack.

C. Ensemble model

For the approach proposed in this model, a basic idea was to build independent classificator for each type of the shilling attacks and then create an ensemble model from all of those clasificators and implement single-vote voting classification, where it is enough if one of the classificators would mark the profile as fake, for the ensemble model to classify the profile as fake. This idea has some shortcomings, major one is that it is impossible to cover all the attacks with different classifiers and for each attack out there train a specific classifier, so after closer examination and experimenting, we can notice that it is possible to group some attacks together. For example, random attack is just a special case of the bandwagon attack when k is set to 0, and both variants of push and nuke random attacks can be put in that same group.

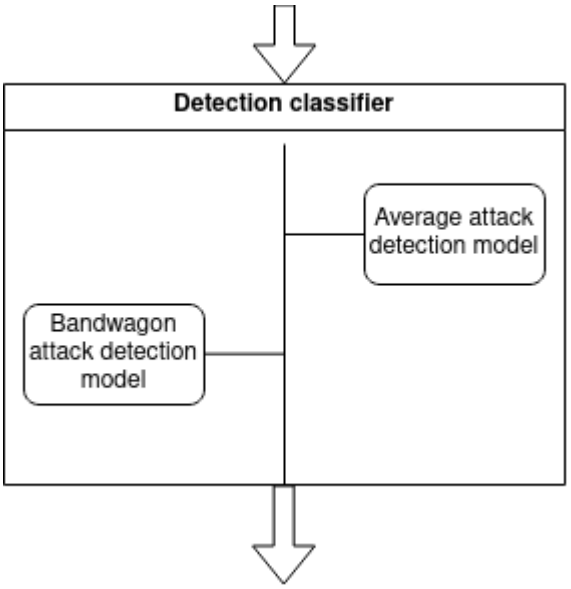


Fig. 1: Architecture of proposed classifier model with single-vote voting classification

Also, using the same logic, both push and nuke average attacks can be put together in the same group and only one classifier will be enough to detect fake profiles. For each group of the attacks, the random forest [16] classifier is used to detect fake profiles, and after grouping the attacks mentioned in this paper, only two detection classifiers are needed. Each random forest classifier is trained independently, without knowing about the existence of the other classifier(s).

Features mentioned above were used to train random forest clasifiers, and classifiers have been trained using the cross-validation [17] to pick the best model for that group of the attacks and avoid overfitting.

Architecture of the proposed model can be seen in the picture 1.

IV. RESULTS

To measure the efficiency of the attack detector, we have picked a standard measure, the F_1 score, which is the harmonic mean of the precision and recall.

As a base testing and training dataset, user ratings from MovieLens 100k [18] were used. The dataset was split in 70-30 train-test ratio, and then two different types of training datasets were created. The first type of datasets contained random push and nuke attacks, and bandwagon attacks (push attack). The second type of datasets contained push and nuke average attacks.

For all the created attacks, filler items size and attack size were varied for each attack. This was done to improve generalization of the detector and to make it as robust as possible when detecting attacks. For filler set and attack sizes, sizes of 1%, 3%, 5%, and 10% were selected in order to create attack datasets, and then all of those attack

	F1	Precision	Recall
Test set	0.91	0.83	1.0

TABLE II: Results F_1, recision, and recall scores for the proposed detection model

datasets were combined with base dataset to create final training set.

The training set for each model contained 25% of authentic profiles, and 75% of fake profiles coming from different variations of attacks. That decision made the training set unbalanced, but it was chosen to put more emphasis on recall since we wanted to create a classifier which would detect as much as possible fake profiles to not let the attack influence the recommender system at all. If the classifier trained on the training set like this one is used in the production, it should not delete the profiles marked as fake ones because it might result in authentic profiles being deleted from the system.

In table II F1, precision and recall scores are shown for the proposed model, it is possible to notice a very high recall, equal to 1, but this is the consequence of the detection system design. Recall of 1 is not unusual to see when measuring the efficiency of the recommender systems and in many papers [10] [19] [20] it is possible to find the recall values close or equal to 1.

Table III shows how each attack influences the same recommender system with and without attack detection model. It is possible to see how big of a difference relatively simple detection system can make and why it is needed to have a proper attack detection systems.

In table IV the comparison between proposed and two other models is shown. In [21] Bayes classifier is used to detect shilling attacks, while in [22] authors proposed unsupervised PCA model to detect shilling attacks. The [21] was selected for comparison because in 2018. it has outperformed state-of-the-art models, and [22] was selected because it shares multiple detection attributes with the model proposed in this paper.

Comparison between models is done on the test set in which attack profiles have filler set size set to 5%, and attack size is also set to 5%, and comparisons were made for the three different attacks: random attack, average attack and bandwagon attack. When comparing the models, proposed method has higher F_1 score for two out of three chosen attacks. It is worth mentioning that for some other attack parameters (different filler set size or attack size), Bayes classifier has greater F1 score than the proposed model, and that unsupervised PCA model has much higher F_1 scores when the attack size is greater (10% or 20%).

		Mean predicted rating	Hit ratio
Clear test set		3.16	6.9
Random attack (push)	Without detector	4.47	207.2
	With detector	3.18	6.7
Average attack (push)	Without detector	4.55	218.9
	With detector	3.17	6.8
Bandwagon attack	Without detector	4.49	206.4
	With detector	3.16	6.7
Random attack (nuke)	Without detector	1.74	0.0
	With detector	3.18	6.8
Average attack (nuke)	Without detector	1.62	0.1
	With detector	3.18	6.8

TABLE III: Display of the influence of the attacks on the recommender system with and without attack detector

	Proposed model	Bayes classifier	Unsupervised PCA
Random attack	**0.91**	0.89	0.76
Average attack	0.91	**0.95**	0.66
Bandwagon attack	**0.94**	0.86	0.75

TABLE IV: Comparison of F1 score between proposed model, Bayes classifier [21] and Unsuprivised PCA classifier [22] when both filler set and attack size are 5%

V. Conclusion and future work

Without the proper attack detection system, even the simplest attacks can have very high influence on the recommendations of the system. In this paper, novel approach has been proposed with the idea of grouping similar attacks together and training specific classifiers for that group of the attacks. The proposed method uses random forest algorithm to detect fake profiles for each group of the attacks, and has shown promising results.

For the future work, more groups of attacks should be added and some other classifier algorithms could be explored and tested.

References

[1] J. Bennett and S. Lanning, "The netflix prize," in *Proceedings of the KDD Cup Workshop 2007*. New York: ACM, Aug. 2007, pp. 3–6. [Online]. Available: http://www.cs.uic.edu/~liub/KDD-cup-2007/NetflixPrize-description.pdf

[2] B. Smith and G. Linden, "Two decades of recommender systems at amazon.com." *IEEE Internet Comput.*, vol. 21, no. 3, pp. 12–18, 2017. [Online]. Available: http://dblp.uni-trier.de/db/journals/internet/internet21.html#SmithL17

[3] Y. Koren and R. M. Bell, "Advances in collaborative filtering." in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer, 2011, pp. 145–186. [Online]. Available: http://dblp.uni-trier.de/db/reference/rsh/rsh2011.html#KorenB11

[4] I. Gunes, C. Kaleli, A. Bilge, and H. Polat, "Shilling attacks against recommender systems: A comprehensive survey." *Artificial Intelligence Review*, vol. 42, no. 4, 2014.

[5] P.-A. Chirita, W. Nejdl, and C. Zamfir, "Preventing shilling attacks in online recommender systems," in *Proceedings of the 7th annual ACM international workshop on Web information and data management*, 2005, pp. 67–74.
[6] F. Rezaimehr and C. Dadkhah, "A survey of attack detection approaches in collaborative filtering recommender systems," *Artificial Intelligence Review*, vol. 54, pp. 2011–2066, 2021.
[7] R. Bhaumik, B. Mobasher, and R. Burke, "A clustering approach to unsupervised attack detection in collaborative recommender systems," in *Proceedings of the International Conference on Data Science (ICDATA)*. Citeseer, 2011, p. 1.
[8] W. Zhou, J. Wen, M. Gao, L. Liu, H. Cai, and X. Wang, "A shilling attack detection method based on svm and target item analysis in collaborative filtering recommender systems," in *Knowledge Science, Engineering and Management: 8th International Conference, KSEM 2015, Chongqing, China, October 28-30, 2015, Proceedings 8*. Springer, 2015, pp. 751–763.
[9] L. Yang, W. Huang, and X. Niu, "Defending shilling attacks in recommender systems using soft co-clustering," *IET Information Security*, vol. 11, no. 6, pp. 319–325, 2017.
[10] C. Li and Z. Luo, "Detection of shilling attacks in collaborative filtering recommender systems," in *2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR)*. IEEE, 2011, pp. 190–193.
[11] J. Cao, Z. Wu, B. Mao, and Y. Zhang, "Shilling attack detection utilizing semi-supervised learning method for collaborative recommender system," *World Wide Web*, vol. 16, pp. 729–748, 2013.
[12] C.-Y. Chung, P.-Y. Hsu, and S.-H. Huang, "βp: A novel approach to filter out malicious rating profiles from recommender systems," *Decision Support Systems*, vol. 55, no. 1, pp. 314–325, 2013.
[13] Z. Yang, Z. Cai, and X. Guan, "Estimating user behavior toward detecting anomalous ratings in rating systems," *Knowledge-Based Systems*, vol. 111, pp. 144–158, 2016.
[14] S. K. Lam and J. Riedl, "Shilling recommender systems for fun and profit," in *Proceedings of the 13th international conference on World Wide Web*, 2004, pp. 393–402.
[15] R. Burke, B. Mobasher, C. Williams, and R. Bhaumik, "Classification features for attack detection in collaborative recommender systems," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2006, pp. 542–547.
[16] T. K. Ho, "Random decision forests," vol. 1, pp. 278–282, 1995.
[17] A. Ng, "Preventing "overfitting" of cross-validation data," *Proceedings of the Fourteenth International Conference on Machine Learning*, 01 1998.
[18] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *Acm transactions on interactive intelligent systems (tiis)*, vol. 5, no. 4, pp. 1–19, 2015.
[19] F. Zhang and Q. Zhou, "A meta-learning-based approach for detecting profile injection attacks in collaborative recommender systems." *J. Comput.*, vol. 7, no. 1, pp. 226–234, 2012.
[20] Q. Zhou and F. Zhang, "A hybrid unsupervised approach for detecting profile injection attacks in collaborative recommender systems," *Journal of Information & Computational Science*, vol. 9, no. 3, pp. 687–694, 2012.
[21] F. Yang, M. Gao, J. Yu, Y. Song, and X. Wang, "Detection of shilling attack based on bayesian model and user embedding," in *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 2018, pp. 639–646.
[22] F. Zhang, Z.-J. Deng, Z.-M. He, X.-C. Lin, and L.-L. Sun, "Detection of shilling attack in collaborative filtering recommender system by pca and data complexity," in *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 2. IEEE, 2018, pp. 673–678.

High-Performance Serverless Request Generator: Capable of Generating a Hundred Thousand Requests Per Second

D. Mileski* and M. Gusev**
* Innovation Dooel, Skopje, North Macedonia
** Sts Cyril and Methodius University in Skopje, Faculty of Computer Science and Engineering, Skopje, North Macedonia
E-mail: dimitar.mileski@ieee.org, marjan.gushev@finki.ukim.mk

***Abstract*—Depending on the scope, performance testing systems are distributed and scalable systems that can use many distributed instances to generate workload. Server-based systems used in performance testing can not cope with flexible workloads and specifications that may generate high costs for an extensive number of generated requests. Our goal was to develop a Virtual Patient Generator for testing the electrocardiogram streaming system for monitoring thousands of patients. To deliver a fully functional workload generator of a hundred thousand requests per second, in this paper, we developed a High-Performance Serverless Request Generator as an implementation of a Distributed Serverless Workload Generator concept. The system architecture comprises Serverless services, Pub/Sub, and Cloud Storage capable to deliver a cheaper version of a flexible number of requests, being scalable to an even higher extend. We present limitations, performance, and advantages by comparing the presented serverless load testing system with other workload generators available as SaaS or distributed load testing systems based on servers.**

Keywords— high performance computing; performance testing; serverless; FaaS; distributed systems; cloud computing; ECG

I. Introduction

The performance of scalable and elastic systems is tested with workload generators. Performance testing [1] is the process of evaluating a system or application's performance characteristics, such as responsiveness, stability, and scalability, under a variety of workloads. The goal of performance testing is to identify bottlenecks and other issues that may affect the system's ability to meet its performance requirements and to ensure that the system can handle the expected levels of usage. This can include testing the system's response time, throughput, and resource utilization under different loads and conditions, as well as simulating real-world usage scenarios.

The generation of a considerable workload requires the deployment of scalable, elastic, and high-performance systems. Distributed systems may be necessary to accommodate the demands of such evaluation and testing procedures. The most common performance tests types are:

- *Load testing* [2] is a type of performance testing used to evaluate a system or application's behavior under a specific load or user demand.
- *Stress testing* [1], also known as endurance testing, is a type of performance testing that is used to evaluate a system or application's behavior under extreme loads or conditions. The goal of stress testing is to identify the upper limits of a system's capabilities and to determine how the system behaves when it is pushed beyond its normal operational limits.
- *Soak testing* [1], also known as endurance testing or reliability testing, is a type of performance testing that is used to evaluate a system or application's behavior over an extended period of time. The goal of soak testing is to identify any issues that may arise after prolonged usages, such as memory leaks or resource exhaustion, and to ensure that the system remains stable and responsive over an extended period of time.
- *Spike testing* [1] is a type of performance testing that is used to evaluate a system or application's behavior under sudden and significant changes in load or usage. The goal of spike testing is to identify how the system behaves under unexpected and extreme conditions to ensure that the system can handle sudden and significant increases in load or usage without impacting performance.

During our research, we have addressed embarrassingly parallel problems and evaluated the corresponding speedup performance [3] experimenting with virtual machine approaches versus parallel serverless threads in the CardioHPC use case. CardioHPC [4] project aims at realizing an experiment and simulating real-time processing and monitoring of electrocardiogram (ECG)s. To evaluate the performance of the ECG Streaming system for monitoring 10,000 patients for the CardioHPC project, we developed a High-Performance Serverless Request Generator (HPSRG) capable of generating hundred thousand requests per second. Evaluation of serverless ECG stream processing in federated cloud implementations was also addressed in our earlier research [5]. We have already evaluated the automatic scalability of serverless services focusing on FaaS serverless [6] that served as the basis for the development of HPSRG.

Systems for performance testing are often server-based distributed systems, and serverless architecture can provide automatic scalability as an infrastructure for the devel-

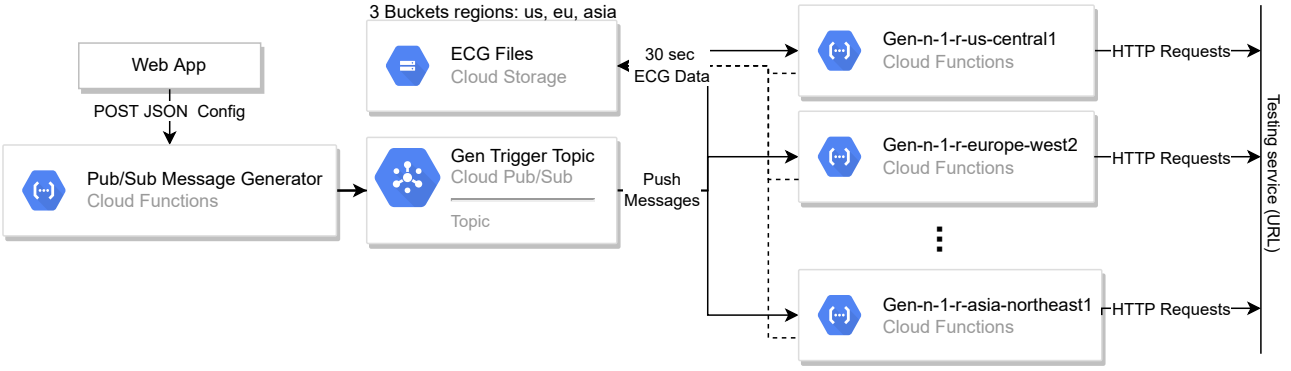


Fig. 1: System architecture: High-Performance Serverless Request Generator

opment of performance testing systems being able to generate such a large workload. Performance tests require flexible workloads and specifications, configured by a corresponding user interface.

Performance testing of CardioHPC and ECG Streaming system for monitoring of 10,000 patients uses the new HPSRG system that generates a workload in the form of JSON files containing ECG data. These are then streamed to a production system for monitoring of ECG streams [7]. HPSRG architecture can be implemented and deployed in any cloud that supports serverless services FaaS or serverless containers, object storage, and implementation of Pub/Sub - Publisher Subscriber model.

The paper follows the next structure. Related work and analysis of the state-of-the-art are presented in Section II. System architecture, Experiments, and Evaluation methodology are described in Section III illustrating the approach that will achieve generating a hundred thousand requests per second. Results are evaluated in Section IV. Finally, results are discussed in Section V and conclusions presented in Section VI along with future work directions.

II. State of the art

A variety of software and tools are used for testing the performance of different applications, most of which are server-based rather than serverless. A survey of performance testing software [8] focusing on well-known performance testing systems, such as WebLOAD [9] [10], LoadNinja [11] [12], HP Load Runner [13] [14], Apache J Meter [15] [14], Selenium [16] [14], NeoLoad [17] [10], WebLoad Professional [10], LoadUI [18] [10], WAPT [19] [20], Loadster [21] [10], and LoadImpact (Load Impact is now rebrended as k6) [22] [8], highlights the significance of performance testing, including scalability, crucial for accommodating increased numbers of users, and stability, ensuring the application works reliably at all times.

WebLOAD [9] [10], NeoLoad [17] [10], WebLoad Professional [10], and WAPT [19] [20] are load testing tools for performance testing of web applications, APIs, and mobile applications supporting a wide range of protocols, technologies, real-time monitoring and result analysis. These tools identify performance bottlenecks and scalability issues.

HP LoadRunner [13] [14] is a load testing tool developed by Hewlett Packard Enterprise (HPE) to simulate user traffic and monitor the performance of web applications and APIs, supporting multiple protocols and technologies, including various web, mobile, and cloud-based applications.

Apache J Meter [15] [14] is an open-source load testing tool that can simulate user traffic and monitor the web applications and API's performance. A wide range of supported protocols and technologies provide a highly extensible environment to create custom test cases and plug-ins.

Selenium [16] is an open-source testing framework that supports functional testing of web applications and APIs. It can also be used for load testing, but it is primarily used for functional testing.

SAAS load testing tools offer cloud-based performance and scalability testing for applications through simulated user traffic [23] [24]. Some well-known SAAS load testing tools are BlazeMeter [25] [26], LoadImpact [22] [8], LoadRunner Cloud [13] [26], Loader.io [27], Gatling FrontLine [28] [12], Neoload [17] [10], Flood.io [29], OctoPerf [30], LoadUI Pro Cloud [18] [8], and Sauce Labs [31] [32].

BlazeMeter [25] [26], LoadNinja [11] [12], LoadUI [18] [8], Loadster [21] [10], LoadImpact [22] [8], Gatling FrontLine [28] [12], Neoload [17] [10], Flood.io [29], OctoPerf [30], LoadUI Pro Cloud [18] [8], and Sauce Labs [31] [32] are all cloud-based load testing tools that evaluate the performance of websites, mobile applications, and APIs by simulating large amounts of user traffic, providing real-time monitoring and results analysis with drag-and-drop interfaces to create and manage load tests. LoadRunner Cloud [13] [26] is a cloud-based performance testing solution offered by HPE. Loader.io [27] is a cloud-based load testing platform that offers an easy-to-use interface for load testing websites and APIs.

The HPSRG system offers a distinct advantage over the conventional distributed testing systems through its use of serverless technology, which eliminates the need for server management and significantly reduces deployment

time. The utilization of a serverless pay-as-you-go system offers a more precise payment model as users are solely charged for resources used per request, while in the case of virtual machines, the users are charged for the entire duration of usage, even during periods when performance tests are not executed.

III. Methods

In this section, we present the system architecture of a new High-Performance Serverless Request Generator, and then specify the experimental and evaluation methodology.

A. Solution Architecture

HPSRG is a solution based on serverless technology, and the specific implementation in this paper is based on Google Cloud [33]. In implementing the solution architecture presented in Fig 1, we have utilized Google Cloud Functions Generation 1 [34] Function-as-a-Service (FaaS) offering, in conjunction with Google Cloud Storage [35], an object storage solution, and Google Cloud Pub/Sub [36], which represents an implementation of the Publisher-Subscriber model within the Google Cloud ecosystem.

HPSRG generates requests to a Testing service (invoked by a URL) as presented in Fig 1. The Web App allows users to input configuration parameters for HPSRG through a dedicated client form. Upon submission of the form, the Web App sends an HTTP Post request containing JSON Config to start the HPSRG.

Configurations for HPSRG include Pub/Sub Message Generator (PMG) execution time, Pub/Sub topic name, and Google Cloud project ID. The configuration also includes parameters for individual Generators (Gen) (eg: Gen-1-r-us-central1, Gen-n-1-r-europe-west2, ... Gen-n-1-r-asia-northeast1) such as a number of requests per second, execution time in seconds, and Testing service (URL).

The web application makes an HTTP POST request to FaaS serverless function, which initiates the Pub/Sub Message Generator with Cloud Pub/Sub Topic. All Gen generators (Gen-n-1-r-us-central1, Gen-n-1-r-europe-west2 ...) have a subscription to a topic named "Gen Trigger Topic" being a Push or Pull request in Google Cloud Pub/Sub.

Table I presents the configuration of each FaaS serverless function. The automatic scalability of a FaaS serverless function is an important feature of the HPSRG architecture. Performance testing will reveal a sufficient number of active FaaS instances for each generator. According to Google Cloud Function Generation documentation [34], one FaaS can activate up to 3000 instances and each instance will generate requests to Testing service (URL), currently a high limit of all other serverless providers.

Note that the Testing service (URL) in Fig. 1 is a no-op function, which is a dummy placeholder in computer programming that performs no action, usually used in situations where a real function is to be implemented later or when a function must be specified but no operation is required. Although, the Testing service (URL) is a FaaS deployed in the Google Cloud europe-west-1 region but it can be deployed anywhere since the request generator developed and tested in this paper can send requests to any service.

TABLE I: FaaS Config

	PMG	Generator	Testing service (URL)
Memory	512MiB	256MiB	256MiB
vCPUs	8	1	1
Min number of instances	0	0	0
Max number of instances	1	3000	3000
FaaS Timeout	540s	540s	5s

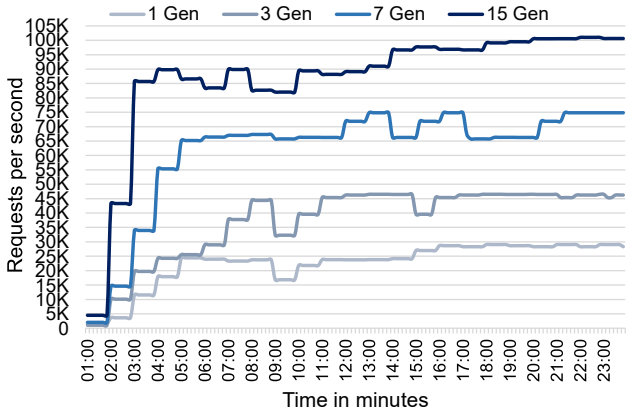
TABLE II: Cloud Regions

FaaS	Region
PMG	europe-west-1
Testing service (URL)	europe-west-1
1 Gen	europe-west-1
3 Gen	europe-west-1, asia-north-east-1, us-east-1
7 Gen	europe-west-1, asia-north-east-1, us-east-1 europe-central-2, asia-south1, us-west4, southamerica-east1
15 Gen	europe-west-1, asia-north-east-1, us-east-1, europe-central-2, asia-south1, us-west4, southamerica-east1, europe-west2, australia-southeast1, asia-southeast1, europe-west6, us-west1, us-central-1, asia-east2, northamerica-northeast1

B. CardioHPC Experiments

Our CardioHPC experiment aims at testing more FaaS serverless function generators: Gen-n-1-r-us-central1, Gen-n-1-r-europe-west2 ... Gen-n-1-r-asia-northeast1. In our use case, we use push messages, which means that when a message is published on that topic, everyone who has subscribed to that topic will receive the message following to the push principle. Generators (Gen) use ECG data prerecorded and stored in the Google Cloud Storage, as a type of Object Storage. Data in Google Cloud Storage is organized in buckets, and for the specific architecture we created three different buckets (ecg_files_eu, ecg_files_us, ecg_files_asia). The ECG data originates from the MIT-BIH [37] [38] ECG benchmark dataset, being resampled from 360 to 125 Hz and rescaled from 11 to 10-bit resolution. Each Generator selects a random ECG recording and sends a 30-second ECG signal (with 3750 ECG samples) to the specified Testing service (URL). The no-op function returns success or error based on the successful parsing of the JSON with ECG.

The HPSRG JSON Config parameters were experimentally determined to maximize the number of instances for Gen, then all 4 experiments were conducted with these parameters using a reasonable number of request generators, 1, 3, 7, and 15. Each experiment lasts 25 minutes, during which the number of requests stabilized after reaching a certain point. The deployment cloud regions of all Gen are listed in Table II.



Generators	req/s	req/min	req/h	req/day	req/month
1 Gen	30K	1.8M	108M	2.592B	77.76B
3 Gen	45K	2.7M	162M	3.888B	116.64B
7 Gen	75K	4.5M	270M	6.48B	194.4B
15 Gen	100K	6M	360M	8.64B	259.2B

Fig. 2: Number of Generators (Gen) and Requests per second hitting the Testing service (URL).

C. Evaluation methodology

To evaluate the performance of the service, various tests such as Load Tests, Stress Tests, Soak Tests, and Spike Tests require the HPSRG to generate a high volume of requests to the service. In our CardioHPC experiment we use the following evaluation metrics to test the generator:

- number of **Requests Per Second** sent to the Testing service (URL),
- number of **Active FaaS Instances** per Generator, which is used to assess if the maximum number of active instances specified by the cloud platform is being utilized,
- **cost** of running experiments on serverless services in the Google Cloud public cloud is the third evaluation metric.

The achieved results for these evaluation metrics are collected by the Google Cloud Monitoring service of Google Cloud Function Generation 1.

IV. Results

Fig. 2 presents the request rate per second directed toward the specified Testing service (URL) of the service being evaluated, as a function of varying numbers of Generators. The CardioHPC experiment was conducted for several distinct scenarios - test cases, and in this paper, for clarity of visual presentation, we present results only for a single generator ($n = 1$), $n = 3$, $n = 7$, and $n = 15$ generators.

Fig. 3 presents the active instances of each service for the HP serverless request generator system architecture (Fig. 1). Specifically, the number of active instances is provided for the following services:

- Pub/Sub Message Generator,
- one of the Generators (such as the Gen-n-1-r-europe-west1), and

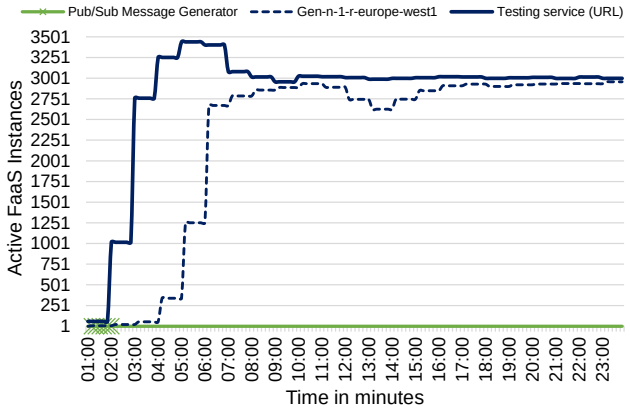


Fig. 3: Active FaaS Instances

- Testing service (URL)

V. Discussion

A. Performance analysis

The test cases within the CardioHPC experiment were conducted over a duration of 25 minutes. The initiation of the experiment (Fig. 3) for Gen-n-1-r-europe-west1 is specially marked. We observe an increase in the number of requests per second from zero to a stabilized value for all cases, upon activation of the instances from the Generators. The number of active instances increased from zero to the specified maximum of 3000, as outlined in the Google Cloud Function Generation 1 documentation.

Fig. 2 displays fluctuations of the number of requests per second until the value is stabilized, which can be attributed to the specified timeout of 540 seconds (9 minutes) for each of the Generators in Table I. This requires a new instance to be initiated upon termination, with a maximum running time of 540 seconds, as defined in the JSON Config.

These fluctuations can be mitigated through the utilization of Google Cloud Function Generation 2, which operates on the Google Cloud Run service with a maximum instance timeout of one hour. Additionally, oscillations may also occur due to network factors and the fact that the generator is a distributed system generating requests from various global locations (such as London, Las Vegas, Sao Paulo, Sydney, and Mumbai). The values are presented in terms of requests per second, but it may be more appropriate to consider them as requests per minute or hour, depending on the system being evaluated.

Fig. 2 presents the rate of requests per unit of time for different number of Generators within the HPSRG system. The values are indicated in the form of thousands, millions, and billions, with the respective units colour-coded in grey, orange, and red.

B. Limitations

The number of active instances of the Pub/Sub Message Generator (Fig. 2) shows only one active instance

which generates the messages to act as triggers for the FaaS Generators. The FaaS Generators, while generating requests for the Testing service (URL), exhibit a range of active instances from 2700 to 3000, with one instance falling below 2700. The Testing service (URL) represents the service under test, to which the HPSRG sends requests.

The stability of the number of active instances around 3000, which is the theoretical maximum according to the documentation, is noteworthy. Initially, the platform starts up to 3500 instances, which could be in response to sudden spikes in requests. However, the maximum number of instances in Google Cloud Function Generation 1 cannot be specified to be greater than 3000, despite the platform occasionally surpassing this number.

C. Scalability of HPSRG

Solutions to accommodate an increase in demand for higher rates of requests include:

- Increasing the number of Pub/Sub Message Generators and Gen Trigger Topic.
- Deploying new generators (ex: Gen-n-x-r-us-central1 ...) in the same or in other public or private clouds. Each generator uses the same codebase written in Python that can be deployed to other serverless services in other public or private clouds.

Potential limitations to scaling HPSRG, as suggested, could include restrictions from the public cloud provider in terms of resource consumption rates and quotas. However, it should be noted that such limitations are not inherent to the HPSRG architecture itself.

D. Advantages

The HPSRG system has an advantage over distributed testing systems by utilizing serverless technology, eliminating the need for server management, and reducing deployment time. It can handle any workload as HTTP Post requests and does not require the management of virtual machines to achieve high request throughput. However, HPSRG lacks result visualization tools and its Web App is only used for configuring test setups. Apache JMeter, on the other hand, is a more comprehensive system with a range of tools and plugins for visualizing performance testing results.

In the CardioHPC project, results from the HPSRG testing were obtained through the Google Cloud Monitoring [39] service and visualized using external tools specialized for ECG monitoring.

VI. Conclusion and Future Work

In this paper, we present the High-Performance Serverless Request Generator as a new solution for service performance testing. The HPSRG can generate a hundred thousand requests per second, millions of requests per minute, and billions of requests per day. This demonstrates the potential of the serverless model and automatic scalability in cloud-based performance testing. The HPSRG is exploited to evaluate the serverless ECG Streaming system for monitoring 10,000 patients in the CardioHPC experiment for real-time processing and monitoring of ECG streams for more than 10.000 patients simultaneously.

As a potential future work, we aim at offering the HPSRG as a software-as-a-service (SaaS) publicly available solution and support other configurable performance testing systems for a lower cost and scalable flexible increased throughput. It is essential to examine the operational cost of such a generator and determine if a private cloud that supports serverless services would be more advantageous.

Acknowledgement

The experiment "CardioHPC - Improving DL-based Arrhythmia Classification Algorithm and Simulation of Real- Time Heart Monitoring of Thousands of Patients" has received funding from the European High-Performance Computing Joint Undertaking (JU) through the FF4EuroHPC project under grant agreement No 951745. The JU receives support from the European Union's Horizon 2020 research and innovation program and Germany, Italy, Slovenia, France, and Spain.

References

[1] I. Molyneaux, *The art of application performance testing: from strategy to tools*. " O'Reilly Media, Inc.", 2014.

[2] S. Pradeep and Y. K. Sharma, "A pragmatic evaluation of stress and performance testing technologies for web based applications," in *2019 Amity International Conference on Artificial Intelligence (AICAI)*. IEEE, 2019, pp. 399–403.

[3] D. Mileski and M. Gusev, "Serverless implementations of real-time embarrassingly parallel problems," in *2022 30th Telecommunications Forum (TELFOR)*, 2022, pp. 1–4.

[4] M. Gusev, S. Ristov, A. Amza, A. Hohenegger, R. Prodan, D. Mileski, P. Gushev, and G. Temelkov, "Cardiohpc: Serverless approaches for real-time heart monitoring of thousands of patients," in *2022 IEEE/ACM Workshop on Workflows in Support of Large-Scale Science (WORKS)*, 2022, pp. 76–83.

[5] S. Ristov, M. Gusev, A. Hohenegger, R. Prodan, D. Mileski, P. Gushev, and G. Temelkov, "Serverless ecg stream processing in federated clouds with lambda architecture," *IEEE Computer*, vol. to be published, 2023.

[6] D. Mileski and M. Gusev, "Serverless faas scalability evaluation: An ecg signal processing use case," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 853–858.

[7] S. Nastic, T. Rausch, O. Scekic, S. Dustdar, M. Gusev, B. Koteska, M. Kostoska, B. Jakimovski, S. Ristov, and R. Prodan, "A serverless real-time data analytics platform for edge computing," *IEEE Internet Computing*, vol. 21, no. 4, pp. 64–71, 2017.

[8] N. Srivastava, U. Kumar, and P. Singh, "Software and performance testing tools," *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, vol. 2, no. 1, pp. 1–12, 2021.

[9] RadView, "Webload overview," Apr 2022, last accessed 06 February 2023. [Online]. Available: https://www.radview.com/webload-overview/

[10] M. Sharma, S. Vaishnavi, S. Sugandhi, and S. Abhinandhan, "A comparative study on load testing tools," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 2, pp. 1906–1912, 2016.

[11] SmartBear, "No learning curve. no correlations. just real-world accuracy." last accessed 06 February 2023. [Online]. Available: https://loadninja.com/

[12] A. Kołtun and B. Pańczyk, "Comparative analysis of web application performance testing tools," *Journal of Computer Sciences Institute*, vol. 17, pp. 351–357, 2020.

[13] MicroFocus, "Load testing software: Loadrunner professional," last accessed 06 February 2023. [Online]. Available: https://www.microfocus.com/en-us/products/loadrunner-professional/overview

[14] R. Abbas, Z. Sultan, and S. N. Bhatti, "Comparative analysis of automated load testing tools: Apache jmeter, microsoft visual studio (tfs), loadrunner, siege," in *2017 international conference on communication technologies (comtech)*. IEEE, 2017, pp. 39–44.
[15] Apache, "Apache jmeter," last accessed 06 February 2023. [Online]. Available: https://jmeter.apache.org/
[16] ThoughtWorks, "The selenium browser automation project," last accessed 06 February 2023. [Online]. Available: https://www.selenium.dev/documentation/
[17] Tricentis, "Tricentis neoload for enterprise performance testing," Feb 2023, last accessed 06 February 2023. [Online]. Available: https://www.tricentis.com/products/performance-testing-neoload
[18] SmartBear, "Ensure high performance apis in less time," last accessed 06 February 2023. [Online]. Available: https://www.soapui.org/tools/readyapi/api-performance-testing/
[19] SoftLogica, "Test the performance of web applications under load," last accessed 06 February 2023. [Online]. Available: https://www.loadtestingtool.com/
[20] S. Kundu, "Web testing: tool, challenges and methods," *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 2, pp. 1694–0814, 2012.
[21] I. Loadster, "Load stress testing for high-performance websites," last accessed 06 February 2023. [Online]. Available: https://loadster.app/
[22] G. Labs, "Open-source load testing tool for developers: K6 oss," last accessed 06 February 2023. [Online]. Available: https://k6.io/open-source/
[23] Q. Gao, W. Wang, G. Wu, X. Li, J. Wei, and H. Zhong, "Migrating load testing to the cloud: a case study," in *2013 IEEE Seventh International Symposium on Service-Oriented System Engineering*. IEEE, 2013, pp. 429–434.
[24] M. Yan, H. Sun, X. Liu, T. Deng, and X. Wang, "Delivering web service load testing as a service with a global cloud," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 3, pp. 526–545, 2015.
[25] I. Perforce Software, "The complete continuous testing platform," last accessed 06 February 2023. [Online]. Available: https://www.blazemeter.com/
[26] P. Memon, T. Hafiz, S. Bhatti, and S. S. Qureshi, "Comparative study of testing tools blazemeter and apache jmeter," *Sukkur IBA Journal of Computing and Mathematical Sciences*, vol. 2, no. 1, pp. 70–76, 2018.
[27] "Loader.io." [Online]. Available: http://docs.loader.io/api/intro.html
[28] G. Corp, "Gatling enterprise - load testing tool for business," Jan 2023, last accessed 06 February 2023. [Online]. Available: https://gatling.io/enterprise/
[29] "Flood io." [Online]. Available: https://guides.flood.io/overview-of-flood/what-is-flood
[30] [Online]. Available: https://doc.octoperf.com/
[31] S. L. Inc, "Cross browser testing, selenium testing, mobile testing," last accessed 06 February 2023. [Online]. Available: https://saucelabs.com/
[32] N. Kilinc, L. Sezer, and A. Mishra, "Cloud-based test tools: A brief comparative view," *Cybernetics and Information Technologies*, vol. 18, no. 4, pp. 3–14, 2018.
[33] "Google cloud documentation," last accessed 25 April 2023. [Online]. Available: https://cloud.google.com/docs
[34] "Google cloud functions generation 1 documentation," last accessed 25 April 2023. [Online]. Available: https://cloud.google.com/functions/docs
[35] "Google cloud storage documentation," last accessed 25 April 2023. [Online]. Available: https://cloud.google.com/storage/docs
[36] "Google cloud pub/sub documentation," last accessed 25 April 2023. [Online]. Available: https://cloud.google.com/pubsub/docs
[37] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE engineering in medicine and biology magazine*, vol. 20, no. 3, pp. 45–50, 2001.
[38] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
[39] Last accessed 25 April 2023. [Online]. Available: https://cloud.google.com/monitoring/docs

Future Dynamic Computing and User Profile Execution Considering User Migration in Social Media Contexts

A.A. Periola *, A.A. Alonge ** and K.A. Ogudo ***
* Cape Peninsula University of Technology/Electrical, Electronic and Computer Engineering, Cape Town, South Africa
** University of Johannesburg/Electrical and Electronic Engineering Technology, Johannesburg, South Africa
*** University of Johannesburg/Electrical and Electronic Engineering Technology, Johannesburg, South Africa
periolaa@cput.ac.za*,aalonge@uj.ac.za**,kingsleyo@uj.ac.za***

***Abstract* - Asymmetricity in user subscription aboard social media platforms leads to a case where emerging social media platforms have more users than existing and established social media platforms. This leads to a scenario where the computing capacity aboard the data centers of established social media platforms is unutilized or underutilized. In addition, emerging social media platforms are often capital constrained and require access to computing resources to host and execute user profiles. The presented research proposes a network architecture that enables emerging social media platforms to make use of the idle computing resources of established social media platforms. The power usage effectiveness and computing resource utilization are enhanced by 31.2% and (32.8-50.5)% on average, respectively.**

Keywords* - *cloud computing; social media networks; computing utilization; power usage effectiveness

I. INTRODUCTION

The use of computing platforms plays an important role in enabling service delivery by social media-related firms and providers. Data centers in cloud computing platforms host data that describe the status and profile of multiple subscribers. In addition, servers host computing resources enabling the execution of functionalities embodied in algorithms that deliver services to the user(s).

Therefore, the subscribers and users of social media providers (SMPs) execute functions that lead to resource usage and power consumption in servers that constitute data centers and cloud computing platforms. Currently, there are many SMPs that operate their own data center facilities. Examples of SMPs in this regard are LinkedIn, Meta (Facebook, Instagram, WhatsApp), and TikTok (which utilizes Oracle, Google, and Amazon Web Services). These SMPs host different and distinct interfaces and front-end platforms to interact with their users.

The independent operation of data centers by SMPs results in the use of power resources by each SMP to maintain user profiles and store user data. This manner of resource usage does not consider that SMPs may have profiles for the same individual in their cloud computing platform. Hence, SMPs consume a significant amount of resources in making data center services available. These resources are expended even when a profile is not being accessed by the associated user.

The current usage of resources i.e., power and computing by SMPs does not consider the active and inactive state of the user profile for a given SMP. However, existing research shows that there is an asymmetricity in the user subscription to different SMPs [5–7]. This implies that the data center of some SMPs host user profiles with high activity ratios while the data centers of other SMPs have low activity ratios (LAR). The occurrence of high activity ratio (HAR) and LAR.

The occurrence of HAR and LAR influences the power used in operating and cooling servers aboard data centers. This also results in an influence on power usage effectiveness (PUE) and computing resource utilization (CRU). In addition, it is important to limit inefficient resource usage. This occurs when an SMP with a reduced activity ratio profile hosts a user profile on a high PUE cloud computing platform. The research being proposed addresses this challenge.

The main contribution of the presented research is the design of a network architecture that enables emerging SMPs to utilize the computing resources aboard the data centers and cloud computing platforms of established SMPs.

The discussion in the paper addresses the design and presentation of a system focused on the emerging SMP. In this case, the emerging SMP has significant capital constraints and is unable to operate its own data center. In the proposed solution, emerging SMPs (EMSPs) do not have sufficient capital resources to engage in the ownership of cloud computing platforms comprising networked data centers. Nevertheless, EMSPs have an increasing number of onboarding users and subscribers. Established SMPs (ETSPs) have data centers and cloud computing platform but have a reducing subscriber profile. This reduction is being experienced due to the migration to EMSPs by a significant number of users. The PUE, and CRU of ETSPs with EMSP migration is also formulated and investigated.

The discussion here is organized in this manner: Related background work receives focus in Section II. The proposed network architecture is in Section III. The performance formulation and modelling are addressed in

Section IV, and Section V, respectively. The discussion is concluded in Section VI.

II. BACKGROUND WORK

Appel et al. [8] identify the high level of subscription to social media platforms by users. It recognizes social media as a communication channel. The disposition in [8] identifies that social media has a pervasive role in computing with focus on its influence in marketing. The discussion has not considered the variation in the number of users on different social media platforms.

Sawicki et al. [9] recognize the increasing usage and subscription to TikTok by users. The discussion recognizes the ubiquity of social media platforms. However, the supporting and underlying architecture and its operational details have not been considered. An analysis of the algorithm underlying TikTok has been presented in [10]. The analysis is conducted from the perspective of self – making and self–representation. The concept of networking receives consideration from the linking of self–representation and identity management. This link provides the capacity for social connection. The focus of the discussion in [10] is the analysis of user experience due to interaction with the TikTok platform.

Zulkifil [11] addresses the challenges associated with data privacy in TikTok. The concern in [11] recognizes the challenges associated with conforming to regulations on data sovereignty and retention. The realization of data sovereignty requires the design of data center network architecture to realize a distributed cloud with controlled data access [12–14]. The discussion in [14] examines the approach of managing data centers by Facebook, an SMP. The analysis of the efficient operation of Facebook data centers is done from the perspective of providing low latency services to Facebook users. The efficiency has not been considered with a view to examine the role of user immigration of emigration in relation to Facebook. This transition is implicitly and not explicitly considered. Hence, the context of resource over–provisioning with the aim of supporting a significant number of onboard users can be implied. However, the event of user immigration or emigration can be implied from the discussion in [15–17].

The transition from one SMP to another SMP as seen in [15–20] does not reduce or eliminate the need for social media providers to have operational data centers as a computing and communication base. However, new social media platforms may have an increasing subscriber profile without a corresponding data center/computing platform operational base. However, a declining number of users can be observed to be associated with an SMP having an established data center network. In addition, an emerging SMP with increasing subscriber profile may not necessarily have the required data center platform support base. For example, ByteDance and Mastodon are emerging social media platforms with a smaller data center operational base in comparison to social media platforms like LinkedIn (Microsoft), Twitter and Meta (Facebook). However, ByteDance and Mastodon have an increasing subscriber base. Hence, a scenario wherein EMSPs such as Mastodon and ByteDance (TikTok) can make use of the computing resources and assets of ETSPs like Facebook require additional research consideration.

The presented research alongside the proposed mechanism is deemed to be beneficial to the case of capital constrained social media platforms. Such platforms are unable to acquire large scale and distributed data centers and cloud computing platforms. This is due to significant capital constraints. It is recognized that such media platforms can make use of existing cloud computing service providers. However, such a scenario limits the choice of these emerging social media platforms in the absence of the proposed mechanism and associated network architecture. Furthermore, the presented research decouples the computing platform and the social media platform which has been closely coupled in the existing case as seen in the case of LinkedIn (Microsoft owned and uses Microsoft Azure), Meta (having own data centers), Twitter (having own data centers) and Telegram (having own data centers). The case of the Fediverse describes the context of the proposed computing paradigm. However, Fediverse based social media platforms such as Mastodon make use of the non–social media cloud computing platforms only [21]. The proposed research advances the case of the Fediverse by incorporating more client cloud computing platforms for the emerging social media platform.

III. PROBLEM DESCRIPTION

The context being considered is one in which there are multiple existing and established social media platforms. Each of these platforms has own data centers and users. The set of ETMPs is denoted α such that:

$$\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_A\} \quad (1)$$

The number of subscribers associated with the ath social media platform $\alpha_a, \alpha_a \boldsymbol{\epsilon}\, \boldsymbol{\alpha}$ at the epoch $t_y, t_y \epsilon\, t, t = \{t_1, \dots, t_Y\}$ is denoted as $N_1(\alpha_a, , t_y)$. The maximum and minimum number of subscribers associated with α_a are also dynamic and denoted as $N_{max}(\alpha_a, , t_y)$ and $N_{min}(\alpha_a, , t_y)$, respectively. The ath social media platform α_a requires more servers (data centers) when:

$$A_1 > B_1 > C_1 \quad (2)$$

$$A_1 = \sum_{y=1}^{p} \left(N_{max}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (3)$$

$$B_1 = \sum_{y=p+1}^{q} \left(N_{max}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (4)$$

$$C_1 = \sum_{y=q+1}^{Y} \left(N_{max}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (5)$$

The scenario in (2) is described as the case of high activity ratio. This arises because of a significant increase in the number of users. Server (data centers) experience a low activity ratio when:

$$A_2 > B_2 > C_2 \quad (6)$$

$$A_2 = \sum_{y=1}^{p} \left(N_{min}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (7)$$

$$B_2 = \sum_{y=p+1}^{q} \left(N_{min}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (8)$$

$$C_2 = \sum_{y=q+1}^{Y} \left(N_{min}(\alpha_a, , t_y) - N_1(\alpha_a, , t_y)\right) \quad (9)$$

The set of ESMPs is denoted β and given as:

$$\beta = \{\beta_1, \beta_2, \dots, \beta_B\} \quad (10)$$

The ESMP $\beta_b, \beta_b \in \beta$ is deemed to have an increasing number of users and does not have sufficient resources to acquire a server based and capacity. The ESMP β_b can utilize the server capacity of α_a when (6) holds true (low activity ratio of the ETMP). However, an architecture and solution to enable this has not received sufficient research consideration. This is the focus of the proposed mechanism and network architecture.

IV. PROPOSED SOLUTION

The proposed solution is presented in this section that has two aspects. The conceptual background for the proposed solution is considered in the first aspect. The second aspect discusses the proposed network architecture.

A. *Conceptual Background*

The proposed network architecture essentially increases the amount of potential computing resources that can be utilized by ESMPs. In the existing case, ETMPs such as Facebook, Twitter and Telegram have own data centers that enable the provisioning of service to users and subscribers. In this case, the ETMP and the computational resource base have a close coupling and are not separated. Instead, they co-exist as a single entity. The scenario of this close–coupling context is in Figure 1. In Figure 1, the social media platform and the computing platform are indistinguishable and are enclosed in the context of providing the social media service.

A second scenario is one in which the ETMP such as Twitter is perceived to be a server capacity organization. Such a perspective can be derived from the view of Elon Musk on Twitter being a server capacity provisioning company at its core as found in [22]. This implies that ETMPs can potentially provide server capacity and associated software to other companies. However, this aspect is yet to receive further research consideration. In this case, the social media platform–computing platform has a less close–coupling than the case of the close–coupling context and is presented in Figure 2.

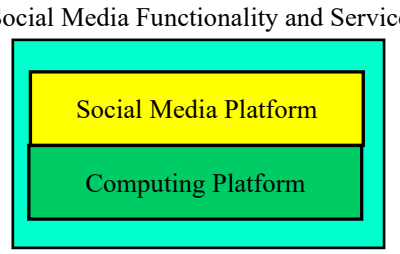


Figure 1. Relations between the social media platform and the computing platform in the closely coupled context of an ETMP.

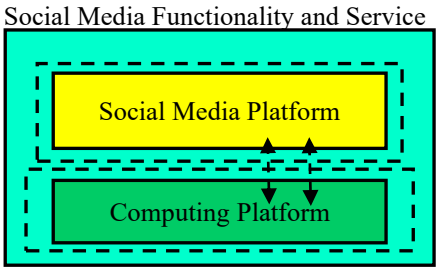


Figure 2. Relations between social media platform and the computing platform in a less closely coupled context of an ETMP.

The proposed case is one in which the computing platform provides services to other applications besides those related to social media platforms. In this case, the additional applications being considered is the ESMP. This is the context in Figure 3.

Figure 3 shows the case where the established social media provider (ETMP) and the emerging social media provider (ESMP) share the use of the computing platform as proposed. In this case, the computing platform is a data center owned and operated by the ETMP. However, a sharing of the computing platform has been presented in Figure 3 to indicate the operational context of the proposed solution. In this case, the ETMP can be an established social media platform such as Twitter. The ESMP is a social media platform which is newly established and in which the operators have significant capital constraints with limited capability of owning and operating computing platform.

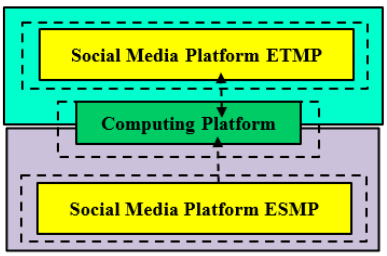


Figure 3: Relations between the social media platform and the computing platform in the considered context.

B. *Proposed Network Architecture*

The proposed network architecture considers that ESMPs have significant capital constraints thereby limiting their acquisition of data centers (significant number of networked servers). The ETMPs are not subject to these constraints in our consideration. The ESMP hosts entities enabling the monitoring and determining the occurrence of large user subscription. In addition, the ESMP hosts

computing entities that interact with the network. These computing entities send the server and computing resource requirement of ESMPs to a cloud-based database entity (CBDE).

The CRDE also hosts repositories that receives information on the availability of computing resources when ESMP have a low activity ratio. The CBDE has two main repositories i.e., the ESMP repository (ESMT) and the ETSP repository (ETMT). The ETST holds data on the computational resources (memory and networking capabilities) available from the ETSPs having a low activity ratio.

In realizing its functionality, the CRDE is hosted aboard multiple data centers in the cloud computing platform. The concerned entities i.e., the EMST and ETST interact with the ESMP and ETMP, respectively. In this case, the interaction is essentially between the gateway entities of ESMP and ETMP. The relations between the interacting entities enabling the sharing of computing platform between the ESMP (EMST) and the ETMP (ETST) are presented in Figure 4.

The aim and novelty of the architecture presented in Figure 4 is to describe how emerging social media providers and established social media providers determine the usage (request for and utilization of computing resources) alongside the execution of the required and associated communication.

In Figure 4, EMST and ETST access the cloud computing platform via own gateway entities. The gateway entity for the ETST and EMST are ETGW, and EMGW, respectively. Each of the EMGW and the ETGW engages in communication with the computing platform via the network entity (NE).

Figure 4 shows the relations between the ESMP, ETMP and the CBDE. The ESMP comprises the EMSP gateway (EMGW) and the computing resource determining entity (CRDE). In a similar manner, the ETMP comprises the ETMT gateway (ETGW) and the CRDE. The CRDE determines the required and idle computing resources by the ESMP and ETMP, respectively.

In Figure 4, the EMGW and ETGW communicate with each other via the cloud gateway (C–GW). This communication enables the ESMP to use available computational resources of the ETMP. The concerned computational resources are used to host data related to the ESMP user profile and execute algorithms associated with the delivery of different functionalities associated with the ESMP.

The C–GW incorporates a high-speed network with a low latency backbone enabling the rapid transfer of data in a bi- directional fashion. The exchange from the EMGW to the ETGW enables the ESMP to make use of accessible (idle) computing resources at the ETMP's cloud computing platform. In addition, an increase in the number of subscribers (users) of the ETMP necessitating an increased demand for computing resources that was previously suitable for use by ESMPs.

In addition, the C–GW engages in communications with external cloud computing platforms (data centers) via the external cloud gateway (ECGW). The ECGW is connected to other data centers whose use is deemed necessary when there is need to access additional computing resources.

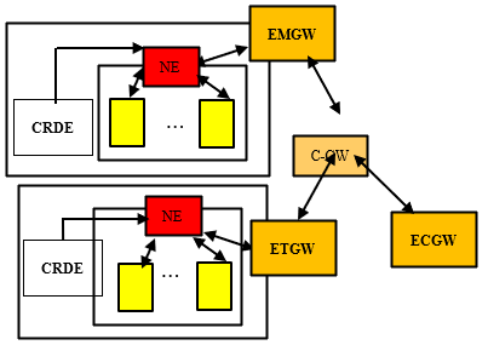


Figure 4: Relations between the ETMP and ESMP.

The CRDE determines the amount of computing resources and sends this to the concerned repository. Information in either repository i.e., ESMT and ETMT is sent to the network entity (NE). The NE interacts with the ESGW or the ETGW as deemed necessary. The flowchart for the architecture is presented in Figure 5.

In Figure 5, the occurrence of a low activity ratio (LAR) in the ETMP implies the low usage of computing resources aboard the concerned computing platform. This arises when (1) there is an overprovisioning of computing resources in the cloud platform, and (2) massive emigration of users due to the adoption of unfriendly policies. Emigration is deemed to have occurred when a user profile does not experience significant activity for a pre-determined duration. A high activity ratio (HAR) is deemed to have occurred when there is increasing usage of computing resources aboard a concerned computing platform. Such a case arises when an EMSP experiences increased subscription from many subscribers. In this case, there is a consistent increase in the utilization of computing resources aboard the cloud platform.

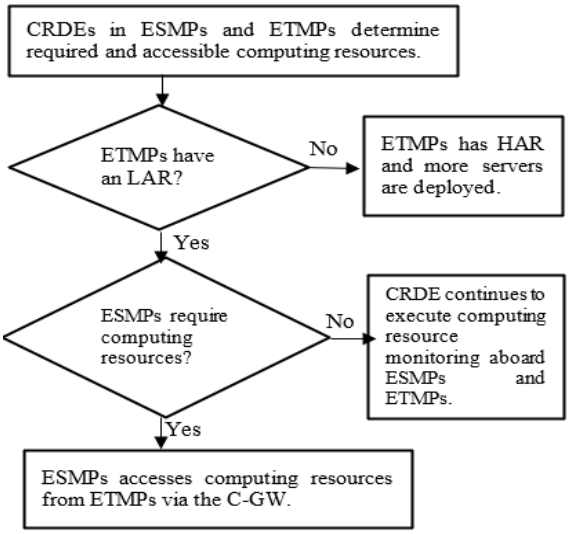


Figure 5: Flowchart showing the relations between ETMP, ESMP and associated entities.

V. PERFORMANCE FORMULATION

The formulation of the performance metrics i.e., the PUE and CRU is considered in this section. This is done for cases with and without the incorporation of the proposed architecture.

The performance formulation recognizes that there are multiple aspects associated with the description of a data center. However, the focus in the presented research is on formulating the PUE. In addition, the power components are those concerned with operating the server and the cooling system in the data center.

Let $\gamma(\alpha_a)$ be the set of servers in the ESMP datacenter α_a such that:

$$\gamma(\alpha_a) = \{\gamma_1(\alpha_a), \dots, \gamma_M(\alpha_a)\} \quad (11)$$

The servers in $\gamma(\alpha_a)$ are networked to realize a data center and computing platform.

Let $P_1(\gamma_m(\alpha_a), t_y), \gamma_m(\alpha_a) \in \gamma(\alpha_a)$ and $P_2(\gamma_m(\alpha_a), t_y)$ denote the power required to operate and achieve the cooling of the mth server of the ath ESMP at the epoch t_y, respectively. The activity state of $\gamma_m(\alpha_a)$ at the epoch t_y is denoted $I_A(\gamma_m(\alpha_a), t_y) \in \{0,1\}$. The server $\gamma_m(\alpha_a)$ is active and inactive at the epoch t_y when $I_A(\gamma_m(\alpha_a), t_y) = 1$ and $I_A(\gamma_m(\alpha_a), t_y) = 0$, respectively.

The PUE in the existing case θ_1 for multiple servers in the data center being used by the ath ESMP at multiple epochs can be given as:

$$\theta_1 = \sum_{m=1}^{M}\sum_{y=1}^{Y}\sum_{k=1}^{2}\frac{P_1(\gamma_m(\alpha_a), t_y)}{P_k(\gamma_m(\alpha_a), t_y)} \quad (12)$$

In (12), the numerator is a single term and describes the power required to operate the server (computing payload). The denominator describes the sum of the power required to operate and cool the server. Considering the influence of the activity status $I_A(\gamma_m(\alpha_a), t_y)$, θ_1 can be re–written as (within the context of for multiple servers in the data center being used by the ath ESMP) θ_1' and given as:

$$\theta_1' = \sum_{m=1}^{M}\sum_{y=1}^{Y}\sum_{k=1}^{2}\frac{P_1(\gamma_m(\alpha_a), t_y) I_A(\gamma_m(\alpha_a), t_y)}{P_k(\gamma_m(\alpha_a), t_y)} \quad (13)$$

$$\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| < \left|I_A(\gamma_m(\alpha_a), t_y = 0)\right| \quad (14)$$

The relations in (13) – (14) is for an ETMP with a low activity ratio (LAR). Hence, the number of idle epochs exceeds the number of active epochs. This is because more users are emigrating from the ETMP. In this case, more energy is expended storing and maintaining the profiles of users who are no longer onboard the ETMP. These are non–active user profiles that utilize cloud computing resources.

In the case of the proposed network architecture, the number of active epochs is increased as ESMPs make use of ETMP's servers (data centers).

The PUE θ_2 (formulated under the same context as (12)) is given as:

$$\theta_2 = \sum_{m=1}^{M}\sum_{y=1}^{Y}\sum_{k=1}^{2}\frac{P_1(\gamma_m(\alpha_a), t_y) I_A(\gamma_m(\alpha_a), t_y)}{P_k(\gamma_m(\alpha_a), t_y)} \quad (15)$$

$$\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| > \left|I_A(\gamma_m(\alpha_a), t_y = 0)\right| \quad (16)$$

In addition, the computing platforms of the ETMP can be utilized in a manner wherein the number of idle and active epochs are equal. In this case, the PUE θ_2' can be re–written as:

$$\theta_2' = \sum_{m=1}^{M}\sum_{y=1}^{Y}\sum_{k=1}^{2}\frac{P_1(\gamma_m(\alpha_a), t_y) I_A(\gamma_m(\alpha_a), t_y)}{P_k(\gamma_m(\alpha_a), t_y)} \quad (17)$$

$$\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| = \left|I_A(\gamma_m(\alpha_a), t_y = 0)\right| \quad (18)$$

The CRU metric is also formulated for the proposed and existing mechanisms. The CRU before and after incorporation of the proposed architecture are denoted as θ_3 and θ_3' , respectively.

$$\theta_3 = \frac{\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right|}{\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| + \left|I_A(\gamma_m(\alpha_a), t_y = 2)\right|} \quad (19)$$

$$\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| < \left|I_A(\gamma_m(\alpha_a), t_y = 0)\right| \quad (20)$$

θ_3' is the same as (19) but with the relation in (20) changed to $\left|I_A(\gamma_m(\alpha_a), t_y = 1)\right| > \left|I_A(\gamma_m(\alpha_a), t_y = 0)\right|$.

VI. PERFORMANCE EVALUATION

The performance evaluation is done for a context when there is an ETMP with a reducing number of subscribers. The ETMP has a datacenter comprising the multiple servers. An ESMP has multiple users and requires additional computing resources. The ETMP has a centralized cooling system comprising fans chillier, pipes, pumps, and capital-constrained a central management system. Two contexts are considered. In the first context, servers that are unutilized have their data moved to other servers and are switched off. In the second context, there are servers having a low utilization due to users migrating from the ETMP to the ESMP. The simulation parameters are in Table 1.

TABLE 1: PARAMETERS USED IN THE SIMULATION PROCEDURE

S/N	PARAMETER	VALUE
1	Total Number of Servers	40
2	Maximum Server Operational Power	380.6 W
3	Minimum Server Operational Power	1.78 W
4	Mean Server Operational Power	192.2 W
5	Maximum Proportion of Power used in Cooling	92.7%
6	Minimum Proportion of Power used in Cooling	3.3%
7	Mean Proportion of Power used in Cooling	48.1 %
FIRST CONTEXT		
8	Number of Active Servers	23
9	Number of Inactive Servers	17
SECOND CONTEXT		
10	Maximum Server Utilization	16%
11	Minimum Server Utilization	0.41%
12	Mean Server Utilization	7.8%

The results for the PUE, CRU (first context), and CRU (second context) as regards the performance simulation and evaluation are in Figure 6, Figure 7, and Figure 8, respectively. The results in Figure 6 show that the PUE exceeds a value of unity (ideal). This is because the simulation has not focused on improving the PUE. However, the proposed architecture has a lower PUE than the existing approach. Hence, the PUE performance is enhanced.

Analysis shows that the incorporation of the proposed mechanism and architecture enhances the PUE by an average of 31.2%. The use of the proposed mechanism improves the CRU as seen in Figures 7 and 8. Analysis shows that the incorporation of the proposed mechanism in the first context and second context enhances the CRU. The CRU is enhanced due to the use of the proposed mechanism in the first and second contexts by an average of 32.8% and 50.5%, respectively.

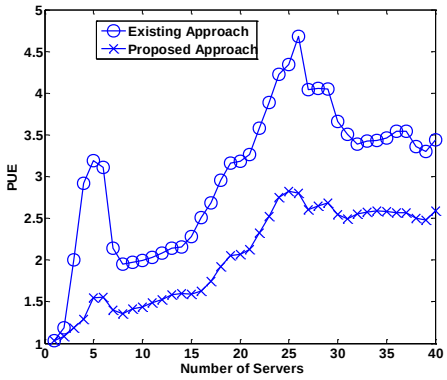


Figure 6: Performance evaluation results of the PUE obtained via simulation for the existing approach and proposed approach.

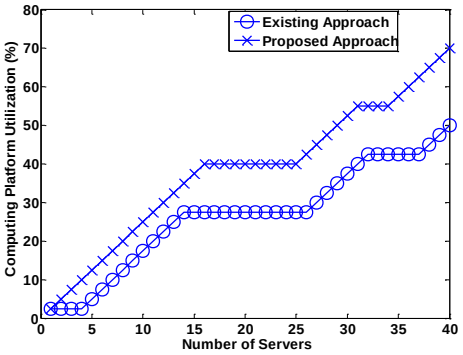


Figure 7: Performance evaluation results of the CRU in the first context obtained via simulation for the existing approach and proposed approach.

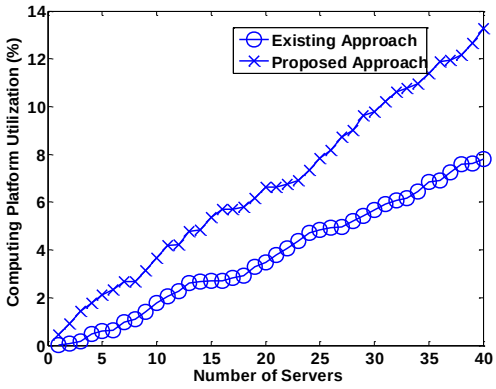


Figure 8: Performance evaluation results of the CRU in the second context obtained via simulation for the existing approach and proposed approach.

VII. CONCLUSION

The presented research recognizes that user emigration and immigration influence the resource utilization of computing resources in cloud computing platforms and data centers of social media platforms. It is recognized that this transition provides an opportunity for emerging social media platforms to make opportunistic use of idle and unutilized computing resources in the computing platforms of established social media platforms. Such a use of computing resources is beneficial to established social media platforms owning and operating cloud computing platforms i.e., data centers. In the performance analysis, the use of the computing resources in the proposed manner is noted to enhance the power usage effectiveness and the computing resource utilization. Furthermore, the established social media platforms can realize benefits by providing an additional service (for emerging social media platforms). However, it is noted that the security of information associated with user profile is a challenge. This can be addressed by ensuring that the emerging social media providers are hosted in a partition where there is no access to established social media provider operational related data. In addition, the proposed architecture enables emerging social media platforms with an increasing number of users to deliver services and required functionality to their subscribers. Furthermore, this sharing is shown to enhance the power usage effectiveness and computing resource utilization for computing platforms of established (existing) social media platforms experiencing user emigration to emerging social media platforms. Nevertheless, the concern of designing robust security systems for the concerned application is a subject of future research. Furthermore, the design of an intelligent network architecture considering subscriber profiles is also a subject of future work.

ACKNOWLEDGMENT

The support of the Cape Peninsula University of Technology and University of Johannesburg from the Departments of Electrical, Electronic and Computer Engineering, Electrical and Electronic Engineering Technology, respectively is acknowledged.

REFERENCES

[1] J. Saunavaara, A. Laine, and M. Salo, "The Nordic societies and the development of the data centre industry: Digital transformation meets infrastructural and industrial inheritance," Technology in Society, vol. 69, p. 101931, May 2022, doi: 10.1016/j.techsoc.2022.101931.

[2] M. Hashemi, "The Infrastructure Behind Twitter: Scale," blog.twitter.com, Jan. 19, 2017. https://blog.twitter.com/engineering/en_us/topics/infrastructure/2017/the-infrastructure-behind-twitter-scale

[3] S.Faramazi, H.H Tabrizi, and A. Chalak, 'Telegram: An Instant Messaging Applications to Assist Distance Language Learning (App Review)', Teaching English with Technology, 19(1), 132-147

[4] M. Octavianti, S. Mulyana, and D. Sugiana, "The Optimization of Telegram Bot in Improving the Cohesiveness of Bandung Raya Confero Community," Journal of Physics: Conference Series, vol. 1114, p. 012077, Nov. 2018, doi: 10.1088/1742-6596/1114/1/012077.

[5] A.Budree, K.Fietkiewicz, and E.Lins, 'Investigating usage of Social Media Platforms in South Africa', African Journal of Information Systems, Vol. 11, No. 4, 2019, pp 314 – 336.

[6] M. Limniou, Y. Ascroft, and S. McLean, "Differences between Facebook and Instagram Usage in Regard to Problematic Use and Well-Being," Journal of Technology in Behavioral Science, Oct. 2021, doi: 10.1007/s41347-021-00229-z.

[7] T. Aichner, M. Grünfelder, O. Maurer, and D. Jegeni, "Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019," Cyberpsychology, Behavior, and Social Networking, vol. 24, no. 4, Oct. 2020, doi: 10.1089/cyber.2020.0134.

[8] G. Appel, L. Grewal, R. Hadi, and A. T. Stephen, "The Future of Social Media in Marketing," Journal of the Academy of Marketing Science, vol. 48, no. 1, pp. 79–95, Oct. 2019, doi: 10.1007/s11747-019-00695-1.

[9] E. E. Zawacki, W. Bohon, S. Johnson, and D. J. Charlevoix, "Exploring TikTok as a promising platform for geoscience communication," Geoscience Communication, vol. 5, no. 4, pp. 363–380, Nov. 2022, doi: 10.5194/gc-5-363-2022.

[10] A. Bhandari and S. Bimo, "Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media," Social Media + Society, vol. 8, no. 1, p. 205630512210862, Jan. 2022, doi: 10.1177/20563051221086241.

[11] A. Zulkifli, "TikTok in 2022: Revisiting Data and Privacy," in Computer, vol. 55, no. 6, pp. 77-80, June 2022, doi: 10.1109/MC.2022.3164226.

[12] M. Lukings and A. Habibi Lashkari, Understanding Cybersecurity Law in Data Sovereignty and Digital Governance. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-031-14264-2.

[13] E. Celeste and F. Fabbrini, "Competing Jurisdictions: Data Privacy Across the Borders," Palgrave Studies in Digital Business & Enabling Technologies, pp. 43–58, Oct. 2020, doi: 10.1007/978-3-030-54660-1_3.

[14] K. M. U. Ahmed, M. H. J. Bollen and M. Alvarez, "A Review of Data Centers Energy Consumption and Reliability Modeling," in IEEE Access, vol. 9, pp. 152536-152563, 2021, doi: 10.1109/ACCESS.2021.3125092.

[15] S. Hong and S. K. Oh, "Why People Don't Use Facebook Anymore? An Investigation Into the Relationship Between the Big Five Personality Traits and the Motivation to Leave Facebook," Frontiers in Psychology, vol. 11, Jul. 2020, doi: 10.3389/fpsyg.2020.01497.

[16] M. Slot and S. J. Opree, "Saying no to Facebook: Uncovering motivations to resist or reject social media platforms," The Information Society, pp. 1–13, May 2021, doi: 10.1080/01972243.2021.1924905.

[17] S. Dang, "Exclusive: Twitter is losing its most active users, internal documents show," Reuters, Oct. 26, 2022. [Online]. Available: https://www.reuters.com/technology/exclusive-where-did-tweeters-go-twitter-is-losing-its-most-active-users-internal-2022-10-25/

[18] "Mastodon: Social media for people, your Twitter replacement | MPI," MPIWeb. https://www.mpi.org/blog/article/mastodon-social-media-for-people-your-twitter-replacement (accessed Dec. 13, 2022).

[19] C. Stokel-Walker, "Should I join Mastodon? A scientists' guide to Twitter's rival," Nature, Nov. 2022, doi: 10.1038/d41586-022-03668-7.

[20] L. La Cava, S. Greco, and A. Tagarelli, "Understanding the growth of the Fediverse through the lens of Mastodon," Applied Network Science, vol. 6, no. 1, Sep. 2021, doi: 10.1007/s41109-021-00392-5.

[21] D.Trautwein, A.Raman, G.Tyson, I.Castro, W.Scott, M.Schubotz, B.Gipp and Y.Psaras, '"Design and evaluation of IPFS," Proceedings of the ACM SIGCOMM 2022 Conference, Aug. 2022, doi: 10.1145/3544216.3544232.'

[22] "Elon Musk tells Twitter staff to commit to 'hardcore' culture or leave," Financial Times, Nov. 16, 2022. [Online]. Available: https://www.ft.com/content/082f3860-37fc-4e45-ab92-0da9ea3c508b

Brief Introduction to Active Noise Control

Dubravko Miljković
Hrvatska elektroprivreda, Zagreb, Croatia
dubravko.miljkovic@hep.hr

***Abstract* - This paper is a very brief introduction to the field of active noise control. The concept of active noise control based on destructive interference of sounds is presented. The need for high amplitude and phase accuracy of cancelling sound for efficient cancelation is shown. Considered is the possibility of noise cancelation in open and closed space. Quiet zone formed in the local noise cancellation related to the noise wavelength is illustrated. Most important approaches used for active noise control such as the feedforward, feedback and hybrid method are laid out. Adaptive feedforward and feedback structures based on adaptive filtering are outlined considering the influence of secondary path transfer function and acoustic feedback. Suitability of particular methods for different sound characteristics like broadband and narrowband noise and availability or absence of the reference signal is considered. Synchronized waveform synthesis is also presented as a solution for active control of periodic noise commonly produced by rotating machinery. General directions for digital implementations such as choice of sampling rate, adaptive filters length, computational cost, circuits for pre-filtering and post-filtering of reference, error, and secondary signal are given. Some typical applications are listed.**

Keywords - active noise control; feedforward; feedback; hybrid; waveform synthesis

I. INTRODUCTION

Quite often in living and work environments, one encounters annoying sounds. The common goal is to reduce noise to acceptable levels. Traditionally, noise reduction is achieved by passive methods, [1]. This involves the use of sound isolating and sound absorbing materials and reactive devices such as Helmholtz resonators or its combination like in reactive dissipative mufflers. However, passive methods are mainly effective at higher frequencies. Active Noise Control (ANC), [2]-[8] is a method for reducing of unwanted noise by generating an acoustic signal of same amplitude but the opposite phase that interferes with the originating noise. Terms Active Noise Cancellation and Active Noise Reduction (ANR) are also used. It is best suited for noise dominated by low frequencies, where passive methods are not efficient. History of developments in an ANC field is given in [9]. Operating range of active and passive methods is shown in Fig. 1. Passive and active methods should not compete but complement each other.

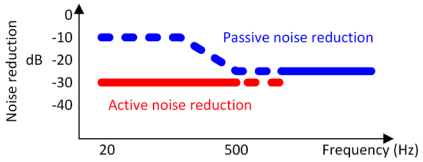


Figure 1. Operational range of active and passive noise reduction

II. ACTIVE NOISE CONTROL

Operation of ANC systems is based on the principle of destructive interference between sound fields of noise and controlling sound produced by the secondary source, [2].

A. *Destructive Interference*

Destructive interference is illustrated in Fig. 2. Primary noise and cancelling sound from the ANC system should cancel each other. Noise and cancelling sound should be of equal amplitude, but the opposite phase, [2]-[8].

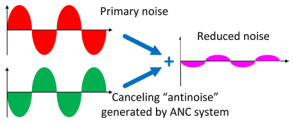


Figure 2. Destructive interference of acoustic signals

B. *Active Noise Control System*

The ANC system consists of these main components: acoustic sensors (microphones), acoustic actuator (speaker) and controller, [2]-[8], Fig. 3. Noise is picked by a reference microphone and processed by a controller that produces a secondary signal for the actuator. For adaptation controller uses residual noise signal from the error microphone. Important issues are the positions of the reference and error microphone, canceling loudspeaker and adaptation process.

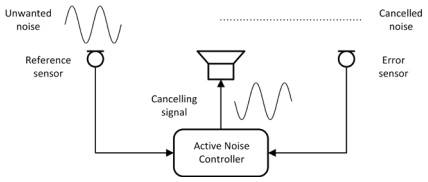


Figure 3. Principle of Active Noise Control

C. *Zone of Silence*

Significant noise cancelation by a typical ANC system is achieved in a localized region around error microphone, called zone of silence or quiet zone, Fig. 4. Dimension of zone is λ/10, where λ is the wavelength of a noise. Hence, it is easier to form a sizable zone at lower frequencies, [2,4].

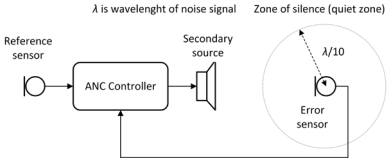


Figure 4. Zone of silence

D. *Required Accuracy – Amplitude and Phase Precision*

The ANC canceling signal must be very precisely determined to achieve significant noise reduction. If noise pressure is described by (1) and canceling sound pressure by (2), its vector sum is given by (3), where *A* is amplitude,

$$p_1 = Asin(\omega t) \quad (1)$$
$$p_2 = -(A + a)sin(\omega t + \theta) \quad (2)$$

α amplitude error and θ phase error of canceling sound, [2]. Data in Table I (dB) determined by (3) illustrates how small amplitude and phase deviations of ideal cancelling signal quickly diminish the noise reduction of ANC system, [2].

$$\Delta = 10log\left[4\left(1 + \frac{a}{A}\right)sin^2\frac{\phi}{2} + \left(\frac{a}{A}\right)^2\right] \quad \text{(dB)} \quad (3)$$

TABLE I. NOISE REDUCTION VS. CANCELING SIGNAL ACCURACY

a/A	Φ (°)						
	0	1	3	10	20	30	36 (λ/10)
0	∞	35.0	25.5	15.2	9.2	5,7	4,2
0,005	46	33.7	24.4	15.1	9.2	5,7	4,2
0,05	26	25.7	22.6	14.2	9.0	4,8	3,9
0,10	20	19.8	18.6	13.6	8.6	4,0	3,7
0,20	14	14.0	13.6	11.0	7.6	2,8	3,4

E. Cancellation of Periodic and Broadband Noise

Periodic noise consists of periodically or quasi-periodically distributed repetitive waveform patterns. It has a harmonic structure and is generally deterministic. An example of such noise is a sound field generated by rotating machinery, engines, electricity humming noise and devices driven beyond their capability. Broadband (or wideband) noise, has sound energy that is distributed over a wide part of the audible range as opposed to narrowband noise. Such noise is often stochastic, has a wide range of wavelengths including short ones and as such is much more difficult to attenuate. In the real world, both types of noise come together with various relative contributions. Broadband noise superimposed on periodic noise is shown in Fig. 5.

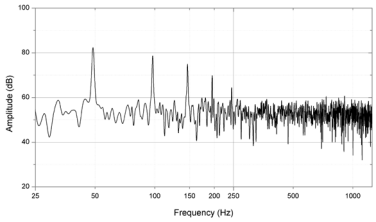


Figure 5. Tonal noise with superimposed broadband noise

F. Cancellation at Source

The preferable method of noise reduction is the reduction of noise at source, as shown in Fig. 6. [10]. Although, not cancelled at the very point of generation, noise at the output of exhaust duct is reduced before radiated into the environment. However, in most cases, cancelation at (or close to) source is not possible and reduction is achieved in small target zones of silence.

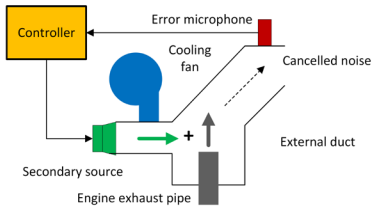


Figure 6. Noise cancellation at source

G. Active Noise Control in Closed and Open Space

1) ANC In Small Closed Space

It is relatively easy to reduce noise within a small closed space like an ear cup of a headphone, Fig. 7, as one doesn't have to deal with spatial sound considerations, [11].

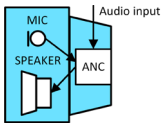


Figure 7. ANC in closed small space (headphones)

Dimensions of an ear cup are small, and zone of silence related to $\lambda/10$ is quite effective up to mid frequency range.

2) ANC Within a Large Closed Space

This task is much more complicated and require more secondary sources. One way is to construct numerous smaller zones of silence at points of interest within a space. Another way is to position secondary sources in corners of the space to counteract acoustical modes, [4].

3) ANC in Open Space

In open space it is possible to use an active acoustic barrier that reduces noise at its path, [2,12,13]. Active control of broadband sound through the open aperture of a full-sized domestic window is illustrated in Fig. 8, [12].

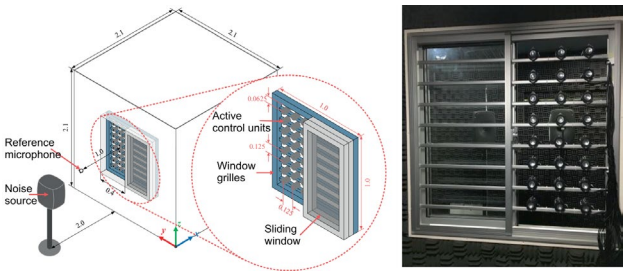


Figure 8. ANC applied to window, [12]

III. ADAPTIVE DIGITAL FILTERING

With the exception of some ANC headphones, almost all ANC implementations today are digital. Algorithms for ANC vary for the particular cancellation strategy, but generally use adaptive digital filters. Block diagram of finite impulse response (FIR) filter, [14] is shown in Fig. 9.

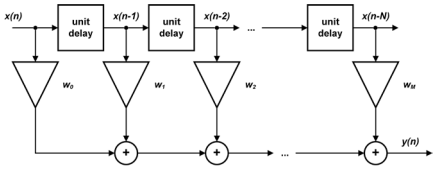


Figure 9. FIR filter

Output of the filter $y(n)$ is given by (4), where w_i are filter coefficients, $x(n)$ input signal, and N filter length.

$$y(n) = \sum_{i=0}^{N-1} w_i\, x(n-i) \quad (4)$$

There also exists infinite impulse response (IIR) filter that is a recursive and is more complicated. Its output depends on the current and previous inputs and previous outputs. However, FIR filters are more widely used. An adaptive filter is a filter with self-adjusting characteristics. The filter coefficients are adapted according to an adaptation algorithm. The most common algorithm is the Least Mean Square (LMS) algorithm, [14], Fig. 10. Adaptation of filter coefficients w_i at the moment n+1 is given by (5) and (6), where μ is a step size, $x(n)$ signal, $y(n)$ filter output, $d(n)$ desired filter output and $e(n)$ and error at the moment of the sample n.

$$e(n) = d(n) - y(n) \quad (5)$$
$$\mathbf{w}_i(n+1) = w_i(n) + \mu x(n)e(n) \quad (6)$$

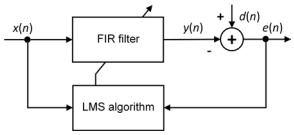


Figure 10. FIR filter with adaptive algorithm

Adaptive filters are relatively simple, but their real-time adaptation requires significant processing resources.

IV. Active Noise Control Structures

The active noise control structures are classified as feedforward, feedback and hybrid structure, [2]-[8].

A. Feedforward Active Noise Control

The feedforward ANC approach uses a reference microphone to pick up the noise signal. It also uses the error microphone to measure the residual noise that remains after cancellation. This error signal is used for adaptation in ANC controller.

1) Broadband Feeforward Active Noise Control

This approach can be used for cancelling both of periodic and broadband noise. Simplified block diagram of feedforward system, [2]-[8] is shown in Fig. 11.

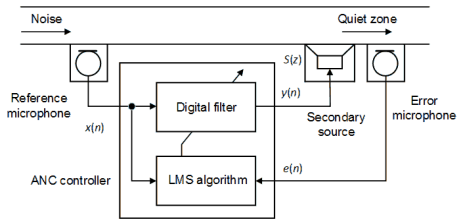


Figure 11. Broadband feedforward ANC

2) Synchronized Waveform Generator

This system belongs to the category of feedforward systems. It doesn't use the reference microphone. Instead, tacho signal is picked up by optical or magnetic sensor, Fig. 12. As there is no reference microphone, there is no acoustic feedback and the need for its cancelation. The system is capable of canceling a periodic components of narrowband noise while random broadband component is not canceled. It is suitable for noise with pronounced harmonic content. Waveform samples are stored in a waveform table, (7), [4], L is the period length. These samples are sent in a sequence to the D/A converter to produce a secondary signal.

$$\{w_l(n), l = 0, 1, \dots, L-1\} \tag{7}$$

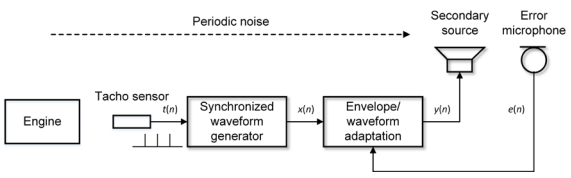


Figure 12. ANC using synchronized waveform generator

B. Feedback Active Noise Control

The feedback ANC doesn't have the luxury of reference signal picked by the reference microphone. Instead, it tries to reconstruct the reference signal using the feedback predictor, [2]-[8], as shown simplified in Fig. 13. It can attenuate only the predictable components of the noise. In non-adaptive ANC, filter is fixed. Adaptive system is more

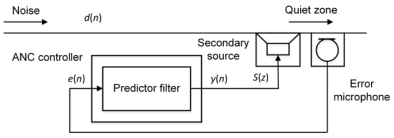


Figure 13. Feedback ANC

complicated and reference signal is estimated as a sum of filtered output and error signal (more details in subsection *D*). The feedback ANC is less robust than feedforward ANC. However, it is more compact and cost effective.

C. Hybrid Active Noise Control

Hybrid ANC combines feedforward and feedback control structures, Fig. 14. The feedforward ANC cancels noise correlated with a reference signal. The predictable component of the residual noise that is not present in a reference signal is canceled by the feedback ANC, [4]. Lower order filters may be used in hybrid implementation.

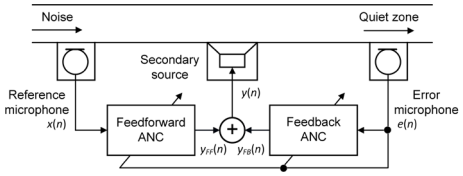


Figure 14. Hybrid ANC

D. Secondary Path Modeling

Operation of the ANC system is influenced by the transfer function of the secondary path $S(z)$. This includes D/A converter, reconstruction low pass filter, amplifier, secondary source, acoustic path between the secondary source (sound propagation delay), an error microphone, it's preamplifier, antialiasing filter and A/D converter, [4]. Frequency and phase characteristic of a typical speaker is shown in Fig. 15 and of a typical microphone in Fig. 16. It is necessary to compensate the influence of the $S(z)$, [2,15].

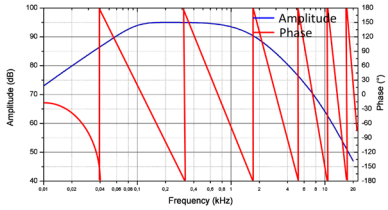


Figure 15. Typical speaker frequency and phase response

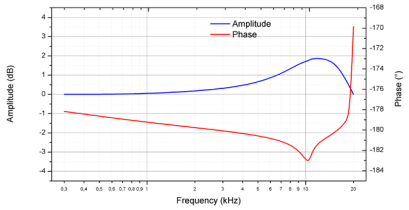


Figure 16. Typical microphone frequency and phase response

To address the influence of the secondary pat, the Filtered-x LMS (FxLMS) algorithm is developed. Instead of signal x, its filtered version x' is used produced by the filter with transfer function $\hat{S}(z)$ that approximates the transfer function of the secondary path $S(z)$, as shown in Fig. 17 for feedforward system and Fig. 18 for feedback system. Adaptation is given by (8), [4]. Estimated reference signal for use in feedback ANC is given by (9), [4]:

$$w_i(n+1) = w_i(n) + \mu x'(n) e(n) \tag{8}$$

$$x(n) \equiv \hat{d}(n) = y'(n) + e(n) \tag{9}$$

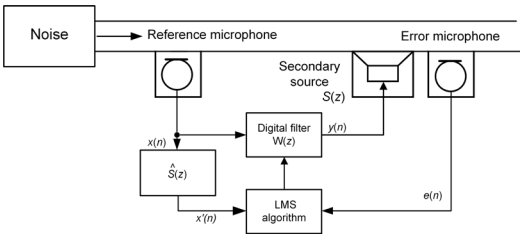


Figure 17. Feedforward ANC using FxLMS algorithm

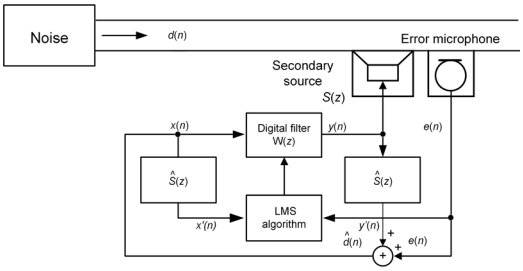


Figure 18. Feedback ANC using FxLMS algorithm

If $S(z)$ doesn't change a lot, it can be estimated off-line, Fig. 19. It is also possible to determine $\hat{S}(z)$ online, by use of additional random noise that is injected to a system after W(z), [4]. Additional LMS algorithm is used to adjust $\hat{S}(z)$.

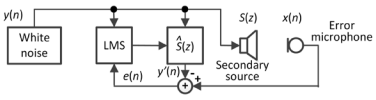


Figure 19. Off-line estimation of secondary path

E. Feedback Cancelation

Canceling "antinoise" signal from the secondary source propagates upstream to the reference microphone and corrupts the reference signal, resulting in an undesired acoustic feedback that may cause instability, Fig. 20. It may be necessary to add a special feedback neutralization filter with transfer function $\hat{F}$ for counteracting the feedback, [4].

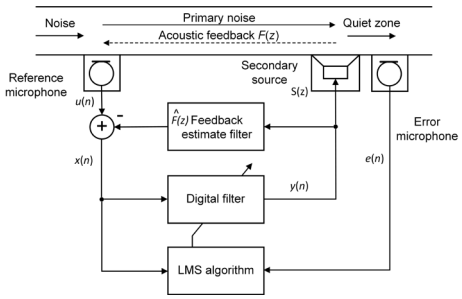


Figure 20. Feedback cancellation

F. Multi-Channel Active Noise Control

Multi-channel ANC employs multiple secondary sources, multiple error sensors and sometimes even multiple reference sensors, as illustrated in Fig. 21. Such system can create multiple zones of silence and is feasible for the attenuation of low-frequency noise fields in three-dimensional space. It tries to minimize the total potential acoustic energy across the space as picked by M error microphones, by minimizing cost function ξ, (10), [4].

$$\xi = \frac{V}{4\rho c^2 M} \sum_{m=1}^{M} {p_m}^2 \tag{10}$$

Fx-LMS algorithm for this system can be expressed in matrix notation, (11),(12). $\boldsymbol{E}$ is error signal, $\boldsymbol{D}$ noise, $\boldsymbol{S}$ secondary path, $\mathbf{W}$ adaptive filter and $\boldsymbol{I}$ identity matrix, [4].

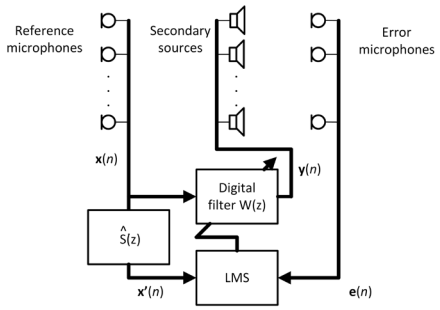


Figure 21. Multi-channel ANC

$$\boldsymbol{E} = \boldsymbol{D} - \boldsymbol{SW} \tag{11}$$

$$\boldsymbol{W}(n+1) = (\boldsymbol{I} - \boldsymbol{\mu S}^H \boldsymbol{S})\boldsymbol{W}(n) + \boldsymbol{\mu S}^H \boldsymbol{D} \tag{12}$$

G. Adaptive-Passive Noise Control

Adaptive-passive (semi-active) noise control uses passive devices whose parameters can be varied to achieve optimal noise attenuation over a narrow band of operating frequencies. Helmholtz resonators (passive devices) are used to used to damp/trap narrow frequency band of noise. They can be, within limits, actively tuned to a different frequency. Two approaches can be used, one, mechanically complex, that modify the geometry of the Helmholtz resonator using the stepping motor, and another that use time-varying feedback Proportional-Integral (PI) controller to tune the natural frequency and damping, [16], Fig. 22.

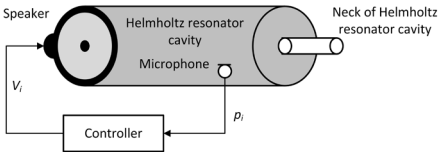


Figure 22. Adaptive Helmholtz resonator

V. IMPLEMENTATIONS

Diagram with the steps for ANC implementation is shown in Fig. 23. Noise from the reference and error sensor is amplified by preamplifiers, filtered by low pass filters and fed to A/D converters. Digital Signal Processor (DSP) implements an ANC algorithm. Values for secondary signal are fed to D/A converter. The analog signal is filtered by a low pass filter and amplified by an amplifier that drives the secondary source. One DSP board with necessary elements for ANC implementation is shown in Fig. 24, [17]. More modern DSP solutions exist, as well as micro-controllers with integrated low latency ADCs and DACs, [18].

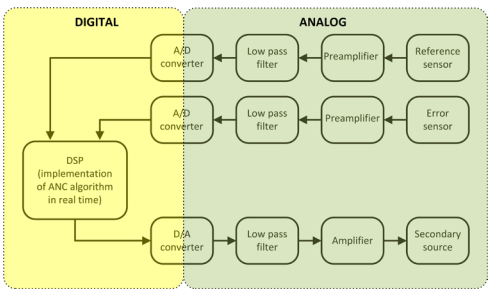


Figure 23. Block diagram of ANC implementation

A. Sampling Rate

Signal processing systems, including ADCs and DACs, must conform to the Nyquist criteria, which states that the signal sample rate must be at least two times greater than the highest frequency component present in the signal.

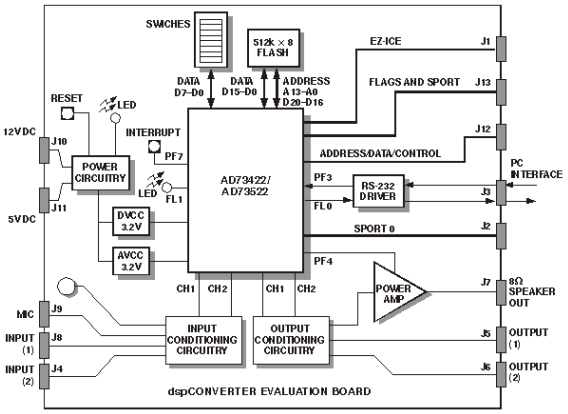


Figure 24. Example of DSP board suitable for ANC (Analog Devices)

However, it is recommended that sampling rate is ten times the target frequency. ANC systems are effective for low frequency noise, hence the required sampling rate may remain low, with manageable processing requirements.

B. *Filter Length*

How long FIR filter to use for $W(z)$ and $S(z)$ in FxLMS implementation? Start with 4-20 taps per tone/harmonic, up to several hundred for cancelling broadband noise, [19]. More taps provide better accuracy, but slows adaptation.

C. *Computational Cost*

Multichannel FxLMS ANC system with I reference inputs, J outputs, K error inputs, L taps for control filter and L_S taps for secondary path estimate must perform the N multiplications and additions per sample, (13), [20].

$$N = IJL + IJK(L+1) + IJKL_S \tag{13}$$

D. *Filtering of Reference, Error and Output Signal*

Input and output signals should be band-limited by anti-aliasing low pass filters. Classical IC designs with resistor-capacitive (RC) filters are sensitive to component tolerances and temperature stability. The switched capacitor filter allows for very accurate and tunable analog filter to be manufactured without using resistors, [21].

E. *Loudspeaker Sound Ouput*

Speaker used as a secondary source has its sensitivity, a measure of decibel output when supplied with a specified amount of amplifier power. The achieved dB output from an amplifier and speaker combination should match the dB level of the offending noise.

F. *Environmental Influences*

Many practical ANC systems must endure significant vibrations, temperature, humidity and chemical (acid) influences (e.g. engine exhaust mufflers).

VI. APPLICATIONS

There is a large gap between academic research and effective practical applications. Practical robust ANC implementations that can be commercially exploited are difficult and require synergy of numerous skills: electronics, software, transducer technology and physical acoustics. Key to success is picking applications where it can work.

A. *Active Noise Control Headphones*

The canceling signal is transmitted from the secondary source within the ear cup producing the zone of silence around the ear drum. Analog solution is shown in Fig. 25.

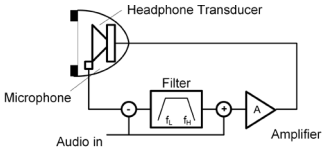


Figure 25. ANC headphones (analog)

There exist open and close loop systems, [11] with microphones positioned outside or inside the ear cup.

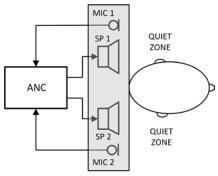


Figure 26. Active headrest

B. *Active Noise Control for Headrests*

The concept can be extended to a headrest, Fig. 26 with secondary sources integrated in a vehicle headrest, [22].

C. *Ventilation Ducts/Heat Ventilation Air Conditioning*

Ventilation ducts and Heat Ventilation Air Conditioning (HVAC) systems must blow enough air through the air ducts. Single channel broadband feedforward ANC system can be efficiently used in ventilation duct, [23]-[25], Fig. 27. There exist two different approaches to installing an ANC in duct applications: (a) microphones and the loudspeaker installed centrally inside the duct, and (b) microphones and the loudspeaker in a separate modules outside the duct.

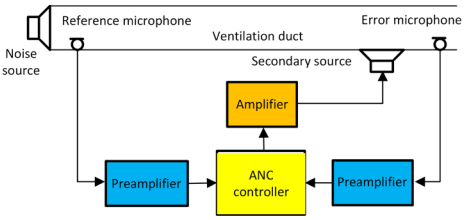


Figure 27. Ventilation duct ANC

D. *Car Interiors*

This includes silencing of car interiors and active engine exhaust mufflers for improved passenger comfort. Car interior noise is affected by engine, road, wind, HVAC and exhaust noise. Ford's ANC solution available on Explorer Limited Hybrid uses strategically placed microphones throughout the cabin to pick up unwanted noise, Fig. 28, [26]. System produce opposing sound distributed through the car audio system. It provides noise attenuation across low and mid-frequency noise.

E. *Interiors of Turborop Aircrafts*

Similarly, ANC is used for reduction of propeller tones (blade pass frequencies), [27]. Achieved results for NVX system applied to King Air aircraft is shown in Fig. 29, [28].

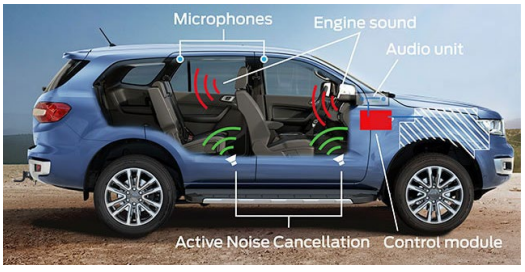


Figure 28. ANC in a SUV vehicle (Ford)

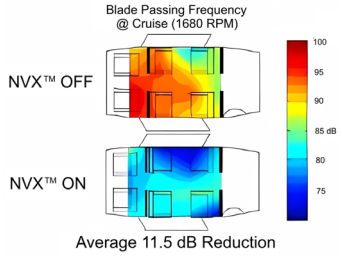


Figure 29. ANC applied to King Air 90 turboprop aircraft (Lord NVX)

Secondary sources in an aircraft are not necessary loudspeakers. Solution with a piezo stack that influences aircraft window is shown in Fig. 30, [29].

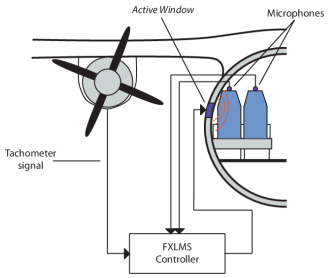


Figure 30. ANC using active window using piezo stack, [29]

F. Household Appliances

ANC applied to a vacuum cleaner is shown in Fig. 31, [30]. Passive methods disrupt the airflow and reduce efficiency. Reduction of tonal components have been achieved (about 10 dB), however much less for the entire spectrum (about 2 dB).

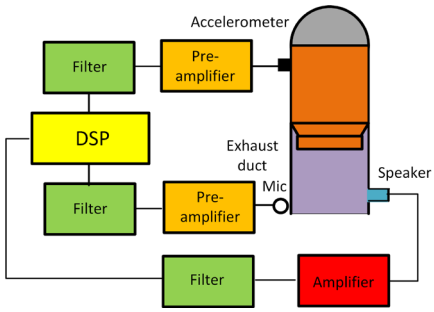


Figure 31. ANC applied to vacuum cleaner

VII. CONCLUSION

Active noise control has achieved considerable success at attenuating low frequencies, particularly tonal components and forming small local zones of silence. Main ANC control structures are feedforward, feedback and hybrid. These approaches are suitable both for broadband and narrowband noise. The influence of the secondary path in ANC systems has to be addressed and Fx-LMS algorithm was developed for that purpose. It is sometimes desirable to cancel the acoustic feedback as well. The synchronized waveform generator is suitable for cancelling periodic noise, and because of nonacoustic sensor it is not susceptible to acoustic feedback. Multichannel system are suitable for vehicle interiors. Some success has been achieved with semiactive solutions (tunable Helmholtz resonators) as well. Due to the low cost of DSP and microcontrollers, with the exceptions of some headphones, most ANC implementations of today are digital.

REFERENCES

[1] ATCO Structures and Logistics, Environmental Noise Control, ATCO, 2001

[2] H. G. Leventhall and L. Wong, “A review of active attenuation and development of an active attenuator 'Open Refuge’,” HSE Contract Res. Rpt. No. 4/1988, W S Atkins Engin. Sciences, 1988

[3] S. J. Elliott and P. A. Nelson, “Active sound control,” IEEE Signal Processing Magazine, October 1993

[4] C. R. Fuller and A. H. von Flotow: “Active control of sound and vibration,” IEEE Control Systems, December 1995

[5] S. M. Kuo and D. R. Morgan, “Active noise control: a tutorial review,” Proceedings of the IEEE, Vol. 87, No. 6, June 1999

[6] S. D. Snyder, Active Noise Control Primer, Springer, 2000

[7] C. N. Hansen, Understanding Active Noise Cancellation, CRC Press, 2002

[8] I. L.Ver, L. L. Beranek (Eds.), Noise and Vibration Control Engineering, Principles and Applications, John Wiley, NJ, 2006

[9] D. Miljković, “Active Noise Control: From Analog to Digital – Last 80 Years,” Proc. MIPRO 2016, Opatija, Croatia, 2016, pp. 1358-1363

[10] INC Engineering Co. Ltd., Construction Sites Got Silence via a Novel Noise Reduction Technique, IHI Eng. Rev., Vol. 48, No. 2 2015, Available: www.ihi.co.jp/inc/laneng/ [Accessed 20 Feb. 2023]

[11] C. Moy, Active Noise Reduction Headphone Systems, [Online]. Available: https://headwizememorial.wordpress.com/2018/03/15/active-noise-reduction-headphone-systems/ [Accessed 20 Feb. 2023]

[12] B. Lam, D. Shi, Woon-Seng Gan, S. J. Elliott and M. Nishimura, “Active control of broadband sound through the open aperture of a full-sized domestic window,” Sci. Rep. 2020 Jul 9;10(1)

[13] B. Lam and Woog-Seng Gan, “Active acoustic windows: towards a quieter home,” IEEE Potentials, Jan. 2016, Vol. 35, No. 1, pp. 11-18

[14] B. Widrow and S. D. Stearns, Adaptive Signal Processing, Englewood Cliffs, NJ, Prentice Hall, 1985

[15] R. Lacoste, Group Delay Basics – More Filter Fun, Circuit Cellar, [Online]. Available: https://circuitcellar.com/research-design-hub/group-delay-basics-more-filter-fun/ [Accessed 20 Feb. 2023]

[16] Qibo Maoa, Shengquan Lib and Weiwei Liua, “Development of a sweeping Helmholtz resonator for noise control,” Applied Acoustics, Vol. 141, Dec. 2018, pp. 348-354

[17] P. Minogue, N. Rankin and J.Ryan, “Adaptively Canceling Server Fan Noise, Principles and Experiments with a Short Duct and the AD73522 dspConverter,” Analog Dialogue 34-2, 2000

[18] Analog Devices, Data Sheet ADuC7019/20/21/22/24/25/26/27/28/29 Precision Analog Microcontroller, ARM7TDMI MCU, 2018

[19] S. D. Synder, C. H. Hansem and G. Vokalek, EZ-ANC II Users Guide, Causal Systems 2000

[20] T. Murao, C. Shi, W. S. Gan and M. Nishimura, “Mixed-error approach for multichannel active noise control of open windows,” Applied Acoustics, Vol. 127, Dec. 2017, pp. 305–315

[21] A. Pini, Understand How to Use Switched Capacitor Filters to Save Space and Improve Filter Performance, Digi-Key, 13 Dec. 2018

[22] S. Johansson, A. Siswanto, Cheng-Yuan Chang, S. M. Kuo, “Active noise control for headrests,” In Proc. 2015 APSIPA, Hong Kong, China, 2015, pp. 688-692

[23] M. Larsson, “Active noise control in ventilation systems, practical implementation aspects,” Licentiate Dissertation Series No. 2008:12, Blekinge Institute of Technology, Sweden, 2008

[24] L. J. Gelin, “Active noise control: a tutorial for HVAC designers,” ASHRAE Journal, 1 Aug. 1997

[25] V. Ratnam and S. Ramaiah, “Development of Active Noise Control (ANC) System for an Acoustic Duct,” In Proc. Advanced Research in Engineering and Technology 2013, Andhra Pradesh, 8-9 Feb 2013

[26] Ford, First-for-Ford innovation makes all-new 2020 Ford Explorer interior the quietest yet, Ford Media center, 2020

[27] S. Johanssonn, P. Sjosten, S. Nordebo and I. Claesson, “comparison of multiple- and -single-reference MIMO active noise control approaches using data measured in Dornier 328 Aircraft,” International Journal of Acoustics and Vibration, vol. 5, No. 2, 2000 (pp. 77-88)

[28] Lord, NVX Active Noise Control Systems for the King Air Series of Aircraft, [Online]. Available: http://educinno.intec.ugent.be/oipproject5/LordMPD.htm [Accessed 20 Feb. 2023]

[29] I. Dimino, A. Concilio, M. H. Aliabadi and Andrea Pierré, “A feedforward control system for the active noise control of a triple-pane aircraft-type window,” ICSV17, Cairo, Egypt, 18-22 July 2010

[30] R. Paurobally, “Active control of noise from a vacuum cleaner,” In Proc. 13th International Congress on Sound and Vibration. Vienna, Austria, 2-6 July 2006.

Active Noise Control Using Waveform Synthesis with Improved Convergence and Tracking

Dubravko Miljković
Hrvatska elektroprivreda, Zagreb, Croatia
dubravko.miljkovic@hep.hr

***Abstract* - Active noise control of periodic sounds, as commonly produced by rotating machinery, can be efficiently implemented using a synchronized waveform generator that produces a secondary signal necessary for noise cancellation. Synchronization of synthesized waveform is achieved by the use of reference tacho signal picked up by the magnetic or optical sensor. Waveform adaptation is performed in accordance with the residual noise signal after noise cancelation. Such approach for active noise control works well when a spectrum of offending noise is not changing much. However, during initial machine spool up, varying load and final spool down significant adaptability from waveform synthesizer is required to counteract changes in machine noise. In this paper, methods for improving convergence and tracking properties of waveform generator are presented and analyzed. An adaptive coefficient of waveform generator is modified in accordance to elapsed time during startup, varying frequency of the reference tacho signal and residual noise level. Results of simulations with different noise spectrum corresponding to various machinery rotational speeds are included.**

Keywords - active noise control; periodic noise; waveform synthesis; convergence; tracking

I. INTRODUCTION

When noise is dominated with low to mid frequency range where passive methods for noise attenuation lose its effectiveness, active noise control may be applied for its cancelation [1],[2]. Often, the offending noise is of periodic nature, e.g. noise produced by rotary machinery like motors and engines. This also includes humming noise generated by transformers and other devices and appliances that use electricity main. Such type of noise can be attenuated using active noise cancelation based on the synchronized waveform generator. Synchronization of the waveform synthesizer is achieved using reference tacho signal, [3]. There exist few methods for implementing synchronous waveform synthesizer. Here, coherent averaging, [4]-[6] of the residual noise waveform is used. Output waveform needs to be adapted to the change of the rotational speed and load conditions. This adaptation process uses adaptive coefficient, also called a convergence factor. To facilitate faster convergence and better tracking of changing acoustic noise spectrum and level it is beneficial to relate value of this factor to changes in rotational speed and residual noise level. This paper proposes methods for calculating variable adaptive coefficient. They are motivated by the behavior of rotating machinery during rotational speed change, RPM, and load change, influencing spectrum and noise level.

II. ACTIVE NOISE CONTROL USING WAVEFORM SYNTHESIS

A. General System Using Waveform Generator

Block diagram of the waveform generator applied to active noise control is illustrated in Fig. 1, [2]. It does not use a reference microphone and doesn't depend on measurement of the original noise. System is based on the synchronized waveform generator that uses synchro signal picked from non-contact inductive or optical tacho sensor. As there is no reference microphone, there is no possibility for unwanted acoustic feedback between secondary source and reference microphone.

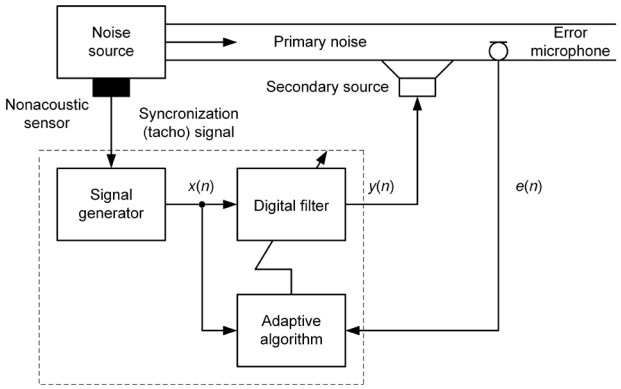


Figure 1. Block diagram of ANC using waveform synthesis

B. Secondary Path

Secondary path includes D/A converter, reconstruction low pass filter, amplifier, secondary source, acoustic path between the secondary source (sound propagation delay), an error microphone, it's preamplifier, antialiasing filter and A/D converter, [2]. The transfer function of the secondary path is not ideal and it has its frequency and phase characteristics. Most components in secondary path have reasonably linear frequency and constant phase characteristics. The cancelling speaker is generally the greatest contributor to frequency and phase deviations. The transfer function of the microphone has amplitude and phase characteristics that are in the frequency range of interest much closer to ideal one, so no additional consideration is given to their influence on the operation of the ANC system. There exists methods for addressing influences of the secondary path, inspired by Filtered-x LMS (Fx-LMS) algorithm, [1,2]. However, it requires off-line or on-line identification of the secondary path and together with filtered-*x* signal adds significantly to the complexity of ANC system.

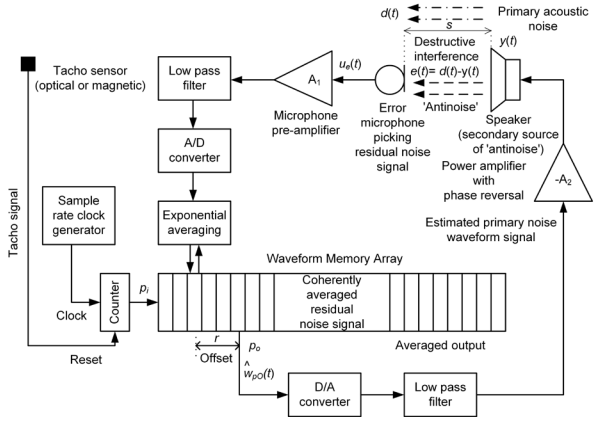


Figure 2. ANC using waveform synthesis by coherent averaging

C. *Coherent Averaging of Residual Signal*

Block diagram of the waveform generator applied to active noise control that uses coherent averaging of residual noise is shown in Fig. 2. Waveform generator uses a memory table, where waveform samples w_i are stored, N is max period, (1).

$$\boldsymbol{w} = [w_0, w_1, \ldots, w_{N-1}] \tag{1}$$

For a signal of period L, samples w_0 to w_{L-1} are streamed out to D/A converter with appropriate sample rate f_s. After D/A converter, low pass filter is applied to smooth output signal. There is an offset of r samples between input and output address pointer, due to the time necessary for the propagation of sound the distance s from secondary source to the error microphone, v is the speed of sound, (2). Some variable delay is also present due to the frequency dependent phase shift in secondary source.

$$r = \frac{s f_s}{v} \tag{2}$$

It separates positions of input pointer p_i and output pointer p_O. If p_O is larger than L it wraps around, (3).

$$p_O = (p_i + r) \bmod L \tag{3}$$

In the coherent averaging process (also known as vector averaging), multiple sets of signal plus noise samples are collected and averaged over the successive periods, as illustrated in Fig. 3. After averaging process remains only periodic signal, while random signal components cancels out over the successive averagings. To perform this averaging very precise alignment of successive periods of the originating signal must be provided. For this purpose, tacho signal is used. Time synchronous averaging acts as a low pass filter.

Let the $x(t)$ represent periodic acoustic noise signal, (4). It is a sum of deterministic periodic component $s(t)$, (5) with the period T and nonperiodic or random noise component $n(t)$.

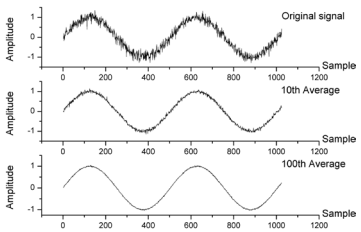


Figure 3. Coherent averaging of signal

$$x(t) = s(t) + n(t) \tag{4}$$

$$s(t+T) = s(t) \tag{5}$$

The idea behind coherent averaging performed over the period of the signal is that random noise $n(t)$ will cancel itself out and only deterministic periodic signal $s(t)$ will remain. Two types of averaging exist, arithmetic and exponential averaging. Arithmetic averaging of signal over N periods is described by (6), $\hat{x}_i$ is an averaged signal of the i-th sample in the waveform (one period), x is the input signal and L is the period length in samples:

$$\hat{x}_i(t) = \frac{1}{N}\sum_{k=0}^{N-1} x(t - kL) \tag{6}$$

In practice, not all summations are necessary for each period, just subtract oldest sample and add the newest sample to the sum that is already present. Exponential averaging uses for averaging just sample from the last period, giving more emphasis on new input signal sample. It is described by (7).

$$\hat{x}_i(t) = (1-\alpha)\hat{x}_i(t-1) + \alpha x(t) \tag{7}$$

The period it takes for the exponentially smoothed response of a unit step function to reach ≈ 63,2% of the original signal is called time constant, τ. If sampling interval ΔT is small compared to τ, it can be estimated from (8), giving info about relevant period for averaging.

$$\tau \approx \frac{\Delta T}{\alpha} \tag{8}$$

Synchronized waveform generator used in experiment uses exponential coherent averaging to adapt tits waveform. Waveform synthesis is relatively simple and intuitive concept. Not to add complexity with the complex modeling of the secondary path using digital filter, simpler solution for the secondary path may be used, particularly if one can find a speaker whose phase characteristics doesn't deviate too much in a frequency range of interest for noise cancellation. Then main problem is to determine the time necessary for acoustic propagation of signal form secondary source to the error microphone, around which the zone of silence is to be formed. Hence, phase shift introduced by the secondary path, $S(z)$, must be within the interval (9) for all frequency components of the canceling signal. One solution for dealing with phase changes related to change of tacho signal is described in [7].

$$\theta \in [-60°, +60°] \tag{9}$$

D. *Adaptation Using Residual Noise*

One approach toward the adaptation of synchronized waveform generator makes use of coherent averaging of residual noise signal. At the beginning of the operation of the ANC system waveform is empty and a residual noise is present and signal $e(n)$ is picked by error microphone. This waveform is coherently averaged with the waveform of residual noise. After first averaging, there are now some nonzero sample values in the waveform and this waveform is used to cancel the offending noise. This time the residual noise should be a little weaker, because it is a little bit cancelled by the waveform from the waveform generator. The remaining noise waveform is again averaged with the stored waveform, whose samples are increasing. In next iteration residual noise signal is again lower, etc. until

convergence is achieved and residual noise remains low. Residual noise signal $e(t)$ is determined as a difference between noise signal $d(t)$ and signal from the secondary source $y(t)$, in acoustical domain, (10).

$$e(t) = d(t) - y(t) \tag{10}$$

In detail, more parameters are considered: the gain of the error signal preamplifier, A_1, gain of output amplifier, A_2, the electroacoustic transfer factor of the secondary source B_S (Pa/V), microphone sensitivity S_M (V/Pa), distance s, so $e(t)$ is given by (11), $\widehat{w}_{p_O}$ is averaged waveform output at pointer p_O, (3), Fig. 2. In the near field term $1/s^2$ can be neglected.

$$e(t) = d(t) - \frac{1}{s^2} A_2 B_S \widehat{w}_{p_O} \tag{11}$$

Due to error signal preamplifier adaptation is given by (12).

$$\widehat{w}_i(t)(1-\alpha)\widehat{w}_i(t-1) + \alpha A_1 S_M e(t) \tag{12}$$

E. Adaptation to RPM Changes

For a system to be adaptive to changes in rotational speed, two approaches can be used. One is the use of variable sampling rate that is a multiple of the tacho signal (e.g. using a wheel with numerous markings to be detected by tacho signal). Another is to use fixed sampling rate and to add zero sample(s) to extend the waveform or skip sample(s) of the waveform to shrink the waveform. The remaining part of the waveform is not ideally suited for cancelation and must be adapted.

III. Periodic Noise of Machinery and Equipment

Periodic noise has a waveform of repetitive nature. Such noise has pronounced harmonic structure. Pure periodic noise is rare, it is often superimposed by broadband noise. Example of the periodic nose with superimposed broadband random noise is shown in Fig. 4.

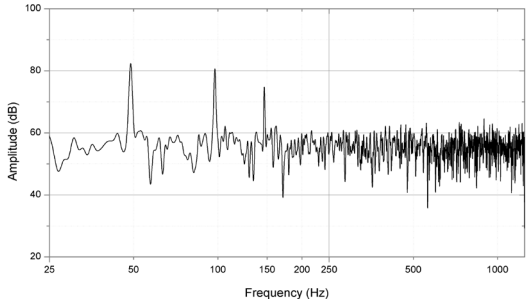


Figure 4. Periodic noise with superimposed broadband random noise

As already mentioned in the previous paragraph, ANC with synchronized waveform generator is suitable for reducing periodic noise. In an ideal case, all harmonic components should be cancelled and only the broadband noise remains.

A. Noise of Rotating Machinery and Tranformers

Motivation for modification of adaptation process stems from noise of rotating machines that are used in many engineering applications, [8]-[10]. Examples are pumps, compressors, motors and engines. Such machines have to operate over a range of speed and load conditions producing acoustic noise. Spectrogram of such noise has pronounced harmonic structure common in periodic noise.

Here are listed main causes that change the spectra of noise generated by rotating machinery.

1) Spool Up

Spool up is the process of increasing the rotational speed of an engine.

2) Spool Down

Spool down is the process of reduction of the rotational speed of an engine.

3) Other Speed Variations

Speed variations are not only limited to spool up and spool down, sometimes machine operates under various regimes with corresponding rotational speeds.

4) Load Variation

Noise of machinery is to an extent dependent on its load. Load variations can cause changes in rotation speed changing frequency spectrum and the overall noise level.

B. Transformer Noise

Transformer noise contains harmonics at odd multiples of twice the power grid frequency that is caused by core magnetostriction, [11,12]. However, noise level changes when the transformer is turned on and during load variations, [13]. Instead of tacho sensor, tacho signal is derived from zero crossings of power grid voltage.

IV. Adaptation Improvements

Proposed adaptation improvements are motivated by operation of machines that generate noise to be cancelled by ANC using waveform generator. Value of adaptive coefficient is changing according to regime of operation. Range of change R, (13) is bounded by use of min and max functions.

$$R = \frac{a_{max}}{a_0} \tag{13}$$

A. Basic System

In basic system, with no improvements, adaptive coefficient is held constant, (14), α_0=0.001.

$$\alpha_B(t) = a_0 \tag{14}$$

B. Initial Convergernce Period

A greater adaptive coefficient is used at the beginning of adaptation, and then gradually decreasing as adaptation progresses over the time t, (15). Essentially, after initial period of 10000 samples adaptive coefficient decays to α_0.

$$\alpha_I(t) = \alpha_0 \, max[1, R(1-kt)] \,, \; R{=}10, \, k{=}0.0001 \tag{15}$$

C. RPM Based

Value of adaptive coefficient depends on the rate of change of shaft rotational speed v_R. When speed changes, (16), producing noise of different spectrum, for the i-th signal period greater adaptation coefficient is used, (17).

$$\Delta v_R = |v_R(i) - v_R(i-1)| \tag{16}$$

$$\alpha_R(i) = \alpha_0(1 + \Delta v_R) \tag{17}$$

Not to complicate with calculation of v_R from the period between tacho impulses T_R that is already available, computationally simpler approach is used, (18) and (19).

$$\Delta T_R(i) = |T_R(i) - T_R(i-1)| \tag{18}$$

$$\alpha_R(i) = \alpha_0 max[1, min(R, M|\Delta T_R(i)|)], \; R{=}10, \, M{=}5 \tag{19}$$

D. RPM Based Plus Delay

This is an improvement of the previous RPM method. It acts same during RPM changes (18), (19), with the addition of end delay after the last RPM change (i.e. period T_R change as an inverse proxy) that keeps α higher for some time, e.g. K=2000 samples or several periods, (20).

$\alpha_D(t)=M\alpha_0$ for K samples after last RPM change, M=5 (20)

E. Residual Noise Level

If adaptive coefficient is small, the convergence is slow. Hence, when the residual noise level is large, greater adaptive coefficient is used. Once, when residual noise is significantly reduced, fine convergence could be achieved by smaller values of adaptive coefficient. Noise level L is determined by exponentially averaging absolute values of a residual noise signal $e(t)$, (21).

$$L(t) = (1-\alpha_L)L(t-1) + \alpha_L|e(t)|\ ,\ \alpha_L=0.005 \quad (21)$$

Value of adaptive coefficient is given by (22), where L is the noise level and M constant of proportionality.

$$\alpha_{RNL} = \alpha_0 max\left(1, min\left(R, ML(t)\right)\right),\ R=10,\ M=30 \quad (22)$$

V. Results of Simulated ANC System

For the purpose of running tests of convergence and tracking, test signals are generated and fed to ANC system.

A. Test Signals

Generated test signals are composed of three sinusoidal harmonic components and superimposed broadband noise, (23), (24), (25) and (26). Signal is normalized to an interval [-1,1]. During the synthesis of signals, tacho signal is also generated marking the beginning of each signal period.

$$x(t) = a_1 sin(2\pi f_1 t + \theta_1) + a_2 sin(2\pi f_2 t + \theta_2) + a_3 sin(2\pi f_3 t + \theta_3) + n(t)\ ,$$

$$a_1=1.0,\ a_2=0.5,\ a_3=0.2 \quad (23)$$

$$f_2 = 2f_1 \quad (24)$$

$$f_3 = 3f_1 \quad (25)$$

$$n(t) = M(1 - 2 * rand(t)),\ M=0.1 \quad (26)$$

Motivated by previous consideration regarding noise changes in rotating machinery and transformers, four test signals have been generated (electroacoustic loop gain is 30, involving A_IS_M = 10, and A_2B_S = 3).

B. Convergence Tests

Convergence tests were performed by applying a test signal at the start of ANC system operation. Continuous and spool up signal was used. Convergences, as mean squared error, MSE, are shown in Fig. 5, Fig. 6 and Table I.

1) Continuous Signal

Continuous signal of constant frequency and amplitude was applied, Fig. 5. Best convergence was achieved with an initial convergence, followed by noise level method, Fig. 6.

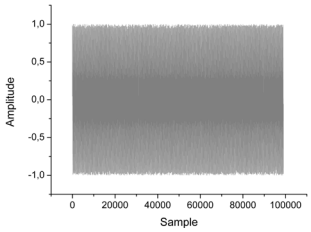


Figure 5. Continuous signal

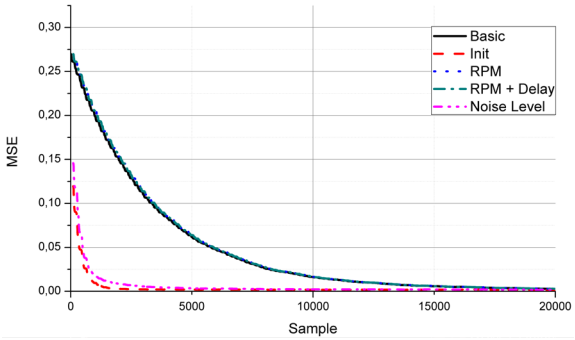


Figure 6. Convergence applied to continuous signal

2) Spool Up Signal

Signal with rising frequency (1:2 ratio) and amplitude was applied, Fig. 7. Significant advantage (faster convergence) was achieved with methods based on residual noise level and rotational speed with delay, Fig. 8.

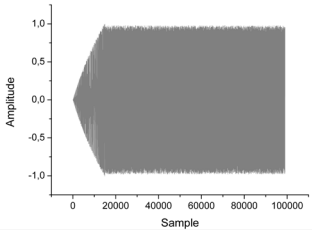


Figure 7. Spool up signal

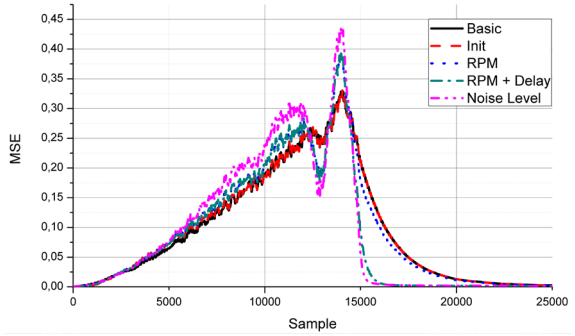


Figure 8. Convergence applied to spool up signal

TABLE I. Convergence Times

Method	Final MSE	90 % convergence (samples)
Continuous signal		
Basic	0.00196	7969
Initial period		782
RPM based[1)]		7969
RPM based + delay[1)]		7969
Residual noise level		1086
Spool up signal[2)]		
Basic	0.00193	18199
Initial period		18198
RPM based		17910
RPM based + delay		15273
Residual noise level		15041

[1)] RPM is fixed in continuous signal [2)]end of spool up at sample 15000

C. Tracking Tests

Tracking tests were performed by applying time-varying test signal to the ANC system that has already converged and operated in a steady state for some period. It shows the dynamic behavior of an ANC system when there is a change in the spectrum or noise level of the test signal (departure from the steady state). These tests verify

adaptation after rotating machinery speed change and residual noise level change after load change. RPM change and load change test signals were used.

1) *RPM Change*

Signal of constant amplitude with the change in RPM (1:3 ratio) between the samples 40000 and 50000, Fig. 9 was applied. Best result was achieved with noise level method, very closely followed by RPM combined with delay, Fig. 10.

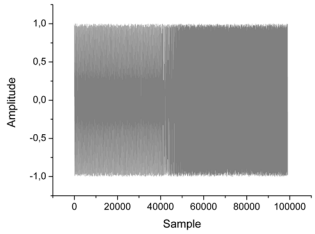


Figure 9. Signal with RPM change after steady section

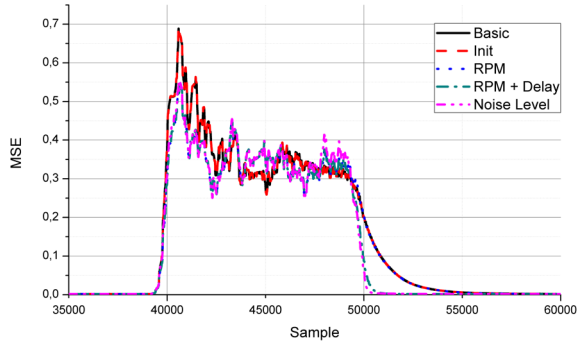


Figure 10. Tracking of RPM change after convergence

2) *Level Change Signal*

Signal with constant RPM but with the abrupt level change was applied. Drop in level of 50% was inserted between the samples 40000 and 60000, Fig. 11. Best result was achieved with noise level method, Fig. 12.

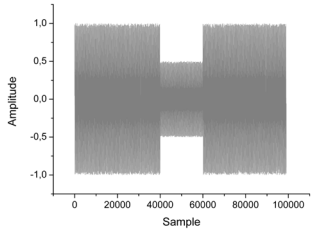


Figure 11. Level change signal

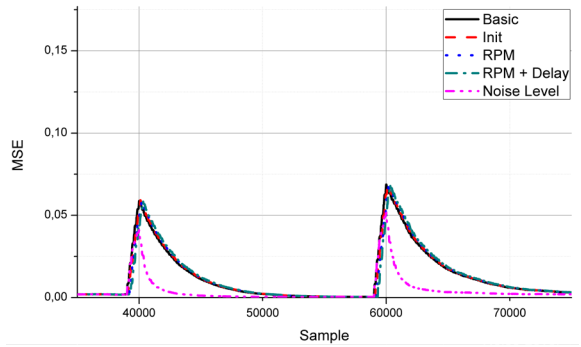


Figure 12. Tracking of level change after convergence

VI. Discussion

Results depend to an extent on the parameters used for defining adaptive step of each method: α, α_L, M, R, initial convergence period, the rate of change of RPM and the noise level in test signal. Simple, basic and initial period methods yield acceptable results, with initial period providing support for spool up. RPM based method with delay and noise level method have close results on spool up and RPM change tests. Keeping adaptive coefficient at a higher level for some time after the last RPM change is a significant improvement to the raw RPM method. Due to the mild rate of change of RPM in test signals, there is no great difference between results of noise level and RPM method. The noise level method is slightly more complex.

VII. Conclusion

ANC by the synchronized waveform generator is a viable solution for attenuation of periodic noise present in rotating machinery and transformers. Convergence and tracking properties can be improved by introduction of variable adaptive coefficient. Motivated by operation of rotating machinery (spool up, spool down, change of rotational speed and load change) and transformers (load change), the variable adaptive coefficient that depends on elapsed time (initial period), RPM with variant (delay) and the residual noise level has been investigated. Greatest improvements have been achieved with adaptive coefficient related to noise level and RPM with delay. For rotating machinery best approach are methods based on the residual noise level and RPM with delay or a combination (maximum of both). For transformers (fixed frequency), the best approach is a method based on a residual noise level.

References

[1] H. G. Leventhall and L. Wong, “A Review of Active Attenuation And Development of an Active Attenuator 'Open Refuge,” HSE Contract Res. Rpt. No. 4/1988, WS Atkins Engin. Sciences, 1988

[2] S. M. Kuo and D. R. Morgan, “Active Noise Control: A Tutorial Review,” Proceedings of the IEEE, Vol. 87, No. 6, June 1999.

[3] G. B. B. Chaplin and R. A. Smith, “Waveform synthesis - the Essex solution to repetitive noise and vibration,” Proc. Inter-Noise 83, pp. 399–402, 1983.

[4] Crystal Instruments, “Introduction of Time Synchronous Averaging,” CI Product Note No. 017, 2015

[5] R. G. Lyons, Understanding Digital Signal Processing, 2nd Ed., Prentice Hall, 2004

[6] O. Rompelman and H. H. Ros, “The coherent averaging technique, a tutorial review: Part 1: Noise reduction and the equivalent filter,” J. Biomed. Eng., Vol. 8, pp. 24-29, Jan. 1986.

[7] D. Miljković, “Simple Secondary Path Modeling for Active Noise Control Using Waveform Synthesis,” Proc. of MIPRO 2018, 21-25 May 2018, Opatija, Croatia

[8] R. Bigret, Rotating machinery, essential features, In: Encyclopedia of Vibration, Accademic Press, 2001, pp. 1064-1069

[9] W. K. Blake, Noise From Rotating Machinery, In: Mechanics of Flow-Induced Sound and Vibration, Vol. 2 (2nd Ed.), Accademic Press, 2017, pp. 505-658

[10] nCode Software, Vibration of Rotation Machinery, [Online]. Available: https://www.youtube.com/watch?v=SJw2WcjrWL4&ab_channel=nCodeSoftware [Accessed 20 Feb. 2023]

[11] Understanding Transformer Noise, Federal Pacific, Feb. 2016

[12] L. Solanki, “Transformer Noise,” Transformers Magazine, Special Edition: Digitalization, 2020, pp. 50-55

[13] 120 MVA transformer switched ON, [Online]. https://www.youtube.com/watch?v=VKLsWW4gym8&ab_channel=RODALCO2007 [Accessed 20 Feb. 2023]

Higher-Level Experimental Prototype of a Control Device for Dynamic Beverage-Cooling Process

T. Špoljarić*, B. Vuletić Komljen*, Lj. Cvitaš*, G. Malčić*, J. Matuško**, M. Miletić*
*Department of Electrical Engineering, Zagreb University of Applied Sciences, Croatia
**Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
tomislav.spoljaric@tvz.hr, bvuletick@tvz.hr, ljubivoj.cvitas@tvz.hr, goran.malcic@tvz.hr, jadranko.matusko@fer.hr, marko.miletic@tvz.hr

***Summary* - In this paper development and testing of an electronic controller for operation of a beverage cooling device is described. This development is a part of a project research that includes a cooling device with a water-submerged evaporator and its adequate regulation system. Authors have developed a controller device, mathematical models, and algorithms "from scratch" and in this paper an optimization of a regulation device is described: selection of a smaller and more effective microprocessor, necessary reduction in number of inputs and outputs, new ECB development and an optimized algorithm for the dynamic cooling process.**

Keywords - beverage cooling, control device prototype, electronic controller, electronic circuit board

I. INTRODUCTION

Cooling system with water-submerged evaporator uses a concept of direct evaporation of the refrigerant on the evaporator coils, thus avoiding additional heat losses. In this type of cooling a total heat transfer coefficient is increased, as is the cooling capacity, when compared to standard cooling systems, such as air evaporator cooling systems. When circulating beverages are being cooled dynamically these advantages come in hand, since water-submerged system provides more direct cooling: heat is being transferred from liquid to liquid, and intermediate heat losses are being reduced to a minimum.

IRI2 project (with reference code KK.01.2.1.02.0036) team in its experimental phase proposed, developed, and tested a mechanical design solution comprising of water-submerged cylindrical evaporator prototype, as described in [1], [2], [3]. Additionally, a developed cooling system has fixed speed condenser fan, compressor with implemented speed control and several operating solenoid valves for refrigerant flow control. Refrigerant used for cooling is R290, which is ecologically acceptable with higher rate of energy efficiency compared to standard refrigerants.

In further development of a prototype from the mechanical and thermal perspective number of solenoid valves was reduced to three: one on the suction side, one for the oil recharge, and a bypass valve for prevention of pipe frosting. Cooling system elements (compressor, evaporator, condenser, control unit) were repositioned to make a more compact layout (fig.1) and an adequate housing was manufactured.

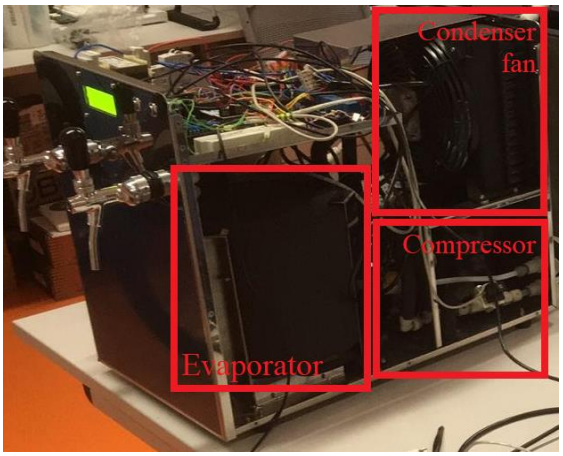


Figure 1. Final prototype of a beverage-cooling system with water-submerged evaporator

With these changes a needed progress towards commercialized unit was made, and only a development of a control system remained [4]. Control system adjustment was done accordingly with changes in mechanical and thermal element disposition.

II. CONTROL SYSTEM FOR A HIGHER LEVEL BEVERAGE-COOLING PROTOTYPE WITH WATER-SUBMERGED EVAPORATOR

In the industrial research phase, a reduction of several input and output elements was made for the purpose of better unit commercialization and more compact element layout. Also, a simpler control unit was made, with main change in selection of a microprocessor unit. Since a reduction in number of signals and actuators was made, a simpler microprocessor unit was selected and with it a new circuit board project for was designed. Subsequently, a new ECB was made that was compatible with new microprocessor unit and external I/O elements.

A. Sensors –input signals

In an experimental phase prototype several sensors for pressure, temperature and liquid flow signals were installed. Pressure sensors were used for refrigerant pressure measurement on evaporator and condenser side. Temperature and liquid flow sensors were used for measurement of beverage state, with option for feedback control of a cooling system according to temperature. Pressure sensors were connected to evaporator's and condenser's pipeline entry points.

In industrial phase prototype pressure sensors have remained, and are used for control of condenser fan and compressor speed. From the beverage side, two different flow sensors are used for monitoring of two separate cooled beverages' flows, thus making an enhancement by adding a second cooled beverage pipeline.

Inputs have been assigned to its numbered connector pins, soldered on a new ECB. Assigned pin numbers are shown in Table I.

B. *Actuators – output signals*

For an adequate operation of a higher-level prototype's cooling system SSR relays are selected as actuating devices for compressor and condenser fan. SSR relays are more reliable than standard electromechanical contactors (used in previous prototype level) and are "noise-free". Four additional relays for solenoid valves have remained from previous prototype structure, with three mandatory for cooling operation: suction valve, bypass pipeline valve and oil drain valve.

Outputs have been assigned to its numbered connector pins, soldered on a new ECB. Assigned pin numbers are shown in Table I.

TABLE I. INPUT/OUTPUT SYSTEM FOR HIGHER-LEVEL PROTOTYPE

Type	***Manufacturer***	***Model***	***Range***	***El. signal***	***Pin address (STM32 ECB)***
Flow sensor 1	Digmesa AG	938-2540/01	0.041 ÷ 15 l/min	236 pulse/l	X13
Flow sensor 2	Digmesa AG	938-2540/01	0.041 ÷ 15 l/min	236 pulse/l	X14
Pressure sensor	Emerson	PT5N-30M	0 ÷ 30 bar	4 ÷ 20 mA	X16
Pressure sensor	Emerson	PT5N-07M	-0.8 ÷ 7 bar	4 ÷ 20 mA	X15
Solenoid valve 1	Castel	CM2 9110/RA2	OFF/ON	0/24VAC	X53
Solenoid valve 2	Castel	CM2 9110/RA2	OFF/ON	0/24 VAC	X56
Solenoid valve 3	Castel	CM2 9110/RA2	OFF/ON	0/24 VAC	X54
Solenoid valve 4 - optional	-	-	OFF/ON	0/24 VAC	X57
Compressor	SECOP	NLV12.6CN	40 ÷ 155 Hz	2000÷4500 rpm	X59
Condenser fan	ELCO	VNT 12/20 334	OFF/ON	0 ÷ 230 VAC	X58

C. *Hardware interface*

After analysis of the lower level prototype, it was necessary to make certain improvements on the circuits and on the program support of the microcontroller. The optimization included new selection of microcontroller and SSR relays that control main cooling elements on the device. Due to reduced number of analog/digital I/O, the main change compared to the lower-level prototype is the replacement of the microcontroller. New unit is from the same series (fig. 2) – *"Blue Pill"* STM32F103C8T6.



Figure 2. Microprocessor unit "Blue Pill" STM32F103C8T6

In the previous phase of the development project, three different transmitters were connected to a control system of a prototype: temperature and pressure sensors' MBED transmitter, flow sensor's MBED transmitter, and frequency control MBED transmitter [1]. However, due to the reduction requirements of an electronic device (spatially and in I/O number) a new ECB was designed and made to accommodate new microprocessor unit and I/O connector kits. On fig 3. a block schematic of a new control system is shown.

The signals are 4÷20mA current loop pressure sensors P1 and P2, and a PWM type flow control sensors F1 and F2 shown on fig. 3. Their appropriate signals are converted into digital form [5] and are displayed locally on the LCD display of a device, shown on fig. 4.

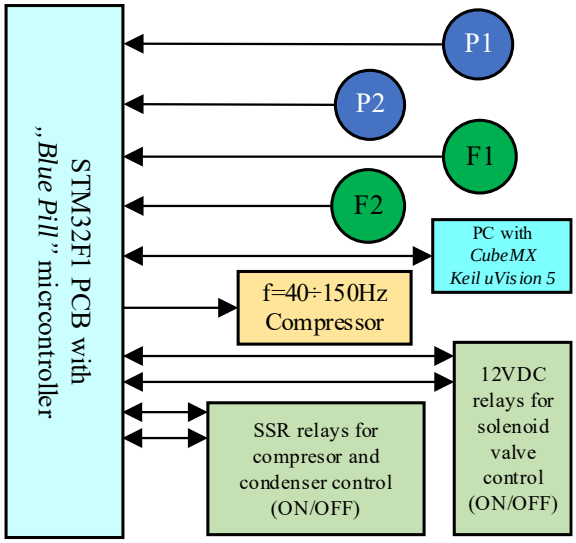


Figure 3. Block schematic of a higher-level prototype's control system



Figure 4. LCD display on a front panel of a prototype

D. ECB design for a higher level prototype

Compact control device development phase had a goal of eliminating unreliable connections and raising the reliability of the control system. Soldering of contacts and shielding the lower-level analog signals managed to eliminate these physical flaws. A printed circuit board, that would be used to connect microprocessor unit and all the necessary sensors and actuators was made as a result of ECB development phase. Requirements for this development phase were:

- the NUCLEO STM32F103C8T6 microprocessor control unit [6] is connected to the main ECB by a strip connector for modular attachment or detachment,
- output for main elements is done with external SSR relays, while output for valves is done with relays soldered on the board,
- the connection of external elements is done with soldered terminal blocks,
- 12 VDC power supply is derived with required voltage levels for all circuits,
- external LCD display is connected via I2C protocol to appropriate connector, and it is mounted on the housing of the prototype [7].

The ECB (fig. 5) and the schematic (shown on fig. 6) were designed via Autodesk EAGLE CAD tool. The ECB is double-sided. Connecting traces are laid out on both sides of the ECB and it is designed in THT technology with no SMD components for surface mounting, as was done with ECB in previous phase of experimental research.

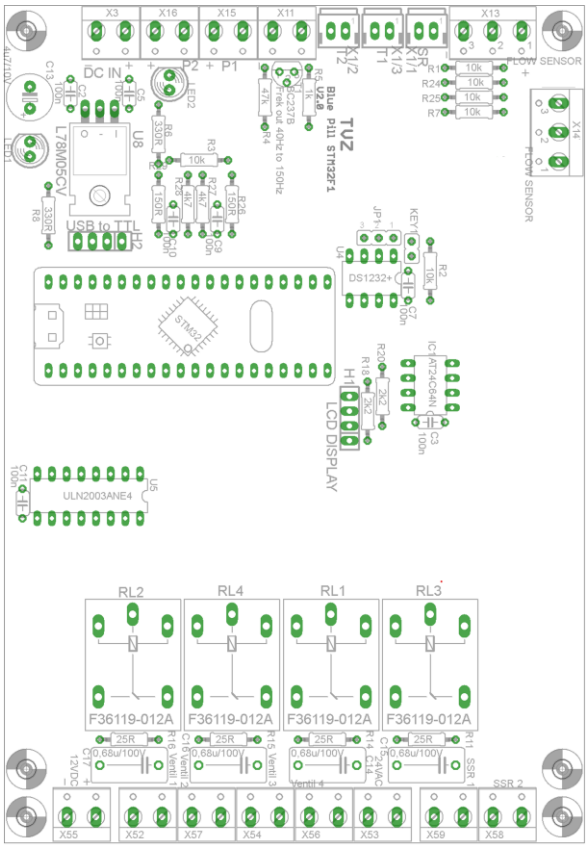


Figure 5. ECB design layout for control system

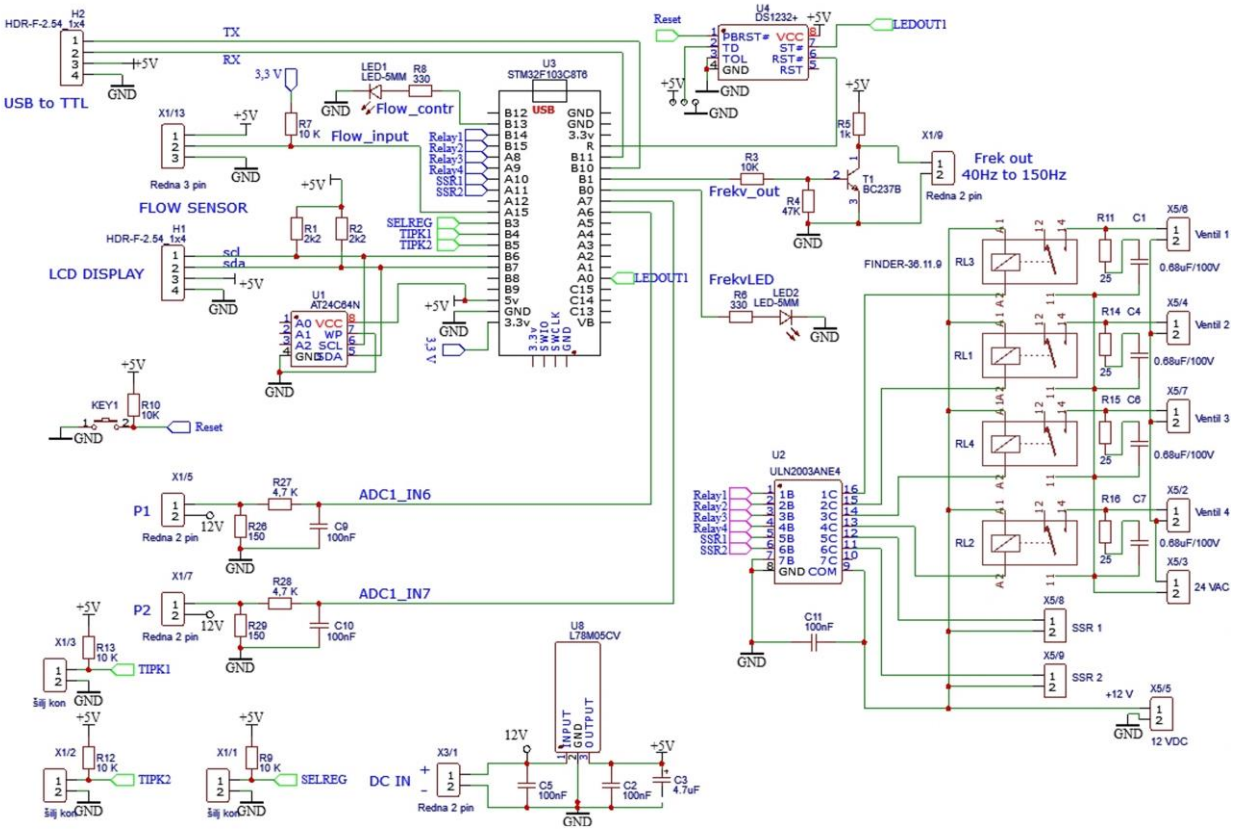


Figure 6. Detailed schematic of a ECB for higher-level prototype's control system

E. *Control concept and code development for control*

For control reference a temperature is selected. Since there are no temperature probes in the system, a temperature is obtained through pressure measurement since it is directly linked to a pressure of the refrigerant [8]. Scaling of the temperature is non-linear and it is described by a fourth order model, depending solely on the thermodynamical analysis of the refrigerant. The scaling equation is as follows:

$$p(T) = -27.2037 + 0.35656 \cdot T - 1.29068 \cdot 10^{-3} \cdot T^2 - 5.5771 \cdot 10^{-7} \cdot T^3 + 7.58418 \cdot 10^{-9} \cdot T^4 \quad (1)$$

Control concept (and its appropriate code) is divided into two different parts:

1. Beverage flow hold – preparation state, and
2. Beverage flow state.

In the first part a concept for regulation includes dead-zone band for temperature maintenance of a static or quasi-static state of beverage, where no flow or very slow flow is being implemented. A discontinuous type of regulation is needed in this static environment, due to possibility of pipeline freezing if a continuous type is being used. In the second part, a continuous PI regulation type is implemented. While a flow of the beverage is above a limit defined for the static state, the continuous concept takes place, and a frequency of the compressor is being regulated according to the temperature/pressure setpoint.

Package *Keil uVision5*, along with *"Blue Pill"* microcontroller, enables project management, source code editing, and targeted microcontroller code loading and execution. Module *uVision Debugger* is used for code optimization of the applied application. With an integrated ST-LINK USB on the "Blue Pill" microcontroller, debugging and code manipulation is simple to implement. Software interfaces *STM32CubeMX* and *Keil uVision5* with *Keil C-compiler* use a layer labeled HAL (*Hardware Abstraction Layer*) that connects the microcontroller software libraries and applications in one project. Direct approach to measurements is also provided to these instances via excel add-in *Data Streamer*. With this add-in, acquisition is made simpler than the one used in previous research phase, because a smaller set of real-time data. A MATLAB/Simulink environment [9] is no longer needed and has been discarded in this phase.

F. *Control system placement in a device prototype*

Control system, along with its actuators is placed on top side of a device prototype. An inox tray is designed to accommodate developed ECB with microprocessor, DC power supplies (for SSR's and ECB power input), actuating devices, transformer (solenoid valves' power supply) and protection devices. Disposition of elements placed on a tray is shown in fig. 7. Tray with elements has adequate dimensions, for an outer housing to be finely adjusted so whole device can be covered appropriately.

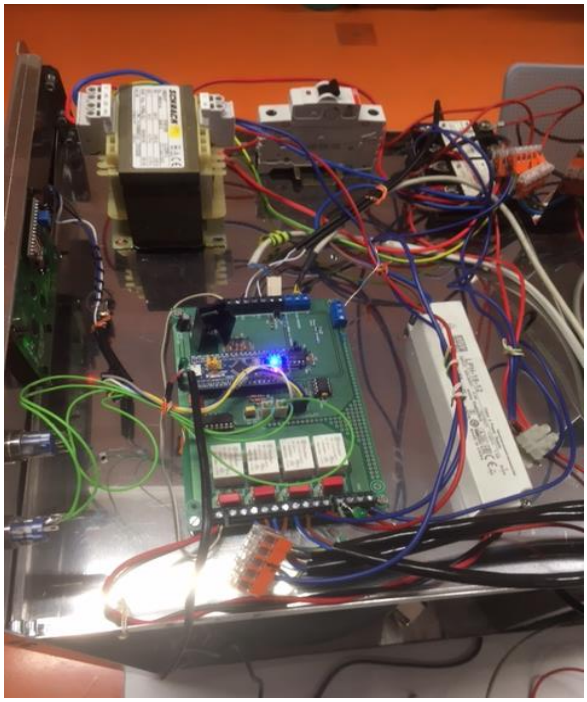


Figure 7. Higher-level prototype's control system – placement in housing

III. CONTROL SYSTEM'S PERFORMANCE – TESTING RESULTS

For the purposes of testing the prototype and a control system, a heating system for preparing test liquid has been installed. This system consists of water (testing liquid) tank, electric heater module and a simple dead-zone controller for switching the heater according to desired temperature of the liquid that is to be prepared for testing. The heating/preparation system is shown on fig. 8.

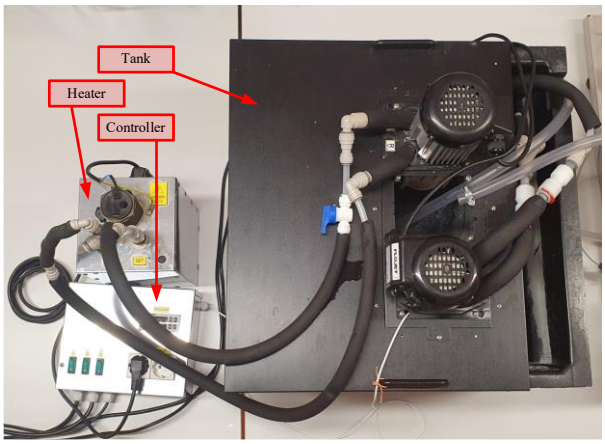


Figure 8. Heating system for preparation of testing liquid

Testing was done for different liquid flows, with subsequent stop of liquid after a tap is opened for some time. Values that are being observed are cooled liquid flow/stop, refrigerant pressures, temperature setpoint, compressor's drive frequency and compressor switching state. Four different cases were observed:

1. Test case with no liquid flow
2. Test case with liquid flow equal to $0.44\ l/min$,

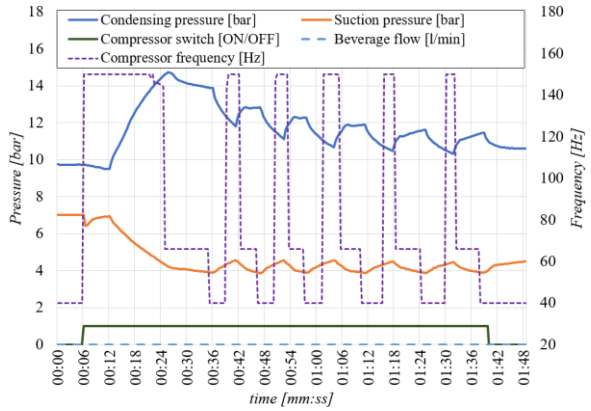


Figure 9. Test case 1 – cooling the beverage in preparation state with no flow

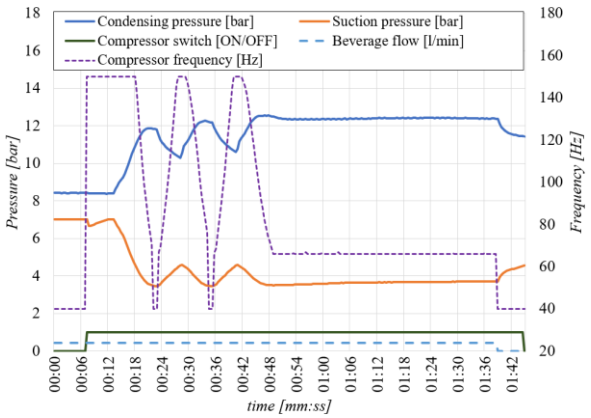


Figure 10. Test case 2 – cooling the beverage, flow of 0.44 l/min

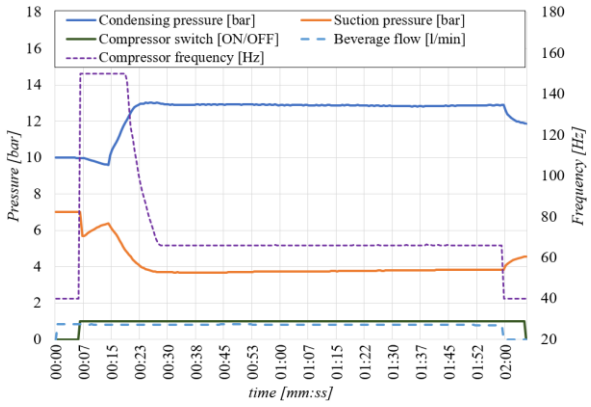


Figure 11. Test case 3 – cooling the beverage, flow of 0.82 l/min

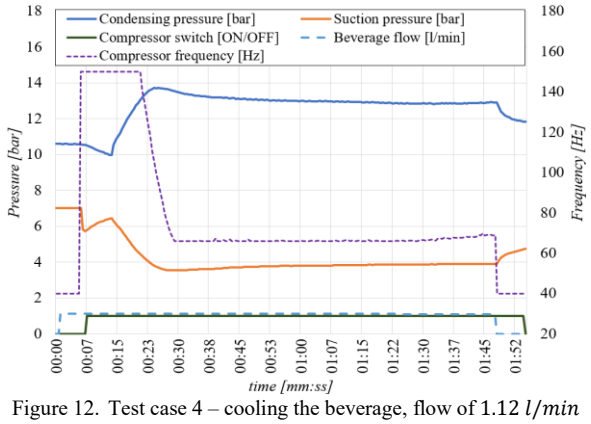


Figure 12. Test case 4 – cooling the beverage, flow of 1.12 l/min

3. Test case with liquid flow equal to 0.82 l/min,
4. Test case with liquid flow equal to 1.12 l/min.

In all four cases a test beverage temperature of 3°C corresponds to suction pressure of 3.9 bar of a refrigerant. Results from measurements are shown on fig. 9 to fig. 12, for described test cases, respectively.

Linear regulation of compressor's motor speed is accustomed to frequency range from 66 Hz to 150 Hz, and below 50 Hz compressor stops. Between 50 and 66 Hz compressor's motor runs at minimum speed. Control curve characteristic is shown on fig. 13.

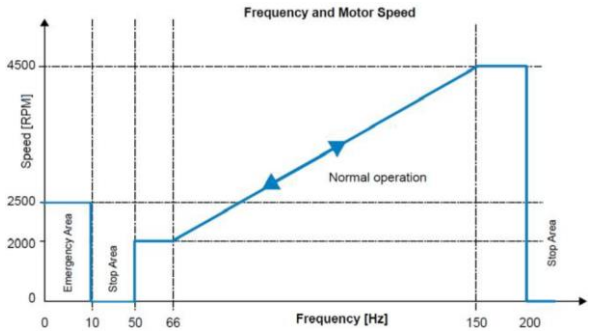


Figure 13. Frequency vs. speed characteristic for compressor control [10]

While compressor reaches its operating point it runs at minimum speed, until a preprogrammed delay for minimum compressor on time elapses. This delay is implemented so the compressor doesn't switch off too soon. This compressor's behavior can be observed in all four cases. Compressor switch state observes the switching state of the SSR relay of the compressor, and with it a suction solenoid valve. This state is not to be confused with actual state of the motor speed of the compressor which may reach a 0 rpm, when a 40 Hz frequency state occurs, when still a suction solenoid valve is opened.

In the first case, a dead-zone type regulation is applied, and this occurs in a repeated switching of the compressor. These 40 Hz switching states occur in lesser numbers once a refrigerant (and beverage) reaches its desired temperature/pressure. In the second case, when a reduced beverage flow is applied, a compressor also repeatedly switches, but with continuous type of pressure regulation. This behavior is a consequence of a compressor strong enough to endure faster exchange of beverage. In more static environment it reaches operating point much sooner with an undershoot of pressure as a consequence. This causes a compressor to switch from 0 to 1500 rpm repeatedly until a desired pressure is reached without an undershoot.

In the tests 3 and 4 operating point is reached quickly (during a "first glass pouring") and with only one peak of speed at 1500 rpm, after which a fine continuous regulation takes place. No undershoots of suction pressure are present, as well as no repeated compressor switching. In conclusion, this observation acknowledges the correct choice of the compressor for the appropriate dynamic operation of the cooling beverage tap device, while adequately enduring dynamic thermal load of a beverage flowing.

IV. Conclusion

In the experimental research phase for the tap device prototype with water-submerged evaporator a control algorithm with various measurements and actuators was implemented, tested, and described in [1], [2], [3]. This phase was a waypoint for observing the commercial aspect of the project. Since a more aesthetic and practical product needed to be made for possible successful commercial launch, a mechanical, electrical, and automatic control elements were reduced spatially and numerically.

In the industrial research project phase following changes were made:

- device prototype gained a new look: it was more compact in size and more acceptable for transport,
- all parts of the power supply and automatic control systems were adequately placed on a tray designed for such a purpose, and for easy access for the device maintenance personnel,
- direct temperature control with pressure monitoring was replaced with more appropriate pressure control (better response time, smaller number of sensors),
- a new microcontroller unit was selected, lesser in size and in number of ports, and with adequate performance,
- a new ECB was made to optimally accommodate microcontroller, actuators and connecting pins,
- a new control program was implemented for the final control solution that corresponds to described changes in the device,
- functionality of the device was tested with satisfactory results.

As a result, following goals of the industrial research phase were achieved: more compact device was made with lesser potential expenses for the commercial product manufacturing process and a simpler use of the device for the end-user is provided.

V. References

[1] L. Cvitaš, B. V. Komljen, T. Špoljarić, M. Miletić and G. Malčić, "Electronic unit for controlling the operation of the device with a flooded evaporator," Proceedings of 45th International Convention on Information, Communication and Electronic Technology (MIPRO), 2022, pp. 907-912

[2] N. Šare, B. Vuletić Komljen, K. Osman, J. Matuško, "Control of Flooded Evaporator in a Beer Dispenser Unit", Proceedings of 45th International Convention on Information, Communication and Electronic Technology (MIPRO), 2022, pp. 913-918

[3] N. Šare, B. Vuletić Komljen, J. Matuško, K. Osman, "Modelling and Experimental Verification of Device with a Submerged Evaporator for Beer Cooling", Proceedings of 45th International Convention on Information, Communication and Electronic Technology (MIPRO), 2022, pp. 919-924

[4] Lj. Cvitaš, S. Đonlić, "Projektno upravljanje razvojem novog proizvoda", 11. Hrvatska konferencija o kvaliteti (2. znanstveni skup Hrvatskog društva za kvalitetu), 2011

[5] M. Miletić, "Izbor mikroupravljača", Lectures on Microcontrollers course, Zagreb University of Applied Sciences, 2017

[6] "STM32F103x8 datasheet", STMicroelectronics, www.st.com

[7] "20x4 Character LCD datasheet", Vishay Intertechnology, Inc., www.vishay.com

[8] "NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP): Version 10", National Institue of Standards and Technology, www.nist.gov

[9] B. Palašek, P. Mesarić, M. Kukec, "Ostvarivanje sučelja između miroupravljača i MATLAB-a", Tehnical journal vol. 8, no. 3, 2014

[10] "Controllers for NLV-CN Compressors – operating instructions", SECOP, www.secop.com

Artificial intelligence-based predictive fuel blending control for flare gas mitigation

J. Kir Hromatko*, Š. Ileš*, R. Huljev†, V. Vučković‡
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
E-mail: {josip.kir.hromatko, sandor.iles}@fer.hr
† Genuyn, Office 9, Dalton House, 60 Windsor Avenue, London, United Kingdom, SW19 2RR
E-mail: rube@genuyn.com
‡ Popravak Brodskih Motora d.o.o., Svilno bb, 51219 Čavle (Rijeka), Croatia
E-mail: velibor.vuckovic@pbm.hr

***Abstract*—This paper describes a fuel blending algorithm based on artificial intelligence and model predictive control. A gas-fired power plant was modeled using physical laws and on-site measurements. A neural network is used to calculate the methane number of the fuel and determine the fuel blending ratio limits so that the methane number is within the limits specified by the engine manufacturer. A model predictive controller adjusts the final blending ratio to meet safety requirements and minimize operating costs. The algorithm was tested in simulations with different scenarios and a reduction in both the operating costs and amount of flaring was observed.**

***Keywords*—fuel blending, methane number, neural network, model predictive control, flare mitigation**

I. Introduction

Flares are important safety devices that can burn off unwanted gases due to their volatile nature concerning the amount, availability, and quality. However, if operated improperly, flaring can emit vast amounts of pollutants, and pose significant environmental and health risks. It is highly desirable to achieve zero flaring via flare minimization and flare gas recovery. By integrating the flared gases into a fuel gas network, it is possible to reduce emissions, as well as conserve energy in refineries [1]. A comparison of economic profitability for various scenarios for using flare gases is investigated in [2]. According to the paper, the flare gases with higher flow rates and better gas composition in terms of the amount of hydrogen and hydrogen sulfide highly influence the profitability of using such gases in various applications.

In this paper, we investigate the possibility of using model predictive control and neural networks to use flare gases in a dual-fuel engine.

Model predictive control (MPC) is a control strategy that uses a model of a system to predict its response to a given control input. An MPC algorithm repeatedly solves an optimization problem and computes the optimal control action that minimises the cost function and controls the states of the system to the desired operating point while satisfying the control input and state constraints. MPC is widely used in industrial and process control applications, where it is valued for its ability to handle constraints. For example, MPC is used to control nonlinear systems with constraints, such as internal combustion engines where multiple objectives must be satisfied [3], or in applications such as fuel blending of six different feed gases where the objective is to achieve fuel blends with specific limits on the different composition properties [4].

Instead of industrial fuel blending, in this paper we investigate the possibility of blending the flaring gas with natural gas in order to use it with a dual-fuel engine (or a gas-powered generator - genset). However, one of the main quality requirements of natural gas as an engine fuel is the methane number (MN). This parameter indicates the fuel's capability to avoid knocking (spontaneous combustion) in the engine. A higher MN value indicates a better fuel quality for gas engines. The required minimum MN value usually depends on the engine manufacturer. More and more engine manufacturers or require fuel with a minimum methane number of 80, while some still maintain methane number limits below 80 (most often 70) [5]. In [6], a fuel blending system was used to vary the MN. With smaller MNs, NOx and CO emissions increased while combustion stability remained unaffected.

The use of artificial neural networks for methane number prediction is reported in [7], where the model was trained using 1202 different gaseous fuel compositions. The neural network model was compared to the AVL (Anstalt für Verbrennungskraftmaschinen List) model and it was reported that the model was able to predict the MN accurately (R = 0.999).

In this paper, the artificial neural network is used to predict the MN of different fuel blends based on the tail gas and natural gas with assumed known gas composition. The results from the neural network are used to find the limits on the fuel blending ratio which achieve the MN according to the motor specification. The model predictive control algorithm is used to minimize the price of the fuel blend and the amount of flaring while respecting several constraints on the variables of interest.

This work has been supported in part by the European Regional Development Fund under the project KK.01.2.1.02.107.

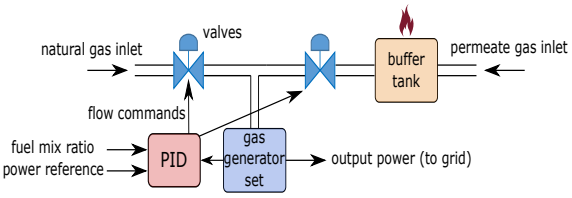


Fig. 1. Illustration of the plant operation.

II. Methods

A. Plant description

The industrial process of interest in this work is a gas-fired power plant. The main component of the power plant is a generator set which runs on gas and produces approximately 1 MW of electrical power. Two gases are available for fueling the generator: natural gas, which is in constant supply and has to be paid for, and permeate gas, whose supply varies, but is available for free. The generator runs on a mixture of the two fuel gases and the mixing ratio is set manually. The permeate gas is usually of lower quality (i.e., contains less methane), which can affect the generation of power and should be considered when setting the mixing ratio.

Additionally, a gas tank was installed on the permeate gas stream to act as a buffer and mitigate the effect of variations in availability. Since the pressure in the tank needs to be limited, a safety mechanism which flares excess gas in case of high tank pressure was implemented.

A simplified schematic of the power plant is given in Figure 1.

B. Plant model

The power output of a fuel engine that runs on methane depends on the flow rate of the methane, as well as the engine's design and operating conditions. Generally, increasing the flow rate of methane to the engine will result in an increase in power output. However, there is a limit to how much methane can be safely and efficiently combusted within the engine, and at high flow rates, the engine may not be able to utilize the methane, resulting in decreased power output. Additionally, the engine's compression ratio, ignition timing, and other factors can also affect the engine's power output.

The focus of the work presented here was on controlling two variables of interest, the power output of the gas generator (engine) and the pressure in the gas tank (buffer). The actual power plant contains several control loops which ensure stable operation and allow for certain simplifications in control design.

In modelling the gas generator, it was assumed that the output power is proportional to the amount of methane in the fuel, with a certain response time. The resulting model can be written as a differential equation:

$$\frac{dP}{dt} = \frac{1}{\tau}\left(k_n q_n + k_p q_p - P\right) \tag{1}$$

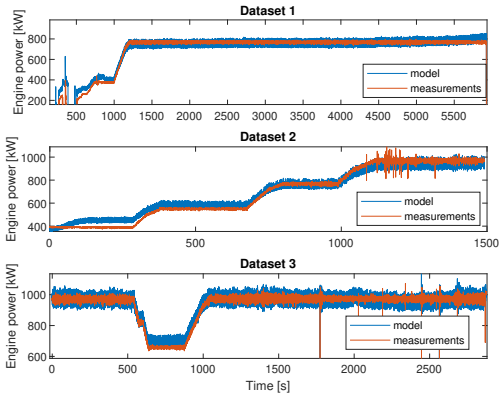


Fig. 2. Recorded data and model fit.

where P is the generated power, q_i the mass flow of fuel component i and τ the time constant of the power output. The notation $(\cdot)_n$ and $(\cdot)_p$ refers to natural and permeate gas quantities. The gains from fuel components to generated power, k_n and k_p, generally depend on the fuel characteristics and the operating point (engine speed and torque), as well as other factors such as the engine efficiency. However, the on-site measurements enabled calculating the gains by a least squares fit and approximating them as constants. Figure 2 shows the measurements and the power output calculations using the obtained gains, which were in the order of 10 ± 1 $kW/(L/s)$ for different datasets.

The tank pressure is affected by the incoming and outgoing permeate gas. The conservation of mass principle states that [8]:

$$\frac{dM}{dt} = \dot{m}_{in} - \dot{m}_{out} \tag{2}$$

where M is the total mass of a substance in a tank and $\dot{m}_i$ are the incoming and outgoing mass flows. Following the procedure in [9], the derivative of the tank pressure can be expressed as:

$$\frac{dp}{dt} = \frac{p}{\rho V}(q_i - q_p) \tag{3}$$

where p, ρ and V denote the gas pressure, density and volume, q_i the incoming mass flow and q_p the outgoing mass flow of permeate gas.

The available measurements indicate that the gas density is approximately constant and equal to 1 kg/m^3 for both gases so volumetric and mass flow can be used for calculations interchangeably.

C. Neural network-based ratio limit calculation

Gas engine manufacturers usually specify the desired limits for the methane number of fuel. Several methane number calculators are available as web [10] or mobile phone [11] applications, but using them in a control system requires a hardware implementation. The European Association of Internal Combustion Engine and Alternative Powertrain Manufacturers provides an open-source

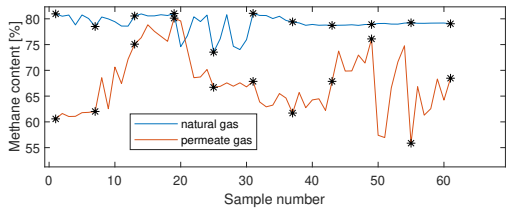


Fig. 3. Gas methane contents dataset (stars indicate the measurements).

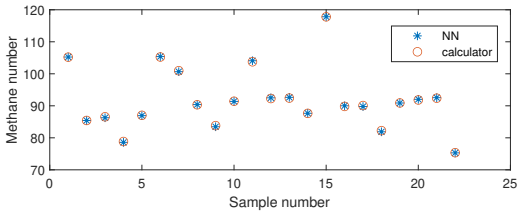


Fig. 4. Neural network performance on the test set.

methane number calculator developed using Microsoft Excel and BASIC [12].

Since the methane number is a nonlinear function of the gas components and running on an embedded system is required, a neural network was trained for calculating the methane number. For this purpose, on-site gas contents measurements were used. The permeate gas measurements obtained during one year are given in Table I. The natural gas measurements from the same period are approximately constant and similar to those of permeate gas in April 2021. With this data, an augmented dataset of 110 samples was created with random linear combinations of adjacent samples, as shown in Figure 3.

A relatively small network, with one hidden layer and two neurons with a radial basis activation function, was trained using Bayesian regularization and an 80/20 train/test split. It proved to be complex enough and performed well on the test set (MSE=0.0547), as shown in Figure 4. The main difference between this network and the one presented in [7] is the network size; in this case, a much smaller network is sufficient since there is less variation in the dataset.

Given the contents of the natural and permeate gases, the neural network calculates the methane number for each one. The methane number of the engine fuel, however, depends on the mixing ratio of the two components. For simplicity, it was assumed that the contents of the mix can be calculated as:

$$c_{mix} = rc_p + (1 - r)c_n \tag{4}$$

where c_i are the contents of component i and r is the mixing ratio. The methane number of the mix is then calculated for different values of the mixing ratio and the boundary ratio values which satisfy the methane number limits (set by the manufacturer or the operator) are found.

Since the methane number is a nonlinear function of the components, it is possible that there are several mix ratio

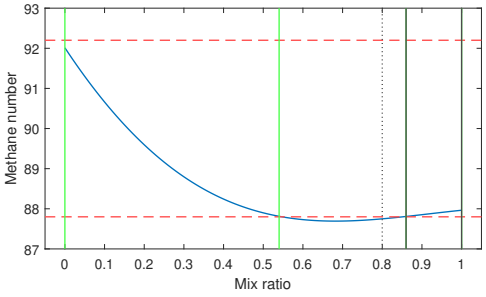


Fig. 5. Ratio limit selection.

intervals which satisfy the methane number constraint. However, only one interval can be included in the optimal control problem formulation presented in the next section. In that case, the interval which contains the current mix ratio is chosen to avoid oscillations in plant operation. If the current mix ratio is not in any of the allowed intervals, the closest permitted value is chosen. Additionally, if the gas contents are such that the methane number constraint cannot be satisfied for any mix ratio, the ratio which will produce the best possible methane number (although outside the specifications) is selected.

An illustration is given in Figure 5, where the green line indicates the permitted mix ratio intervals, the dotted black line shows the current mix ratio and the full black lines indicate the selected mix ratio limits.

D. Model predictive fuel mix ratio controller

In order to minimize the use of gas flares and the operating costs, while limiting the tank pressure and power tracking error, a model predictive controller was used to determine the optimal fuel mix ratio. The reason for choosing the mix ratio instead of the individual gas flows as the output variable was the already existing engine controller, which determines the final flows according to the desired mix ratio and generated power.

Generally, a model predictive controller seeks the control sequence such that the user-defined cost function (which contains, e.g., energy usage, reference tracking error, safety limit violations) is minimized. The total cost function is calculated over a prediction horizon of open-loop behavior, consisting of several discrete time steps. Also, the choice of optimal control values is subject to state and control constraints (e.g., safety boundaries and actuator limitations).

1) Discrete-time model: Firstly, a discrete-time model of the plant is needed for MPC. Expressions (1) and (3) lead to the following state-space representation:

$$\begin{bmatrix} P \\ p \end{bmatrix}^+ = \begin{bmatrix} 1 - \frac{T_s}{\tau} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P \\ p \end{bmatrix} + \begin{bmatrix} \frac{T_s}{\tau} k_n & \frac{T_s}{\tau} k_p \\ 0 & -\frac{T_s p}{\rho V} \end{bmatrix} \begin{bmatrix} q_n \\ q_p \end{bmatrix} \tag{5}$$

where T_s denotes the chosen sampling time. Note that the incoming flow of permeate gas, q_i, is not included in this model since it is not measured/controlled and can be treated as an unknown disturbance. Also, it is assumed

TABLE I
PERMEATE GAS CONTENTS SAMPLED DURING ONE YEAR

component	8.1.2021.	3.2.2021.	3.3.2021.	6.4.2021.	5.5.2021.	7.6.2021.	6.7.2021.	10.8.2021.	2.9.2021.	4.10.2021.	3.11.2021.	mean value
C1 [mol%]	60.58	62.02	75.08	80.13	66.71	67.86	61.71	67.86	76.1	55.866	68.471	67.49
C2 [mol%]	1.28	1.44	1.63	4.21	1.26	1.36	8.23	2.07	6.53	10.159	6.751	4.08
C3 [mol%]	0.32	0.13	0.15	0.51	0.15	0.17	2.31	0.24	0.96	1.715	0.972	0.69
i-C4 [mol%]	0.1	0	0.05	0.12	0.05	0.05	0.45	0.04	0.17	0.294	0.161	0.14
n-C4 [mol%]	0.28	0	0	0.04	0.02	0.09	0.59	0.02	0.11	0.22	0.085	0.13
i-C5 [mol%]	0.16	0	0	0	0	0.04	0.18	0	0.05	0.082	0.033	0.05
n-C5 [mol%]	0.24	0	0	0	0	0.06	0.13	0	0	0.07	0	0.05
C6+ [mol%]	0.72	0.07	0.12	0.31	0.03	0.59	0.31	0.05	0.23	0.305	0.06	0.25
N2 [mol%]	4.12	4.58	6.1	6.48	7.45	7.65	1.58	7.66	4.77	2.312	4.077	5.16
CO2 [mol%]	32.21	31.77	16.87	8.2	24.33	22.14	24.51	22.08	11.09	28.977	19.39	21.96

that both the generator power and the tank pressure are measured. Finally, the required output, optimal mixing ratio, can be obtained from the optimal gas flows, q_n and q_p, as $r = q_p/q_n$. The resulting model (5) is a discrete linear time-varying system:

$$x^+ = Ax + B(x)u, \quad x = [P\ p]^\top, \quad u = [q_n\ q_p]^\top \quad (6)$$

which can be used in the MPC design.

2) The objective function: The objective function indicates which variables are important, i.e., should be minimized in the optimization. In this instance, the primary aim of the controller is to minimize deviations from the reference and the operating cost. The objective function can then be formulated as:

$$V(x, u, r) = \sum_{i=0}^{N} \|x_i - r\|_Q^2 + \|u_i\|_R^2 \quad (7)$$

where N is the prediction horizon, $x = [x_0, \ldots, x_N]^\top$ and $u = [u_0, \ldots, u_N]^\top$ the state and control sequences, r the state reference and Q and R positive semi-definite weight matrices. The notation $\|v\|_M^2$ denotes the quadratic form $v^\top Mv$.

Reference tracking errors of engine power and tank pressure can be penalized individually by choosing Q to be diagonal. Also, operating cost minimization can be achieved by setting:

$$R = \begin{bmatrix} r_n & 0 \\ 0 & r_p \end{bmatrix}, \quad r_n \gg r_p \quad (8)$$

which will also result in mix ratio maximization, i.e., permeate gas will be used as much as possible and flaring will be minimized.

3) Constraints: Ideally, the optimal control problem should be formulated as a linear or quadratic program, since there are efficient methods and solvers for solving them [13]. To achieve this, the control input matrix in (5) is assumed to be constant during the horizon, using the current tank pressure measurement. If the tank pressure does not change significantly during the prediction horizon, this simplification should not affect the results. Another option would be to use the controller predictions from the previous time step. The system dynamics can then be included as linear constraints in the optimization problem.

Additionally, the gas flows q_n and q_p should be constrained between zero and maximum flows, $q_{n,max}$ and $q_{p,max}$. The mix ratio constraints from the neural network can be added with the following expression:

$$r_{min} \leq \frac{q_p}{q_n} \leq r_{max} \quad (9)$$

Since the natural gas flow is nonnegative, this can be rewritten as a linear inequality:

$$q_n r_{min} \leq q_p \leq q_n r_{max} \quad (10)$$

If the optimization argument includes the state, an initial state constraint is needed such that the predictions start from the current measurement. Finally, the engine power and tank pressure can also be constrained to lie between the minimum and maximum allowed values, $x_{min} \leq x \leq x_{max}$.

4) The optimization problem: The resulting optimization problem can be formulated as:

$$\min_{x,u} \quad V(x, u, r) = \sum_{i=0}^{N} \|x_i - x_r\|_Q^2 + \|u_i\|_R^2 \quad (11a)$$

$$\text{s.t.} \quad x_i^+ = Ax_i + Bu_i \quad (11b)$$

$$x(0) = x_0 \quad (11c)$$

$$x_{min} \leq x_i \leq x_{max} \quad (11d)$$

$$u_{min} \leq u_i \leq u_{max} \quad (11e)$$

$$u_{1,i} r_{min} \leq u_{2,i} \leq u_{1,i} r_{max}, \quad \forall i \in [1, N] \quad (11f)$$

The cost function is quadratic in the optimization variables and all constraints are linear. Therefore, the problem can be written as a standard quadratic program:

$$\min_x \quad \frac{1}{2} x^\top Px + q^\top x \quad (12a)$$

$$\text{s.t.} \quad l \leq Ax \leq u \quad (12b)$$

and the optimal solution can be obtained using a QP solver such as OSQP [14] or MATLAB's quadprog [15].

E. The simulation model

A simulation of the plant and the control algorithm was developed using MATLAB and Simulink. In the simulation, gas contents are periodically sampled from the dataset used to train and test the neural network, i.e., the on-site measurements. The neural network uses these measurements to calculate the allowed fuel mix ratio. The mix ratio limits are given to the MPC, which outputs the optimal mix ratio that satisfies the imposed constraints and minimizes the specified cost. Finally, this ratio is used by

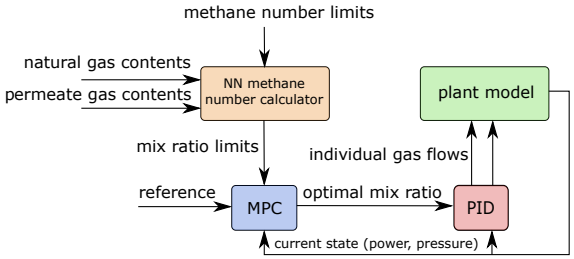


Fig. 6. The simulation scheme.

TABLE II
SIMULATION PARAMETERS

Parameter	Value	Unit	Description
q_{max}	300	L/s	maximum fuel flow
$q_{i,max}$	0.1667	L/s	maximum incoming permeate gas flow
τ	20	s	power response time constant
V	35	m^3	buffer tank volume
N	20	–	MPC prediction horizon
Q_P	$1/1000^2$	$1/(kW)^2$	MPC power tracking error weight
Q_p	$1/60^2$	$1/(kPa)^2$	MPC pressure tracking error weight
R_n	$1 \cdot 10^{-4}$	$1/(L/s)^2$	MPC natural gas flow cost
R_p	$1 \cdot 10^{-6}$	$1/(L/s)^2$	MPC permeate gas flow cost
T_s	1	s	control loop sampling time

a lower-level PID controller that sets the individual flows of the gases in order to achieve the desired engine power output.

The control loop runs with a sampling time of 1 second, while the engine power and tank pressure dynamics are simulated using continuous-time models (1) and (3). An upper bound on the tank pressure is not imposed since the excess gas can be flared. Also, the incoming permeate gas is modeled as a uniformly distributed noise of moderate amplitude.

The simulation scheme is shown in Figure 6 and the most important parameters are given in Table II.

III. RESULTS

The primary objective of the described control scheme was to automatically set the mixing ratio of the two gases. Therefore, closed-loop simulations were compared with the case where the mix ratio is set manually and constant. For calculating the cost of fuel, it was assumed that the unit price of natural gas is 1 and that of the permeate gas is 0.3. Two different scenarios were considered.

In the first scenario, the natural and permeate gas contents were obtained by blending the points obtained from on-site measurements (Figure 3). In this case, the natural gas content varies slowly and the methane content in the natural gas is approximately constant, while the permeate gas is more volatile.

In the second scenario, the gas contents are randomly selected from the combined dataset, and this scenario represents a mixing of two volatile gasses where it is more difficult to meet the methane number limits.

Figures 7 and 8 show the simulation results for the first scenario with the gas contents dataset shown in Figure 3. In both cases, engine power, tank pressure and methane number are within the specified limits. However, in the

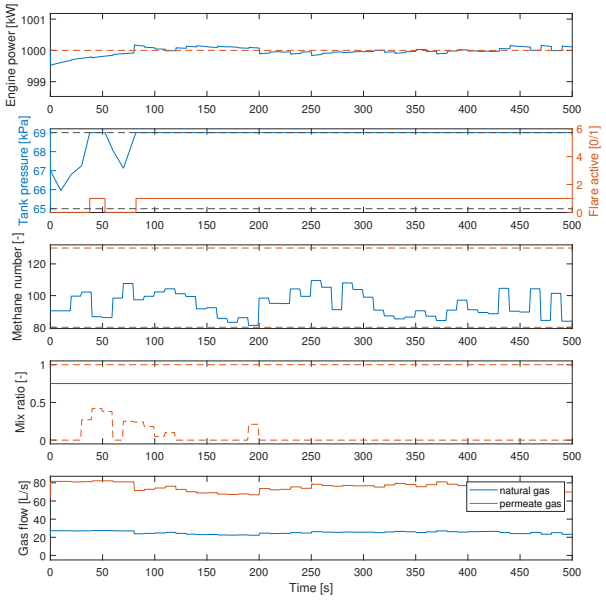


Fig. 7. Simulation results with a manually set mix ratio and sequentially chosen gas contents.

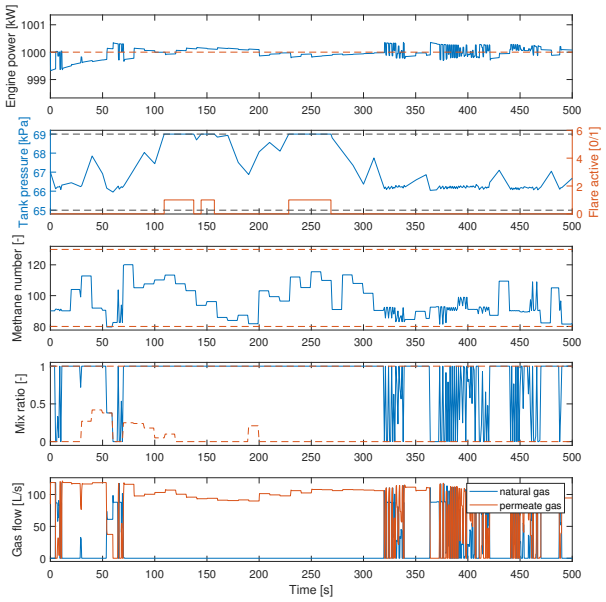


Fig. 8. Simulation results with an automatically set mix ratio and sequentially chosen gas contents.

second case, the control input oscillates more, while the total fuel cost is reduced by 16% and the flaring time is reduced by 81%.

Figures 9 and 10 show the results for the second scenario when the gas contents are chosen randomly from the combined dataset, instead of sequentially from separate datasets for each gas. In the first case, the power output is fairly stable, but the tank pressure is too high and the flares are activated more often. Additionally, the methane number limits are exceeded several times as the manually set mix ratio does not fall within the allowed interval. In the second case, the control inputs again oscillate more, but the duration of flaring is reduced by approximately

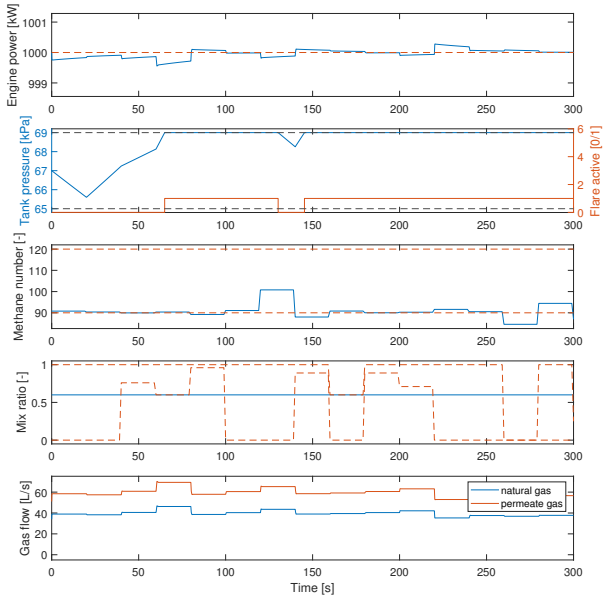


Fig. 9. Simulation results with a manually set mix ratio and randomly chosen gas contents.

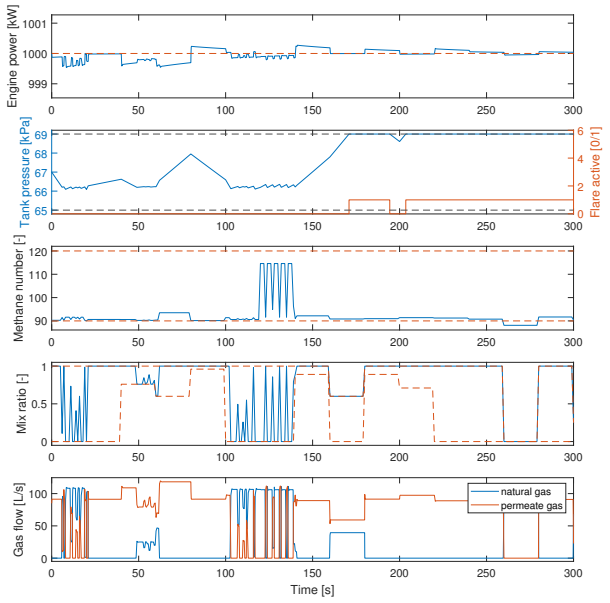


Fig. 10. Simulation results with an automatically set mix ratio and randomly chosen gas contents.

45%. The methane number limits are slightly violated only once, when the gas contents are such that the methane number constraints cannot be met for any mix ratio. Also, there is no significant difference in the power output compared to the first case. Finally, the total cost of fuel is reduced by 19% in the second case.

IV. Conclusion

This paper presents artificial intelligence-based predictive fuel blending control for flare gas mitigation. Since the methane number is a nonlinear function that depends on gas composition, a neural network was trained to calculate it for different fuel blends and find the fuel mixing ratio values that satisfy the limits on methane number provided by the motor manufacturer. With these constraints on the mix ratio, a finite-horizon optimal control problem is solved where the main goal is to output the desired power while minimizing both the cost of the fuel and gas flaring for all possible control inputs. The optimal solution is then applied to the plant and the whole process is repeated. The algorithm was tested in two scenarios where the natural and permeate gas contents were sampled sequentially or randomly from the real data obtained from the site. The simulation results show that the proposed algorithm is capable of producing the desired power while mitigating costs and flaring, but also respecting the constraints on the fuel methane number.

Acknowledgment

The authors would like to thank Bojan Spahija for his contribution to this work, especially in data analysis and plant modelling.

References

[1] N. Tahouni, M. Gholami, and M. H. Panjeshahi, "Integration of flare gas with fuel gas network in refineries," *Energy*, vol. 111, pp. 82–91, 2016.

[2] M. Nezhadfard and A. Khalili-Garakani, "Power generation as a useful option for flare gas recovery: Enviro-economic evaluation of different scenarios," *Energy*, vol. 204, p. 117940, 2020.

[3] A. Norouzi, H. Heidarifar, M. Shahbakhti, C. R. Koch, and H. Borhan, "Model predictive control of internal combustion engines: A review and future directions," *Energies*, vol. 14, no. 19, p. 6251, 2021.

[4] C. J. Muller, I. K. Craig, and N. L. Ricker, "Modelling, validation, and control of an industrial fuel gas blending system," *Journal of Process Control*, vol. 21, no. 6, pp. 852–860, 2011.

[5] S. Kuczyński, M. Łaciak, A. Szurlej, and T. Włodek, "Impact of liquefied natural gas composition changes on methane number as a fuel quality requirement," *Energies*, vol. 13, no. 19, p. 5060, 2020.

[6] C. J. Kreutzer, D. B. Olsen, and R. J. Bremmer, "Evaluation of a lean-burn natural gas engine operating on variable methane number fuel," in *Internal Combustion Engine Division Fall Technical Conference*, vol. 44427, 2011, pp. 159–166.

[7] S. K. Gupta and M. Mittal, "Predicting the methane number of gaseous fuels using an artificial neural network," *Biofuels*, vol. 12, no. 10, pp. 1191–1198, 2021.

[8] Y. A. Çengel, M. A. Boles, and M. Kanoğlu, *Thermodynamics: An engineering approach*, 9th ed. New York: McGraw-Hill Education, 2019.

[9] C. Long, C. Holland, and E. Gatzke, "Experimental air-pressure tank systems for process control education," *Chemical Engineering Education*, vol. 40, pp. 24–31, 12 2006.

[10] Wärtsilä, "Wärtsilä Methane Number Calculator." [Online]. Available: https://www.wartsila.com/marine/products/gas-solutions/methane-number-calculator

[11] Cat, "Cat Methane Number Calculator." [Online]. Available: https://www.cat.com/en_US/articles/solutions/oil-gas/Cat-Methane-Number-Calculator.html

[12] J. Torrens, "Methane Number EN16726," 2021. [Online]. Available: https://github.com/TorrensJoaquin/MethaneNumber_EN16726

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[14] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: an operator splitting solver for quadratic programs," *Mathematical Programming Computation*, vol. 12, no. 4, pp. 637–672, 2020. [Online]. Available: https://doi.org/10.1007/s12532-020-00179-2

[15] MATLAB, "Quadprog package - functions to solve quadratic programming problems." [Online]. Available: https://www.mathworks.com/help/optim/ug/quadprog.html

Program Solution for Regulation and Monitoring of Evaporator Section in Individual Quick-Freezing Process

Tomislav Špoljarić, Davor Gadže, Ivan Šulekić
University of Applied Sciences, Zagreb, Croatia
tomislav.spoljaric@tvz.hr, davor.gadze@tvz.hr, ivan.sulekic@tvz.hr

***Abstract* - In this paper a program solution for control and monitoring of evaporator section in cooling tunnel is proposed. Control strategy for freezing process is designed to work according to temperature probes used for measurement of temperature of frozen goods. Defrost control is designed to work with active probes used for ice thickness measurement thus providing more efficient defrosting of pipeline section in evaporator. Monitoring of certain values of significance is maintained by using pressure observation in condenser circuit and temperature observation expanded to room/tunnel temperature along with probes used in control (goods' temperature). Also, significant actuator states are also monitored: evaporator fan speed and defrosting heater switching state.**

Keywords* - *Dynamic defrost control, Evaporator control, Individual quick freezing

I. INTRODUCTION

Individual quick-freezing process (IQF) is a dynamic cooling process that takes place in cooling tunnels and is used for freezing and storage of perishable foodstuff in food processing industry such as fish and meat.

A time-maintained control of such freezing process was introduced in [1] and [2] and which included the use of time blocks in LADDER program solution, as opposed to standard simple cooling process which implement mostly one cycle of cooling and optional defrost [3]. This control strategy uses heuristic approach by observing room temperature as indirect indicator on how the goods are being refrigerated and how the icing on the evaporator section is being defrosted. Since it is mainly time-controlled, a strategy defined in [1] is also limited to one cycle of goods' preparation after which a storage of frozen goods must take place in another cold room, thus making the manipulation of goods more complex. This manipulation's time limit is described as follows [1]:

$$T_p = \sum_{i=1}^{n} \left(T_{pr,i} + T_{r,i} + T_{pd,i} + T_{d,i} \right) \qquad (1)$$

where:

- $T_{r,i}$ – refrigeration period,
- $T_{d,i}$ – defrost period,
- $T_{pr,i}$ – pause period before refrigeration,
- $T_{pd,i}$ – pause period before defrosting.

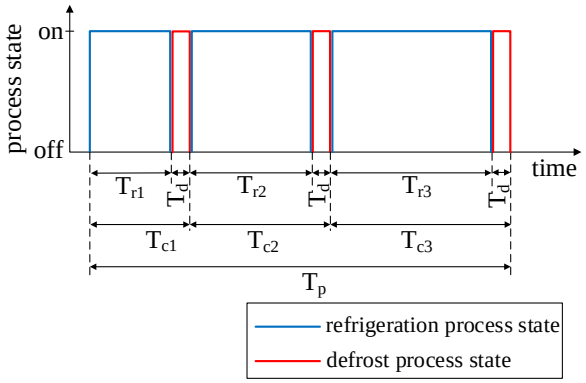


Fig 1. Time flow diagram of a three-cycle IQF process with manipulation time limit [1]

In this paper a novel control strategy is proposed for individual quick-freezing process, with regards to optimal control of defrosting heaters and regulation of cooling process with much more accurate measurement with several additional features such as room temperature and condenser pressure monitoring.

II. CONTROL STRATEGY FOR IQF PROCESS AND STORAGE OF GOODS IN COOLING TUNNEL

A. Cooling process with evaporator defrosting

A cooling effect is obtained in a cold room (in this example a cooling tunnel) by use of refrigerant and various elements of a refrigeration system such as compressor, condenser, and evaporator. In this process a refrigerant is pressurized at different pressure levels in these elements. Heat is absorbed from cold room through evaporator and transported to condenser where it is being released. To reinforce this cooling effect fans on evaporator and condenser are mounted for circulation of air in the environment [4].

Since the icing on the pipeline of the evaporator can be a problem in environments below 0°C, defrosting heaters are mounted in its body with sole purpose to melt the ice and secure the effective circulation of air.

Control of switching of evaporator heaters in a complex cooling process can be demanding to implement, since in dynamic freezing processes the moisture release is very non-linear. Thus, a non-standard control approach by means of ice-thickness monitoring is implemented and tested on a cooling tunnel.

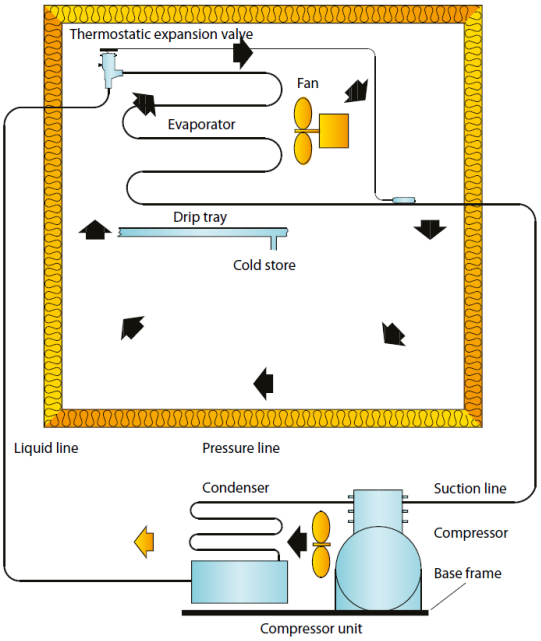


Fig 2. Refrigeration process – principle [3]

B. *Direct measurement used for regulation*

Several improvements are being introduced in this control strategy regarding control that is described in [1], most importantly a use of aseptic and hygienic temperature probe *iTherm TM412* (Endress+Hauser) with 4-20 mA integrated transmitter for direct measurement of goods' temperature and defrost sensor *HBDF-MK2* (HB Products) for dynamic defrost control.

Use of these sensors ensures more dynamic control of cooling/defrost cycles, with program solution being much more efficient, since there is no need to implement timing blocks. With these products cooling tunnel can be more efficient and can be also used as a cold room storage once the IQF process finishes.

C. *Control and monitoring equipment selection*

Siemens S7-1500 modular PLC is used as a primary regulation device. Modules used for implementation, programing and testing are:

- CPU module *6AG1516-3AN01-2AB0*,
- AI module *6ES7531-7KF00-0AB0*,
- DIO module *6ES7523-1BL00-0AA0*.

Also, SINAMICS AC drives are used for regulation of fans' speed according to refrigerant condensing pressure and these devices are connected via PROFINET communication with the CPU module. Condensing pressure is thus connected to analog input of AC drives and this information, along with switching/alarm states of condenser fans, is forwarded via PROFINET to the CPU.

HMI device used for monitoring and reference input is Siemens's touch panel, type *TP700 Comfort* and is connected to PLC via PROFINET for easy user control and observation of measured signals in a controlled system.

A pressure probe is indirectly used for monitoring of condenser's pressure (Danfoss *DST P110, 0-30bar).*

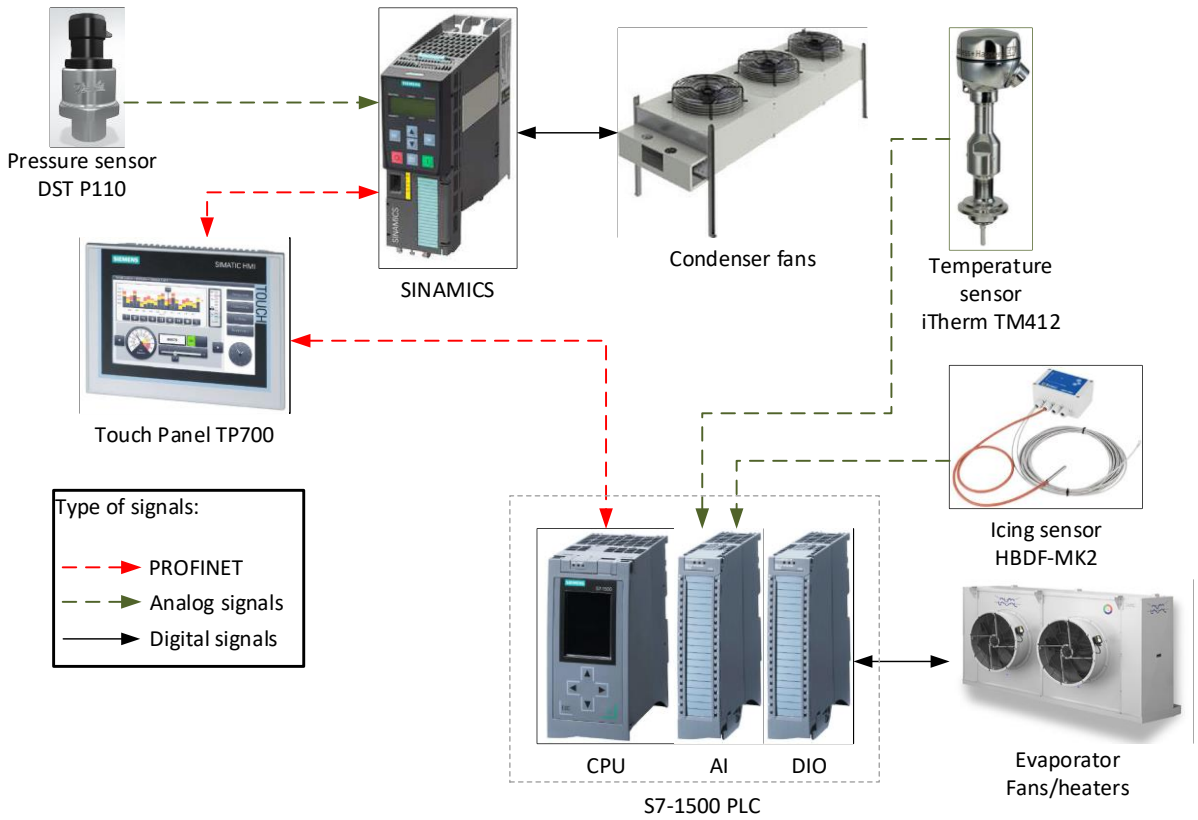


Fig 3. Block schematics of a control system and refrigeration system's elements

III. PROGRAM SOLUTION FOR CONTROL

A. PLC workflow and addresses

TIA portal package was used for development of program solution for control and monitoring of a cooling tunnel. For control of a cooling process two temperature sensors are being used, and temperature reference is obtained by:

$$T_{ref} = \frac{T_{probe1} + T_{probe2}}{2} \quad (2)$$

These sensors measure temperature of goods in cooling tunnel during refrigeration and room temperature. With two sensors a more accurate measurement is being performed. This reference is used for switching the cooling elements on by use of user set temperature and a dead zone (usually set to 2°C). Tested system uses reference of -35°C with dead zone of +2°C, making the dead zone in between -35°C and -33°C, for better assurance of quality of goods [5, 6].

Icing sensor is used as a measuring element for evaporator defrost control, and thickness of the ice is set from 0 mm to size of the fin spacing in evaporator's body (tested system uses evaporators with 12mm fin spacing). Regulation reference is set to 10% with dead zone of 20% of the fin spacing range. Measurement scaling is:

$$\langle 0 \div fin\ space \rangle \xrightarrow{AI} \langle 4 \div 20mA \rangle \xrightarrow{PLC} \langle 0 \div 100\% \rangle \quad (3)$$

Condenser fans are being switched on after a cooling state is set to ON and are being regulated by use of AC drives with pressure as a reference for speed control of fans.

Since the icing of evaporator is primarily used as a reference for switching between states of refrigeration and defrost, a control strategy developed in observed system uses icing measurement as primary instrument in the alternation of cooling and defrosting cycles. These dynamics that are being introduced in control strategy make the whole system work in optimal state with regards to effectiveness of cooling and energy consumption.

Workflow diagram is shown on fig. 4. On this diagram one cycle of PLC is observed. First cycle starts with setting the process in the state ON by user. Process is being maintained in this state until end user turns it off via user interface on HMI device.

Addresses used for observed states and measurements are shown in table I. These signals are being processed in PLC throughout one cycle of PLC's operation and forwarded to HMI device for user observation and control.

TABLE I. – SIGNAL ADDRESSES

Signal	*Type*	*Communication*	*Data type*	*Address*
Cooling state ON	Digital	DIO module	Bool	Q0.0
Defrost state ON	Digital	DIO module	Bool	Q0.1
Ice thickness	Analog	AI module	Real	IW4
Temperature 1	Analog	AI module	Real	IW6
Temperature 2	Analog	AI module	Real	IW8
Evaporator fan state	Digital	DIO module	Bool	I1.0
Evaporator heater state	Digital	DIO module	Bool	I1.1
Evaporator fan alarm	Digital	DIO module	Bool	I1.2
Evaporator heater alarm	Digital	DIO module	Bool	I1.3
Condenser/Compressor fan state	Digital	PROFINET	Bool	PI80.0
Condenser fan alarm	Digital	PROFINET	Bool	PI80.1
Condenser ON	Analog	PROFINET	Real	PQ60.0
Condenser pressure	Analog	PROFINET	Real	PID84
Condenser fan speed	Digital	PROFINET	Real	PID88

Digital outputs are used for switching the evaporator fans and electromagnetic valve (Q0.0) and for switching defrost heaters (Q0.1). These signals are mutually exclusive. As a feedback information for these actuators, digital inputs I1.0 and I1.2 are used for evaporator fans while I1.1 and I.3 are used for evaporator heaters.

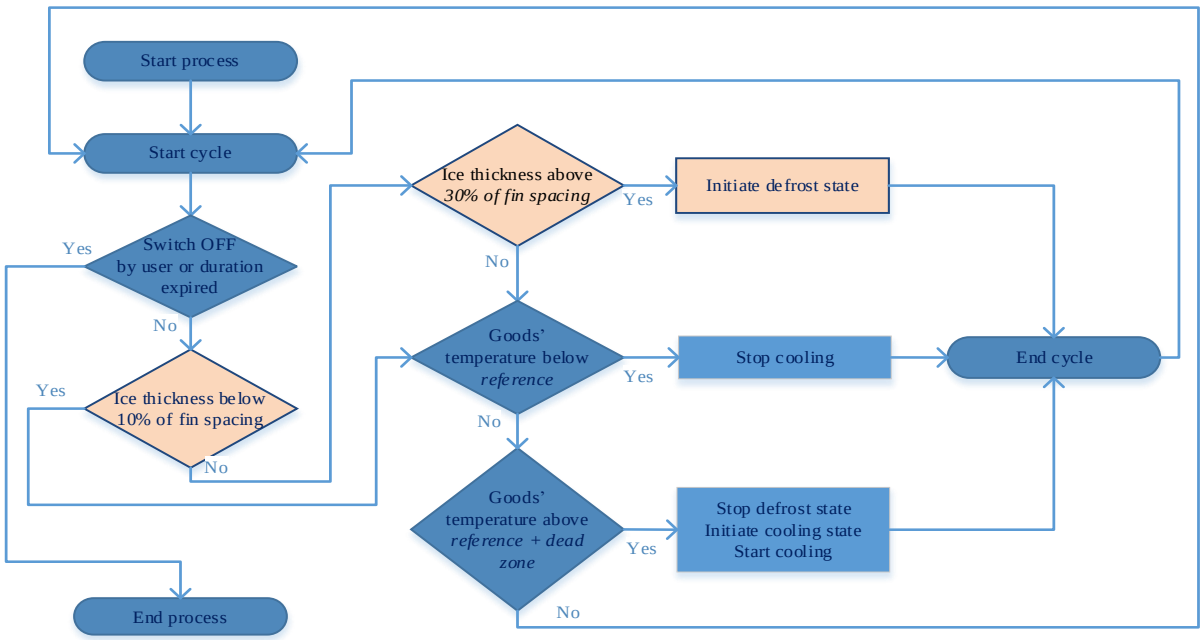


Fig 4. Work flow diagram of one cycle in control process of defrost and cooling states of evaporator

B. HMI user interface

User interface consists of four different screens which implement the possibility of parameter change, monitoring of significant refrigeration values and observation of process finishing. These screens are navigated through according to logic presented on fig. 5.

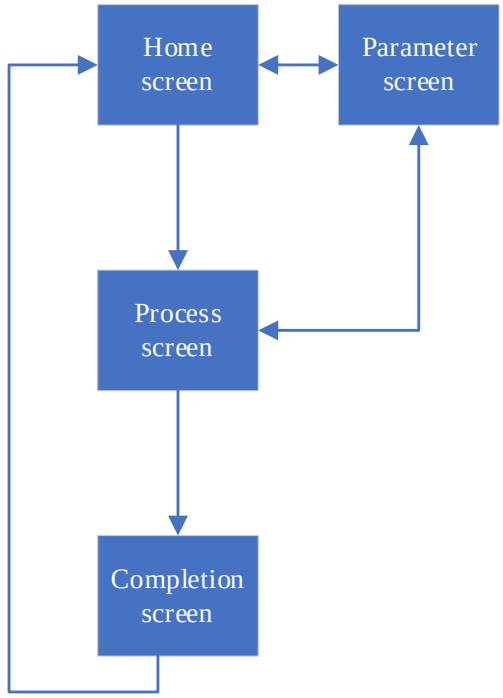


Fig 5. Refrigeration process – principle [3]

Home screen is a starting screen for user, and it contains possibilities to navigate to parameter screen or start the pre-parameterized quick-freezing process, as is shown on fig. 6.

Parameter screen is designed to set the cooling temperature reference and relative thickness of icing according to evaporator fin spacing, as is shown on fig. 7. These parameter settings are to be set according to limits: maximum value of reference temperature is -5°C and ice thickness is 90% of absolute fin spacing. These maximum values are defined for the greater flexibility of user's choice of frozen goods and defrost control according to ice thickness. Also, an IQF process duration can be defined.

Once the parameters are set and the process is initiated by the user, a process screen appear in which a user can observe significant process values such as goods' temperature, room temperature, evaporator temperature, refrigerant pressure in suction pipeline and relative ice thickness on evaporator. Parameters can be changed also by navigating to parameter screen, and process can be also stopped by the user. Process screen is shown on fig. 8.

Last screen that can be obtained after a freezing process is done is a completion screen, shown on fig. 10. After a process completes after defined duration (set on parameter screen) a single message appears, noting the user that a quick-freezing process is complete and that the further manipulation of goods may be performed.

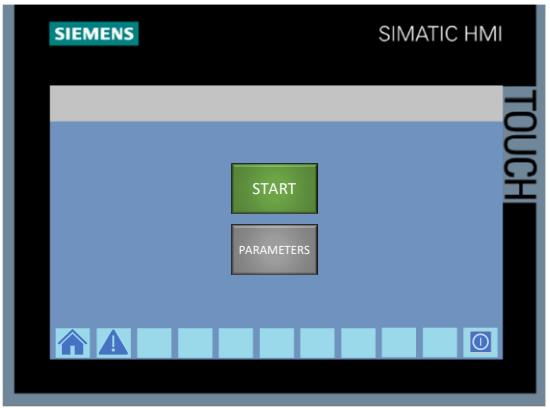


Fig 6. HMI - home screen

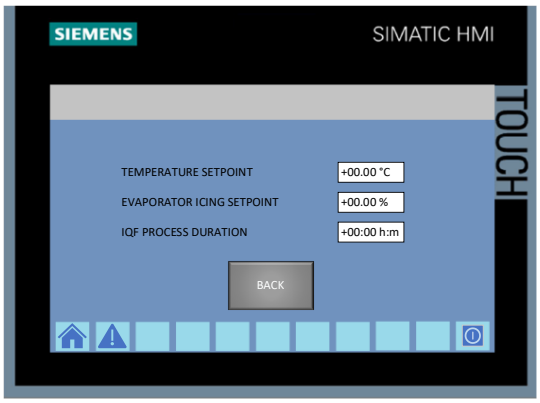


Fig 7. HMI – parameter screen

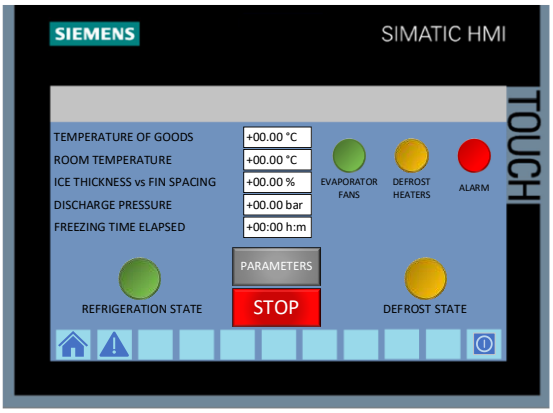


Fig 8. HMI – process screen

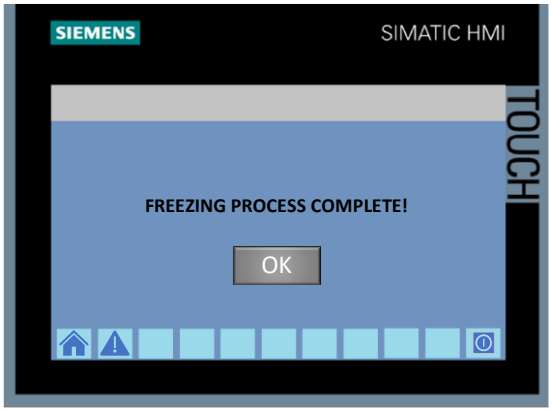


Fig 9. HMI – completion screen

IV. DEFROST CONTROL EVALUATION – TESTING RESULTS

Observations for the testing case were made for a cooling tunnel used for applying IQF process in freezing a fresh fish. Nominal carrying capacity of a tested cooling tunnel is 15 tons of goods in a single intake.

Cooling system of a tested tunnel has the same principle of operation as described in [1], with the main changes being made to defrost regulation of the evaporator, used by means of icing thickness dead-zone control. Also, changes in the principle of condenser fan regulation have been made. Condenser fans are operated by the SINAMICS devices, thus giving a finer approach to pressure and condensing temperature regulation.

Three test cases were observed, and these vary according to carrying capacity of a cooling tunnel: 70%, 100% and 120% of a nominal capacity (15 tons of fresh fish). Temperature median was obtained as mean value of room and goods' temperature readings and was used for evaporator fan switching state control by means of a dead zone regulation. Icing thickness readings are used for defrosting heater switching state control.

While defrost mode is on, heaters are always on. When appropriate icing thickness is achieved, heaters are switched off, cooling state control takes place, and evaporator fans are switched on according to temperature median.

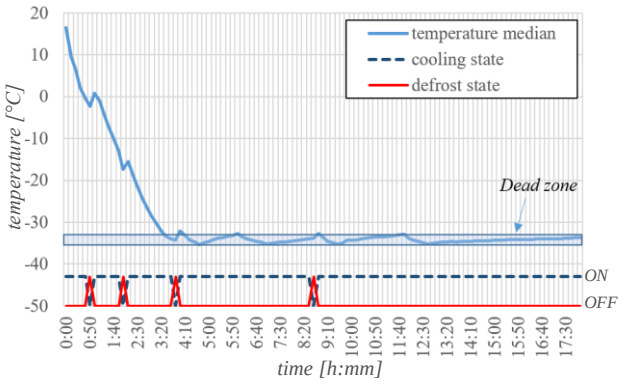


Fig 10. Temperature median graph of an IQF process in a tested cooling tunnel – 70% of nominal carrying capacity

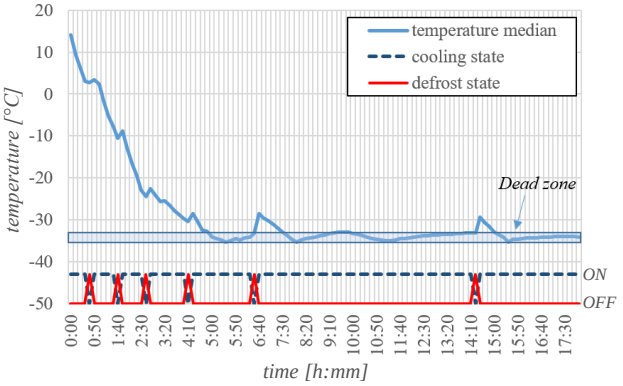


Fig 13. Temperature median graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

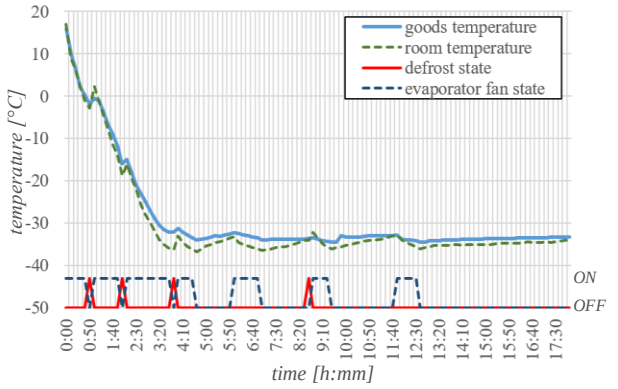


Fig 11. Room and goods' temperature graph of an IQF process in a tested cooling tunnel – 70% of nominal carrying capacity

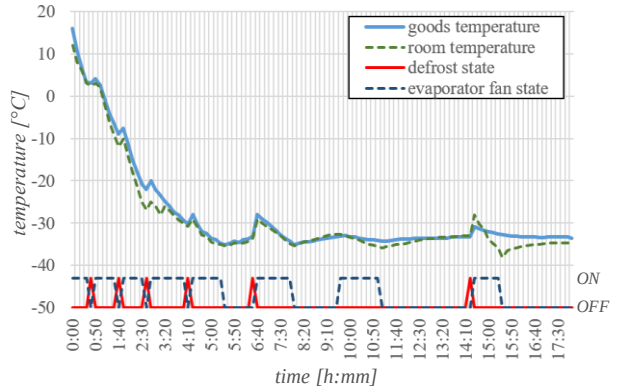


Fig 14. Room and goods' temperature graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

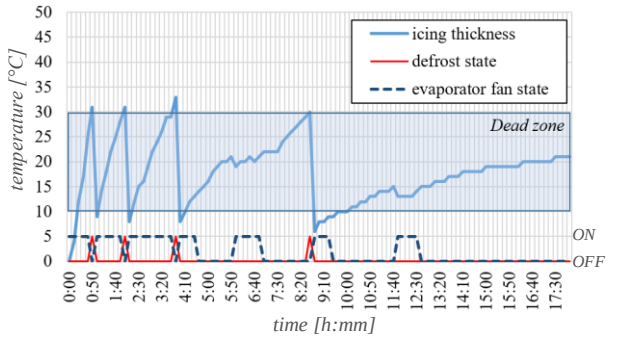


Fig 12. Icing thickness graph of an IQF process in a tested cooling tunnel – 70% of nominal carrying capacity

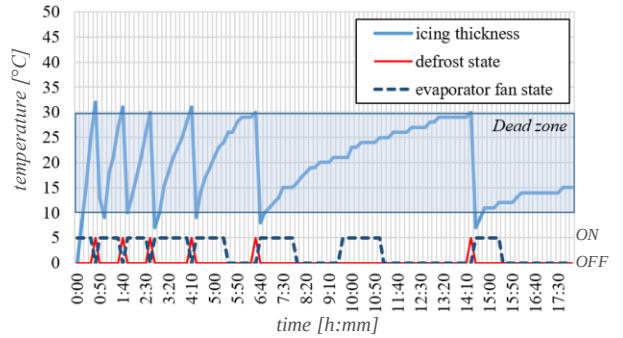


Fig 15. Icing thickness graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

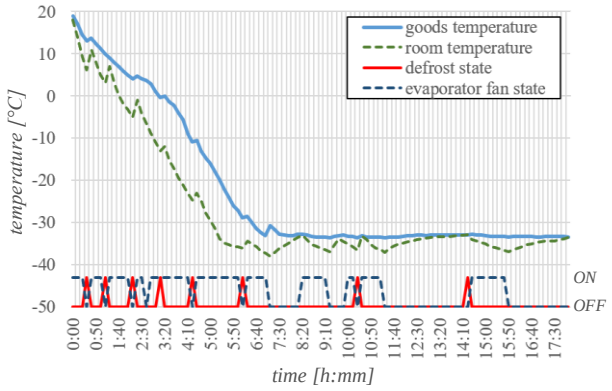


Fig 16. Temperature median graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

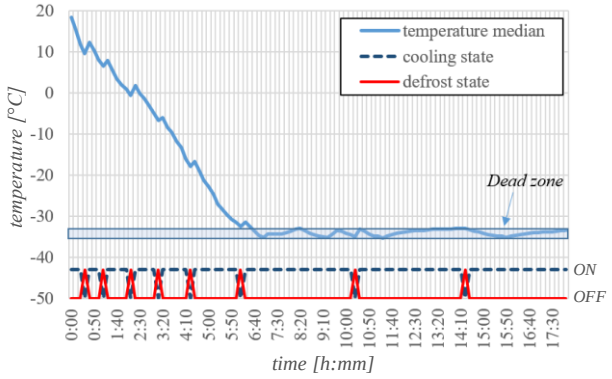


Fig 17. Room and goods' temperature graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

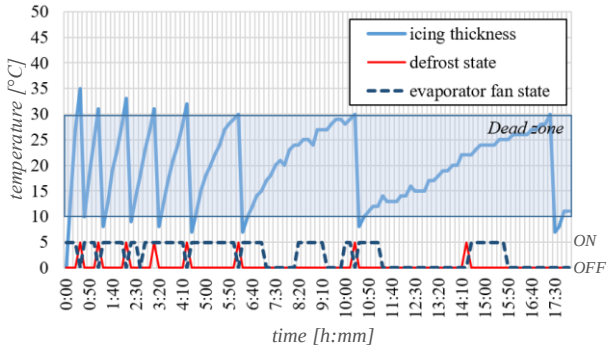


Fig 18. Icing thickness graph of an IQF process in a tested cooling tunnel – 100% of nominal carrying capacity

Results of testing the icing thickness control strategy for 15-ton fresh fish cooling tunnel show that in each of the test cases a referenced median temperature is obtained in the first half of the freezing process duration cycle. Number of defrost cycle increases according to tonnage of the load being frozen. More accurate measurements were made by two probes: one used for direct measurement and placed in goods and one used for room temperature monitoring. A deviation gap between these two measurements was slightly more significant in a case of a slightly overloaded tunnel, but in general, it can be concluded that these gaps are of acceptable value for effective freezing of goods.

When compared to control strategy presented in previous research tests [1] for the same tunnel, control strategy via icing thickness outperforms time-maintained control strategy for the cases of 100% and 120% carrying capacity of a tested tunnel. As a consequence lesser and variable defrost duration times prove to be a better solution for overall refrigeration system performance than the previously implemented control strategy. Also, an ice-thickness control strategy is not that much dependent on an amount of goods being frozen, as much as is the strategy described in [1] and [2], which is a main problem in quick achieving a desired temperature and a quality of frozen goods in an overloaded tunnel.

V. CONCLUSION

Dynamic control of defrost duration and switching was implemented with the idea of upgrading an existing refrigeration process control. The main reason lies in the modernization process of an existing refrigeration system with the purpose of more flexible use of a cooling tunnel. With the more dynamic control of evaporator section (mainly, a defrosting process), an end-user can utilize the tunnel for freezing various amounts and types of goods, not only the ones that it was initially designed for. Also, a flexibility is manifested through lesser time that a tunnel performs a freezing to a certain goods' temperature point.

An HMI interface was made as simple as possible, since only temperature and ice thickness variables are parameters of interest to the evaporator control strategy process.

Although the process control with static defrosts and fixed number of defrost cycles is a more cheap and sufficiently effective control strategy for monitoring and regulation of an IGF process in cooling tunnels, a described dynamic control in this paper seriously outperforms the former method. This evaporator control method optimizes the use of heaters for defrost through ice thickness monitoring, while also switching evaporator fans in a more effective manner during a cooling state. This effectiveness is manifested through a better circulation of a cold air in a tunnel during cooling, since there is a strict control of an ice volume on the body of an evaporator.

With the more effective use of all the actuating devices in an evaporator section, and with the SINAMICS on a condenser section, an electric power consumption is more effective, thus giving a possibility of an investment return.

REFERENCES

[1] T. Špoljarić, M. Fruk, G. Vujisić, "Program Solutions for the Complex Cooling Process via Ladder Logic" 37th International convention MIPRO Proceedings, pp. 1221–1226, May 2014

[2] G. Jagetić, M. Habazin, T. Špoljarić, "PLC/SCADA systems in automation control design for individual quick-freezing process in cooling tunnels" 38th International convention MIPRO Proceedings, pp. 1896–1903, May 2015

[3] T. Špoljarić, M. Špoljarić, "Program Solution for Multi-probe Monitoring and Control in Simple Cooling Process", 39h International convention MIPRO Proceedings, pp. 1339–1346, May 2016

[4] Refrigeration – an introduction to the basics, lecture handbook Danfoss, 2007.

[5] D. R. Heldman, R. P. Singh, Food Process Engineering, 2nd ed., Springer, 1981

[6] R. T. Toledo, Fundamentals of Food Process Engineering, Springer 2007.

Electrical Profile of User with Photovoltaic System

I. Petrović*, D. Koprivanac** and E. Cirković**
* Parpar d.o.o. / Electrical Engineering, Bjelovar, Croatia
** Virovitica University of Applied Sciences / Electrical Engineering, Virovitica, Croatia
igor@parpar.hr, danijel.koprivanac@vuv.hr, enes.cirikovic@vuv.hr

Abstract **- Energy and fuel-distribution management performed by suppliers is usually based on daily predictions created by distributors. Those predictions are based on data measured the day before; weather predictions such as ambient temperature, wind, precipitation along with some general information such as day of the week, whether it is weekend or not, are there any preannounced events within the distribution area etc. There's been a rapidly growing trend in demand for micro photovoltaic systems in single user installation for the last few years, which implies that classic user profiles are becoming obsolete for more users, both new and existing ones. This research focuses on a single example representing a consumer with an integrated micro photovoltaic system. In order to keep electrical energy management up to date with latest developments, it is essential to create a user profile for consumers with micro photovoltaic systems. This paper focuses on a classic user profile combined with estimated energy production prediction for a micro photovoltaic system, which results in a new profile of a user with a micro photovoltaic system. Measured data will express current user energy consumption as a consumer, compared to the same user energy consumption/production when the consumer is upgraded with a photovoltaic system. With the new profile, it is possible to analyze main characteristics of transformation of the user profile.**

Keywords – electrical energy; energy planing; energy consumption; modeling

I. INTRODUCTION

Energy becomes most precious asset when geopolitical situation introduces unstable energy market. In these times it is most necessary to be able to manage energy production wisely and as cheep as possible, as described in [1] and [2]. In order to do so, one must have access to information and mathematical models which will allow rather accurate prediction of energy consumption in time ahead. Mathematical models of energy consumption behavior is usually generated from monitoring and recording energy consumption in past. From these information it is possible to generate simple or complex presumption of energy consumption for one day, week or month in advance, called profile. Such profiles, as described in [3] can even use ambiental temperature, weather forecast, etc. In this paper simple mathematical model and measured data is used for transformation of electrical energy consumption of classic user profile to user profile with photovoltaic system, like used in [4]. Results are compared using single day measured data and basic statistical analysis. All models are based on Kirchhoff's junction rule using standard measured data.

II. ELECTRICAL CONSUMPTION LOAD PROFILE

Classic energy consumption user profile is most simple production-consumption system where all energy claim of such user is provided from distribution network. Therefore, the electrical energy passed through electrical counter is actual electrical energy consumed by user, as shown in Figure 1.

Energy consumption user with photovoltaic system is introducing second energy source working in parallel with distribution network. User is described as passive block which can only take energy, and never produce any of it. Such models are presented in [5] and [6]. At the same time photovoltaic system is described as active block which can only produce energy, and never consuming any of it. Distribution network is transformed from active block for classic user profile into reactive block for user with photovoltaic system. Reactive block is described as part of system which can adjust to system demands, and therefore can produce energy in the system when there is energy shortage or can absorb energy from system when there is extra produced energy in system, as shown in Figure 2.

Distribution network can also be described as large battery bank with endless capacity. Therefore, one can take reduced cost of photovoltaic system working in parallel to distribution network.

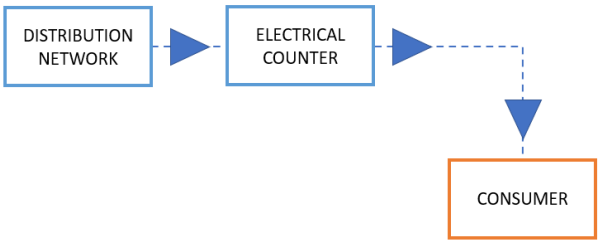


Figure 1. Energy flow in Classic consumption user profile

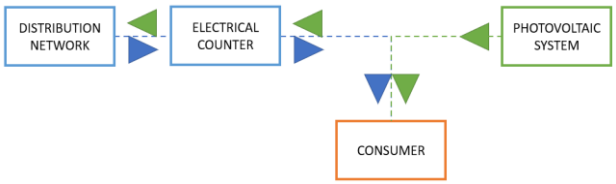


Figure 2. Energy flow in Consumption user profile with photovoltaic system

III. MODELING OF USER PROFILE WITH PHOTOVOLTAIC SYSTEM

TABLE I. LIST OF ENERGY FLOWS FOR USER PROFILE

No.	*Name*	*Description*	*Symbol*
1	Consumer instalation	Only consumes electrical energy	$\mathbf{E_I}$
2	Photovoltaic system	Only produces electrical energy	$\mathbf{E_{PV}}$
3	Electrical energy counter - positive	Energy taken from distribution network	$\mathbf{E_+}$
4	Electrical energy counter - negative	Energy given into distribution network	$\mathbf{E_-}$

When transferring from classic user profile to user with photovoltaic system profile, there is only one installation point that is important. It is the point of connection between distribution network and consumer installation. Photovoltaic is introduced as third party, integrating into system at that very point.

In order to describe classic user profile and user with photovoltaic system profile, one must consider all possible electrical energy flows. From Figure 1. and Figure 2. such list of energy flows can be observed. The list is presented in Table 1.

A. *Classic user profile*

When energy flows list from Table 1. is introduced to Figure 1., it can be seen that there are three major statements describing classic user profile. They are listed in Equation 1., Equation 2. and Equation 3.

$$E_I = E_+ \tag{1}$$

Since there is no photovoltaic system, there is no active production of electrical energy inside consumer installation, or possibility to give any energy to distribution network.

$$E_{PV} = 0 \tag{2}$$

$$E_- = 0 \tag{3}$$

B. *User with photovoltaic system profile*

When energy flows list from Table 1. is introduced to Figure 2., it can be seen that there is only one statement describing user with photovoltaic system profile. It is expanded Equation 1. with additional energy coming in system from photovoltaics (E_{PV}), but part of that energy can pass through into distribution network (E_-). Expression is presented in Equation 4.

$$E_I = E_{PV} + E_+ - E_- \tag{4}$$



Figure 3. Sample building with integrated photovoltaic system

All presumptions made in chapters III.A. and III.B. will be analyzed using measured data from distribution network and photovoltaic system, all supplying electrical consumption of sample building with integrated photovoltaic system.

This photovoltaic system is 103 kW peak power of production, which is provided with 4 inverters with 25 kW nominal production power each. Figure 3. shows placement of photovoltaic modules on the roof of sample building.

IV. MEASURED DATA AND PROFILE RESULTS

In order to transform classic user profile to user profile with photovoltaic system it is necessary to introduce exact photovoltaic system information and classic consumption measured values. Corresponding energies from Equation 1. are measured and presented in Figure 4.

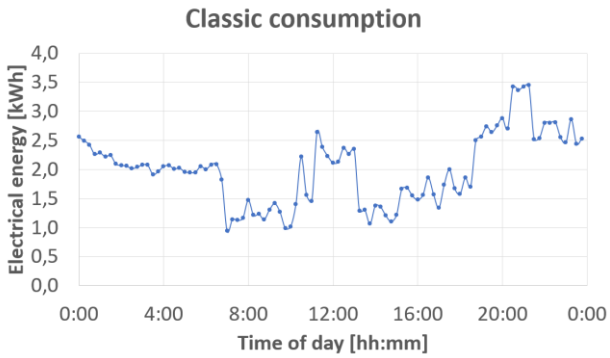


Figure 4. Daily Classic consumption for Sample building, November 2022

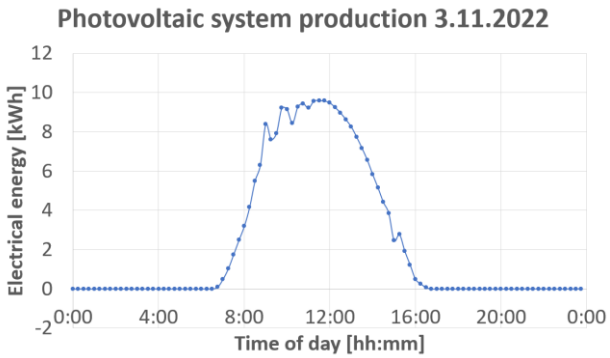


Figure 5. Daily Photovoltaic system energy production, November 2022

Transformation to user with photovoltaic system requires measured data from photovoltaic system, given with energy production regardless of whether it will be consumed inside installation or given into distribution network. Photovoltaic system production can be planed with simulations, but real-time measured data is used in this research in order to get valid results and conclusions. Photovoltaic system production is presented in Figure 5.

Electrical energy counter is split in two separate measurements, one for energy taken from distribution network to consumer and second one for energy given from photovoltaic system to distribution network. Separation is done for financial purposes only, so in this research counter values of E_+ and E_- can be presented as joined value. Measured values presented in Figure 6. and Figure 7. are selected for days where weather conditions were very similar to days of measured values in Figure 4. and Figure 5. Same dates, 3rd November in two separate years are presented. First in Figure 6. real measured data is presented for selected date in 2021, and then in Figure 7. same date in 2022. Therefore, all given measured values in Figures 4-7 are valid comparative information.

If placing measured values of electrical counter energy flow and photovoltaic system production in Equation 4., one can calculate generic classic user consumption by combining Counter energy flow (E_+ - E_-, Figures 6. and 7.) with photovoltaic system energy production (E_{PV}), Figure 5.).

Final results in Figure 8. and Figure 9. confirm that user energy consumption is still the same as it was without photovoltaic system. It also confirms that classic user profile is still valid. Introduction of photovoltaic system only adds new direction of energy flow, so while planning one must take into consideration specific properties of all three participants in energy exchange. Main goal is to get the same energy distribution quality as it was before integration of photovoltaic system, but at the same time disburden distribution network.

TABLE II. DAILY AVERAGE ELECTRICAL POWER VALUES

No.	*Name*	*Value*	*Unit*
1	Profile 2022	1,99	kW
2	3. 11. 2021.	3,30	kW
3	3. 11. 2022.	1,88	kW

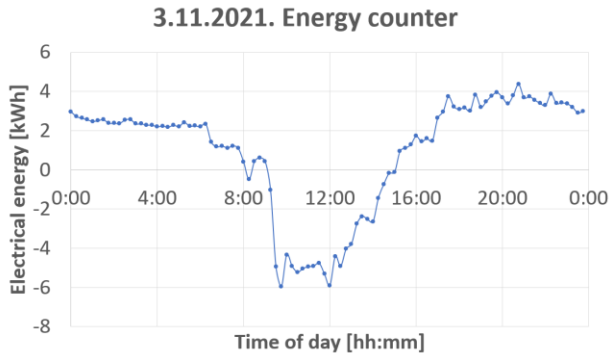


Figure 6. Daily energy flow trough Counter, 3rd November 2021

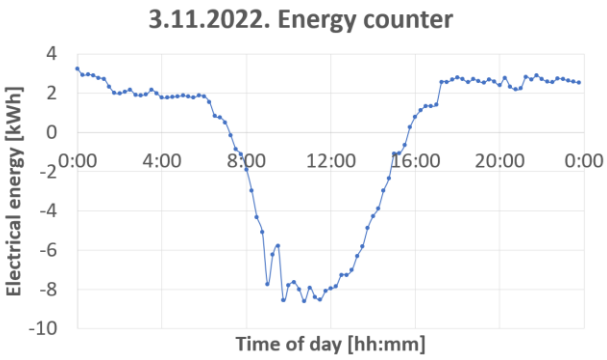


Figure 7. Daily energy flow trough Counter, 3rd November 2022

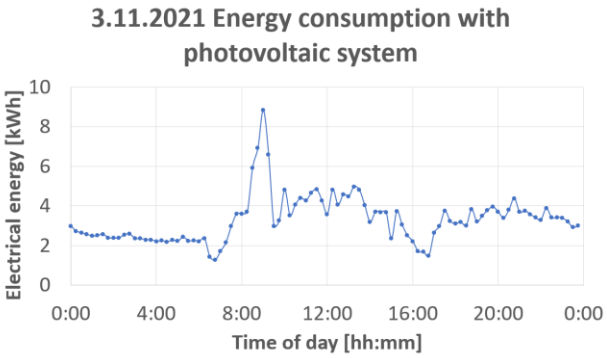


Figure 8. Generic user energy consumption, 3rd November 2021

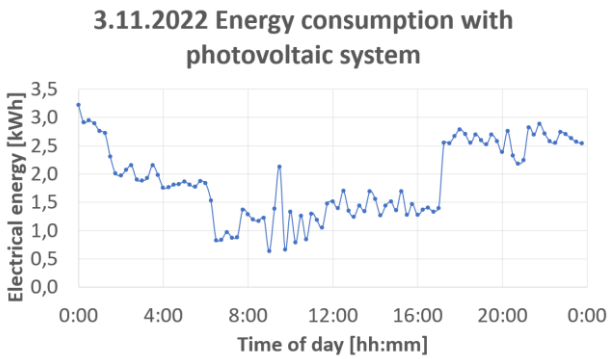


Figure 9. Generic user energy consumption, 3rd November 2022

In order to be able to compare given results presented in Figure 4. (Profile 2022) with generic user energy consumption it is necessary to transform each of them in per unit values. In this research transformation is done using daily average value. Daily average values for electrical power are presented in Table 2.

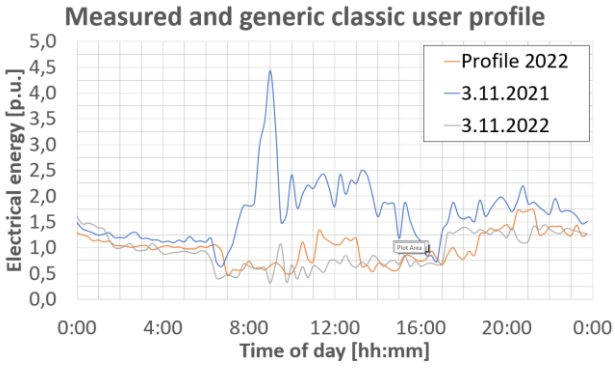


Figure 10. Transformed daily consumptions for energy comparison

Transformed daily energy consumptions are presented in in Figure 10. When comparing each result to another one can conclude that:

- Daily averages of Profile 2022 and 3. 11. 2022. are relatively close in values, while 3. 11. 2021. has some significant deviation.
- Night consumptions for all daily diagrams are relatively close in values, while day consumption has some significant deviation.
- Daily consumptions for Profile 2022 and 3. 11. 2022. have relatively small gap between daily minimum and daily maximum values, while 3. 11. 2021. has relatively large gap between daily minimum and daily maximum values

V. CONCLUSION

In order to maintain good energy management and reasonable energy resource usage it is necessary to have usable model of distribution network and electrical energy consumers. Implementation of photovoltaic system in classic user consumption installation changes behavior this user energy flow. In this research model of transformation from classic user to user with photovoltaic system. Transformation model is derived from profile data while user did not have installed photovoltaic system, and energy production planning for photovoltaic system. Results are presented, compared and analyzed. Transformation model is verified using data of measured classic profile and generic consumption load profile after photovoltaic system is implemented.

After analyzing and comparing results one can conclude that transformation model from classic user profile to user with photovoltaic system is more accurate if user profile matches real daily consumption. Therefore, transformation model will be at least equal in profile accuracy to classic user profile, and will not introduce new errors in energy management and energy resource usage.

REFERENCES

[1] M. H. Radhi, E. J. Mahdi, and A. K. Mftwol, "Design and Performance Analysis of Solar P.V. System Size 2.56 kWp", 4th Scientific International Conference Najaf (SICN), pp. 70–73, 2019

[2] A. F. Meyabadi, M. H. Deihimi, „A review of demand-side management: Reconsidering theoretical framework", Renewable and Sustainable Energy Reviews, Vol. 80, pp. 367-379, 2017

[3] O. Alrawi, I. S. Bayram, S. G. Al-Ghamdi, M. Koc, "High-Resolution Household Load Profiling and Evaluation of Rooftop PV Systems in Selected Houses in Qatar", Energies, Nr. 12, Vol. 20, 3876

[4] M. Kamran, M. R. Fazal, M. Mudassar, S. R. Ahmed, M. Adnan, I. Abid, F. J. S. Randhawa, H. Shams „Solar photovoltaic grid parity: a review of issues, challenges and status of different PV markets", International Journal of Renewable Energy Research, Nr. 9, Vol. 1, pp. 244-260, 2019

[5] U. B. Tayab, F. Yang, M. El-Hendawi, J. Lu, "Energy management system for a grid-connected microgrid with photovoltaic and battery energy storage system", Australian & New Zealand Control Conference, pp. 141-144, 2018

[6] H. J. Jabir, J. Teh, D. Ishak, H. Abunima, "Impacts of demand-side management on electrical power systems: A review", Energies, Nr. 11, Vol. 5, 1050.

Analiza rada fotonaponske elektrane radi optimizacije potrošnje energenata

Danijel Koprivanac*, Igor Petrović ** i Ivan Heđi*
* Veleučilište u Virovitici, Virovitica, Hrvatska
** Parpar d.o.o., Bjelovar, Hrvatska
danijel.koprivanac@vuv.hr, igor.petrovic@parpar.hr, ivan.hedi@vuv.hr,

Energetska politika Europske unije naglašava razvoj obnovljivih izvora energije kao alternativu klasičnim skupim izvorima energije. Geopolitička situacija u svijetu i rast cijena energenata u 2022. godini, dovodi do bržeg povrata investicija u obnovljive izvore energije. Zadatak ovog rada analizira fotonaponsku elektranu kao jedan primjer obnovljivog izvora energije, te provodi analizu same potrebe za električnom energijom. Prikazani su mjereni podaci potrošnje i proizvodnje električne energije građevine čija je električna instalacija opremljena fotonaponskom elektranom snage 100 kW u paralelnom radu sa distribucijskom mrežom. Također, prikazani su rezultati indirektno izračunatih parametara i veličina za dva slična dana u uzastopnim godinama čiji su hidrometeorološki podaci slični, te analizirana energetska slika građevine koja se može koristiti u predviđanju proizvodnje fotonaponske elektrane. Zaključci ovog istraživanja su prikazani grafičkim putom i omogućavaju lakše planiranje i ekonomično korištenje energenata, ali i planiranje investicija i nadogradnje sustava koji troše energiju i energente.

Ključne riječi - Fotonaponska elektrana, proizvodnja električne energije, potrošnja električne energije, energetska bilanca

I. UVOD

U situacijama nestabilne geopolitičke situacije, pri nestabilnom energetskom tržište, energija postaje ekonomski sve značajniji element. U takvim okolnostima, proizvodnja energiju na ekonomičan način postaje imperativ, zbog kojeg to je nužno imati pristup informacijskim i matematičkim modelima koji će omogućiti preciznije predviđanje buduće potrošnje energije [1].

Pohranjivanje podataka o prethodno potrošenoj energiji je uobičajen temelj koji se analizira pri predviđanu buduće potrošnje. Jedan od načina je izrada vlastite baze podataka i njegova prilagodba za izradu izvještaja [2]. Pristup tim podacima te njihovo prikupljanje može se olakšati postavljanjem baze podataka u cloud radi dostupnosti samih podataka drugim aplikacijama bez potrebe izrade vlastitog rješenja (AWS cloud) [3]. Ujedno, treba uzeti u obzir da opskrba energijom je ključna za normalno odvijanje ljudskog razvoja i održavanje životnog standarda te dostupnost podataka u slučajevima prirodne nepogode ili pandemije [4].

Stoga, promatrani uzorkovani dani su prikazani očekivanim grafovima proizvodnje i potrošnje u odnosu na same navike korisnika građevine.

Ovaj rad govori o mogućnosti stvaranja obrasca potrošnje energije, dobivenih analizom uzorkovanih podataka, koji se može iskoristiti u predviđanju potrošnje energije hidrometeorološki sličnih dana te koristeći postojeću foto naponsku elektranu optimizirati raspodjelu potrošnje [5]. Osim toga, dijagrami potrošnje, koji su prikazani, trebali bi prikazati na nedostatke sustava, te ukazati na investiciju u dodatne izvore energije koji bi nadopunjavali postojeće.

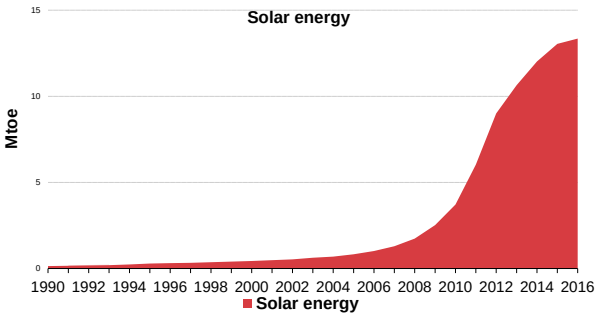
Ujedno je objašnjen tok električne energije za slučaj studentskog doma sa pripadajućim potrošačima, koji promatranu građevinu ne promatra kako standardnog potrošača, već kao reaktivnog koji se prilagođava zahtjevima mreže.

II. OPIS PREDMETA ANALIZE

Porast proizvodnje električne energije iz obnovljivih izvora energije u Hrvatskoj u 2020. godini iznosila je gotovo 3000 GWh što je šestorostruko povećanje u odnosu na 2010. godinu, a primjetno povećanje se može vidjeti i u odnosu 2020.g. i zadnje obrađene 2021.g., što je i prikazano na slici 1. Gledano na ukupnu proizvodnju (uz izuzetak velikih hidroelektrana) iznosi 22,3%. Udio proizvodnje električne energije dobivene iz solarnih izvora u Hrvatskoj u odnosu na ukupnu proizvodnju električne energije obnovljivih izvora energije (OIE)

Vrsta izvora	Proizvodnja električne energije (GWh)	
	2020.g.	2021.g.
Sunce	95,50	148,90
Vjetar	1720,70	2061,80
Biomasa	558,90	659,60
Bioplin	419,30	440,20
Male hidroelektrane	102,00	115,20
Geotermalna	93,70	89,70
Ukupno	2990,10	3515,40

Slika 1. Proizvodnja električne energije iz obnovljivih izvora energije u Hrvatskoj u 2020. i 2021. godini



Slika 2. Primarna proizvodnja solarne energije, EU-28, 1990.-2016.

iznosila je za 2020. godinu 3,19%. [6], te je povećanje prikazano na slici 2 [7]. Samom povećanju su doprinijele poticajne cijene za isporučenu električnu energiju koja su za 2021. godini prikazane na slici 3.

Analiza rada foto naponske elektrane (FNE) jedan je od važnih koraka u optimizaciji potrošnje energije. Cilj ove analize je identificirati načine za povećanje učinkovitosti i smanjenje troškova energije.

U okviru ove analize je usporedba dva slična hidrometeoroloska dana s ciljem identificiranja sličnosti u potrošnji energije te informacije koje proizlaze iz toga. To omogućava da se utvrdi je li bilo nekih varijabli koje su utjecale na potrošnju energije tijekom tog razdoblja.

Uz to, potrebno je pratiti i duži vremenski period kako bi se utvrdile trendove potrošnje energije i na temelju toga donijeti odgovarajuće mjere. Praćenje proizvodnje je omogućeno preko web portala, koji ujedno služi kao arhiva.

III. RAZRADA ANALIZE

Na slici 4. prikazana je fotonaponska elektrana (FNE) snage 100 kW instalirana na krovu građevine s kojom čini jednu energetsku cijelinu u smislu potrošnje električne energije.

Građevinu u kojoj se sva energetska potraživanja osiguravaju iz distribucijske mreže te je električna energija na brojilu stvarna električna energija koja je

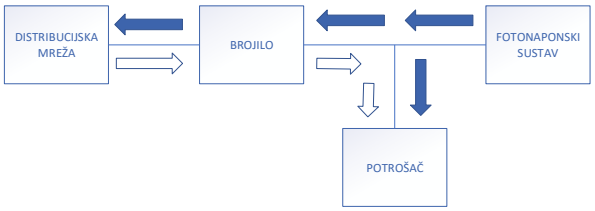
Sunčane elektrane instalirane snage ≤ 1 MW	Korekcija poticajne cijene za integrirane SE instalirane snage ≤ 300 kW				
Tip postrojenja	Poticajna cijena	Korektivni koeficijent za integrirane SE	Korekcija fiksne tarifne stavke	Korektivni koeficijent za korištenje toplinskog sustava	Korekcija fiksne tarifne stavke
	C (kWh)	k1	Ck = (C x k1)	k2	Ck = (C x k1 x k2)
1.a.1. Sunčane elektrane instalirane snage do uključivo 10 kW	1,10	2,39	2,63	1,20	3,16
1.a.2. Sunčane elektrane instalirane snage veće od 10 kW do uključivo 30 kW	1,10	2,03	2,23	1,10	2,45
1.a.3. Sunčane elektrane instalirane snage veće od 30 kW	1,10	1,50	1,65	1,03	1,70
1.a.3. Sunčane elektrane instalirane snage veće od 30 kW do 1 MW	1,10	-	-	-	-

Slika 3. Poticajne cijene za isporučenu električnu energiju iz sunčanih elektrana instalirane električne snage ≤ 1 MW za 2021.



Slika 4. Prikaz građevine sa instalacijom FNE

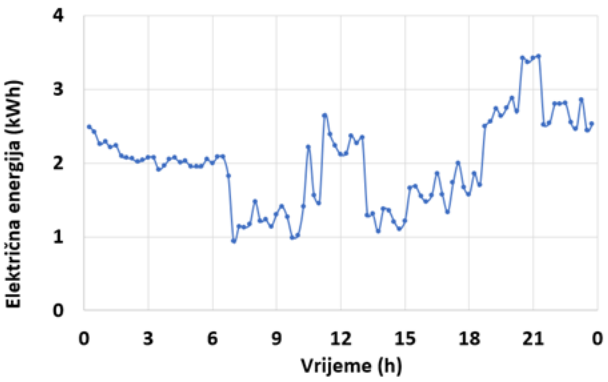
potrošena u građevini. Ukoliko je paralelno s distribucijskom mrežom uveden s fotonaponski sustav, kao sekundarni izvor, tada je potrošače moguće opisati kao pasivni blok koji može samo uzimati energiju, a nikada je ne proizvodi. U isto vrijeme fotonaponski sustav se prikazuje kao aktivni blok koji isključivo ima mogućnost proizvodnje energije. Distribucijska mreža se transformira iz aktivnog bloka u reaktivni blok za potrošače s fotonaponskim sustavom. Ujedno, reaktivni blok se prikazuje kao dio koji se može prilagoditi zahtjevima sustava, te stoga može proizvoditi energiju u sustavu kada postoji manjak energije ili može apsorbirati energiju iz sustava kada postoji višak proizvedene energije u sustavu (slika 5).



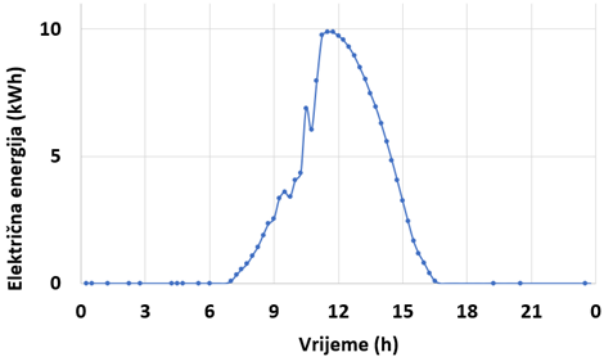
Slika 5. Tok električne energije

A. *Karakteristična dnevna potrošnja*

Prikaz karakteristične dnevne potrošnje građevine (karakteristične srijede) u kojoj se vanjska temperatura zraka kretala od 4° do 16° C, moguće je vidjeti na slici 6. [8]



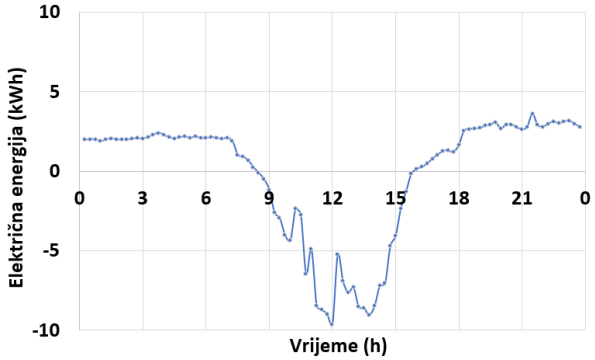
Slika 6. Dnevni dijagram potrošnja građevine



Slika 7. Proizvodnja FNE 01.11.2022.

Na samom grafu se može primjetiti relativno ujednačena potrošnja tijekom dana, koja odgovara namjeni samog objekta - studentskog doma. Ujedno možemo reći da je ista krivulja primjenjiva na svaki radni dan sa izuzetkom ljetnog odmora u kojem se potrošnja smanjuje zbog odsutnosti samih korisnika građevine.

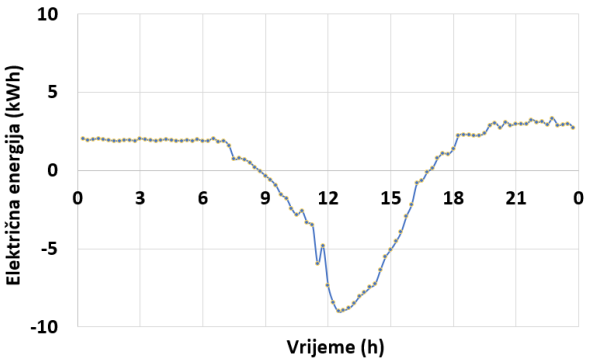
Potrebno je napomenuti da glavni energent vezan za grijanje je plin u sklopu centralnog plinskog sustava grijanja.



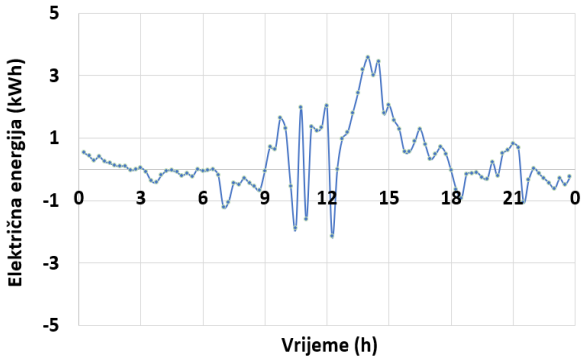
Slika 8. Razlika proizvodnje i potrošnje u 01.11.2021.

B. *Proizvodnja fotonaponske elektrane*

S konkretnim podacima stohastične proizvodnje i utjecaja zasjenjivanja u prijepodnevnim satima prikazana je krivulja proizvodnje dnevnog uzorkovanja na slici 7. Prikazana krivulja odgovora karakterističnoj proizvodnji



Slika 9. Razlika proizvodnje i potrošnje u 01.11.2022.



Slika 10. Prilagođena dnevna potrošnja energije za 01.11.2021

standardne FNE u Republici Hrvatskoj [9].

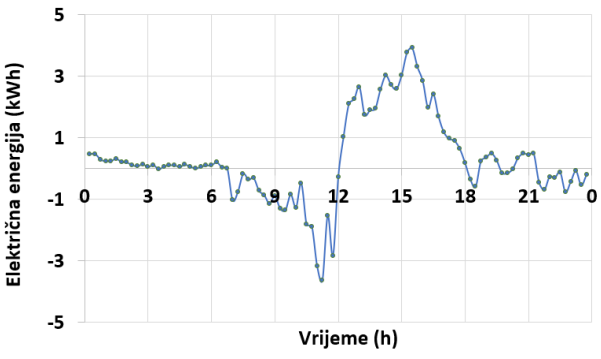
C. *Razlika proizvodnje i potrošnje*

Daljnje razmatranjem analizira razliku proizvodnje i potrošnje uzorkovanih dana sa ciljem dobivanja približnih vijednosti oba dana. Moguće je primjetiti stvaranja „viškova“ električne energije koje s predaju mreži. Iste karakteristike se je moguće vidjeti na slici 8. i 9. vezane za 2021. i 2022. godinu. Krivulja pokazuje isključivo mjerenu potrošnju preuzetu od mreže od koje je oduzeta izmjerena proizvodnja FN elektrane.

D. *Ekonomska bilanca proizvodnje*

Analizirana energetska slika građevine pokazuje smjernice za optimiziranje potrošnje. Višak proizvedene električne energije potrebno je preusmjeriti u samu potrošnju građevine zbog neisplativosti prodaje električne energije.

Vrijednosti prikazane na slikama 10 i 11 izračunate po principu prikazanom na slici 5. Željena vrijednost obje krivulje maksimalno približavanje nuli, što dovodi do optimizacije omjera proizvodnje i potrošnje instalacije na građevini



Slika 11. Prilagođena dnevna potrošnja energije za 01.11.2022

IV. ZAKLJUČAK

Planiranje potrošnje energije i energenata je više nego ikad od primarne važnosti za planiranje budućih troškova. Prikazanim uzorkovanjem pokazano je predvidljivo ponašanje korisnika, te ujedno prikazala grafički veću proizvodnju od same potrošnje. Ujedno, usprkos „režimu štednje“ koje je obuhvaća uzorkovani dan u 2022. godini, krivulje potrošnje su slične uzorkovanom danu u 2021. godini što nam olakšava predviđanja za budućnost i buduća ulaganja u infrastrukturu same građevine.

Analiza i usporedba rezultata prikazuju mogućnosti prilagodbe grijanja na električnu energiju postavljanjem inverterskih klima komora visoke učinkovitosti u samu građevinu, koje bi iskoristile proizvedenu električnu energiju za grijanje. Ujedno, postavljenjem automatizacije na izbor energenata, mogućnost trošenja „viška“ električne energije bi smanjila troškove koji trenutno postoje sa centralnim plinskim sustavom grijanja. Druga mogućnost je postavljanje baterijskog spremišta koje bi spremale višak električne energije za potrošnje nakon prestanka proizvodnje FNE, koja bi najviše došla do izražaja u ljetnim mjesecima kada potrošnja studentskog doma je niska a postoji prozvodnja FN.

REFERENCE

[1] L. Ciabattoni, G. Ippoliti, M. Benini, S. Longhi, and M. Pirro, “Design of a home energy management system by online neural networks,” 11th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing, Caen, France, pp. 677–682, July 2013.

[2] O. Jukić and I. Heđi, "The use of call detail records and data mart dimensioning for telecommunication companies," 2012 20th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2012, pp. 292-295

[3] O. Jukić, I. Heđi and E. Ciriković, "IoT cloud-based services in network management solutions," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 419-424

[4] I. Heđi, E. Ciriković and O. Jukić, "Improving city infrastructure resilience capacity: resilience management guidelines in mobile network," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 421-427

[5] R.Meenal, A.Immanuel Selvakumar, E.Rajasekaran “Review on mathematical models for the prediction of Solar radiation”, Indonesian Journal of Electrical Engineering and Computer Science Vol. 15, No. 1, July 2019, pp. 54~59 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v15.i1.pp54-59

[6] https://eihp.hr/wp-content/uploads/2023/01/Energija%20u%20HR%202021_WEB_LR.pdf (Pristup 08.05.2023)

[7] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Renewable_energy_statistics#Share_of_renewable_energy_more_than_doubled_between_2004_and_2021 (Pristup 08.05.2023.)

[8] https://www.accuweather.com/hr/hr/virovitica/116063/november-weather/116063?year=2022 (Pristup 01.02.2023.)

[9] https://oie.hr/elektroenergetska-kretanja-u-hrvatskoj-u-2022/ (Pristup 08.05.2023)

AIS

International Conference on

ARTIFICIAL INTELLIGENCE SYSTEMS

Steering Committee

Chairs: Darko Huljenić, Ericsson Nikola Tesla, Zagreb, Croatia
Alan Jović, University of Zagreb, Croatia

Members: Andrea Budin, Ericsson Nikola Tesla, Zagreb, Croatia
Bojan Cukic, UNC Charlotte, United States
Marko Đurasević, University of Zagreb, Croatia
Marina Ivašić-Kos, University of Rijeka, Croatia
Domagoj Jakobović, University of Zagreb, Croatia
Ruizhe Ma, University of Massachusetts, Lowell, United States
Neeta Nain, MNIT Jaipur, India
Stjepan Picek, University of Zagreb, Croatia
Slobodan Ribarić, University of Zagreb, Croatia
Vitomir Štruc, University of Ljubljana, Slovenia

Program Committee

Members: Mario Brčić, University of Zagreb, Croatia
Karla Brkić, University of Zagreb, Croatia
Marko Čupić, University of Zagreb, Croatia
Marko Đurasević, University of Zagreb, Croatia
Marko Horvat, University of Zagreb, Croatia
Ivo Ipšić, University of Rijeka, Croatia
Domagoj Jakobović, University of Zagreb, Croatia

PAPERS

ARTIFICIAL INTELLIGENCE THEORY

Binary Dynamic Models of Structural Synthesis of Programs

G.A. Oparin, V.G. Bogdanova and A.A. Pashinin
Matrosov Institute for System Dynamics and Control Theory of SB RAS, Irkutsk, Russia
prn51@icc.ru, bvg@icc.ru, apcrol@gmail.com

***Abstract* – In this paper, we propose a matrix-vector Boolean differential model for constructing plans of computational actions in solving non-procedural problems on a computational model of a modular software system. The conditions of the problem of structural synthesis of a program from pre-implemented modules are formulated as a system of Boolean differential equations, the solutions of which, under given initial conditions (consistent with the non-procedural formulation of the problem), determine the solvability of the problem and give constructive plans for its solution (including parallel ones). The proposed method of structural synthesis is focused on high-dimensional models. It allows highly efficient software implementation due to the high parallelism of performing vector-matrix operations on binary vectors at the level of machine instructions. The developed model is used for planning computations in packages of applied microservices based on the HPCSOMAS-MSC platform.**

Keywords – computational model; Boolean differential equations; abstract program; action planning

I. INTRODUCTION

There is a wide class of complex subject areas in which the process of computer solving of problems is represented as a step-by-step execution of modules-procedures from the set *M*, working on a common field of variables *Z*, which are the actual parameters of these procedures. Typical representatives of software systems with such an organization are computing software packages with functional content in the form of libraries of autonomously compiled modules (subroutines/functions written in the traditional sequential programming languages Fortran/C), as well as tool environments for organizing distributed computing, which allows integrating geographically remote heterogeneous autonomous computational resources into one resource-intensive multidisciplinary task.

A distinctive feature of this kind of applied computation is as follows: the variables from *Z* in these computations act as distinguished concepts of the theory and concurrent parameters of the problem statement and therefore have applied independent semantics. Many problems within some applied theories consist of the fact that it is required to find the required quantities from $B_0 \subset Z$, using other quantities $A_0 \subset Z$ whose values are considered known. The solution to such a computational problem $T = (A_0, B_0)$ (if it exists) is the composition of functions extracted from the computational model as a network of functional relations of the computability of quantities in the considered theory. It is not required to know how its constituent functions (i.e., modules from *M*) are calculated to prove the existence of a solution - it is enough to know only the names of these functions (the set of operation symbols *F*) and associated with these names the sets of input and output parameter names from *Z*. The proof of existence must be constructive. It means that in the course of the proof a problem solving plan must be constructed (an abstract program).

The transformation of a plan into a specific program depends on the specific capabilities of the hardware and software of the computational environment and is not considered in this article.

The article is organized in the following way. Section II provides a brief overview of the problem of structural program synthesis. Section III describes the knowledge base of the abstract program planner. Section IV is devoted to the static Boolean planning model. Section V contains the equations of the dynamic planning model. Section VI discloses the technology of applying the developed theoretical results on the example of a small computational Boolean network. Section VII describes the means of automating the management of computations based on a dynamic model. Section VIII contains a conclusion on the results achieved during the study.

II. RELEATED WORK

Automated generation of plans for solving problems in various subject areas is one of the main areas of research in artificial intelligence, which has been discussed for a long time in foreign and domestic literature [1]. The problem of obtaining abstract programs or, in other words, plans for computational actions on a computational domain model is known as the planning problem in the structural synthesis of programs.

A new wave of increased interest in the idea of program synthesis is mainly associated with the understanding of the possibility of representing a synthesis problem in the form of a Boolean satisfiability problem [2, 3], for which efficient algorithms for its solution are developed.

Structural program synthesis is a method of constructing programs from pre-implemented modules that are treated as functions. For the first time, the idea of structural synthesis of programs was presented in a report

[4] at a conference on algorithms in modern mathematics and computer science, organized by A. Ershov and D. Knuth in 1979. More recently, this idea stimulates the creation of new modern integrated environments for the development of domain-specific software based on the structural synthesis of programs. For example, CoCoVila [5] supports visual and simulation software development and uses structural program synthesis to translate declarative specifications of simulation tasks into executable code. The HPCSOMAS platform [6] provides automated development of applied microservice packages (AMP) for distributed scientific computations. The CLAVIRE framework [7] represents a set of high-level tools for solving e-Science problems within the data-driven approach. Based on iPSE (Intelligent Problem Solving Environment), it allows one to interact with the user based on domain-specific knowledge.

The traditional planning technique in the structural synthesis of programs is deductive means of proving theorems based on the methods of mathematical logic [8]. Our approach to the synthesis of abstract programs (both parallel and sequential) is based on the systematic use of binary dynamical systems as the main formalism of the internal representation for the computational domain model. The search for a plan under these conditions is reduced to constructing particular solutions for such dynamical systems. AI planning languages are not considered in this paper.

III. Planner Knowledge Base

As the knowledge base (KB) of the planner, the computational model $KB = (F, Z, In, Out)$ is used, where Z is a finite set of symbols (names) of parameters (attributes, values of the subject area), and *F* is a finite set of symbols of *k*+*p* arity operations (*k* and *p*, generally speaking, are different for different operation symbols). $In \subset F \times Z$ and $Out \subset F \times Z$ define relationships between operations and parameters at the input and output, respectively. Each symbol of the operation $F_i \in F$ of arity $k_i + p_i$ is connected to the set $in(F_i) \subset Z$ from k_i of input parameters and a set $out(F_i) \subset Z$ from p_i of output parameters associated with it. In terms of content, the operation $F_i \in F$ implies the possibility of calculating variables $out(F_i)$ from variables $in(F_i)$ using some program module M_i from the library of modules *M*.

It is assumed that the KB has a high level of internal parallelism and is redundant in the sense that only a part of the operations from *F* is used to solve the problem, and (or) the problem *T* has several alternative solution plans.

Relationships of both *In* and *Out* are conveniently specified as two Boolean matrices *A* and *B* with dimensions $n \times m$, whose elements are formed as follows: $a_{ij} = 1$ ($b_{ij} = 1$) if the parameter Z_j is input (output) for the operation F_i. The matrices *A* and *B* will be referred to as the input and output matrices, respectively. Rows of matrices A_i and B_i are the binary representation of the set of input parameters $in(F_i)$ and output ones $out(F_i)$ of the operation F_i, respectively. Similarly to the relations *In* and *Out*, it is also convenient to represent the required input-output relation of the problem statement $T = (A_0, B_0)$ in the form of *m*-dimensional Boolean strings a_0 and b_0, assuming that the *j*-th element of these strings takes the value equal to 1 if the parameter Z_j is input (output) in the problem statement *T*.

IV. Static Boolean Planning Model

Let us associate the parameter vectors *Z* and the operation *F* with the Boolean vectors *z* and *f*, the values of the components of which (z_i and f_i) will be determined as follows: $z_i = 1$, if the value of the parameter Z_i is known (given or calculated), and $z_i = 0$ - otherwise, $f_i = 1$, if the operation F_i is executed and $f_i = 0$ - otherwise.

Let us assume that over Boolean matrices and vectors, the elementwise operations of disjunction $\vee$, conjunction $\wedge$ (hereinafter, this sign is replaced by a dot or omitted completely), negation $\neg$, addition modulo two $\oplus$, and implication $\rightarrow$ are defined.

Let A^T be the matrix transposed to *A*. Let *E* be the column, all elements of which are equal to one. The symbols $\otimes$ and $\circ$ will also mean the $\vee-$ and $\wedge-$product of two matrices:

$$C = A \otimes B, \text{ where } c_{ik} = \bigvee_j a_{ij} b_{jk},$$
$$C = A \circ B, \text{ where } c_{ik} = \bigwedge_j (a_{ij} \vee b_{jk}).$$

Taking into account the introduced notation, the relations reflecting the relationship of operations with input and output parameters can be written as a system of two matrix Boolean equations

$$\begin{aligned} &\bar{A} \circ z \rightarrow f = E, \\ &B^T \otimes f \rightarrow z = E. \end{aligned} \tag{1}$$

A distinctive feature of system (1) is its large dimension and strong sparseness of the matrices *A* and *B*. Along with the original system we need the derivative from system (1), called the inverse:

$$\begin{aligned} &f \leftarrow B \otimes z = E, \\ &z \leftarrow A^T \otimes f = E. \end{aligned} \tag{2}$$

In the vector-matrix form, the statement of the problem $T = (A_0, B_0)$ corresponds to a Boolean equation of the form

$$\bar{a}_0 \circ z \rightarrow \bar{b}_0 \circ z = 1. \tag{3}$$

Due to the equivalence of the calculus of functional dependencies to a fragment of the propositional calculus, it is easy to show that the problem $T = (A_0, B_0)$ is solvable in the KB if and only if equation (3) is a logical consequence of the system of equations (1) in the sense that all solutions of system (1) are solutions of equation (3).

The vector-matrix Boolean formalism of representation of the computational model KB proposed in this section allows:

- Introduce a certain standardization in this class of models;
- Create on a regular basis methods for analyzing the structural features of models;
- Obtain (due to the introduction of two types of variables – Boolean vectors *z* and *f* into the model) a plan for solving the problem directly in terms of operation symbols;
- Create efficient planning algorithms due to the high parallelism of matrix-vector operations on binary vectors.

Systems of Boolean equations (1) and (2) can be used in logical derivation in implementing a strategy for moving from initial states to target states ("forward wave" - system (1)) and from target states to initial ones ("reverse wave" - system (2)). However, the representation of the problem conditions in the form of systems of Boolean equations (1, 2) allows the creation of a new approach to constructing plans for solving non-procedural problems based on dynamic Boolean planning models.

V. Dynamic Boolean Planning Model

The vector-matrix model of the subject area considered in Section IV accurately reflects the statics of the modular system (its structure and functional relationships between the input and output actual parameters of operations) but does not allow describing the dynamics of the system behavior in time in solving non-procedural problems.

A well-known mathematical apparatus for describing the dynamics of the behavior of logical systems is Boolean differential calculus [9, 10]. This section proposes new logical models of modular software complexes in the form of systems of nonlinear Boolean differential equations. Such systems of equations allow effective numerical methods for constructing the necessary particular solutions consistent with the problem formulation (3).

The transition from static Boolean equations (1, 2) to dynamic ones is carried out by reducing the implication operation to the time domain, namely: interpreting the implication $x \to y$ as a cause-and-effect dependence of *y* on *x* and considering the variables *x* and *y* as functions of discrete time $t \in \{0,1,2,...\}$, the logical equation $x \to y = 1$ is put in correspondence a pair of Boolean differential equations:

$$\begin{aligned} \dot{x}(t) &= 0, \\ \dot{y}(t) &= \bar{y}(t) \cdot x(t), \end{aligned} \tag{4}$$

where $\dot{y}(t)$ is the time derivative of the function $y(t)$ defined by the following relation:

$$\dot{y}(t) = y(t) \oplus y(t+1) .$$

For $\dot{x}(t) = \dot{y}(t) = 0$, we receive the original equation $x \to y = 1$, whose solutions determine the set of possible equilibrium states for (4). Replacing cause-and-effect relationships with temporal ones makes it possible to more clearly present the structural features of the functioning of a modular computational system and dispense with the use of inference tools in solving non-procedural problem formulations. In a certain sense, equations of type (4) combine the original logical relations and the means of logical inference for them.

Following the methodology just considered, we assign to the system of matrix Boolean equations (1) a system of matrix Boolean differential equations of the form:

$$\begin{aligned} \dot{f}(t) &= \bar{f}(t) \cdot (\bar{A} \circ z(t)), \\ \dot{z}(t) &= \bar{z}(t) \cdot (B^T \otimes f(t)). \end{aligned} \tag{5}$$

For system (2), the matrix Boolean differential equations will be as:

$$\begin{aligned} \dot{f}(t) &= \bar{f}(t) \cdot (B \otimes z(t)), \\ \dot{z}(t) &= \bar{z}(t) \cdot (A^T \otimes f(t)). \end{aligned} \tag{6}$$

Systems of equations (5) and (6) describe the dynamics of the main and auxiliary bipartite computational Boolean networks, respectively.

In (5, 6) the column ($f(t)$, $z(t)$) is the state vector of these Boolean dynamical systems at time *t*.

Taking into account the definition of the derivative, we construct partial solutions $z(t, a_0)$, $f(t, a_0)$ of the systems (5) and $z(t, b_0)$, $f(t, b_0)$ of the systems (6) under the initial conditions $z(0) = a_0^T$, $f(0) = 0$ and $z(0) = b_0^T$, $f(0) = 0$, respectively. With the passage of time ($t \le n + m$), these particular solutions, as nondecreasing functions of time, reach equilibrium states (that is, there are no cyclic attractors). Let us denote these equilibrium states as $z_b(a_0)$, $f_b(a_0)$ for system (5) and $z_a(b_0)$, $f_a(b_0)$ for system (6). Let us introduce the notation $z^* = z_b(a_0) \cdot z_a(b_0)$ and $f^* = f_b(a_0) \cdot f_a(b_0)$. Then the following assertions can be made:

1. If $z_b(a_0) \ge b_0$, then the problem $T = (A_0, B_0)$ is solvable on the computational model KB and the time derivatives $\dot{f}(t, a_0)$, $\dot{z}(t, a_0)$ of particular

solutions $f(t,a_0)$, $z(t,a_0)$ represent the so-called preliminary plan for solving problem T deployed in time. The components of the vector-functions $\dot{f}(t,a_0)$, $\dot{z}(t,a_0)$ are either identically equal to zero, or represent elementary functions of the form:

$$\delta_j(t)=\begin{cases}1, t=j\\ 0, t\neq j\end{cases}.$$

From a meaningful point of view, the equality $\dot{f}_i(t,a_0)=\delta_j(t)$ means that the operation F_i is activated at time $t=j$, and from equality $\dot{z}_i(t,a_0)=\delta_k(t)$ it follows that the parameter Z_i receives a value at time $t=k$.

Generally speaking, a vector-function $\dot{f}(t,a_0)$ defines a network of parallel computations, namely: if several vector $\dot{f}(t,a_0)$ components are determined by the same function $\delta_j(t)$, then the operations corresponding to these components can be activated and executed in parallel.

2. In the general case, the preliminary plan is redundant in the sense that it may contain operations, the execution of which in the end "gives nothing" to achieve the objective B_0. Their elimination is possible due to the formation of Boolean vector functions of the form:

$$\begin{aligned} y_z(t)&=\dot{z}(t,a_0)\cdot z^*,\\ y_f(t)&=\dot{f}(t,a_0)\cdot f^*, \end{aligned} \qquad (7)$$

which determine the so-called complete problem solving plan.

3. If the solution to problem *T* is not unique, then the complete plan is also redundant in the sense that it is a "union" of possible plans for solving problem *T*.

Ultimately, we are interested in non-alternative non-redundant plans, the plans for which the exclusion of any operation (included in the plan) from them leads to the loss of solvability of the original problem by this plan. The presence of the complete plan allows us to construct a set of non-alternative non-redundant plans according to the following algebraic method:

a) For each column of the matrix *B* with the number *j*, such that $z_j^*=1$, we write out an elementary disjunction from the variables f_i, for which $b_{ij}=1$;

b) Write down the conjunction of the obtained elementary disjunctions;

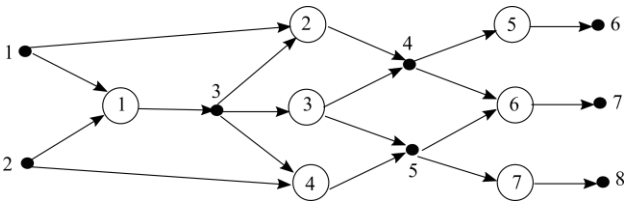


Figure 1. Main computational Boolean network. The dots denote the variables of the vector *z (m*=8*)*, the circles denote the variables of the vector *f (n*=7*).*

c) The resulting expression, taking into account the laws of idempotency and absorption of the algebra of logic, is transformed into a disjunctive normal form. The terms of this form determine the set of all alternative non-redundant plans for solving the problem $T=(A_0,B_0)$.

VI. ILLUSTRATIVE EXAMPLE

Let us consider the technology of synthesis of plans for solving the "given-to-find" problem on the example of a computational Boolean network shown in Fig. 1.

Matrices *A* and *B* corresponding to this network have the form:

$$A=\begin{bmatrix}1&1&0&0&0&0&0&0\\1&0&1&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&1&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&1&1&0&0&0\\0&0&0&0&1&0&0&0\end{bmatrix};$$

$$B=\begin{bmatrix}0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&1&1&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{bmatrix}.$$

In the scalar representation, matrix differential equations (5) and (6) have the form (8) and (9), respectively:

$$\begin{aligned}
\dot{f}_1&=\bar{f}_1 z_1 z_2 & \dot{z}_1&=0\\
\dot{f}_2&=\bar{f}_2 z_1 z_3 & \dot{z}_2&=0\\
\dot{f}_3&=\bar{f}_3 z_3 & \dot{z}_3&=\bar{z}_3 f_1\\
\dot{f}_4&=\bar{f}_4 z_2 z_3 & \dot{z}_4&=\bar{z}_4(f_2\vee f_3)\\
\dot{f}_5&=\bar{f}_5 z_4 & \dot{z}_5&=\bar{z}_5(f_3\vee f_4)\\
\dot{f}_6&=\bar{f}_6 z_4 z_5 & \dot{z}_6&=\bar{z}_6 f_5\\
\dot{f}_7&=\bar{f}_7 z_5 & \dot{z}_7&=\bar{z}_7 f_6\\
& & \dot{z}_8&=\bar{z}_8 f_7
\end{aligned} \qquad (8)$$

$$\begin{aligned}
\dot{f}_1 &= \bar{f}_1 z_3 & \dot{z}_1 &= \bar{z}_1(f_1 \vee f_2)\\
\dot{f}_2 &= \bar{f}_2 z_4 & \dot{z}_2 &= \bar{z}_2(f_1 \vee f_4)\\
\dot{f}_3 &= \bar{f}_3(z_4 \vee z_5) & \dot{z}_3 &= \bar{z}_3(f_2 \vee f_3 \vee f_4)\\
\dot{f}_4 &= \bar{f}_4 z_5 & \dot{z}_4 &= \bar{z}_4(f_5 \vee f_6)\\
\dot{f}_5 &= \bar{f}_5 z_6 & \dot{z}_5 &= \bar{z}_5(f_6 \vee f_7)\\
\dot{f}_6 &= \bar{f}_6 z_7 & \dot{z}_6 &= 0\\
\dot{f}_7 &= \bar{f}_7 z_8 & \dot{z}_7 &= 0\\
 & & \dot{z}_8 &= 0
\end{aligned} \quad . \qquad (9)$$

Let us assume that the problem statement T in the Boolean representation is given as follows: $a_0 = (11000000)$, $b_0 = (00000010)$.

After seven steps ($t = 7$) particular solutions $z(t, a_0)$, $f(t, a_0)$ of system (8) and particular solutions $z(t, b_0)$, $f(t, b_0)$ of system (9) reach equilibrium states

$$\begin{aligned}
z_b(a_0) &= col(11111111),\\
f_b(a_0) &= col(1111111),\\
z_a(b_0) &= col(11111010),\\
f_a(b_0) &= col(1111010).
\end{aligned}$$

So, we have

$$\begin{aligned}
z^* &= col(11111010),\\
f^* &= col(1111010).
\end{aligned}$$

Since $z_b(a_0) \geq b_0$, then the given problem is solvable on a Boolean computational network and the preliminary plan represents the derivative $\dot{f}(t, a_0)$ of a particular solution $f(t, a_0)$ of system (7), presented in Table 1.

In terms of elementary functions $\delta_j(t)$ will be:

$$\begin{aligned}
&\dot{f}_1(t, a_0) = \delta_0(t),\ \dot{f}_2(t, a_0) = \dot{f}_3(t, a_0) = \dot{f}_4(t, a_0) = \delta_2(t),\\
&\dot{f}_5(t, a_0) = \dot{f}_6(t, a_0) = \dot{f}_7(t, a_0) = \delta_4(t)
\end{aligned} .$$

A complete problem solving plan is formed according to equations (7) and is presented in Table 2.

To verify the resulting plan for uniqueness, we write out the logical product of elementary disjunctions

$$f_1(f_2 \vee f_3)(f_3 \vee f_4) f_6 = f_1 f_3 f_6 \vee f_1 f_2 f_4 f_6,$$

which means that the task has two solution plans, and in the second plan, operations F_2 and F_4 can be executed in parallel.

VII. Automation Tools for Computation Control Based on the Dynamic Planning Model

The HPCSOMAS-MSC platform provides tools for creating AMP agents and microservices, tools for automating this process, and tools for multi-agent control

TABLE I. Preliminary Plan

t	$\dot{f}$						
	1	**2**	**3**	**4**	**5**	**6**	**7**
0	1	0	0	0	0	0	0
2	0	1	1	1	0	0	0
4	0	0	0	0	1	1	1

of computations in a hybrid microservice computational infrastructure (HMCI). The basic version of HPCSOMAS-MSC [11] supports the following control schemes:

- Decentralized control of a group of distributed computational agents (DCA), self-organizing according to a non-procedural problem statement (NPS) on a distributed computational model KB (knowledge base) of the subject area;
- Hierarchical control of the composition of microservices formed according to the procedural problem statement (PPS), performed by the computational agents launching these microservices.

In the modified version of HPCSOMAS-MSC (Fig. 2), computation-planning tools are integrated into a separate subsystem that provides an additional possibility of centralized planning and control for the case of a centralized KB formed when creating an AMP based on a library of applied modules. To obtain a plan for solving the problem, a wizard for the abstract program creation uses the dynamic Boolean model proposed in section V. The wizard has WEB and API interfaces and works both alone and within AMP. In the latter case, the centralized planning and control agent (CPC) using this wizard receives a plan, according to which it deploys the computational microservice agents (CMA) required to solve the problem in the HMCI and launches them in the order specified by the plan as soon as the data is ready. When finished, the CMA returns and passes the output to the CPC. At this point, the CMA life cycle ends, and the occupied by this agent resource is released. With a decentralized approach, due to the presence of a decentralized KB, the life cycle of the DCA ends (the resource occupied by the agent is released) after receiving a message about the completion of the task solution from the last completed DCA. Thus, the CPC agent saves resources by distributing them dynamically, unlike the agent UDA (User Dew Agent, [12]), which manages decentralized computing (Fig. 2). So, for the example given in Section VI, the CPC will allocate three

TABLE II. Complete Problem Solving Plan

t	y_f						
	1	**2**	**3**	**4**	**5**	**6**	**7**
0	1	0	0	0	0	0	0
2	0	1	1	1	0	0	0
4	0	0	0	0	0	1	0

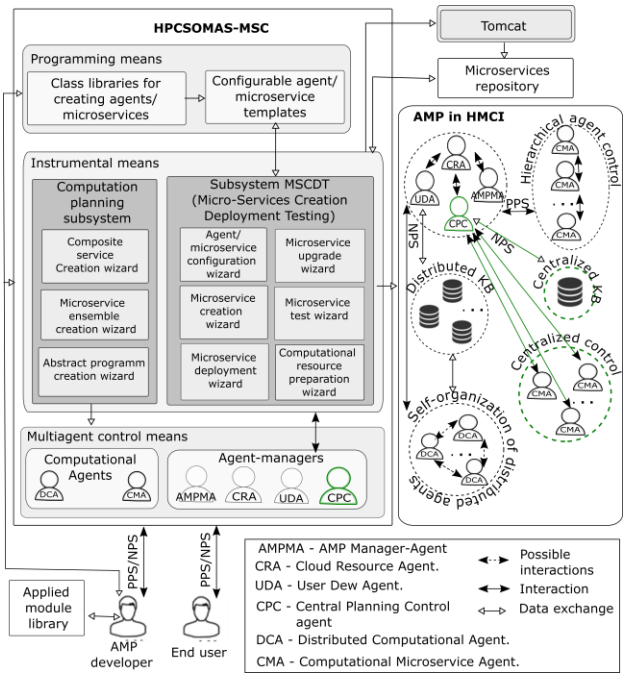


Figure 2. HPCSOMAS-MSC tools for working with AMP

computing resources for the first plan or four resources for the second one. The UDA will allocate seven resources. If necessary, the CPC agent is duplicated and receives the CMA launch monitoring protocol from the main CPC agent to increase the reliability of computations.

VIII. Conclusion

This article proposes new logical models of modular software complexes in the form of systems of nonlinear Boolean differential equations. Such systems allow both the creation of efficient methods for constructing the necessary particular solutions for high-dimensional models and the use of analytical and qualitative methods for studying the required dynamic properties of their solutions.

From a practical point of view, it is important to note that the logical models of modular software complexes in the form of systems of Boolean differential equations allow effective both software and hardware implementation. In this case, it becomes possible to use them to build high-speed programmable logic control devices. Another important area of application of the presented models is the construction of real-time intelligent control systems for moving objects, which provide dynamic reconfiguration of the structure and synthesis of control programs during the object functioning.

Acknowledgment

The study was supported by the Ministry of Science and Higher Education of the Russian Federation, project no. 121032400051-9. The authors would like to thank Irkutsk Supercomputer Center of SB RAS for providing access to HPC-cluster “Akademik V.M. Matrosov” [13].

References

[1] M. Ghallab, D. S. Nau, and P. Traverso “Automated planning: theory and practice,” Morgan Kaufmann, May 2004.

[2] A. Solar-Lezama, “Program synthesis by sketching,” Diss. UNIVERSITY OF CALIFORNIA, BERKELEY, 2008.

[3] G. A. Oparin, A. P. Novopashin, “Boolean models and planning methods for parallel abstract programs,” Autom. Remote Control, vol. 69, no. 8, 2008, pp. 1423–1432.

[4] E.H. Tyugu, “The structural synthesis of programs,” in Algorithms in Modern Mathematics and Computer Science, LNCS, vol. 122, A.P. Ershov and D.E. Knuth, Eds., Berlin, Heidelberg: Springer, 1981, pp. 290–303.

[5] V. Kotkas, A. Ojamaa, P. Grigorenko, R. Maigre, M. Harf, and E. Tyugu, “CoCoViLa as a multifunctional simulation platform,” in International ICST Conference on Simulation Tools and Techniques, Brussels: ICST, 2011, pp. 1–8.

[6] G.A. Oparin, V.G. Bogdanova, A.A. Pashinin, and S.A. Gorsky, “Microservice-oriented approach to automation of distributed scientific computations,” in 42nd Intern. Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2019, pp. 236-241.

[7] K.V. Knyazkov, S.V. Kovalchuk, T.N. Tchurov, S.V. Maryin, and A.V. Boukhanovsky, “CLAVIRE: e-science infrastructure for data-driven computing,” Journal of Computational Science, 2012, vol. 3, no. 6, pp. 504–510.

[8] G. Mints, E. Tyugu, “Justification of the structural synthesis of programs,” Science of Computer Programming, 1982, vol. 2, no. 3, pp. 215–240.

[9] D. Bohmann, R. S. Stankovič, Zh. Toshich, V. P. Shmerko, and S. N. Yanushkevich, “Logic differential calculus: achievements, trends, and applications,” Autom. Remote Control, 2000, vol. 61, no. 6, part 2, pp. 1033–1047, https://zbmath.org/pdf/1060.94051.pdf [accessed: March 16 2023].

[10] D. Bochmann, C. Posthoff, “Binäre dynamische systeme,” Berlin: Akademie-Verlag, 1981.

[11] G.A. Oparin, V.G. Bogdanova, and A.A. Pashinin, “Automated tools for the development of microservice compositions for hybrid scientific computations,” in 2nd Information, Computation, and Control Systems for Distributed Environments, 2020, pp. 201–213.

[12] A.A. Pashinin and V.G. Bogdanova, “Application of user dew agent in hybrid-computing environments,” in 1st International Workshop on Advanced Information and Computation Technologies and Systems (AICTS), 2020, pp. 135–145.

[13] “Irkutsk Supercomputer Centre of SB RAS,” http://hpc.icc.ru [online, accessed: January 31 2023].

The Concept of Cognition as Categorization in the Development of New Metaheuristics and Algorithms Inspired by Nature

Domagoj Tuličić* and Nikola Ivković**
*College for Information Technologies/Department of Computer Systems and Networks, Zagreb, Hrvatska
**Faculty of Organization and Informatics/Department of Computing and Technology, Varaždin, Hrvatska
dtulicic@vsite.hr, nikola.ivkovic@foi.hr

***Abstract* - The sensorimotor system is a crucial mechanism that enables organisms to categorize information in their environment, a skill necessary for survival. According to cognitive scientist Steven Harnad's "To Cognize is to Categorize: Cognition is Categorization", all living things are essentially sensorimotor systems. This approach is proposed to serve as a novel concept for the development of algorithms inspired by swarm intelligence. Many authors of such algorithms have drawn inspiration from animal species, attempting to emulate their abilities to interact with and navigate their environment. However, these abilities are essentially rooted in the organisms' categorization skills, which are a fundamental characteristic of all sensorimotor systems. Therefore, when designing new algorithms and metaheuristics, it is unnecessary to seek out new species with specific behaviors; instead, the entire sensorimotor system can be observed, and its sensor capabilities can be incorporated even if they do not exist in nature. In order to justify this, abstraction was used as a way to demonstrate the creation of concepts. The paper offers a new perspective on the development of swarm intelligence algorithms and promotes a holistic approach to algorithm design based on Steven Harnad's viewpoint of sensory motor systems.**

Keywords* - *Sensorimotor system, nature inspired algorithms, swarm intelligence inspired algorithms, natural computing, cognition, categorization, bio-inspireda algorithms, metaheuristic

I. INTRODUCTION

Algorithms that draw inspiration from nature fall under the category of metaheuristics, which distinguishes them from exact algorithms. These algorithms are classified based on specific characteristics [1], with one category being algorithms inspired by swarm intelligence. This category is rapidly expanding, with researchers attempting to model various animal behaviors, ranging from the smallest insects [2] to slightly larger species like squirrels [3], doves [4] and even larger animals such as monkeys [5] and whales [6].

The significant advancements in algorithm development are largely attributed to the implementation of metaheuristics. Metaheuristics encompass a collection of algorithmic concepts that can be applied to various problems [7]. To put it simply, metaheuristics are like metaphors - they have their limitations, despite their accuracy [8]. The limitations of algorithmic concepts (or metaphors) are thoroughly discussed, with the primary concern being the potential for existing algorithms to be concealed or for inefficiencies to be overlooked under new algorithmic concepts or metaphors [8] [9]. Therefore, the problem of metaphors and their limited knowledge is explained in Section II and further elaborated on in Section III. Additionally, recreating such algorithms requires a thorough understanding of animal behavior in their environment and the mathematical models representing both the animals and the environment. This task is daunting, given that the effectiveness of metaheuristic algorithms is problem-specific.

The first swarm intelligence to be named a metaheuristic was ant colony optimization (ACO) [7]. The ACO author couldn't have predicted the vast number of algorithmic concepts related to animals that would emerge, such as bees, bacteria, bats, wolves, bumblebees, cats, cuckoos, eagles, fireflies, fish, krill, glowworms, monkeys, squirrels, whales, and so on. All of these animals are sensorimotor systems that interact with their environment and can communicate through it, much like ants. Therefore, a more feasible option for researchers is to use a named general framework or master strategy consisting of algorithmic concepts or metaphors.

To truly have a master strategy, which is what metaheuristics represent [10], we need a fundamental concept that applies to all living beings. This can be achieved by viewing each animal as a sensorimotor system. For this purpose, in Section III Steven Harnad's work "To Cognize is to Categorize: Cognition is Categorization" is used to determine characteristics of sensorimotor systems applicable to all living beings [11]. These set of characteristics or elements which can be called the "Concept of Cognition as Categorization" define fundamental concept of all living beings. In Section IV this fundamental concept is linked to concept of ant colony as an algorithmic concept for swarm intelligence (collective intelligence). If there is relation between concepts, what is shown in Section IV, then because of transitivity law intermediate concept is unnecessary, which means there is no need for developing algorithmic

metaheuristic or metaphor of specific animal. "Concept of Cognition as Categorization" is sufficient.

The key takeaway is that designers of swarm-inspired algorithms can create new components with a strong theoretical foundation in cognitive science. They do not need to delve deeply into the specific behaviors of various animals, ranging from insects to mammals. Rather, they can focus on the sensorimotor systems that define these animals and their communication abilities within specific environments. By doing so, they can build algorithms that are effective and efficient without getting bogged down in unnecessary details.

II. THE PROBLEM OF METAPHOR

Various fields of study offer differing definitions of metaphors. In this paper, we adopt the definition of metaphor used in cognitive linguistics. In this field, metaphor is viewed as a cornerstone of the human conceptual system, leading to the development of what are called "conceptual metaphors."

A common example of metaphor in everyday life involves people using the concept of a roundabout to describe a situation where they are unable to find a solution to a problem. However, this simple metaphor belies the complexity of the concept of a roundabout, which includes multiple entrances and exits. In other words, while people may use the idea of a roundabout to describe a challenging situation, the full implications of this metaphor go beyond the surface level.

Metaphors are not only prevalent in everyday life, but they also abound in science. The field of metaheuristic algorithms, in particular, is rife with metaphors. While using metaphors in research can have its benefits, it is important to acknowledge their limitations and potential consequences. Some researchers who focus on developing new algorithms through the use of new metaphors fail to recognize that metaphors tend to emphasize certain characteristics of a topic or area while ignoring others [12] [13]. Instead, they should consider the need for new metaphors that highlight previously overlooked characteristics of complex relationships in intelligent swarms. The metaphors used by researchers studying intelligent swarms do not necessarily expand our knowledge about collective intelligence. In fact, they may even obscure it by avoiding a clear definition of an invariant for collective intelligence and describing only its external manifestations.

III. LEVELS OF ABSTRACTION AS A WAY OF CONCEALING METAPHORS

To demonstrate how metaphors can conceal certain aspects of knowledge while highlighting others, this paragraph will utilize the concept of abstraction. Additionally, this paragraph will provide a brief explanation of three metaphors: the ACO metaphor, the swarm intelligence or collective intelligence metaphor, and the metaphor based on Steven Harnad's work "To Cognize is to Categorize: Cognition is Categorization".

A. *Mathematical explanation of abstraction*

For mathematical explanation of abstraction Frege's axiom of abstraction that states:

$$(\exists y)(\forall x)\,(x \in y \Leftrightarrow F(x)), \qquad (1)$$

is used. Here, set X is set of the concept or metaphor "Cognition as Categorization" that contains the elements that make up that concept. Set Y is set of the concept or metaphor of "Collective Cognition" with y elements that make up that concept. F(x) is arbitrary property of X set.

If the axiom of abstraction is true, then it follows that $Y \subseteq X$. Figure 1. illustrates concepts from the perspective of Venn Diagrams. This means that set X contains more elements than set Y since everything in set Y is also included in set X. In other words, all elements of Collective Cognition (Y elements) are encompassed by the concept of Cognition as Categorization (X elements). By reducing a concept's information content and retaining only its essential characteristics, we engage in abstraction and conceal certain knowledge. This also holds true for the ACO metaphor when compared to the metaphor of "Collective Intelligence".

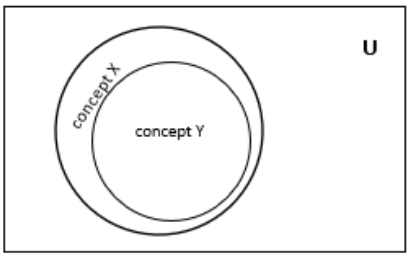


Figure 1. Y ⊆ X therefore an abstraction of Y is possible (author's work)

B. *The Concept of Cognition as Categorization*

The idea or comparison of how the mind works as a categorization process is explained by cognitive and computational neuroscientist Steven Harnad, who is known for his studies on consciousness, philosophy of the mind, and the significance of symbol processing in cognition. Harnad argues that living organisms are sensorimotor systems, making them a unique kind of dynamic system that can categorize information [11]. This sets them apart from other dynamic systems that are governed by the laws of physics. Harnad's definition of categorization highlights the systematic and distinct interaction that occurs between an independent and adaptable sensorimotor system and the surrounding environment.

Table 1. displays the main terms and sub-terms described in his work, serving as the essential foundation of the metaphor known as "Cognition as Categorization," which is established in this paper.

Table 1. The main elements of Harnad's description of cognition as categorization (author's work)

Main terms	Sub-terms
Sensorimotor system	/
Differential interaction	/
Autonomy	/
Adaptivity	/
Invariances in stimulation	/
Memory	Reinforcement Forgetting (amnesia)
Categories	Innate Learned
Types of learning	Unsupervised Supervised Instrumental
Ability of Categorization	Categorical perception Abstraction Absolute Discrimination Features weighting Dimension reduction

C. Ant Colony Optimization and Swarm Intelligence Concepts

After analyzing the concepts presented in the book "Ant Colony Optimization" [7], it can be concluded that there are two main concepts that form its basis. The first concept, which is pervasive throughout the book, is referred to as the Ant Colony Optimization (ACO) Metaheuristic. The second concept is the concept of natural optimization, which is described by the foraging behavior of ants and is essentially a form of swarm intelligence. This concept is based on the binary bridge experiment conducted by Deneubourg et al [14]. The two concepts are interconnected, as the second concept arises from the abstraction of the first.

The first concept, Ant Colony Metaheuristic, is described using pseudocode (as shown in Figure 2). It can be understood as the interplay of three procedures: ConstructAntsSolutions, UpdatePheromones, and DaemonAction [7]. Moving on to the second concept, which is the concept of swarm intelligence, we can refer to it as the concept of collective cognition in order to better understand its similarity to the concept of cognition as categorization. Here, cognition serves as the link between these terms. Since collective intelligence formed the basis for the development of ACO metaheuristics, it is reasonable to assume that ACO is an abstraction of collective intelligence (as depicted in Figure 3). The ACO metaphor simplifies the details involved in collective consciousness, but upon closer examination, these details become apparent once again.

```
procedure ACOMetaheuristic
  ScheduleActivities
    ConstructAntsSolutions
    UpdatePheromones
    DaemonActions          % optional
  end-ScheduleActivities
end-procedure
```

Figure 2. Concept of ACO Metaheuristic written in pseudocode. After Dorigo, Stutzle [3]

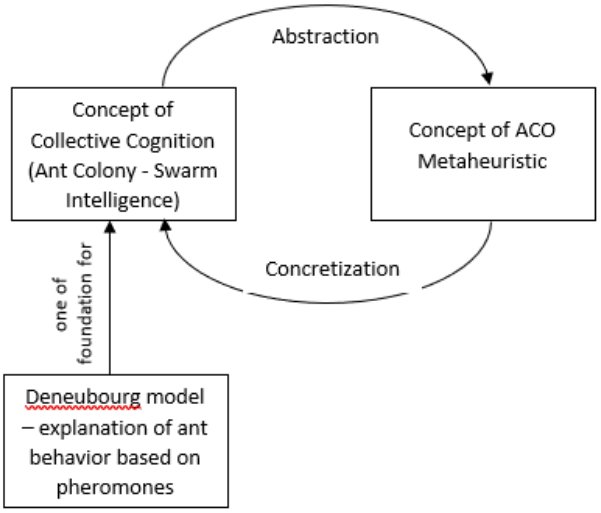


Figure 3. Concepts: Collective Cognition, ACO Metaheuristics and their relations (author's work)

D. Explanation of abstracting the concept of collective intelligence from the concept of cognition as categorization

Concept of Collective Cognition is formed by elements including self-organization and stigmergy. Self-organization relies on four basic ingredients: positive feedback (amplification), negative feedback that is counterbalance of positive feedback, amplification or fluctuations and multiple interactions [15]. Stigmergy refers to an indirect way of communication [15].

As ants, social insects, and other animals are sensorimotor systems, their behavior is not regulated by the laws of self-organization and stigmergy, but rather by the capabilities of each individual's sensorimotor system. To demonstrate this point, let's suppose that their sensorimotor system is incapable of systematic differential interaction with the environment. In such a scenario, ants would be unable to detect pheromone levels or select the most optimal path.

If we examine the elements of self-organization, particularly positive feedback, we can conclude that a sensorimotor system cannot function without feedback. Positive feedback is a property of every individual sensorimotor system, and when an ant colony searches for food, the collective positive feedback is a manifestation of the many positive feedbacks from each sensorimotor system. This positive feedback arises when stimuli undergo the process of cognitive categorization - from learning and memory to adaptation through systematic differential interactions.

The next ingredient in self-organization is negative feedback, which serves to counterbalance positive feedback. At the individual level of the sensorimotor system, this occurs in any situation where the sensorimotor system is engaged in categorization due to stimuli that indicate discrepancies in the environment.

Another ingredient of self-organization is amplification and fluctuation, which account for randomness and errors. At the level of the sensorimotor system, the categorization process is subject to errors such as misclassification and forgetfulness. However, in order for categorization to occur at all, the possibility of incorrect categorization must also exist.

The fourth and final ingredient of self-organization is the possibility of multiple interactions. When considering this ingredient, we must also take into account stigmergy. The reason for this is that multiple interactions occur through the influence of pheromones, which is an indirect form of communication known as stigmergy. At the level of the sensorimotor system, this form of communication represents the interaction between the system and its environment. Therefore, every interaction between individuals that involves the use of pheromones is a process of categorization, where the intensity and existence of the pheromone is categorized. Additionally, stigmergy can be seen as a way for an individual to communicate with hundreds of others from the past and present simultaneously.

IV. CONCEPT OF COGNITION AS CATEGORIZATION: A FUNDAMENTAL CONCEPT FOR ALL SWARMS-BASED METAHEURISTIC ALGORITHMS

As demonstrated in the preceding chapter, the use of metaphor has the capacity to obscure knowledge. **This applies to the metaheuristics of swarm-inspired algorithms as well, which are themselves conceptual metaphors or concepts. One of these concepts is ACO metaheuristics, which builds on the Concept of Collective Cognition, which, in turn, is based on a more fundamental concept. For a concept to be considered fundamental, it must possess a requisite number of constituent elements that enable abstraction.** As demonstrated in the previous section, Stevan Harnad's work can provide a foundation for the creation of a fundamental concept known as the Concept of Cognition as Categorization. Moreover, in Section III.D it has been shown that certain constituent elements of this concept can be utilized to abstract the Concept of Collective Intelligence.

Figure 4. demonstrates how abstraction is possible by starting with a basic concept and progressing to an intermediate concept, and ultimately to the highest level of abstraction. The law of transitivity applies to this process, which means that if we have three abstract concepts A, B, and C, and we know that A is related to B in a particular way, and B is related to C in a specific way, we can deduce that A is related to C in the same way. **This knowledge of the fundamental concept is sufficient for creating the final concept**. **This means that the Concept of Cognition as Categorization is fundamental for all swarm-based metaheuristic algorithms.**

Figure 5. demonstrates that the law of transitivity allows for the omission of the middle concept (Concept of Collective Cognition). However, it is debatable whether this intermediate concept is essential or not. **This paper presents a new perspective that suggests that it may not be essential, as it operates as a conceptual metaphor that obscures more comprehensive knowledge that could be utilized in shaping the final metaheuristic concept.**

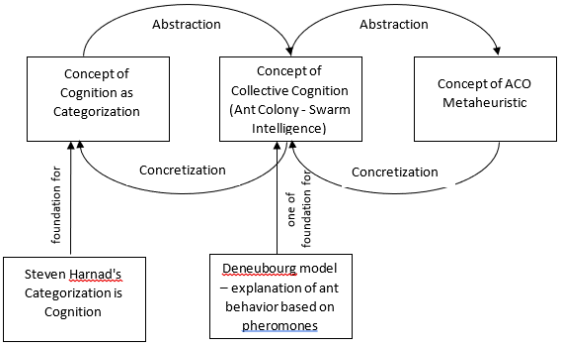


Figure 4. Concepts: Concept of Cognition as Categorization, Collective Cognition, ACO Metaheuristics and their transitive relation (author's work)

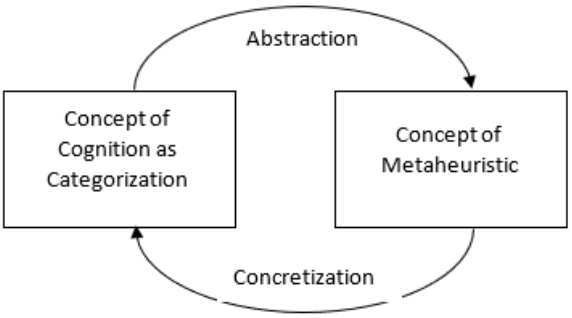


Figure 5. Concept of Metaheuristic is abstraction of fundamental concept Cognition as Categorization in the context of swarms (author's work)

V. CONCLUSION

After considering several algorithmic concepts, it became clear that they all shared a common basis - the concept of the organism as a sensorimotor system. The well-established concept of swarm intelligence and the concept of ACO were used to demonstrate how metaphors can conceal certain features through abstraction. This was applied to all three metaphors discussed in the paper. Ultimately, if algorithm designers can identify a basic metaphor, they can develop new components of algorithms that have a solid theoretical foundation in cognitive science. They do not need to delve deeper into the behaviors of various animals, from insects to mammals, as the abilities of the sensorimotor system define all of them and their communication features in specific environments.

REFERENCES

[1] I. Fister, X.-S. Yang, J. Brest, and D. Fister, "A Brief Review of Nature-Inspired Algorithms for Optimization,", Electrotechnical

Review , 3. vol. 80, pp. 116-122, 2013.
[2] M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: optimization by a colony of cooperating agents,” *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 26, no. 1, pp. 29–41, 1996.
[3] M. Jain, V. Singh, and A. Rani, “A novel nature-inspired algorithm for optimization: Squirrel search algorithm,” *Swarm Evol. Comput.*, vol. 44, pp. 148–175, 2019.
[4] M. Su, J. Chen, A. Utami, C. Chieh, "Dove Swarm Optimization Algorithm", IEEE, vol. 10, 2022.
[5] H. Sharma, G. Hazrati, and J. C. Bansal, “Spider Monkey Optimization Algorithm BT - Evolutionary and Swarm Intelligence Algorithms,” J. C. Bansal, P. K. Singh, and N. R. Pal, Eds. Cham: Springer International Publishing, 2019, pp. 43–59.
[6] S. Mirjalili and A. Lewis, “The Whale Optimization Algorithm,” *Adv. Eng. Softw.*, vol. 95, pp. 51–67, 2016.
[7] M. Dorigo and T. Stützle, *Ant Colony Optimization*. Cambridge, Massachusetts: The MIT Press, 2004.
[8] K. Sörensen, “Metaheuristics-the metaphor exposed,” *Int. Trans. Oper. Res.*, vol. 22, no. 1, pp. 3–18, 2015.
[9] M. Dorigo, T. Stutzle and C. L. Camacho-Villalon, "Grey Wolf, Firefly and Bat Algorithms: Three Widespread Algorithms that Do Not Contain Any Novelty", Int. Conference on Swarm Intelligence, pp. 121-133, 2020.
[10] F. Glover, “Future paths for integer programming and links to artificial intelligence,” *Comput. Oper. Res.*, vol. 13, no. 5, pp. 533–549, 1986.
[11] S. Harnad, *To cognize is to categorize: Cognition is categorization*, Handbook of Categorization in Cognitive Science (Second Edition), Elsevier, pp. 21-54, 2017.
[12] D. Andriessen and M. van den Boom, “East is East, and West is West, and (n)ever its intellectual capital shall meet,” *J. Intellect. Cap.*, vol. 8, no. 4, pp. 641–652, 2007.
[13] C. Bratianu[1] and D. Andriessen, “Knowledge as energy: a metaphorical analysis,” 2008.
[14] J.-L. Deneubourg, S. Aron, S. Goss, and J. M. Pasteels, “The self-organizing exploratory pattern of the argentine ant,” *J. Insect Behav.*, vol. 3, no. 2, pp. 159–168, 1990.
[15] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press; 1st edition, 1999.

Speeding up the Solving of Logical Equivalence Checking Problems with Disjunctive Diagrams

Victor Kondratiev[1], Ilya Otpuschennikov[2], Alexander Semenov[1]
[1] ITMO University, St. Petersburg, Russia
[2] Matrosov Institute for System Dynamics and Control Theory SB RAS, Irkutsk, Russia
Emails: vikseko@gmail.com, otilya@yandex.ru, biclop.rambler@yandex.ru

***Abstract*—In the context of the problem of checking the equivalence of Boolean circuits (LEC), we propose an approach that increases the efficiency of modern SAT solvers on this problem by generating additional constraints of a special kind. To generate such constraints, we use a variant of decision diagrams called disjunctive diagrams. In contrast to well-known Binary Decision Diagrams these diagrams can be constructed effectively for an original formula and can also be used to extend the original CNF formula with new clauses which are its logical consequences. In computational experiments, we show that the resulting formulas, encoding difficult LEC variants, extended by the generated constraints are often significantly easier to solve for state-of-the-art SAT solvers compared to the original formulas.**

***Keywords*—Boolean satisfiability problem; Logical Equivalence Checking; Disjunctive Diagrams; Multipliers**

I. Introduction

Boolean circuits are one of the basic objects of modern computer science in both theoretical [1], [2] and practical [3], [4] senses. Verification of Boolean circuits is considered to be a very important direction in such industrial area as Electronic Design Automation (EDA). In this paper we consider a well-known problem from EDA referred as Logical Equivalence Checking (LEC). One of the main computational tools for solving LEC are SAT solvers. Unfortunately, there are many examples of well-known functions that give LEC problems in SAT form which are very hard for all state-of-the-art SAT solving algorithms. In such situations the ways which provide opportunity to improve the efficiency of SAT solvers on such extremely hard instances are required. For instance, different preprocessing techniques can be applied for this purpose.

In this paper we propose a preprocessing technique based on the notion of Disjunctive Diagrams [5]. The main goal of our paper is to develop a corresponding technique and to argue its efficiency in application to some extremely hard LEC problems in form of SAT instances. We achieve this purpose, describing corresponding algorithms, and using equivalence checking problems for several different algorithms performing multiplication of integer numbers in the role of benchmarks.

Let us give a brief overview of the paper. The next section contains the basic definitions and results necessary for understanding the rest of the paper. Section III presents the description of main novel developed algorithms. In particular, we mean a basic preprocessing algorithm which constructs new additional constraints using which we decrease SAT solver runtime on considered benchmarks.

Also we discuss here, how we can use a special SAT partitioning based technique to estimate the hardness of extremely hard LEC instances in SAT form. In the experimental part we present the results of computational experiments and we can see, that the proposed method is able to cope with extremely hard LEC in SAT form outperforming the standard approach by $5-14\%$. It should be noted that these results are successful, taking into account the well-known extreme hardness of formulas encoding LEC for multipliers.

II. Preliminaries

By $\{0,1\}^K$ we denote the set of all binary words (also called Boolean vectors) of length K. Let F be an arbitrary Boolean formula [6] over set of variables X, $|X| = K$. The formula specifies a Boolean function $\phi_F : \{0,1\}^K \to \{0,1\}$, which is defined on whole set $\{0,1\}^K$. The function ϕ_F is defined by substituting values of Boolean variables from X into F in the standard sense [7]. Vectors from $\{0,1\}^K$, which are sets of values of variables from X, are called assignments of these variables. An arbitrary assignment $\alpha \in \{0,1\}^K$ such that the following holds: $\phi_F(\alpha) = 1$, is called a satisfying assignment for formula F. If F has a satisfying assignment then F is called satisfiable, and the formula is unsatisfiable if does not exist any satisfying assignment for F. The Boolean Satisfiability Problem (SAT) requires for an arbitrary Boolean formula F to determine if this formula is satisfiable. This problem is NP complete and it is NP-hard if, besides checking satisfiability, we would like to find some satisfying assignment in case when the formula is satisfiable.

It is well-known that the problem of checking the satisfiability of an arbitrary Boolean formula F can be reduced in polynomial time to the problem of checking the satisfiability of some Conjunctive Normal Form (CNF) C, using Tseitin transformations [8]. Hereinafter, when mentioning SAT we will assume the problem of checking the satisfiability of an arbitrary formula in a CNF.

A huge amount of combinatorial problems from a variety of industrial areas: program analysis, computer security, cryptanalysis, bioinformatics, combinatorics, etc. are effectively reduced to SAT. However, one of the main applications of modern SAT solvers for more than 20 years continues to be Electronic Design Automation (EDA). Within EDA, SAT

solvers are used to solve the Logical Equivalence Checking (LEC) problem and directly related Automated Test Pattern Generation (ATPG) problem. The LEC problem is usually considered for two Boolean circuits.

Let us remind [9], that Boolean circuit S_f is mathematically interpreted by some directed graph (V, A), V is set of vertices and A is set of arcs. V contains n vertices without parents, these vertices are called inputs of S_f and we denote their set by V^{in}, also V contains m vertices without children, these vertices form set V^{out} and are called outputs of S_f. Each vertex in $V \setminus V^{in}$ is assigned some Boolean connective from some basis. The vertices from $V \setminus V^{in}$ are called gates. Let us further consider the complete basis $\{\neg, \wedge\}$ everywhere and call the graph representing the corresponding circuit the And-Inverter Graph (AIG) [10].

On a circuit S_f we can define a procedure for its interpretation on an arbitrary input vector $\alpha \in \{0,1\}^n$, which consists of successive computation of values of Boolean functions associated with all gates, which are topologically sorted w.r.t. structure of corresponding graph. The result of interpreting S_f on α is a vector $\gamma \in \{0,1\}^m$ and thus S_f specifies a (total) discrete function of kind $f : \{0,1\}^n \to \{0,1\}^m$.

Let us return to LEC problem. This problem is formulated in the following manner: given two circuits S_f and S_h, both with n inputs and m outputs. Answer the question: is it true that S_f and S_h actually specify the same function, i.e. the following point-wise equality holds: $f \equiv h$? In case of answer 'yes' for this question, the circuits S_f and S_h are called equivalent ($S_f \equiv S_h$).

It is well-known fact that LEC problem is efficiently (in polynomial time of the length of descriptions of circuits S_f and S_h) reduced to SAT [4], [11]. To do this, a new circuit is constructed from circuits S_f, S_h: for each pair of identical inputs (with the same natural number) these inputs are glued together, and each pair of identical outputs is combined by additional XOR gate, after which the outputs from all m such XOR gates are fed to the gate linking them by disjunction. The block consisting of XOR gates and the resulting disjunction is called a miter [4], see Figure 1. The constructed circuit, denoted by $S_{M(f,h)}$, defines a Boolean function $\Phi_{M(f,h)} : \{0,1\}^n \to \{0,1\}$. The CNF $C_{M(f,h)}$ is constructed for $S_{M(f,h)}$ using Tseitin transformation. It can be shown that the CNF $C_{M(f,h)}$ is unsatisfiable if and only if the original circuits are equivalent.

Unfortunately, CNF $C_{M(f,h)}$ can turn out to be extremely hard for all state-of-the art SAT solvers. Such situations are typical for a number of circuits defining well-known arithmetic functions, e.g., multiplication functions for natural numbers. Such circuits are called multipliers. It is well-known fact that the LEC problem in SAT form for two multipliers is extremely hard for the best SAT which won in SAT solvers competitions conducted annually. As a consequence, in some recent works instead LEC the problems of identification of multipliers are considered: in the framework of this approach the problem to recognize if the considered circuit implements multiplier is studied. Such an approach gives the possibility to work with

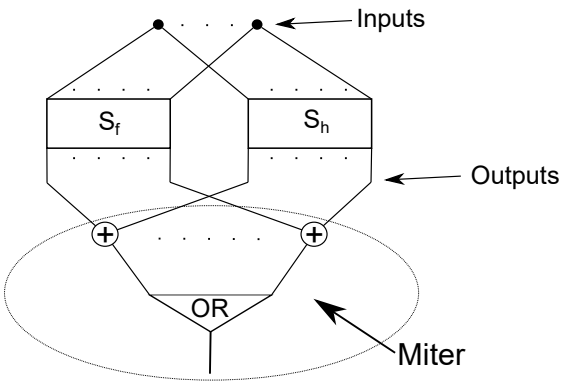


Fig. 1. LEC constructed by two circuits S_f and S_h with miter.

multipliers of significantly higher dimension comparing with LEC approach [12], [13], [14], [15]. In our study we consider only LEC approach to verification of multipliers.

Often when dealing with such hard problems it is not even possible to predict the SAT solver running time a priori, and this is a serious problem of its own. To solve this problem, in [16] was proposed an approach based on decomposition hardness notion. In [17] ideas from [16] were transferred to hardness estimation of hard LEC in form of SAT instances, exploiting information about structure of considered circuits. This technique from [17] will be used further to estimate hardness of LEC in SAT form for two multipliers when we augmented our CNF encoding of considered LEC with some additional information, obtained using preprocessing techniques, based on decision diagrams of special kind.

Specifically, we will use further so-called Disjunctive Diagrams, introduced in [5]. Let us briefly recap their main point. A DisJunctive Diagram (DJD) is constructed by an arbitrary formula in disjunctive normal form (DNF) and the diagram is essentially Shannon bracket representation of considered DNF in form of directed graph. Using arbitrary CNF C, we construct DNF $D = \neg C$, and then using D we construct DJD $R(D)$ representing it. It is important that $R(D)$ is constructed in polynomial time of length D. An arbitrary diagram $R(D)$ has two terminal vertices: vertex '1' and vertex '?'. The diagram $R(D)$ in contrast to the well-known ROBDD [18] and ZBDD [19] is not binary and can have more than one root.

Each path in $R(D)$ from some root to terminal vertex '1' corresponds to a particular conjunction of literals in D. An arbitrary path π from some root to terminal '?' defines an assignment α_π over some set of variables $X_\pi \subseteq X$ (hereinafter, X the set of variables over which the original CNF formula C is specified). Next, we show how the path properties in $R(D)$ in terminal vertex '?' can be used to generate additional constraints using which we can improve the performance of SAT solver in application to extremely hard LEC instances.

Figure 2 shows DNF and the corresponding disjunctive forest (left) and the disjunctive diagram (right) constructed

$$D = (x_1 \wedge \neg x_2 \wedge \neg x_3) \vee (x_1 \wedge \neg x_3 \wedge \neg x_4) \vee (\neg x_2 \wedge \neg x_3) \vee (x_2 \wedge x_4) \vee (\neg x_3 \wedge x_4) \vee x_4$$

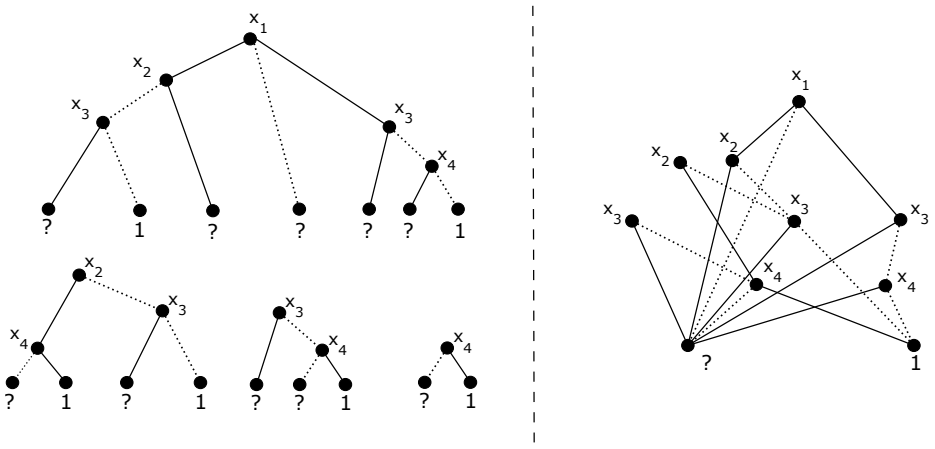


Fig. 2. DNF D with Disjunctive forest $F(D)$ (left) and DJD $R(D)$ (right) for it.

with the order $x_1 < x_2 < x_3 < x_4$.

III. Hardness estimations of Extremely hard LEC instance and disjunctive diagrams based preprocessing

So, let C be a CNF formula encoding some extremely hard LEC instance. As noted above, on such formula a modern SAT solver A can work indefinitely long and, unfortunately, this time cannot be predicted with the necessary accuracy in the general case. However, it is possible to construct an estimate of the hardness of formula C using the ideas from the article [16]. Let us recall that in [16] it was proposed to estimate the hardness of C w.r.t. A and some plain partitioning [20] of formula C, and this approach is close in sense of idea to one from [21] to use the notion of backdoor [22] to evaluate the hardness of arbitrary Boolean formula. Corresponding hardness measure is referred as decomposition hardness (d-hardness) or hardness w.r.t. SAT partitioning. In [17] we described one particular construction of SAT partitioning that can be used to evaluate the hardness of CNF formulas encoding namely some LEC instances. Let us give its brief description.

So, consider LEC for two Boolean circuits S_f and S_h, as well as miter-circuit $S_{M(f,h)}$ and construct for this circuit CNF formula $C_{M(f,h)}$ over set of variables X. Denote by X^{in}, $X^{in} \subseteq X$ set of variables associated with inputs of S_f,S_h and $S_{M(f,h)}$ when transition from $S_{M(f,h)}$ into $C_{M(f,h)}$ happens. For some natural $k : 2 \leq k < n$ let us split X^{in} into $q = \lceil \frac{n}{k} \rceil$ disjoint subsets, where:

$$n = k\lfloor \frac{n}{k} \rfloor + r, r \in \{0, 1, ..., k-1\}$$

Let us denote the corresponding subsets X^j, $j \in \{1, \ldots, q\}$, and suppose that for each $j \in \{1, \ldots, \lfloor \frac{n}{q} \rfloor\}$ we have $|X^j| = k$, and in case $r \neq 0$ the following condition holds: $|X^{\lceil \frac{n}{k} \rceil}| = r$. We associate two Boolean functions with each X^j: some function $\lambda^j : \{0,1\}^{|X^j|} \to \{0,1\}$ and its negation $\neg\lambda^j$. Let ${\phi_1}^j$ and ${\phi_2}^j$ be Boolean formulas in CNF specified functions λ^j and $\neg\lambda^j$ respectively, and ϕ^j be a formula which can be in different moments both formulas ${\phi_1}^j$ and ${\phi_2}^j$. It is obvious, that there exist $2^{\lceil \frac{n}{k} \rceil}$ different formulas of kind:

$$\phi^1 \wedge ... \wedge \phi^{\lceil \frac{n}{k} \rceil}$$

The main fact established in [17] is that this family consisting of $2^{\lceil \frac{n}{k} \rceil}$ formulas is SAT partitioning of formula $C_{M(f,h)}$. Experiments show that picking up functions of kind λ^j and $\neg\lambda^j$ carefully, we can have the situation when formulas of kind:

$$\phi^1 \wedge ... \wedge \phi^{\lceil \frac{n}{k} \rceil} \wedge C_{M(f,h)} \tag{1}$$

are significantly simpler for SAT solver A comparing with original formula $C_{M(f,h)}$.Then hardness of $C_{M(f,h)}$ w.r.t. A and SAT partitioning

$$P = \left\{ (\phi^1 \wedge ... \wedge \phi^{\lceil \frac{n}{k} \rceil})_i \right\}_{i=1}^{2^{\lceil \frac{n}{k} \rceil}}$$

can be specified (by analogy with [16]) as summary runtime of algorithm A for all formulas of kind (1). Let us denote this measure $\mu_{A,P}(C_{M(f,h)})$. Arguing by analogy with [16], it is easy to show that following equation holds:

$$\mu_{A,P}(C_M(f,h)) = 2^{\lceil \frac{n}{k} \rceil} \cdot E[\xi_P]$$

where ξ_P is random variable associated with P in manner similar to one from [17]. But then we can estimate the value $\mu_{A,P}(C_{M(f,h)})$ estimating expectation $E[\xi_P]$ using the Monte Carlo method [16], [17].

The described method of estimating the hardness of formulas turns out to be very useful for understanding the effectiveness of various techniques that speed up SAT solvers in application to extremely hard SAT instances (when we cannot predict the runtime of conventional SAT solver A). We use the method of hardness estimation just described to investigate how efficient the preprocessing method can be for hard formulas of the kind $C_{M(f,h)}$, when one harnesses

disjunctive diagrams to this preprocessing. Let us describe corresponding method.

Let us return to diagram $R(D)$ constructed for DNF formula $D = \neg C$, where C is some CNF formula over variables X, and consider an arbitrary path π which is passed from some root vertex to terminal vertex '?'. Use the notations introduces above: X_π and α_π. Denote Ω_π the set of all assignments of variables from X which coincide with α_π in values of variables from X_π, thus,

$$|\Omega_\pi| = 2^{|X|-|X_\pi|}$$

Let $\Pi_?$ be the set of all paths in terminal vertex '?' in $R(D)$. From the basic properties of disjunctive diagrams it follows that the value $|\Pi_?|$ is bounded above by some polynomial of size of formula C in general case.

Leading by [5], it is not hard to see that formula C is unsatisfiable if and only if formula of kind $C(\alpha_\pi)$ is unsatisfiable for any $\pi \in \Pi_?$, where by $C(\alpha_\pi)$ is denoted formula which is result when substituting assignment α_π of variables from X_π into C. Starting from this point, let us establish the following fact.

Statement 1. *Consider an arbitrary $\pi \in \Pi_?$. If formula $C(\alpha_\pi)$ is unsatisfiable, then clause obtained by negation of α_π is logical entailment of formula C.*

Proof sketch. Let $C(\alpha_\pi)$ be unsatisfiable and $\Delta(\alpha_\pi)$ be the clause obtained by negation of α_π: e.g. let α_π be $x_{10} = 1, x_{18} = 0, x_{27} = 1$, then we represent α_π as the following conjunction of literals: $x_{10} \wedge \neg x_{18} \wedge x_{27}$. Accordingly, $\Delta(\alpha_\pi) = (\neg x_{10} \vee x_{18} \vee \neg x_{27})$. If α is some satisfying assignment for C, then α can not coincide with α_π in corresponding literals (otherwise we have contradiction with the fact that α satisfies C). This fact means that $C \rightarrow \Delta(\alpha_\pi)$. □

The established fact, despite of its simplicity, gives us a simple way to construct additional constraints which can be used to augment the original formula C, hoping to reduce the search space for algorithm A. We can consider formulas of kind $C(\alpha_\pi)$ for all $\pi \in \Pi_?$ (keeping in mind that fact, that $|\Pi_?|$ is bounded by polynomial of formula C description size). To each formula $C(\alpha_\pi)$ we will apply SAT solver A. Corresponding SAT problems can be extremely hard, and consequently we can launch A with some fixed time limit t. And from all said above it follows that clauses of kind $\Delta(\alpha_\pi)$, if it managed to construct them in time $\leq t$, can be very useful due to prohibition by them some prospective search directions for algorithm A.

IV. Computational experiments

This section presents the results of computational experiments on the application of DJD-preprocessing to hard LEC problems.

We conducted computational experiments on the "Academician V.M. Matrosov" cluster of Irkutsk Supercomputer Center [23].

TABLE I
Comparison of decomposition hardness estimates of instance CvK with DJD-preprocessing by different orders

Order	Average time to solve subproblems (s.)	Estimate of d-hardness after preprocessing (s.)
Direct	30.29	1 985 085
Frequency	29.68	1 945 108
Random 1	30.06	1 970 013
Random 2	29.78	1 951 663
Random 3	29.95	1 962 804
Random 4	29.75	1 949 696
Random 5	30.02	1 967 391
Random 6	**29.53**	**1 935 279**
Random 7	29.93	1 961 493
Random 8	30.03	1 968 047
Random 9	29.96	1 963 459
Random 10	29.87	1 957 561

As benchmarks, as mentioned above, we considered problems encoding equivalence of different algorithms for multiplication of pairs of integers. Concretely, the following algorithms were considered: standard "Column" multiplier, "Wallace tree" [24], Dadda algorithm [25], as well as "Karatsuba decomposition" [26]. In all cases, we considered versions of algorithms that multiply two 16-bit numbers. Further we will refer to corresponding benchmarks as to `CvK`, `CvW`, `KvW`, `DvC`, `DvK` and `DvW`, by first letters of algorithms names whose equivalence is tested in each particular problem. All these algorithms were translated to And-Inverter Graphs, using Transalg program tool [27], [28]. Corresponding AIGs and supporting scripts are available in github repository[1].

The main results of the experiments demonstrate that the generating additional constraints using disjunctive diagrams can reduce the estimated time for solving extremely hard LEC instances in SAT form quite significantly. All estimates were constructed in the way described in Section III, and the following parameters of this method were used:

$$q = 2, X = \{x_1, ..., x_{32}\}, X^j = \{x_{2j-1}, x_{2j}\}, j = 1, ..., 16$$

The functions $\lambda^j, \neg\lambda^j$, where

$$\lambda^j = x_{2j-1} \oplus x_{2j}, j = 1, ..., 16$$

were used. Thus, each SAT partitioning consisted of $2^{32/2} = 2^{16}$ subproblems. The SAT solver Kissat [29] was used as Algorithm A when we constructed the estimates.

Let us remind [5] that when constructing some disjunctive diagram one can use different orders of variables. We generated additional constraints of kind $\Delta(\alpha_\pi)$ using different such orders. Specifically, the following orders were used:

- Frequency order — order according to the frequency of occurrence of variables in a formula;

[1] https://github.com/Vikseko/LEC_benchmarks

TABLE II
COMPARISON OF DECOMPOSITION HARDNESS BEFORE AND AFTER DJD-PREPROCESSING

Instance	d-hardness before prep. (s.)	d-hardness after prep. (s.)	% of speedup
CvK	2 178 663	1 929 530	11.4
CvW	1 418 200	1 298 746	8.4
DvC	1 302 856	1 118 082	14.1
DvK	2 301 015	2 165 651	5.8
DvW	1 595 147	—	0
KvW	2 344 223	2 215 808	5.4

- Direct order — the order that coincides with the numbers of variables in the formula and, respectively, gates in the circuit;
- Random order — random order of variables.

All work with disjunctive diagrams was performed using the program PyDJD [30], which uses the software complex PySAT [31]. The SAT solver MapleChrono [32] was used as SAT oracle in the PySAT. The time limit for checking one path was 0.01 seconds.

We constructed estimates of decomposition hardness before and after DJD preprocessing in 12 different ways for each formula under the following considerations: a disjunctive diagram by direct order, by frequency order, and by 10 different random orders. In all our experiments decomposition hardness equals the time of sequential solving of all weakened subproblems in corresponding partitioning.

Estimates were constructed from a random sample of 1000 subproblems. As an example, detailed results for instance CvK are shown in Table I. Decomposition hardness of the instance CvK, constructed by solving all 2^{16} subproblems, was 2178663 seconds, with an average runtime of one subproblem equals 33.24 seconds.

The first column of Table I shows the order in which the disjunctive diagram was constructed to perform preprocessing. The second column of the table gives the average time to solve the subproblems from the decomposition. The third column gives an estimate of the time it would take to completely solve the all sub-problems when performing order-specific preprocessing. As can be seen in the table, the best estimate was obtained for one of the random orders.

A similar study was performed for all of the problems under consideration. In the results for each CNF, except for DvW, an order was found whose preprocessing improves the decomposition hardness estimation more than the other orders.

In the next experiment, a complete solving of all subproblems in corresponding partitioning was performed to obtain the exact decomposition hardness for each problem before and after preprocessing with the best order w.r.t. found estimate for each considered LEC problem.

Let us briefly describe Table II. The first column gives the names of the problems under consideration. The second column gives the decomposition hardness of the problems. In this case, we are no longer talking about estimation, but about the exact time it takes to solve all subproblems in the decomposition (for corresponding calculations computing cluster was used). The third column gives the decomposition hardness of problems after preprocessing, in case it is better than that given in the second column. The last column of the table indicates how much faster the entire partitioning was solved when DJD-preprocessing was applied. As we can see, for 5 of the 6 pairs of multiplication algorithms considered, preprocessing improved the decomposition difficulty of the corresponding equivalence check problem by $5 - 14\%$. And this can be considered as significant achievement for such hard problem as LEC for two multipliers.

V. CONCLUSION AND FUTURE WORK

In the present paper we propose an approach to preprocessing of extremely hard SAT instances using disjunctive diagrams (DJD) introduced in [5]. The key feature of our approach is that we construct new constraints, which augment an original formula C, using different orders to construct DJD for C. Also we use the relatively new technique [17] to estimate the hardness of considered formulas. We demonstrate that proposed preprocessing technique allows to improve the state-of-the-art SAT solvers performance in application to extremely hard formulas encoding the Logical Equivalence Checking problem for two Boolean circuits implementing different algorithms of integer multiplication. In the nearest future we plan to develop preprocessing algorithms which use DJD structure in more complete manner, rearranging the diagram due to exclusion some paths in vertex '?' with help of SAT oracle.

ACKNOWLEDGMENTS

We would like to express our deep gratitude to anonymous reviewers for their valuable comments.

Victor Kondratiev and Alexander Semenov were supported by the Russian Science Foundation, project № 22-21-00583. Their contribution consists of theoretical base of proposed method and development of related algorithms.

Ilya Otpuschennikov was supported by the Ministry of Science and Higher Education of Russian Federation, research project no. 121041300065-9. His contribution is development of benchmarks sets, used in the experimental part of this research.

Victor Kondratiev was also supported by a personal scholarship from Huawei.

REFERENCES

[1] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.
[2] O. Goldreich, *Computational Complexity: A Conceptual Perspective*. Cambridge University Press, 2008.
[3] A. Kuehlmann and F. Krohm, "Equivalence checking using cuts and heaps," in *DAC*, 1997, pp. 263–268.
[4] P. Molitor and J. Mohnke, *Equivalence Checking of Digital Circuits: Fundamentals, Principles, Methods*. Kluwer Academic Publishers, 2004.

[5] A. A. Semenov and I. V. Otpuschennikov, "On one class of decision diagrams," *Automation and Remote Control*, vol. 77, no. 4, pp. 617–628, 2016.
[6] A. Biere, M. Heule, H. van Maaren, and T. Walsh, Eds., *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336.
[7] C.-L. Chang and R. C.-T. Lee, *Symbolic Logic and Mechanical Theorem Proving*, ser. Computer Science Classics. Academic Press, 1973.
[8] G. S. Tseitin, "On the complexity of derivation in propositional calculus," *Studies in Constructive Mathematics and Mathematical Logic, Part II*, pp. 115–125, 1970.
[9] I. Wegener, *The Complexity of Boolean Functions*. John Wiley & Sons, 1987.
[10] A. Biere, "The AIGER And-Inverter Graph (AIG) format version 20071012," Institute for Formal Models and Verification, Johannes Kepler University, Altenbergerstr. 69, 4040 Linz, Austria, Tech. Rep. 07/1, 2007.
[11] R. Drechsler, T. A. Junttila, and I. Niemelä, "Non-clausal SAT and ATPG," in *Handbook of Satisfiability*, ser. Frontiers in Artificial Intelligence and Applications, A. Biere, M. Heule, H. van Maaren, and T. Walsh, Eds. IOS Press, 2009, vol. 185, pp. 655–693.
[12] D. Kaufmann, A. Biere, and M. Kauers, "Verifying Large Multipliers by Combining SAT and Computer Algebra," in *2019 Formal Methods in Computer Aided Design (FMCAD)*, 2019, pp. 28–36.
[13] D. Kaufmann and A. Biere, "Amulet 2.0 for verifying multiplier circuits," in *TACAS*, ser. LNCS, vol. 12652, 2021, pp. 357–364.
[14] D. Kaufmann and A. Biere, "Improving AMulet2 for verifying multiplier circuits using SAT solving and computer algebra," *International Journal on Software Tools for Technology Transfer*, 2023.
[15] C. Yu, M. Ciesielski, and A. Mishchenko, "Fast algebraic rewriting based on And-Inverter Graphs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1907–1911, 2018.
[16] A. A. Semenov, D. Chivilikhin, A. Pavlenko, I. V. Otpuschennikov, V. Ulyantsev, and A. Ignatiev, "Evaluating the hardness of SAT instances using evolutionary optimization algorithms," in *CP*, ser. LIPIcs, vol. 210, 2021, pp. 47:1–47:18.
[17] A. A. Semenov, K. Chukharev, E. Tarasov, D. Chivilikhin, and V. Kondratiev, "Estimating the hardness of SAT encodings for logical equivalence checking of boolean circuits," *CoRR*, vol. abs/2210.01484, 2022.
[18] R. E. Bryant, "Graph-Based Algorithms for Boolean Function Manipulation," *IEEE Transactions on Computers*, vol. C-35, no. 8, pp. 677–691, 1986.
[19] S.-I. Minato, "Zero-Suppressed BDDs for Set Manipulation in Combinatorial Problems," in *Proceedings of DAC '93*. ACM, 1993, p. 272–277.
[20] A. E. J. Hyvärinen, "Grid Based Propositional Satisfiability Solving," 2011, PhD thesis. Aalto University publication series.
[21] C. Ansótegui, M. L. Bonet, J. Levy, and F. Manyà, "Measuring the hardness of SAT instances," in *AAAI*, 2008, p. 222–228.
[22] R. Williams, C. P. Gomes, and B. Selman, "Backdoors to typical case complexity," in *IJCAI*, 2003, pp. 1173–1178.
[23] "Irkutsk Supercomputer Center of SB RAS. URL: http://hpc.icc.ru." [Online]. Available: http://hpc.icc.ru
[24] C. S. Wallace, "A suggestion for a fast multiplier," *IEEE Transactions on Electronic Computers*, vol. EC-13, no. 1, pp. 14–17, 1964.
[25] L. Dadda, "Some schemes for parallel multipliers," *Alta Frequenza*, vol. 34, no. 5, pp. 349–356, May 1965.
[26] D. Knuth, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, ser. Addison-Wesley Series in Computer Science & Information Processing. Addison-Wesley, 1969.
[27] I. Otpuschennikov, A. Semenov, I. Gribanova, O. Zaikin, and S. Kochemazov, "Encoding Cryptographic Functions to SAT Using TRANSALG System," in *ECAI 2016, Frontiers of Arificial Intelligence an Applications*, vol. 285, 2016, pp. 1594–1595.
[28] A. Semenov, I. Otpuschennikov, I. Gribanova, O. Zaikin, and S. Kochemazov, "Translation of Algorithmic Descriptions of Discrete Functions to SAT with Applications to Cryptanalysis Problems," *Logical Methods in Computer Science*, vol. 16, no. 1, pp. 1–42, 2020.
[29] A. Biere, K. Fazekas, M. Fleury, and M. Heisinger, "CaDiCaL, Kissat, Paracooba, Plingeling and Treengeling entering the SAT Competition 2020," in *Proc. of SAT Competition 2020 – Solver and Benchmark Descriptions*, ser. Department of Computer Science Report Series B, vol. B-2020-1. University of Helsinki, 2020, pp. 51–53.
[30] V. Kondratiev, "Using disjunctive diagrams for preprocessing of conjunctive normal forms," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 883–887.
[31] A. Ignatiev, A. Morgado, and J. Marques-Silva, "PySAT: A Python toolkit for prototyping with SAT oracles," in *SAT*, 2018, pp. 428–437.
[32] A. Nadel and V. Ryvchin, "Chronological Backtracking," in *SAT*, ser. LNCS, vol. 10929, 2018, pp. 111–121.

Empirical Analysis of the RC2 MaxSAT Algorithm

Stepan Kochemazov, Victor Kondratiev, Irina Gribanova
ISDCT SB RAS, Irkutsk, Russia
Emails: veinamond@gmail.com, vikseko@gmail.com, the42dimension@gmail.com

***Abstract*—The Boolean satisfiability problem (SAT) and maximum satisfiability problem (MaxSAT) are among the most well-known combinatorial problems in today's computer science. The algorithms for their solving also go hand-in-hand, in that most MaxSAT solvers employ SAT solvers as the so-called oracles. In the present paper we perform a computational study of the RC2 algorithm, which is among the best state-of-the-art algorithms for MaxSAT solving. We view it from the SAT oracle viewpoint and consider how the SAT oracle's runtime is distributed among RC2 procedures and heuristics, and how this statistics differs depending on the SAT solver employed as an oracle. In addition to that we consider the two baseline MSE'18 configurations of RC2, analyze their performance and experiment with blending them together.**

***Keywords*—MaxSAT, SAT, rc2**

I. Introduction

Nowadays, the industrial design poses many challenges that require the use of highly efficient computational apparatus for solving combinatorial problems of large dimension. In particular, many problems can be expressed in a relatively simple language, involving the constraints over the variables that take integer values or even only the values from the set $\{0, 1\}$. They arise in economics, planning, computer-aided design and many other areas. Among the most well known scientific approaches for solving such problems one can point out the ability to express them in form of instances of two classical problems from Computer Science: the Integer Linear Programming problem (ILP) [1] and the Boolean satisfiability problem (SAT) [2].

ILP consists in maximizing the value of a linear function defined over a set of integer variables while satisfying a number of linear inequalities. In SAT the problem is specified by an arbitrary Boolean formula, and the goal is to determine whether there exists an assignment of all variables of a formula on which it evaluates to $True$. While both problems are NP-complete, in practical reality, ILP solving algorithms fit well with large-scale optimization tasks planning, scheduling, logistics, etc. Meanwhile, SAT solving algorithms are usually employed to construct formal proofs or find counterexamples under certain assumptions, and find a lot of uses in hardware and software verification, certain areas of planning, etc.

What makes SAT unique is that the algorithms for its solving are often used as the so-called *oracles* in the algorithms designed for tackling combinatorial problems with the hardness beyond NP. The term "SAT oracle" is inspired by the concept of oracles widely used in computational complexity theory, where they represent some entities capable of solving certain (hard) problems. Two well-known examples of the problems, the algorithms for solving which employ SAT oracles, are Satisfiability Modulo Theories (SMT) [3] – the approach that blends the rich expressiveness of languages of mathematical theories different from propositional logic with the effectiveness of modern SAT solvers, and Maximum Satisfiability [4] – the optimization variant of SAT, that makes it possible to naturally consider instances related to optimization. After all, in day-to-day practice, it is usually more convenient to solve *optimization* instances: find the shortest path from A to B, establish the shipping schedule to minimize losses while maximizing profits, etc., instead of *decision* instances: prove or disprove that there is a path from A to B shorter than k, or find a schedule under which the profits are at least P while losses are at most L or prove that it does not exist.

The latter observation, together with the fast progress and wide availability of fast SAT solving algorithms (compared to the ILP area where commercial solvers are in a league of their own performance-wise) led to a large popularity of MaxSAT and the fast development of algorithms for its solving. In the present paper we study one of the algorithms for solving MaxSAT, called `RC2` [5]. It was proposed in 2018 and took the first place in both weighted and unweighted categories of the MaxSAT Evaluations held in 2018 and 2019[1].

In our study we look at `RC2` from the perspective of the employed SAT oracle in order to better understand the possible venues for improving the interconnection between the higher-order MaxSAT logic and the ground-level SAT oracle. For this purpose we analyze the runtime distributions of the main `RC2` procedures, especially the ones that employ SAT oracles.

Let us give a brief overview of the paper: in the next chapter we introduce the necessary notation regarding SAT and MaxSAT, provide the basic information about the architecture of modern SAT and MaxSAT solvers. In Section III we describe the `RC2` algorithm, the way it uses SAT solvers as oracles, and the related heuristics. After this we present the details about the experimental setting, describe our computational experiments and analyze the obtained data.

II. Preliminaries

A Boolean variable is a variable that takes the values from the set $\{False, True\}$ which in practice usually converts into $\{0, 1\}$, where 0 represents $False$ and 1 represents $True$. A Boolean formula is comprised of Boolean variables connected

[1]The MaxSAT Evaluations are annual competitive events for MaxSAT solvers that promote and chronicle the progress in this area

using braces and the so-called *logical connectives*. The latter are the well-known logical operations that usually form a complete basis, e.g. $\{\wedge, \neg\}$, $\{\wedge, \vee, \neg\}$, etc. In both SAT solving and MaxSAT solving, Boolean formulas are considered in the Conjunctive Normal Form (CNF), so for simplicity let us only introduce this form. A formula in CNF is defined over the basis $\{\wedge, \vee, \neg\}$. CNF is a conjunction ($\wedge$) of clauses, where a *clause* is a disjunction ($\vee$) of literals, and a *literal* is either a Boolean variable or its negation ($\neg$). Literals x and $\neg x$ are called *complementary*, because for any value of x one of them will take the value 1. For convenience, the clauses in CNF are not allowed to contain complementary or duplicate literals. An *assignment* of variables is the set of values from $\{0, 1\}$, assigned to each variable. An assignment is called *satisfying* if it turns formula into 1 in accordance with the standard semantics of logical connectives. In the case of CNF it is especially easy to check that an assignment is satisfying: it is sufficient to ensure that every clause if satisfied. A formula is called *satisfiable* if it has at least one satisfying assignment, otherwise it is called *unsatisfiable*.

A. SAT and SAT Solvers

The Boolean satisfiability problem (SAT) consists in the following: for an arbitrary CNF to answer whether it is satisfiable or not (decision problem). The more practical search variant is to find a satisfying assignment if CNF is satisfiable.

The practical algorithms for solving SAT are called *SAT solvers*. They have a (relatively) long history and are considered one of the success stories in today's Computer Science. The SAT solving algorithms achieved a tremendous progress during the recent 30 years, going from barely being able to tackle formulas with hundreds of variables and thousands of clauses to effectively solving practical SAT instances with hundreds of thousands of variables and millions of clauses. This fact, together with the applications of SAT solvers in formal verification of circuits [6] and software [7] made them into an effective tool widely employed today in many other areas, such as bioinformatics [8] or cryptography [9].

The result of the many years of progress in SAT solving and the currently predominant algorithm for this purpose is called *Conflict-Driven Clause Learning* (CDCL). The CDCL SAT solvers enhance the depth-first search with various methods, such as *clause learning* and backjumping [10], variable selection heuristics [11], restarts, heuristics for manipulating the clauses of a formula mid-search, etc., see [2] for a comprehensive overview. An important feature of some of the state-of-the-art SAT solvers is the ability to work in the so-called *incremental* mode [12], which makes it possible to use a single SAT solver to tackle a sequence of SAT instances, in which each following instance is formed by extending the previous instance by adding to it new clauses and possible new variables. Thanks to this one can reuse the information accumulated by the solver throughout the process to improve its performance.

B. Maximum Satisfiability

The *Maximum satisfiability problem* (MaxSAT) is an optimization variant of SAT. For a Boolean formula in CNF the MaxSAT problem is to maximize the number of satisfied clauses. Due to a number of practical considerations, in practice MaxSAT is usually considered in a slightly different formulation called *Weighted Partial MaxSAT* (WPMS). The *Weighted* part refers to the fact that clauses are allowed to have weights. The *Partial* part instead is aimed at splitting formula in two: the part that must be satisfied (*hard* clauses) and the part that may be (partially) falsified (*soft* clauses). One can view hard clauses in the context of the weighted model as the clauses the individual weight of which exceeds the total weight of all soft clauses.

There are two main classes of MaxSAT solving algorithms: complete and incomplete. Complete algorithms make it possible to construct provably optimal solutions to the optimization problem at hand and it is the complete algorithms that employ SAT solvers. Incomplete algorithms, on the other hand, are designed in such a way as to quickly find a solution which is quite good, and then use the allotted time to improve it (if possible). One notable example of incomplete algorithms for MaxSAT which is also employed in the context of SAT solving are the Stochastic Local Search algorithms. Generally, they do not employ any inference rules like SAT solvers, instead, they explore the space of possible assignments of variables of a formula, but they do it with insane speed, checking up to billions of assignments per second. The recent advancements in MaxSAT solving actually focus mostly on incomplete algorithms and also on ways to efficiently combine fast incomplete methods for finding approximations of a solution with the ability of complete methods to find an optimum.

So, a MaxSAT instance is a weighted CNF: a CNF in which some clauses are hard, e.g. they must be satisfied, and the remaining clauses are soft and are assigned some weights. The goal is to find an assignment of variables that satisfies all hard clauses and maximizes the sum of weights of satisfied soft clauses.

The state-of-the-art MaxSAT solvers rely on two techniques to employ SAT oracles: selector literals and cardinality constraints [13]. An individual *selector literal* is added to each soft clause. For example, if there is an original clause $a \vee b \vee c$, it is modified to $a \vee b \vee c \vee s_1$. By assigning $s_1 = 1$ the algorithm can specify that this clause is already satisfied and can be excluded from consideration (e.g. when invoked on a CNF with this clause, the SAT oracle will search for a solution that does not have to satisfy it). Vice versa, by fixing s_1 to 0 the SAT oracle will have to satisfy this clause. The cardinality constraints technique allows to represent in SAT form the [in]equalities of the kind $x_1 + x_2 + \ldots + x_n [\leq, \geq, =] K$.

It is easy to use the two described techniques to construct the following straightforward SAT-based algorithm for MaxSAT. First, introduce selector literals to all soft clauses in an input weighted CNF. Next, set the value of auxiliary variable k to 1. Encode the cardinality constraint over selector variables so

that it enforces that the sum of weights of all satisfied soft clauses is at least k. Then, invoke a SAT solver on the CNF constructed by appending the encoded cardinality constraint to CNF with selectors, if the answer is "Satisfiable" then increment k, otherwise, k is the maximum weight that we needed to find.

It is clear, that if we exclude SAT oracle calls, then the algorithm is polynomial and works for at most S steps. Thus, the use of SAT oracle constitutes the major part of the complexity of this MaxSAT solving algorithm. Due to the fact that the size of cardinality encodings quickly grows with the increase of the number of variables involved, the presented approach is not as good as more sophisticated algorithms for MaxSAT. For an excellent tutorial on MaxSAT we refer to online resource [2].

III. RC2 Algorithm for MaxSAT

The `RC2` algorithm [5] is one of the state-of-the-art complete algorithms for MaxSAT. It is implemented as a part of the PySAT package [14]. Let us below briefly describe its main features and the ways it interacts with a SAT oracle.

The `RC2` abbreviation means *Relaxable Cardinality Constraints*. It is a *core-guided* MaxSAT solver, meaning that its functioning revolves around the so-called unsatisfiable cores. Such a core is formed by a subset of clauses that are responsible for a SAT instance being unsatisfiable. In particular, `RC2` employs the selector variables in a manner similar to that of the algorithm described above, but constructs cardinality constraints only for selector variables from unsatisfiable cores.

`RC2` heavily relies on the incremental capabilities of an employed `MiniSat`-like [15] SAT oracles: the extraction of an unsatisfiable core is performed via the *assumptions* mechanism (see [12]), the manipulation of cardinality constraints is implemented via the use of iterative cardinality encodings, such as incremental Totalizer encoding [16], [17], etc. Since we do not aim to explain `RC2` in detail, for particulars regarding the inner workings of the algorithm refer to [5]. As it was already mentioned, the goal of the present paper is to study how the `RC2` handles the SAT oracles. Let us consider it in detail in the following sections.

IV. Analysis of RC2 Interaction With SAT Oracle

`RC2` performs SAT oracle calls in 4 procedures [3]:

- `compute_` – the main procedure that seeks to find the solution to the optimization problem
- `trim_core` – the procedure used to reduce the size of a recently extracted unsatisfiable core.
- `exhaust_core` – the procedure used to *exhaust* an unsatisfiable core by attempting to increase its cost
- `minimize_core` – the procedure used to invoke a SAT oracle with a tight conflict budget in order to reduce an extracted unsatisfiable core via a deletion-based algorithm.

[2]https://ecai20-maxsat-tutorial.github.io/

[3]https://github.com/pysathq/pysat/blob/master/examples/rc2.py

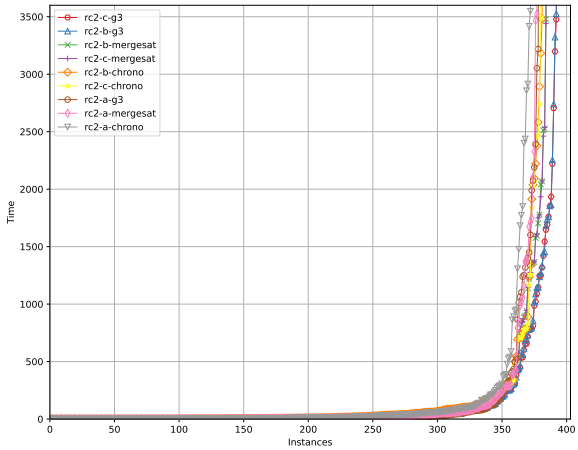


Fig. 1. Cactus plot for considered solvers on unweighted benchmarks

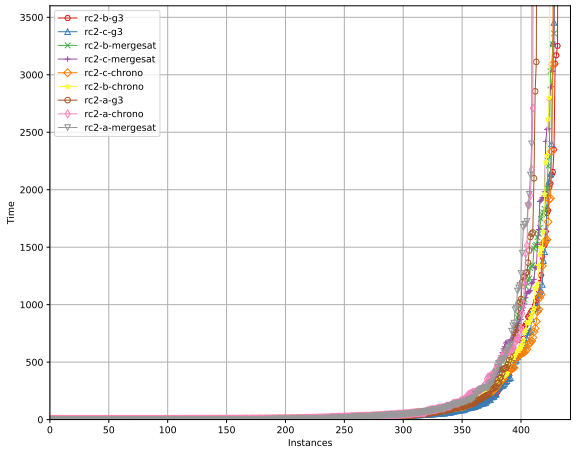


Fig. 2. Cactus plot for considered solvers on weighted benchmarks

Apart from the mentioned methods, one of the major heuristics in the algorithm is the detection of intrinsic AtMost1 constraints, implemented in the `adapt_am1` method.

Note, that there are two baseline versions of `RC2`: `RC2-A` and `RC2-B`, which differ in how they reduce the unsatisfiable cores. Usually, the `RC2-B` version performs slightly better than the other.

A. Experimental Setup

All experiments were run on a PC with AMD Ryzen 5900x CPU (12 cores) and 80 GB RAM, working under the Ubuntu 22.04 OS. In all cases 12 threads were running simultaneously. In the role of benchmarks we initially picked the suite of instances from the weighted and unweighted categories of

TABLE I
STATISTICS ON THE RUNTIME OF PROCEDURES IN RC2 WITH DIFFERENT SAT ORACLES ON **UNWEIGHTED** BENCHMARKS

	Parsing		Compute		Adapt AM1		Minimize		Exhaust		Trim	
	abs	rel	abs	rel	abs	rel	abs	rel	abs	rel	abs	rel
RC2-A-Glucose3												
mean	2.412	0.308	42.815	0.268	1.636	0.145	0.000	0.000	45.771	0.238	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	124.644	1.000	3146.778	1.000	69.196	0.998	0.001	0.003	1994.739	0.991	0.002	0.004
median	0.111	0.139	0.099	0.107	0.009	0.016	0.000	0.000	0.083	0.099	0.000	0.000
RC2-B-Glucose3												
mean	3.093	0.276	47.783	0.219	2.016	0.120	29.067	0.230	35.919	0.132	0.000	0.000
min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	105.170	1.000	2732.359	0.999	70.783	0.998	3298.355	0.996	2967.639	0.990	0.004	0.003
median	0.122	0.104	0.102	0.048	0.009	0.011	0.165	0.082	0.052	0.030	0.000	0.000
RC2-C-Glucose3												
mean	3.066	0.285	47.417	0.217	2.019	0.117	28.459	0.227	35.361	0.131	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	104.207	1.000	2696.451	0.999	71.387	0.998	3255.592	0.995	2852.446	0.989	0.004	0.004
median	0.123	0.108	0.096	0.046	0.009	0.009	0.163	0.081	0.048	0.029	0.000	0.000
RC2-A-Maplechrono												
mean	1.204	0.222	38.834	0.313	1.583	0.113	0.000	0.000	56.555	0.328	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	21.962	1.000	2391.076	0.998	70.024	0.998	0.001	0.002	3175.744	0.996	0.001	0.003
median	0.117	0.033	0.746	0.128	0.008	0.002	0.000	0.000	0.481	0.174	0.000	0.000
RC2-B-Maplechrono												
mean	1.453	0.202	39.783	0.273	1.807	0.102	22.033	0.158	39.587	0.247	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	64.954	1.000	2957.176	0.994	69.389	0.998	3261.089	0.989	2835.435	0.994	0.004	0.003
median	0.118	0.034	0.534	0.061	0.009	0.003	0.210	0.040	0.341	0.065	0.000	0.000
RC2-C-Maplechrono												
mean	1.448	0.203	30.737	0.270	1.830	0.103	21.636	0.160	38.625	0.247	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	67.389	1.000	2283.725	0.993	69.466	0.998	3274.200	0.988	2685.708	0.994	0.004	0.003
median	0.117	0.035	0.475	0.061	0.009	0.003	0.223	0.042	0.304	0.060	0.000	0.000
RC2-A-Mergesat3												
mean	2.316	0.273	37.163	0.307	1.650	0.126	0.000	0.000	45.214	0.266	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	131.545	1.000	3075.583	1.000	68.049	0.998	0.001	0.003	2886.466	0.996	0.002	0.004
median	0.122	0.092	0.209	0.138	0.009	0.007	0.000	0.000	0.159	0.130	0.000	0.000
RC2-B-Mergesat3												
mean	1.975	0.237	48.584	0.248	2.068	0.112	23.784	0.221	29.989	0.163	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	81.939	1.000	2362.645	0.994	72.766	0.998	1445.034	0.995	1685.072	0.989	0.005	0.004
median	0.123	0.061	0.225	0.063	0.009	0.006	0.203	0.075	0.103	0.039	0.000	0.000
RC2-C-Mergesat3												
mean	1.972	0.237	48.436	0.247	2.009	0.111	23.959	0.222	30.120	0.163	0.000	0.000
min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	81.227	1.000	2378.219	0.995	71.547	0.998	1445.101	0.995	1729.190	0.989	0.005	0.003
median	0.122	0.066	0.236	0.062	0.010	0.007	0.195	0.074	0.097	0.040	0.000	0.000

MaxSAT Evaluation 2021 [4]. Then we ran `RC2-B` with Glucose 3 as SAT oracle (which is standard) on these instances with the time limit of 3600 seconds and collected only the benchmarks on which it terminated during this time (either by successfully solving them or due to some unexpected behavior).

B. Evaluating the Contribution of `RC2` Procedures to Runtime

Our main goal is to assess how much each of the procedures contributes to runtime. Of particular interest are the time required for parsing and for extracting intrinsic AtMost 1 constraints, since both procedures are actually implemented in Python, and Python is well known for not being as efficient as e.g. C or C++ when it comes to resource-intensive operations. Apart from that, it is interesting to see how the SAT oracle use, runtime-wise, is split between the main (`compute_`) procedure and the heuristics.

To study the considered details we launched `RC2` in three configurations: standard two: `RC2-A` and `RC2-B` and also `RC2-C` which enables heuristics employed in both configurations. We also decided to see whether the behavior differs depending on what solver is employed as a SAT oracle. Thus, in the role of the latter we considered three solvers supported by PySAT:

- Glucose 3 – the baseline `RC2` solver
- Maplechrono – the MapleLCMDistChronoBT solver that won in SAT Competition 2018
- Mergesat 3 – the MergeSAT solver [18] which is a maintained project aiming at preserving the MiniSat/Glucose legacy.

[4]https://maxsat-evaluations.github.io/2021/

TABLE II
STATISTICS ON THE RUNTIME OF PROCEDURES IN RC2 WITH DIFFERENT SAT ORACLES ON **WEIGHTED** BENCHMARKS

	Parsing		Compute		Adapt AM1		Minimize		Exhaust		Trim	
	abs	rel	abs	rel	abs	rel	abs	rel	abs	rel	abs	rel
RC2-A-Glucose3												
mean	6.707	0.428	36.792	0.173	4.179	0.145	0.000	0.000	49.731	0.228	0.365	0.008
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	400.178	1.000	2894.375	0.982	427.359	0.998	0.007	0.001	2562.246	0.985	42.600	0.157
median	0.148	0.339	0.028	0.061	0.008	0.010	0.000	0.000	0.028	0.051	0.000	0.001
RC2-B-Glucose3												
mean	5.146	0.349	40.586	0.129	3.934	0.121	61.825	0.315	25.773	0.076	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	365.097	1.000	3024.803	0.990	429.336	0.998	3026.639	0.976	1491.174	0.912	0.010	0.001
median	0.134	0.148	0.024	0.026	0.008	0.008	0.192	0.166	0.013	0.023	0.000	0.000
RC2-C-Glucose3												
mean	5.075	0.354	28.677	0.124	4.003	0.121	58.463	0.317	22.642	0.071	0.257	0.004
min	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	417.002	1.000	3028.002	0.961	430.930	0.998	3132.125	0.983	1446.559	0.851	28.443	0.058
median	0.144	0.152	0.022	0.026	0.008	0.007	0.178	0.154	0.011	0.024	0.001	0.001
RC2-A-Maplechrono												
mean	5.857	0.333	36.605	0.205	4.284	0.123	0.000	0.000	42.201	0.318	1.017	0.009
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	390.325	1.000	1776.640	0.992	451.236	0.998	0.009	0.001	1779.656	0.985	105.256	0.186
median	0.135	0.120	0.169	0.070	0.008	0.007	0.000	0.000	0.238	0.150	0.001	0.001
RC2-B-Maplechrono												
mean	5.163	0.294	21.111	0.163	3.939	0.106	51.758	0.265	27.412	0.166	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	368.242	1.000	758.911	0.986	437.650	0.998	2236.522	0.952	1541.358	0.985	0.019	0.001
median	0.139	0.086	0.099	0.030	0.008	0.005	0.318	0.086	0.089	0.052	0.000	0.000
RC2-C-Maplechrono												
mean	5.818	0.293	18.685	0.158	4.098	0.105	49.057	0.260	26.567	0.171	0.455	0.006
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	439.520	1.000	1013.682	0.987	435.952	0.998	2436.954	0.948	1229.961	0.956	39.269	0.154
median	0.144	0.091	0.094	0.031	0.008	0.005	0.337	0.089	0.100	0.049	0.001	0.001
RC2-A-Mergesat3												
mean	6.021	0.374	36.492	0.206	4.295	0.133	0.000	0.000	55.229	0.254	1.548	0.018
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	377.585	1.000	1602.950	0.976	452.126	0.998	0.010	0.001	1941.928	0.986	83.650	0.355
median	0.140	0.197	0.066	0.099	0.009	0.010	0.000	0.000	0.052	0.110	0.003	0.003
RC2-B-Mergesat3												
mean	5.304	0.309	41.213	0.154	3.971	0.112	75.073	0.313	42.450	0.104	0.000	0.000
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	372.215	1.000	2816.565	0.983	444.320	0.998	2693.492	0.976	3079.620	0.955	0.028	0.001
median	0.132	0.107	0.057	0.034	0.009	0.006	0.343	0.146	0.032	0.032	0.000	0.000
RC2-C-Mergesat3												
mean	6.051	0.312	36.214	0.155	4.063	0.110	80.054	0.298	40.218	0.106	2.291	0.010
min	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
max	435.943	1.000	2626.575	0.975	452.325	0.998	2553.352	0.957	2854.320	0.965	263.884	0.323
median	0.136	0.104	0.077	0.032	0.009	0.006	0.300	0.129	0.032	0.031	0.005	0.003

They represent a progression of sorts, since Maplechrono can be viewed as the solver incorporating all Glucose 3 heuristics, and Mergesat3 as the one incorporating all heuristics from both Glucose3 and Maplechrono.

The results of the experiments are summarized in Tables I, II, and also in Figure 1, 2 split in accordance with the class of benchmarks, unweighted for Table I and Figure 1 and weighted for Table II and Figure 2. The columns of the tables contain the measurements of runtime in seconds in *abs* column (absolute) and relative to the total runtime of the solver in table *rel*. The statistics are presented only for the benchmarks that have been successfully solved by each solver variant. As it will be seen later, they solved similar portion of benchmarks overall, thus making the comparison fair.

In fact, the variance in the absolute values in the parsing column is mainly due to slight runtime variations due to the use of different CPU cores, since the parser implementation is independent from the SAT oracle. However, we decided that it will be better to leave it in the Tables to better see the overall picture.

1) Cactus plots: Figures 1 and 2 contain the so-called cactus plots. To draw it, the runtimes of each algorithm over solved benchmarks are ordered in an increasing order (independently from other algorithms) and plotted as a line. It means, that the further to the right the line goes — the more benchmarks have been solved and the closer it is to the horizontal axis — the smaller is the average runtime. From the plots it is clear that indeed, the RC2-A configurations perform the worst, while configurations RC2-B and RC2-C are quite similar. Another interesting fact is that Glucose3 as a SAT

oracle outperforms both Mergesat and Maplechrono, however, not by much, since for the vast majority of benchmarks (about 350 unweighted and 400 weighted) all variants perform more or less the same.

2) General observations: From the median values in both tables it is clear, that typically, the SAT oracle calls in `RC2` are very short, but their number is significant. It means that instead of focusing on heuristics that help solving large and hard SAT instances (as the SAT competition solvers do), the SAT oracles employed by `RC2` are better off with heuristics targeting easy instances, possibly, taking into account the peculiar features of cardinality encodings employed by the algorithm.

3) Parsing: Let us look at the absolute and relative values of the time required to parse the problems. It is quite surprising to see that with the (absolute) mean time 2.4 seconds for unweighted benchmarks and about 5 seconds for weighted benchmarks, in relative terms parsing takes from 20 to 35 percent of overall runtime. One of the reasons for this behavior is that the average runtime of `RC2` on the considered benchmarks is quite small. Nevertheless, it is highly likely that implementing `RC2` in a language that will allow to make a better parser can easily yield a large average runtime reduction.

4) Adapt AM1: It is also quite curious to see how much the adaptation of intrinsic Atmost1 constraints can take: up to 400 seconds, and it takes from 10 to 15 percent on average. Here we can only repeat what was already mentioned with regards to parsing: that the Python implementation is likely hindering the performance of the algorithm in this particular procedure.

5) Compute, Exhaust and Minimize: The most surprising revelation for us was that the main procedure, that aims to solve the problem, actually takes only up to a third of total runtime on average. In fact, the Exhaust procedure in almost all cases appears to take more time than compute. The same goes for the Minimize procedure, which, surprisingly, takes more time on weighted benchmarks and less time on unweighted ones. Overall, treating the Compute procedure as the main one is mostly a misconception, since in the case of `RC2` all three procedures take more or less the same amount of time.

6) Trim: The Trim procedure is not enabled in `RC2-B`, it works only in the other two configurations. Nevertheless, from the tables is is clear that it takes the least amount of time, usually negligible, however, it potentially can improve the performance of the solver in general.

7) Variations depending on the solver used as SAT oracle: From the vast amount of information in the Tables it is hard to quickly grasp the discrepancies that are produced by using a different SAT solver as an oracle, but they are there. For example, while the mean time spent on Compute and Exhaust by Maplechrono is smaller than that by the other two solvers, the median time is actually several times larger, meaning that more often than not the corresponding SAT oracle calls take more time. Apparently, this is the main reason why Maplechrono variants lack in performance as it can be seen from the cactus plots. Overall, the median times for Glucose 3 are the smallest, with Mergesat3 lagging slightly behind. Understanding the reasons behind the better behavior of Glucose 3 and behind the bad one by Maplechrono in the main procedures of `RC2` may well be the key for finding further improvements to the algorithm.

V. Conclusions and Future Work

In this paper we studied the `RC2` algorithm for solving MaxSAT from the SAT oracle perspective. The gathered statistics shed a lot of light on the interconnection between the algorithm and the employed SAT oracles and made it possible to outline a number of possible venues that may enable improving the `RC2` performance on average. These venues include implementing RC2 in C/C++ or at least taking measures towards drastically decreasing the time required to read WCNF; implementing Adapt AM1 heuristic in a SAT solver to reduce its runtime; tuning SAT solvers for good performance on easy instances instead of focus on hard instances characteristic for SAT competition solvers.

Acknowledgments

The research was supported by the Russian Science Foundation, project 22-21-00834.

References

[1] L. A. Wolsey, *Integer programming*. John Wiley & Sons, 2020.
[2] A. Biere, M. Heule, H. van Maaren, and T. Walsh, Eds., *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336.
[3] C. W. Barrett, R. Sebastiani, S. A. Seshia, and C. Tinelli, "Satisfiability modulo theories," in *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336, pp. 1267–1329.
[4] F. Bacchus, M. Järvisalo, and R. Martins, "Maximum satisfiabiliy," in *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336, pp. 929–991.
[5] A. Ignatiev, A. Morgado, and J. Marques-Silva, "RC2: an efficient maxsat solver," *J. Satisf. Boolean Model. Comput.*, vol. 11, no. 1, pp. 53–64, 2019.
[6] R. Drechsler, T. A. Junttila, and I. Niemelä, "Non-clausal SAT and ATPG," in *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336, pp. 1047–1086.
[7] E. Clarke, D. Kroening, and F. Lerda, "A tool for checking ANSI-C programs," in *TACAS*, ser. LNCS, vol. 2988, 2004, pp. 168–176.
[8] I. Lynce and J. Marques-Silva, "Haplotype inference with boolean satisfiability," *Int. J. Artif. Intell. Tools*, vol. 17, no. 2, pp. 355–387, 2008.
[9] G. V. Bard, *Algebraic Cryptanalysis*. Springer US, 2009.
[10] J. a. P. Marques-Silva and K. A. Sakallah, "Grasp: A search algorithm for propositional satisfiability," *IEEE Trans. Comput.*, vol. 48, no. 5, p. 506–521, may 1999.
[11] M. W. Moskewicz, C. F. Madigan, Y. Zhao, L. Zhang, and S. Malik, "Chaff: Engineering an efficient sat solver," in *DAC*, 2001, p. 530–535.
[12] N. Eén and N. Sörensson, "Temporal induction by incremental sat solving," vol. 89, no. 4, pp. 543–560, 2003.
[13] O. Roussel and V. M. Manquinho, "Pseudo-boolean and cardinality constraints," in *Handbook of Satisfiability - Second Edition*, ser. FAIA. IOS Press, 2021, vol. 336, pp. 1087–1129.
[14] A. Ignatiev, A. Morgado, and J. Marques-Silva, "PySAT: A Python toolkit for prototyping with SAT oracles," in *SAT*, 2018, pp. 428–437.
[15] N. Eén and N. Sörensson, "An extensible sat-solver," in *SAT*, ser. LNCS, vol. 2919. Springer, 2003, pp. 502–518.
[16] O. Bailleux and Y. Boufkhad, "Efficient cnf encoding of boolean cardinality constraints," in *CP*, ser. LNCS, vol. 2833, 2003, pp. 108–122.
[17] R. Martins, S. Joshi, V. Manquinho, and I. Lynce, "Incremental cardinality constraints for maxsat," in *CP*, ser. LNCS, vol. 8656, 2014, pp. 531–548.
[18] N. Manthey, "The mergesat solver," in *SAT*, ser. LNCS, vol. 12831. Springer, 2021, pp. 387–398.

Challenges in Collective Intelligence: A Survey

Kristijan Poje†, Mario Brcic†. Mihael Kovač†, Dalibor Krleža†
†Faculty of Electrical Engineering and Computing, University of Zagreb
{kristijan.poje, mario.brcic, mihael.kovac, dalibor.krleza}@fer.hr

***Abstract*—Collective Intelligence (CI) has been gaining significant attention as an effective method for decision-making and forecasting. Prediction Markets (PMs), as a subset of CI, aim to aggregate participants' diverse opinions and knowledge to produce more accurate predictions than any individual could make alone. The unique market-based mechanism of PMs incentivizes participants to reveal their information truthfully, leading to a collectively superior prediction. However, CI and PMs have challenges, including manipulation, fallacies, and group polarization. This paper provides an overview of the challenges facing CI and PMs as tools for collective knowledge aggregation and examines the role of machine learning (ML) models as tools for amplification and hybridization in the future development of CI. Furthermore, the importance of continued research in this field is emphasized.**

***Keywords*—collective intelligence, prediction markets, artificial cognition, challenges**

I. Introduction

Collective Intelligence (CI) refers to the collective behavior of a group of individuals that results in intelligent outcomes. While the concept of CI has been around for a long time, the advent of the global network and the ability to connect people and intelligent bots has led to new and innovative developments [1].

Determining collective knowledge, i.e., reaching a consensus of the collective, is the most important part of properly functioning CI. One of the ways for a collective to reach a consensus is through Prediction Markets (PMs). PMs are defined as virtual markets created to aggregate collective thoughts that operate in a way similar to the stock market [2], [3]. The proof of the effectiveness of PMs is in their widespread usage for sharing knowledge, making a prediction, and solving problems [4], [5].

Despite the many benefits of PMs, they are not immune to challenges and limitations that can negatively impact the accuracy of the predictions. This article will provide a comprehensive overview of PMs, their advantages, and their limitations. In particular, we will delve into the various challenges PMs face, such as manipulations, fallacies, and group polarization. Through this examination, we hope to shed light on the strengths and weaknesses of PMs and how they can be improved for more accurate and reliable predictions.

In this survey paper, we provide a comprehensive overview of the challenges associated with CI from the perspective of PMs. Our main contributions include the following::

- A taxonomy of CI aspects based on interactions and types of agents, with a special focus on the role of Artificial Intelligence (AI) in CI. A clear definition of PMs and a mathematical formulation for future experiments are also discussed.
- A thorough examination of the main challenges in realizing CI and PMs, supported by a review of related literature.
- An exploration of the role of Machine Learning (ML) models and future developments in this field.

The rest of the paper is organized as follows. In Section II, we begin by presenting the definitions and taxonomy of CI. Section III highlights the significance of PMs for decision-making and provides a detailed mathematical model. In Section IV, we provide an overview of the main concepts and challenges involved in achieving CI. Section V provides a detailed analysis of the challenges in PMs, such as fallacies and group polarization. In Section VI we discuss future developments in the field. Finally, conclusions are drawn in Section VII.

II. Overview and Applications of Collective Intelligence

In this section, we first briefly overview the CI definitions and taxonomy and then explain how CI can help solve problems by aggregating peripheral knowledge.

A. Definitions and Taxonomy

CI is a multi-disciplinary concept encompassing various domains, including agent-based models, swarm optimization, artificial intelligence, and human interactions. CI can be categorized into three aspects: *cooperation*, *cognition*, and *coordination* [6]. All three aspects are depicted in Figure 1.

1) Cooperation: When discussing cooperation, primitive agents such as bees and ants acting intelligently are denoted by the term Swarm Intelligence. On the other hand, intelligent agents (i.e., humans) are denoted by the term Wisdom of the Crowds (WoC). WoC refers to a form of CI in which individuals with diverse perspectives and self-interests interact at a macro level to produce more accurate predictions than those made by a single individual [7]. A similar approach for aggregating individual predictions of multiple ML models was introduced in *ensemble learning* [8].

2) Cognition: The cognition type of CI refers to the system of agents with learning abilities. In this case,

an agent can be people and artificially intelligent bots interacting with each other [9], [10].

3) Coordination: In the coordination aspect of the CI, each individual solves a small portion of a global problem where their efforts are combined through the system design of the coordination platform [1]. A typical example of coordination is PMs, where traders, motivated by a profit, sell or buy shares in the outcome of a future event. Through the continuous process of buying/selling, prices are constantly aggregated and updated, yielding a good estimation of future event outcomes [11].

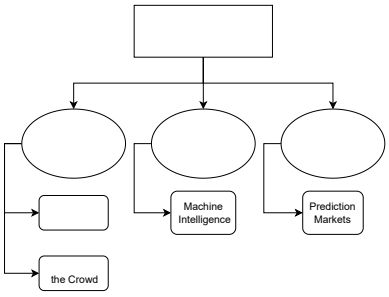


Fig. 1: Categories of Collective Intelligence

Combining the results of individual actions and behaviors achieves a collective outcome. The level of interaction between individuals and the feedback mechanisms used during the aggregation process play a critical role in determining the outcome of CI systems. In order to better understand interactions between the agent in collective, He *et al.* in [12] proposed three main paradigms of CI, namely *isolation*, *collaboration*, and *feedback*. These paradigms reflect the extent to which individuals interact and the role of feedback mechanisms in the aggregation process, thus providing a framework for exploring the properties of CI systems. The paradigms are illustrated in Figure 2.

4) Isolation: In the isolation paradigm, individuals work in isolation, without any interactions or feedback from the environment or other group members. This paradigm can be used for simple statistics (e.g., mean, median), but the results are point estimates, which do not necessarily represent the accurate output like a probability distribution.

5) Collaboration: In the collaboration paradigm, individuals engage in direct interactions and exchange of information without receiving feedback from the environment. Through trial and error, the best-performing agents and collaborative practices emerge. This evolutionary process has led to the development of efficient collaborative techniques and establishment of ethical norms in human societies [13]. In this paradigm, the level of diversity and the impact of social influence (e.g., herding effect) on the decision remains an open research question.

6) Feedback: In the feedback paradigm, individuals engage in both direct interactions with each other and indirect interactions through feedback from the environment [12]. This paradigm is beneficial for solving complex problems and analyzing complex systems, as it facilitates self-organization and synergistic problem-solving [14]. Present studies focus on feedback paradigms with simple rules due to mathematical simplicity, but in reality, such a model encapsulates complex behaviors of human interactions.

This paper focuses on CI involving intelligent agents, such as humans and intelligent bots, as described in [1] following the feedback paradigm, i.e., prediction markets.

B. Finding Solutions by Aggregation

Successful CI aggregates peripheral knowledge to find a solution. *Peripheral knowledge* can be defined as knowledge in the outsourced activity, i.e., core to the specialist that provides an activity but peripheral to the specialist that requires such activity [15]. Therefore, for successful CI following elements must be present [16]:

1) The crowd should comprise a diverse group of individuals to ensure the generation of a rich diversity of information.
2) Individual independence is crucial, both from within the group and from external sources.
3) Each group member should possess specialized knowledge to solve specific subproblems, like in the "divide and conquer" paradigm, to promote a decentralized problem-solving approach. This approach is also used in linear programming by dividing problems into substructures [17].
4) An efficient aggregation mechanism is required to combine local solutions to subproblems and deliver a comprehensive solution to the main problem.

An approach similar to the one proposed for CI has been utilized in machine learning by OpenAI. The Iterated Amplification training strategy was proposed, which incrementally builds training signals for complex problems by aggregating solutions to simpler subproblems. This approach results in a fully automated AI system capable of solving complex tasks without direct training on those tasks, as described in [18].

III. Prediction Markets

This section provides a comprehensive overview of PMs as a subfield of CI. We begin by defining PMs and then proceed to outline their mathematical formulation.

A. Definitions

Collective decision-making has been recognized as an effective method for solving complex problems and predicting future events [19]. While AI techniques have been proposed as an alternative, human intuition and creativity remain challenging to replicate, which has led to the emergence of the concept of hybrid intelligence [20]. This approach combines both intending to achieve superior performance in decision-making and problem-solving tasks.

PMs are a form of CI that leverages the collective wisdom of a group by allowing participants to buy or sell contracts that represent predictions about future events.

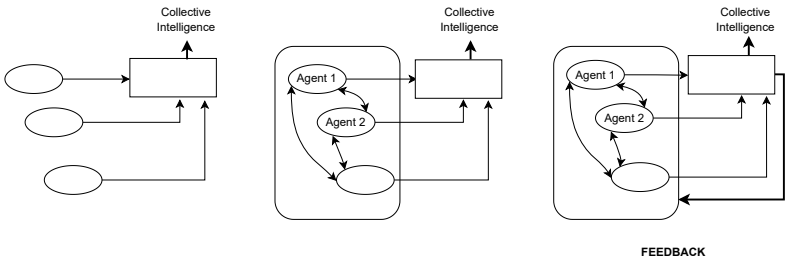


Fig. 2: Three paradigms of Collective Intelligence based on interaction: isolation, collaboration, and feedback

The contract's market price reflects the group's aggregate belief and can be used for a wide range of applications, such as election result forecasting, predicting the launch date of a new product, and more [18].

The mechanism of PMs incentivizes participants to act truthfully and share their knowledge and insights. By buying or selling shares in the outcome, participants are motivated to reveal their genuine beliefs as this strategy is expected to yield the most profit in the long run, [21].

PMs leverage the CI of a group of participants through incentivized interactions in the market. Adhering to the principle of WoC, the PMs enable the aggregation of diverse perspectives and sources of information, thus reducing biases and increasing the accuracy of predictions. This approach is similar to the Random Forest algorithm, which combines multiple decision trees to enhance predictions. [22]. Another technique, boosting algorithms, trains weak models through an iterative process to create a strong learner. Authors [23] demonstrated the effectiveness of these algorithms in creating more accurate predictive models.

Participants are incentivized to be truthful and to share their knowledge and insights by buying or selling shares in the outcome, thereby leading to an effective aggregation of the knowledge. Moreover, the PMs provide opportunities for participants to share knowledge, leading to the continuous improvement of their predictions over time [1].

B. Mathematical Model

Multiple mathematical models have been developed to describe the behavior of PMs, as demonstrated in previous studies such as [24]–[27]. However, due to the complexities inherent in social interactions, a model should encompass all aspects of PMs to provide a comprehensive representation. To model multiple agent interactions, stochastic games are often used. In light of this, we propose the use of a Partially Observable Stochastic Game (POSG) as a model for PMs with unknown environment, as previously demonstrated in [28]–[30]. PMs in POSG are defined as a tuple $\langle N, S, \{b^0\}_{i\in N}, \{A_i\}_{i\in N}, T, \{R_i\}_{i\in N}, \{O_i\}_{i\in N}, \Omega, \{I_i\}_{i\in N}\rangle$. Where

- N represents the finite set of all agents
- S represents the finite non empty set of all states – a quantity of shares being held by trading agent
- b^0 represents the initial distribution of beliefs about the current state of agent i where $b^0 \in B = \Delta(S)$
- A_i represents a final non-empty action space of agent i – a quantity of shares an agent i sells or buys
- O_i represents an observation agent i receives in state k where $k \in S$ joint observation $\overline{o} = \langle o_1, \ldots, o_{|N|}\rangle$
- $T : T(s, \overline{a}, s') = P(s'|s, \overline{a})$ represents the state transition probability of moving from state s to state s' on joint action $\overline{a} = \langle a_1, \ldots, a_{|N|}\rangle$.
- R_i represents the reward for an agent i in state k
- $\Omega : \Omega(s_k, I_{i,k}, o_{i,k}) = P(o_{i,k}|s_k, I_{i,k})$ is the observation probability of agent i receiving observation $o_{i,k}$ in state s_k when information $I_{i,k}$ is received
- $I_i = \bigcup_k I_{i,k}$ is the information agent i receives in state k

The market goes through a set of states duration of a trading period, i.e., how long the market is open. The actual market state at the time is unknown to participating agents. However, each agent has its interpretation of the state, i.e., belief probability distribution $B_{i,t} = \langle b_{1,t}, \ldots b_{|S|,t}\rangle$ where the $b_{s,t}, s \in S, t \in T$ represents the probability of market being in state s at the time t. Belief probabilities are updated using *belief update function* that takes past action past, belief state, and the observation.

Each trading agent aims to make the most profitable decision by choosing the action with the highest expected reward. This objective is often achieved through the truthful revelation of information, as highlighted in [29].

IV. Collective Intelligence Challenges

The realization of CI faces several challenges, which are primarily related to its core components: *diversity*, *independence*, *decentralization*, and *aggregation*. The importance of diversity as a component of CI was emphasized in a study of enterprises, where the authors of [31] found that diversity in answers, such as ideas and creativity, was more crucial than diversity in terms of the individuals themselves.

The challenge of promoting diversity in CI is complex, as diversity encompasses more than just demographic factors such as gender or ethnicity. According to a study [32], diversity in CI is composed of four main elements:

1) diverse perspective – representing problems in a different way
2) diverse interpretation – categorizing problems differently
3) diverse heuristic – solving problems differently

4) diverse predictive methods – a different way to infer cause and effect

Those same principles of diversity can be applied not only in human CI but in AI systems. One typical example of such is *ensamble learning*. Here, we combine diverse classifier systems that train and combine multiple learners to solve a learning problem [8].

Maintaining independence among individuals is crucial for the accuracy of CI but can be challenged by the influence of peer pressure, conformist cultures, and deference to leaders. Peer pressure can lead to group bias or polarization when individuals are asked to provide their opinions in the presence of others. Conformist cultures can result in individuals withholding their genuine opinions to avoid conflict. Deference to leaders, where individuals refrain from sharing negative information to avoid disappointing them, can also undermine the independence and limit the information available to leaders [32].

The decentralization of decision-making processes can pose a challenge, particularly in cultures where centralization is the norm. This can result in neglecting valuable insights and perspectives from individuals at the periphery, who may have different areas of expertise and bring a fresh perspective to problem-solving [32]. The decentralization of decision-making is crucial in leveraging the CI of a group, as it allows for the incorporation of diverse perspectives and expertise, leading to more creative and effective solutions.

Moreover, aggregation methods play a crucial role in CI, as they allow the conversion of individual responses into a single collective response. Currently, the available aggregation methods are limited, and more research is needed to explore different approaches and their impact on individual and group responses. Price mechanisms, as used in prediction markets, have been touted as one of the most effective ways of aggregating dispersed and asymmetrical information sources [33].

V. Prediction Market Challenges

PM results can be manipulated when individuals or groups strategically choose to provide false information, often for personal gain. This can result in biased predictions not representative of the true CI. Moreover, various fallacies can also impact the validity of PMs. Lastly, group polarization can occur when group members become more extreme in their opinions, leading to an increased likelihood of bias in predictions [7]. The challenges associated with aggregation are illustrated in Figure 3.

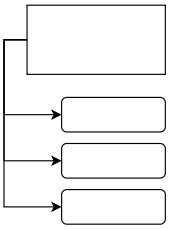


Fig. 3: Collective intelligence challenges

A. Manipulations

The validity of PMs is vulnerable to various challenges, including manipulation by malicious participants, the influence of biased individuals, and the problem of insider information. These challenges can undermine the accuracy of PMs by distorting information, altering incentives, and manipulating the aggregation mechanism.

As a result, it is critical to developing robust mechanisms for detecting and mitigating manipulation and bias in PMs. This can be achieved through technical solutions, such as secure aggregation algorithms, and social solutions, such as transparency, accountability, and community-driven moderation [34], [35]. Additionally, mapping adversarial ML techniques and their defense mechanisms [36], [37] to PMs could provide valuable insights into the systematic classification of manipulations, thereby facilitating their neutralization.

PMs possess a unique ability to self-correct when faced with potential attacks, as demonstrated in [38]. They investigated the effects of manipulations in PMs and found that the market could recover from them independently. However, it took some time for other traders to respond and counteract the manipulation, which created a window for further manipulation closer to the market closing time. This highlights the need for efficient mechanisms to detect and mitigate manipulations in PMs to improve their accuracy and validity.

Additionally, authors [24] found that PMs can be more resilient to manipulation when participants have a high stake in the outcome or when the costs of manipulating the market are high. On the other hand, when incentives are low, or manipulation costs are low, PMs can be more vulnerable. The results of this study emphasize the importance of understanding participants' incentives and designing PMs to align these incentives to obtain accurate predictions. The findings also highlight the need for robust mechanisms to detect and mitigate manipulation in PMs, particularly in cases where manipulation incentives are high and manipulation costs are low [24].

Designing the PMs to mitigate possible manipulations can be achieved through *mechanism design*. Mechanism design, as a subfield of game theory, focuses on designing rules for multi-agent systems to achieve desired outcomes through incentivizing individuals to behave in desired manner [13]. Mechanisms used in PMs are *incentive compatible* where the best strategy is to report truthfully, regardless of other agent's signals [39].

B. Fallacies

The presence of cognitive biases in event forecasting in PMs was explored by authors in a study published in [21]. They discovered that even experienced traders are susceptible to probabilistic fallacies, such as the conjunction fallacy, which can result in missed trading opportunities and affect the outcome of PMs. The authors emphasized

the need to account for these biases in PMs, as they persist despite large trading volumes.

The issue of testimonial injustice in PMs has been the subject of much concern. According to Fricker [40], testimonial injustice occurs when an individual's credibility and the value of their information are unfairly impacted by prejudice related to their social identity. This epistemic injustice can have severe consequences when aggregated in PMs [41]. It is, therefore, crucial to address this challenge to ensure the accuracy and fairness of PMs. One way to achieve that is by implementing solid ethics policies into such systems [13].

The authors in [42] examined the effect of social influence on election forecasting through voting in PMs. Their findings revealed that the influence of various voting groups could result in inaccuracies in prediction outcomes. To mitigate this challenge, the authors suggested implementing a prediction model incorporating social factors' influence.

C. Group polarization

The authors of [43] conducted a study on group polarization in PMs and found that it can significantly impact accuracy. Group polarization occurs when group members become increasingly extreme in their opinions, leading to a biased collective prediction influenced by internal or external factors. This highlights the need to consider social dynamics' impact on PMs and address this challenge to enhance their accuracy and reliability.

To mitigate the harmful effects of group polarization, researchers [44] found that larger groups tend to exhibit lower levels of polarization. Additionally, implementing sequential actions between polarized and non-polarized individuals can produce less polarized outcomes.

The impact of the order of participant entry on the outcome in PMs was explored in a study by Othman *et al.* [25]. The authors discovered that when individuals with firmly held opinions join PMs near the end of the market, they can significantly impact the prices, potentially leading to arbitrary results. However, this study only focused on the influence of changing belief values. It did not consider other factors affecting PMs outcomes, such as information exchange, intrinsic motivation, and multiple groups with extreme views [25].

VI. Future Developments and the AI

PMs can often be underlooked despite their potential to perform as well or better than hyped ML models, according to Cotton's study of PMs for the M6 Financial Forecasting contest [45]. The PMs outperformed 96.5% of ML models used by other contestants, according to [46], highlighting the value of PMs as an alternative to traditional ML models [47]. The superiority of PMs over ML models in prediction contests can be attributed to the distributed nature of human knowledge [47]. While ML models are trained on limited data, human participants in PMs have access to a broader range of information and diverse private data sources, giving them a more comprehensive understanding of the problem. This allows them to make more accurate predictions and reflect the current market conditions.

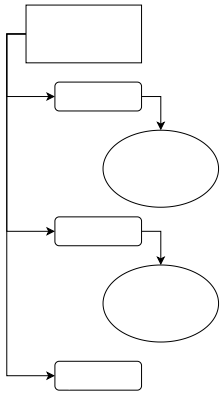


Fig. 4: Categorization of Prediction Markets challenges

Future developments in this area could include the integration of AI into PMs, creating a hybrid approach that combines human intuition and creativity with the analytical reasoning of AI. This integration would enable more accurate predictions and better decisions by leveraging the strengths of both humans and machines [20].

In addition, there is a need for research on new techniques of knowledge aggregation using AI to mitigate the effects of manipulations, fallacies, and group polarization. Such research is crucial for the sustainability and prevalence of PMs. The PMs could be further enhanced by developing techniques that can reduce the negative effects of these challenges. Therefore, future studies should focus on these areas to improve the effectiveness and reliability of PMs. In this regard, utilizing mitigation techniques from adversarial ML attacks to prevent manipulations in PMs, such as those used in [36], [37], might be a promising research direction.

VII. Conclusion

This paper explored the potential challenges faced in implementing Collective Intelligence (CI) through Prediction Markets (PMs). PMs offer a valuable tool for collective decision-making, as participants are incentivized to report accurate information, leading to more reliable predictions. However, like all implementations of CI, PMs are subject to potential manipulations, biases, and polarization due to human nature. Despite these limitations, PMs have been shown to outperform Machine Learning (ML) models due to the broader access to information [47] and human intuition and creativity that severely lacks in analytically oriented Artificial Intelligence (AI). To further enhance the capabilities of CI, integrating human input with ML models to address limitations inherent in human nature could be a promising direction for future research and development.

REFERENCES

[1] T. W. Malone and M. S. Bernstein, *Handbook of Collective Intelligence*. MIT Press, Oct 2015.

[2] R. Palak and N. T. Nguyen, "Prediction markets as a vital part of collective intelligence," in *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. Gdynia, Poland: IEEE, Jul 2017, p. 137–142.

[3] F. A. Hayek, "The use of knowledge in society," in *Knowledge management and organizational design*. Routledge, 2009, pp. 7–15.

[4] F. Petropoulos, D. Apiletti, and et. al., "Forecasting: theory and practice," *International Journal of Forecasting*, vol. 38, no. 3, p. 705–871, Jul 2022.

[5] C. R. Plott and K.-Y. Chen, "Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem," *authors.library.caltech.edu*, 2002.

[6] C. Yu, Y. Chai, and Y. Liu, "Literature review on collective intelligence: a crowd science perspective," *International Journal of Crowd Science*, vol. 2, no. 1, p. 64–73, Jul 2018.

[7] J. Surowiecki, *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*, 1st ed. New York: Doubleday, 2004.

[8] Z.-H. Zhou, *Ensemble Learning*. Singapore: Springer Singapore, 2021, p. 181–210.

[9] R. M. Livingstone, "Models for understanding collective intelligence on wikipedia," *Social Science Computer Review*, vol. 34, no. 4, p. 497–508, Aug 2016.

[10] G. Mulgan, "Artificial intelligence and collective intelligence: the emergence of a new field," *AI & SOCIETY*, vol. 33, no. 4, p. 631–632, Nov 2018.

[11] P. Atanasov, P. Rescober, E. Stone, S. A. Swift, E. Servan-Schreiber, P. E. Tetlock, L. Ungar, and B. Mellers, "Distilling the wisdom of crowds: Prediction markets versus prediction polls," *Academy of Management Proceedings*, vol. 2015, no. 1, p. 15192, Jan 2015.

[12] F. He, Y. Pan, Q. Lin, X. Miao, and Z. Chen, "Collective intelligence: A taxonomy and survey," *IEEE Access*, vol. 7, p. 170213–170225, 2019.

[13] S. Isufi, K. Poje, I. Vukobratovic, and M. Brcic, "Prismal view of ethics," *Philosophies*, vol. 7, no. 6, p. 134, Nov 2022.

[14] H. Haken, *Synergetic Computers and Cognition*, ser. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, vol. 50.

[15] A. Tiwana and M. Keil, "Does peripheral knowledge complement control? an empirical test in technology outsourcing alliances," *Strategic Management Journal*, vol. 28, no. 6, p. 623–634, Jun 2007.

[16] A. W. Lo and R. Zhang, "The wisdom of crowds vs. the madness of mobs: An evolutionary model of bias, polarization, and other challenges to collective intelligence," *SSRN Electronic Journal*, 2021.

[17] R. Bellman, *Dynamic programming*. Princeton, NJ: Princeton Univ. Pr, 1984.

[18] P. Christiano, B. Shlegeris, and D. Amodei, "Supervising strong learners by amplifying weak experts," *arXiv*, 2018.

[19] K. A. McHugh, F. J. Yammarino, S. D. Dionne, A. Serban, H. Sayama, and S. Chatterjee, "Collective decision making, leadership, and collective intelligence: Tests with agent-based simulations and a field study," *The Leadership Quarterly*, vol. 27, no. 2, p. 218–241, Apr 2016.

[20] D. Dellermann, N. Lipusch, P. Ebel, K. M. Popp, and J. M. Leimeister, "Finding the unicorn: Predicting early stage startup success through a hybrid intelligence method," *arXiv*, 2021.

[21] H. C. B. Lee, J. Stallaert, and M. Fan, "Anomalies in probability estimates for event forecasting on prediction markets," *Production and Operations Management*, vol. 29, no. 9, p. 2077–2095, 2020.

[22] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, p. 5–32, Oct 2001.

[23] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer, "An efficient boosting algorithm for combining preferences," *Journal of machine learning research*, vol. 4, no. Nov, pp. 933–969, 2003.

[24] M. Chakraborty and S. Das, "Trading on a rigged game: outcome manipulation in prediction markets," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI'16. New York, New York, USA: AAAI Press, Jul 2016, p. 158–164.

[25] A. Othman and T. Sandholm, "When do markets with simple agents fail?" in *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, ser. AAMAS '10. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, May 2010, p. 865–872.

[26] Y. Chen and D. M. Pennock, "Designing markets for prediction," *AI Magazine*, vol. 31, no. 4, p. 42, Dec 2010.

[27] A. Haider, H. Wang, B. Scotney, and G. Hawe, "Predictive market making via machine learning," *Operations Research Forum*, vol. 3, no. 1, p. 5, Mar 2022.

[28] J. Jumadinova and P. Dasgupta, "Automated pricing in a multiagent prediction market using a partially observable stochastic game," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, no. 4, p. 1–25, Aug 2015.

[29] J. Jumadinova and P. Dasgupta, "A multi-agent system for analyzing the effect of information on prediction markets," *International Journal of Intelligent Systems*, vol. 26, no. 5, p. 383–409, May 2011.

[30] J. Jumadinova and P. Dasgupta, "A multi-agent prediction market based on partially observable stochastic game," in *Proceedings of the 13th International Conference on Electronic Commerce*. Liverpool United Kingdom: ACM, Aug 2011, p. 1–10.

[31] M. Hosseini, J. Moore, M. Almaliki, A. Shahri, K. Phalp, and R. Ali, "Wisdom of the crowd within enterprises: Practices and challenges," *Computer Networks*, vol. 90, p. 121–132, Oct 2015.

[32] K. Matzler, A. Strobl, and F. Bailom, "Leadership and the wisdom of crowds: how to tap into the collective intelligence of an organization," *Strategy & Leadership*, vol. 44, no. 1, p. 30–35, Jan 2016.

[33] P. S. Myers, *Knowledge management and organizational design*. Boston, Mass.: Butterworth-Heinemann, 1996.

[34] B. G. Malkiel, "The efficient market hypothesis and its critics," *Journal of Economic Perspectives*, vol. 17, no. 1, p. 59–82, Feb 2003.

[35] S. Luckner, J. Schröder, C. Slamka, M. Franke, A. Geyer-Schulz, B. Skiera, M. Spann, and C. Weinhardt, *Prediction markets: fundamentals, designs and applications*, 1st ed., ser. Gabler research. Wiesbaden: Gabler, 2012.

[36] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, "A survey on adversarial attacks and defences," *CAAI Transactions on Intelligence Technology*, vol. 6, no. 1, p. 25–45, Mar 2021.

[37] M. Juric, A. Sandic, and M. Brcic, "AI safety: state of the field through quantitative lens," in *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2020, pp. 1254–1259.

[38] P. Buckley and F. O'Brien, "The effect of malicious manipulations on prediction market accuracy," *Information Systems Frontiers*, vol. 19, no. 3, p. 611–623, Jun 2017.

[39] V. Conitzer, "Prediction markets, mechanism design, and cooperative game theory," *arXiv preprint arXiv:1205.2654*, 2012.

[40] M. Fricker, *Testimonial Injustice*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Mar 2019, p. 149–163.

[41] C. D. Mildenberger, "Testimonial injustice and prediction markets," *Social Epistemology*, vol. 36, no. 3, p. 378–392, May 2022.

[42] Y. Zhao and E. Santos, "A failure of collective intelligence," in *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Dec 2018, p. 361–366.

[43] C. R. Sunstein, "The law of group polarization," *Journal of Political Philosophy*, vol. 10, no. 2, p. 175–195, Jun 2002.

[44] G. Iyer and H. Yoganarasimhan, "Strategic polarization in group interactions," *Journal of Marketing Research*, vol. 58, no. 4, p. 782–800, Aug 2021.

[45] "The M6 Financial Forecasting Competition," 2023. [Online]. Available: https://m6competition.com/

[46] Microprediction, "The options market beat 94% of participants in the m6 financial forecasting contest," https://medium.com/geekculture/the-options-market-beat-94-of-participants-in-the-m6-financial-forecasting-contest-fa4f47f57d33, Feb 2023.

[47] P. Cotton, *Microprediction: building an open ai network*. Cambridge, Massachusetts: The MIT Press, 2022.

The Study of the Target Set Selection Problem under Deterministic Linear Threshold Model Using Evolutionary Algorithms

Mikhail Smirnov, Stepan Kochemazov, Alexander Semenov
ITMO University, St Petersburg, Russia
Emails: masmirnov.inf@gmail.com, stepan.kochemazov@itmo.ru, alex.a.semenov@itmo.ru

***Abstract*—In this paper we study the Target Set Selection problem (TSS) – the well-known combinatorial problem associated with collective behaviour in networks. For a given network graph and the information diffusion model that specifies a dynamic process of how vertices activate each other, TSS aims to identify a set of initially active vertices of minimum cardinality such that they manage to activate all the other network vertices. This problem is computationally hard but has numerous applications in practice, thus the development of computational algorithms for TSS is a relevant topic. A surprising fact is that the spectrum of algorithms for TSS presented so far in the literature is quite limited. The main novelty of our study consists in the hybrid approach which combines evolutionary and greedy algorithms for TSS. In more detail, we view TSS as a pseudo-Boolean optimization problem and solve it using simple evolutionary algorithms (in fact several variants of (1+1)-EA). In order to identify a good initial approximate solution we employ a greedy algorithm. In the experiments we apply the developed algorithms to TSS in the context of Deterministic Linear Threshold Model (DLTM) and demonstrate that the presented approach shows good effectiveness when solving TSS on networks with dozens thousands of vertices on personal computer.**

***Keywords*—Boolean satisfiability problem; Target Set Selection; Information Spread**

I. Introduction

Network analysis is one of the most actively developing areas of modern computer science. It is sufficient to say that the references to key papers on random network models e.g. [1], [2] number in tens of thousands. The present paper studies the phenomenon of information spread in networks and the ways how one can solve combinatorial problems associated with this phenomenon. In particular, we consider the Target Set Selection (TSS) problem [3], [4], which is tightly connected to the Influence Maximization Problem (IMP), which in turn was studied in the widely-cited paper [5].

Both TSS and IMP study the spread of activation in networks, where the vertices represent agents conforming to certain behavioral rules. These rules specify the information diffusion process: how active vertices can activate other vertices at the following time moments. The goal of IMP is for a given network and information diffusion model to find a subset of vertices such that if these vertices are active at the initial time moment, the number of vertices at the end of information diffusion process (the so-called *influence spread*) is maximized. Meanwhile, TSS aims to find a subset of vertices of minimum cardinality such that it activates all vertices in a network.

We consider TSS in the context of the so-called Deterministic Linear Threshold Model (DLTM) (see e.g. [5]), with regards to which from the one hand there are known results about its intractability [3], [4], [5], on the other, for relatively small dimensions it can be solved exactly using some combinatorial algorithms (such as the algorithms for solving the Boolean satisfiability problem (SAT) of Integer Linear Programming problem (ILP)).

The main contributions of the paper are as follows. First, we present a natural formulation of TSS as a pseudo-Boolean black-box optimization problem. Second, we combine evolutionary algorithms for pseudo-Boolean optimization with a greedy algorithm from [6]: the latter is used to find a reasonable starting point for the optimization process. Third, in the computational experiments we show the viability of the proposed approach since it makes it possible to solve TSS under DLTM for networks with tens of thousands vertices on a usual PC.

II. Preliminaries

Hereinafter, we consider the representation of some collective $V = \{v_1, \ldots, v_n\}$ via a directed graph $G = (V, A)$ to which we further refer as to *network*. Here V is the set of vertices of G and A is the set of arcs of this graph. For an arbitrary vertex $v \in V$ let us define its *neighborhood* as a set $U_v \subseteq V \setminus \{v\}$ comprised of all vertices u such that $(u, v) \in A$, i.e. the corresponding arc is directed from u to v. For simplicity, we assume that G does not contain loops (however, it is possible to transfer the main results of the paper to take loops into account). Let us associate with network G discrete time moments $t \in \{0, 1, 2, \ldots\}$. With each vertex $v \in V$ associate the function $a_v(\cdot)$ the value of which at time moment $t + 1$ ($t \geq 0$) is defined as the function of values of functions a_u for vertices $u \in U_v$ computed at time moment t:

$$a_v(t+1) = \begin{cases} 1, & \sum_{u \in U_v} a_u(t) \times w_{(u,v)} \geq \theta_v \\ a_v(t), & otherwise \end{cases} \quad (1)$$

In (1) $w_{(u,v)}$ is some weight (which is usually specified by a positive rational number), and θ_v is a positive rational number called *threshold* of vertex v. In addition to that, we assume that

at time moment $t = 0$ the values of all functions of the kind (1) are the numbers from the set $\{0, 1\}$ and that the weights $w_{(u,v)}$ and thresholds θ_v are specified at time moment $t = 0$ and do not change with t.

By $\{0, 1\}^n$ denote the set of all possible binary vectors of length n. It is clear, that for an arbitrary $t \in \{0, 1, 2, \ldots\}$ the values of functions (1) for all vertices of network G can be specified by a binary vector $\alpha_t \in \{0, 1\}^n$. Let us refer to this vector as to a *state* of a Discrete Dynamic System (DDS) specified by network G (or simply the state of G) at time moment t. The described rules specify a function of the kind:

$$F_G : \{0, 1\}^n \to \{0, 1\}^n \tag{2}$$

which defines the transitions between the states of G. Let us refer to state α_0 of G at time moment $t = 0$ as to *initial* state.

It is clear that for any initial state α_0 there exists some t^* such that for all $t > t^*$ it holds that $\alpha_t = \alpha_{t^*}$. Any vector $\alpha \in \{0, 1\}^n$ such that $F_G(\alpha) = \alpha$ is called a *fixed point* of F_G. For a considered DDS it is easy to see that starting from any initial state $\alpha_0 \in \{0, 1\}^n$ the corresponding system will reach a fixed point in at most n time moments.

If there are specified functions of the kind (1) for network G and in addition all the rules mentioned above are satisfied, then let us say that the considered DDS is specified under *Deterministic Linear Threshold Model* (DLTM).

We will call vertex v active at time moment t, if the coordinate of vector α_t which corresponds to vertex v is equal to 1. A *Target Set* (TS) V', $V' \subseteq V$ is a set of vertices which are active in the initial state α_0. Let us define V' by means of some vector α_0 and compute the values of functions (1) for some number s of time moments, $s \leq n$. For every vertex V which is active at time moment t, $t \leq s$ let us say that v was activated by target set V'. Note, that if α_t is not a fixed point, then the number of active vertices at time moment $t + 1$ is greater than that at time moment t. Therefore, if at some moment t the number of active vertices is deemed insufficiently small, it makes sense to consider the state α_{t+1}.

The *Target Set Selection* problem (TSS) is stated as follows: for the predefined number R, $1 \leq R \leq n$ find the smallest target set V' (smallest TS) such that it activates (in the sense above) at least R vertices in a network.

As it is noted in [3] (with reference to [5]) in the formulation described above TSS is NP-hard. Also, based on the results of [5], TSS can not be approximated within any constant factor under assumption that $P \neq NP$. Moreover, TSS is $W[P]$-complete [3] and thus it is fixed-parameter intractable if it is parameterized by the size of the minimum target set. However, TSS can be effectively (in polynomial time in the size of the description of G) reduced to combinatorial problems such as SAT or ILP using quite natural reductions, presented e.g. in [7], [8], [9].

III. TSS as a Pseudo-Boolean Optimization Problem

As we noted above, it is possible to solve both IMP and TSS under DLTM exactly, for example using SAT or ILP, as well as approximately (however, without any guarantees on the approximation ratio). In the present paper we practice the second approach. In particular, we reduce TSS to the pseudo-Boolean optimization problem which we then solve using evolutionary algorithms.

A pseudo-Boolean function (see e.g. [10]) is a function of the following kind:

$$\Phi : \{0, 1\}^n \to \mathbf{R}. \tag{3}$$

The function in equation (3) is not required to be defined analytically. Moreover, it is possible that the values of Φ are produced by some algorithm, the complexity of which is ignored (essentially, it is treated as an oracle). In such case (3) is referred to as a *black-box function*. If (3) is a black-box function, then one can use only metaheuristic algorithms to optimize it [11]. In this situation it is treated as the so-called *fitness function*.

Note, that when solving TSS the exact methods from [10] or [9] are applicable only to networks with relatively small number of vertices (up to several hundred), since the increase of this number leads to the increase of the encoding size and consequently to longer runtimes of the employed combinatorial algorithms. However, for metaheuristic algorithms this fact is inconsequential, since the optimized function can be computed fast.

The search space in the case of TSS is represented by a Boolean hypercube $\{0, 1\}^n$. An arbitrary $\lambda \in \{0, 1\}^n$ specifies a particular target set: the ones in λ correspond to vertices in G which are active at time moment $t = 0$. The fitness function for TSS is the function:

$$\Phi : \{0, 1\}^n \to \{1, \ldots, n\}, \tag{4}$$

which is defined via function (2) as follows: for a particular $\lambda \in \{0, 1\}^n$ first construct the target set represented by the vector corresponding to the initial state $\alpha_0 = \lambda$ of network G. Next, compute the functions

$$\begin{gathered} F_G^1(\lambda) = F_G(\lambda), \\ F_G^2(\lambda) = F_G\left(F_G^1(\lambda)\right), \\ \ldots, \\ F_G^k(\lambda) = F_G\left(F_G^{k-1}(\lambda)\right), \end{gathered} \tag{5}$$

where $F_G^k(\lambda) = \gamma$ is either a fixed point of function F_G or $wt(\gamma) \geq \varepsilon \cdot |V|$ ($wt(\cdot)$ is the function computing the Hamming weight), and $\varepsilon \in (0, 1)$ specifies the portion of vertices that need to be active in the original TSS formulation. As we already mentioned above, in order to enter a fixed point it is sufficient to make at most n calculations of the value of function F_G for any TS λ. The value of Φ is defined as the Hamming weight of vector λ, specifying TSS, in the case if $wt(\gamma) \geq \varepsilon \cdot |V|$. Otherwise, if $wt(\gamma) < \varepsilon \cdot |V|$ and $\gamma = F_G^k(\lambda)$ is a fixed point of F_G then the value of (4) in λ is viewed as undefined, and the fitness function takes the maximal value n.

It is possible to use any metaheuristic algorithm to optimize the function (4) defined above. In the next section we will briefly describe the main algorithms that we will apply to TSS in the formulation above.

IV. Metaheuristic Algorithms Used to Solve TSS

One of the simplest possible metaheuristic algorithms that can be used to minimize (4) is the well-known (1+1)-Evolutionary Algorithm ((1+1)-EA). Apparently, it was first described in [12]. The elementary step in (1+1)-EA is the so-called random (1+1)-mutation: with an arbitrary vector $\lambda \in \{0,1\}^n$ we associate n random independent Bernoulli trials with probability of success (which is referred to as *mutation rate*) $p = 1/n$. If the trial number $i \in \{1,\ldots,n\}$ is successful, then the bit number i in vector λ is flipped (e.g. changed into the complementary bit). Let λ' be the vector resulting from mutation. If $\Phi(\lambda') \leq \Phi(\lambda)$ (in the case of the minimization problem) then move to the point λ' and apply the next random mutation to λ', otherwise mutate λ again (this situation is called *stagnation*). We always store the Best Known Value (BKV) of the considered fitness function.

The (1+1)-EA algorithm despite its simplicity has a lot of nontrivial mathematical properties [13]. In theory, the algorithm is extremely inefficient: in [13] it was shown that the upper bound on its complexity is n^n in the sense of the expected value of the number of random mutations performed to reach a global extremum, and thus in the worst-case scenario the algorithm is even worse than the simple random search. However, in practice (1+1)-EA can show surprisingly good effectiveness. It follows from the fact that on average only a single bit in λ is flipped as a result of random mutation: in more detail, the expected value of the number of flipped bits in λ is 1, when the mutation rate is $p = 1/n$. This fact means that on average (1+1)-EA acts similar to the Hill Climbing algorithm [14], and thus can take into account the features of the landscape of a considered function (e.g. it can make use of the convex areas).

A number of works proposed different approaches to the modification of (1+1)-EA aimed at reducing its worst-case complexity. One of the best approaches of this kind is the (1+1)-Fast Evolutionary Algorithm with parameter β ((1+1)-FEA$_\beta$), described in [15]. Its main idea lies in the use of the variable mutation rate: in particular, before mutating each $\lambda \in \{0,1\}^n$ a random variable μ is observed, the spectrum of which is $S_\mu = \{1,2,\ldots,n/2\}$, and the distribution (the so-called "power-law distribution $D^\beta_{n/2}$") is defined as follows

$$\Pr\{\mu = k, k \in S_\mu\} = \left(C^\beta_{n/2}\right)^{-1} \cdot \frac{1}{k^\beta} \tag{6}$$

In (6) β, $\beta > 1$ is the parameter of the algorithm, $C^\beta_{n/2}$ is a normalizing constant: $C^\beta_{n/2} = \sum_{i=1}^{n/2} i^{-\beta}$ (it is necessary for the normalizing probabilities). A single mutation in the context of (1+1)-FEA$_\beta$ looks as follows: first generate a value $\mu \in \{1,\ldots,n/2\}$ in accordance with the distribution $D^\beta_{n/2}$, then apply the random mutation with mutation rate $p = \mu/n$ to the current $\lambda \in \{0,1\}^n$.

The upper bound on the complexity of (1+1)-FEA$_\beta$ is $O\left(C^\beta_{n/2} \cdot n^\beta \cdot 2^n\right)$, which asymptotically is significantly better than that for (1+1)-EA. However, the expected value of flipped bits of (1+1)-FEA$_\beta$ is a constant independent of n only when $\beta > 2$. For example in the case if $\beta = 3$ (see [14]) we have that $E[d_H(\lambda,\lambda')] \approx 1.3685\ldots$, where $d_H(\lambda,\lambda')$ is the Hamming distance between vectors λ and λ').

In paper [16] the switching variant of (1+1)-EA, denoted as (1+1)-SEA$_\delta$ (from Switched Evolutionary Algorithm) was proposed. It simplifies the original idea underlying (1+1)-FEA$_\beta$. This algorithm also employs variable mutation rate, but the switching is performed between two alternatives: the rate of $p = 1/n$ is chosen with probability $1 - \delta/n$, and the rate $p = 1/2$ is chosen with probability δ/n, for parameter $\delta \in (0,1)$. It is easy to show (similar to how it was done in [14]) that for any $\delta \in (0,1)$ the upper bound on the complexity of (1+1)-SEA$_\delta$ of the following kind holds: $O(n \cdot 2^n)$. For the expected value of the number of flipped bits in the case of (1+1)-SEA$_\delta$ we have (see [15]): $E[d_H(\lambda,\lambda')] = 1 + \delta\left(\frac{1}{2} - \frac{1}{n}\right)$. Thus, e.g. for $\delta = 0.5$ the following estimation holds: $E[d_h(\lambda,\lambda')] = 1.25 - o(1)$. Thus, at least theoretically, (1+1)-SEA$_\delta$ looks better than (1+1)-FEA$_\beta$.

Apart from the evolutionary metaheuristics described above we applied to solving TSS the greedy algorithm which is a slightly modified variant of the algorithm from [6]. We also used evolutionary algorithms to start from the initial solution found by the greedy algorithm. The essence of the latter is as follows. Consider network G with functions of the kind (1) associated with vertices. At an arbitrary time moment with vertex v associate two values: I_v which is called the *influence* of vertex v and $\tilde{\theta}_v$ called *residual threshold*. The residual threshold is computed as follows:

$$\tilde{\theta_v} = \theta_v - \sum_{u \in U_v} a_u \times w_{(u,v)}$$

It is clear that if at some time moment $\tilde{\theta_v} \leq 0$ (while at the previous moment it took place that $\tilde{\theta_v} > 0$) then the vertex v becomes active at this time moment.

Assume, that DDS G starts from some TS and goes into a fixed point, and at the same time the constraint on the number of active vertices is not satisfied. In this case we need to choose the next vertex to be included into TS. In accordance with the greedy strategy, the vertex with the greatest influence is added. To compute the influence I_v of vertex v it makes sense to take into account both the number of vertices that can be made active by adding v to TS and how adding it to TS reduced the residual thresholds of the other vertices. Thus, as the measure of influence of vertex v we consider the following value:

$$I_v = |A_v| + \sum_{w \in W} \frac{\Delta\theta_w}{\theta_w} \tag{7}$$

The value of (7) is computed as follows. Assume that before adding v to TS the system G is in the fixed point λ. We add v to TS and recompute the values of functions (1) for all network vertices until we enter the next fixed point λ'. Denote by A_v the set of vertices that became active and by W the set of vertices the residual thresholds of which changed during the transition from λ to λ' (the situation $\lambda = \lambda'$ is allowed).

TABLE I
EXPERIMENTAL RESULTS FOR SMALL NETWORKS. EVOLUTIONARY ALGORITHMS WERE ALLOWED TO PERFORM UP TO 10^5 ITERATIONS. THE RESULTS OF NONDETERMINISTIC ALGORITHMS ALGORITHMS ARE AVERAGED OVER 10 LAUNCHES AND ARE PRESENTED IN CURSIVE.

	exact-sat (Glucose3)	greedy	(1+1)EA	greedy&(1+1)EA
WS(50,8,0.5)$_{ft}$	18 (4093 s)	26 (0.001 s)	*18.3 (9.93 s)*	*18.3 (10.084 s)*
BA(50,4)$_{ft}$	14 (1575 s)	17 (0.007 s)	*14.2 (10.111 s)*	*14 (10.167 s)*
WS(50,8,0.5)$_{rt}$	7 (3594 s)	11 (0.008 s)	*7.4 (11.066 s)*	*7.3 (11.119 s)*
BA(50,4)$_{rt}$	6 (1614 s)	6 (0.002 s)	*6 (10.973 s)*	*6 (11.145 s)*

By $\Delta\theta_w$ ($\Delta\theta_w \geq 0$) denote the value by which the residual threshold of w is reduced during $\lambda \to \lambda'$.

After checking all inactive vertices in λ we add to TS the vertex v with maximum influence I_v.

V. COMPUTATIONAL EXPERIMENTS

In computational experiments we considered both random networks and fragments of real-world networks. The random networks were constructed in accordance with Watts-Strogatz [1] and Barabasi-Albert [2] models. As for the real-world networks: in particular, we used the 'fb' and 'wiki' networks from the SNAP database [1]. The latter represent the fragments of Facebook (ego-Facebook) and Wikipedia (wiki-vote) respectively. In the original formulation some of the considered graphs (e.g. Watts-Strogatz and Barabasi-Albert) are undirected, thus, we replaced each edge by a pair of arcs with opposite directions, connecting the respective vertices. In all experiments the arc weights were generated as natural numbers chosen randomly and uniformly from $[1, 100]$. As for vertices' thresholds, we considered both the variants where the threshold was specified by a fixed number, as well as ones in which thresholds were generated independently and randomly. In the second case the threshold of an arbitrary vertex v was calculated as a natural number $\theta_v = \lceil \varepsilon \times (\sum_{u \in U_v} w_{u,v}) \rceil$ for ε randomly and uniformly chosen from $[1/2, 1]$. For the fixed thresholds we used the value of $\varepsilon = 3/4$.

All experiments were carried out on a PC with AMD Ryzen 3950x CPU (16 cores) and 64 GB RAM under Ubuntu OS. The main object of our study are the evolutionary and greedy algorithms for solving TSS that were described above.

For random graphs of small size (several dozen vertices) we checked how the solutions found by metaheuristic algorithms are close to the exact ones. In order to obtain the exact solutions of TSS we employed the algorithms for solving the Boolean satisfiability problem (SAT) [17], which are actively applied today to diverse combinatorial problems. To encode to SAT the problems of finding TSS of size K (i.e. containing K vertices) we employed the methods that were described in detail in [7]. In addition to that we used the SAT encoding techniques for pseudo-Boolean constraints of the kind $\sum_{u \in U_v} a_u(t) \times w_{(u,v)} \geq \theta_v$ from the PySAT toolkit [18]. It should be noted that in [7] only the SAT encodings for the cardinality constraints (e.g. if $w_{(u,v)} \in \{0, 1\}$) were used. To find the smallest K we used the dichotomous scheme that employed the Glucose 3 SAT solver.

The results of these experiments are presented in Table I. In particular, in this table we consider 4 networks with 50 vertices each, generated in accordance with Watts-Strogatz (WS) (with parameters $k = 8$, $\beta = 0.5$) and Barabasi-Albert (BA) (with parameter $m = 4$). We considered both the cases with the fixed threshold for all vertices (e.g. WS(50,8,0.5)$_{ft}$) and

[1] http://snap.stanford.edu

TABLE II
EXPERIMENTAL RESULTS FOR LARGE NETWORKS. EVOLUTIONARY ALGORITHMS WERE ALLOWED TO PERFORM UP TO 10^4 ITERATIONS. THE RESULTS OF NONDETERMINISTIC ALGORITHMS ALGORITHMS ARE AVERAGED OVER 10 LAUNCHES AND ARE PRESENTED IN CURSIVE.

	greedy	(1+1)EA	greedy&(1+1)EA	greedy&(1+1)FEA $\beta = 2.5$	greedy&(1+1)FEA $\beta = 3$	greedy&(1+1)SEA $\delta = 0.5$	greedy&(1+1)SEA $\delta = 0.25$
fb	537 (18.2 s)	*968.7 (320.8 s)*	*437.8 (404.8 s)*	*438.3 (380.6 s)*	*437.1 (380.6 s)*	*440.4 (375 s)*	*438.6 (376 s)*
wiki	3089 (38.5 s)	*3716.6 (220.1 s)*	*3004.7 (347.8 s)*	*3001.7 (320.5 s)*	*3002.5 (321.9 s)*	*3004.6 (319.6 s)*	*3005.3 (319.2 s)*
WS(10000,20,0.5)	2278 (117.6 s)	*3909.7 (353.6 s)*	*2005.3 (722.2 s)*	*1991.0 (658.4 s)*	*1993.6 (661.2 s)*	*2000.0 (667.7 s)*	*2006.1 (667.7 s)*
BA(10000,20)	1103 (341.9 s)	*3891.1 (674.4 s)*	*1073.9 (1356.8 s)*	*1073.1 (1275.6 s)*	*1073.8 (1281.5 s)*	*1072.1 (1274.3 s)*	*1072.0 (1284.7 s)*

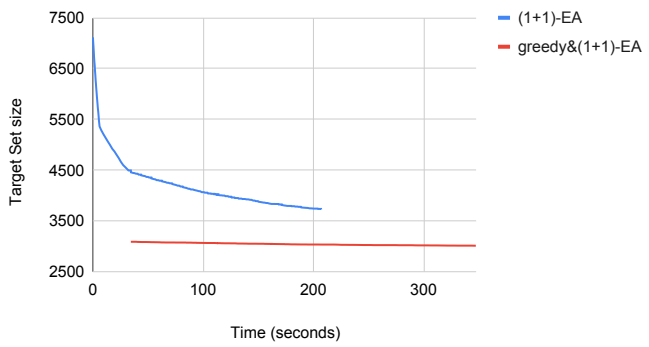


Fig. 1. Convergence plot for a single launch of (1+1)-EA (blue) and of greedy & (1+1)-EA (red) for 10000 iterations on the WIKI problem

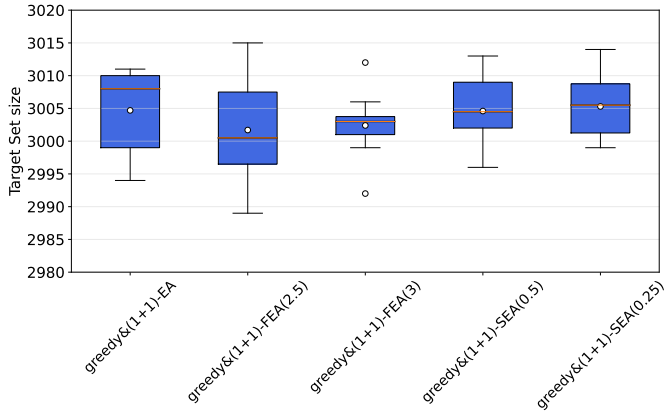


Fig. 2. Box plots of the size of the target set when solving TSS for the WIKI instance using different hybrid algorithms

with thresholds generated randomly according to the scheme outlined above (e.g. BA(50,4)$_{rt}$). The cells of the table present the size of the found TS and the runtime of the employed algorithm. Each evolutionary algorithm in each launch made 10^5 random mutations. As it was noted above, the SAT solver finds an exact solution, but as it can be seen, even for such small networks, it takes a lot of time. The greedy algorithm works very fast, but the found solutions are not always close to the exact one. As for the (1+1)-Evolutionary Algorithm, it performed 10^5 random mutations sufficiently faster than the SAT solver and yields the solutions that are close to the exact ones. The best results have been obtained using the mixed strategy, when the greedy algorithm is used to search for an initial solution which is later improved by using (1+1)-EA: in almost all launches this strategy managed to find an exact solution.

In the next series of experiments we considered the networks with several thousand vertices and employed metaheuristic strategies described above, as well as the hybrid strategy combining greedy and evolutionary algorithms, to solve TSS on these networks. Note, that the SAT solvers can not tackle TSS for networks of such dimension.

In Table II we present the benchmarks 'fb' (4039 vertices) and 'wiki' (7115 vertices) mentioned above. We also considered Watts-Strogatz networks (with $k = 20$ and $\beta = 0.5$) and Barabasi-Albert networks (with $m = 20$) with 10 000 vertices. In all cases the thresholds were generated randomly in accordance with the scheme outlined above. Then we applied to them the greedy algorithm as well as several variants of (1+1)-EA, described above. In each launch each evolutionary algorithm was allowed to perform 10^4 random mutations. What's interesting is that all variants of (1+1)-EA showed more or less the same performance (in the sense of improvement of the found TS). Also note, that the hybrid strategy yielded the best results in the sense of the size of the found TS.

In order to study in detail the behavior of the proposed hybrid algorithm we considered the TSS problem for the WIKI instance in the same configuration as in Table II. We measured the pairs (time, current size of TS) before the first iteration and then after each 100 iterations of the evolutionary algorithm and plotted them at Figure 1. Thus, both red and blue lines have 10000 / 100 = 100 segments. The red plot corresponding to the hybrid greedy&(1+1)-EA algorithm does not start at 0 because the corresponding time is required to construct an initial solution using the greedy algorithm. Also, despite the fact that it seems to not undergo any change, in fact, it steadily decreases from 3089 to 3002.

To analyze the behavior of different variants of hybrid algorithms on average, we performed 20 launches of each algorithm on the same WIKI instance and summed the results of these experiments in the boxplots at Figure 2. Each box represents an interval from 1/4 to 3/4 quartile. The point inside the box is the mean, the line inside corresponds to the median. The whiskers have the length of 1.5 interquartile distance, but since there are only outliers for (1+1)-FEA, in most cases the whiskers just show the interval from the minimum to the maximum. We also performed the Wilcoxon Rank Sum Test between the standard greedy&(1+1)-EA and its four modifications (i.e. 4 tests in total). The hypothesis in each test was that "the target set size in a modified hybrid strategy is smaller than the target set size in greedy&(1+1)-EA" with the significance level $= 0.05$. In all 4 cases the hypothesis was not confirmed.

VI. Conclusion and Future Works

In this paper we study the well-known Target Set Selection (TSS) problem which is associated with a more general problem of influence maximization [5]. We show that the problem of finding the smallest TSS can be viewed in the context of the general pseudo-Boolean optimization problem. To solve the latter we employ metaheuristic algorithms: greedy algorithm (which is a slightly modified algorithm from [6]) as well as several variants of (1+1)- Evolutionary Algorithm. In the role of benchmarks we use both random graphs (Watts-Strogatz, Barabasi-Albert), and fragments of real-world networks. For such graphs over several thousand vertices (up to 10 000) we demonstrate that the best performance is achieved by

the hybrid strategy, in which the greedy algorithm is used to find an initial approximation which is later improved by the evolutionary algorithms. Our plans for the nearest future include the development of parallel variants of evolutionary and genetic algorithms for TSS.

ACKNOWLEDGMENTS

We thank anonymous reviewers for their valuable feedback that made it possible to significantly improve the quality of this paper.

This work was supported by the Ministry of Science and Higher Education of Russian Federation, research project no. 075-03-2020-139/2 (goszadanie no. 2019-1339).

REFERENCES

[1] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, 1998.

[2] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[3] O. Ben-Zwi, D. Hermelin, D. Lokshtanov, and I. Newman, "Treewidth governs the complexity of target set selection," *Discrete Optimization*, vol. 8, no. 1, pp. 87–96, 2011, parameterized Complexity of Discrete Optimization.

[4] E. Ackerman, O. Ben-Zwi, and G. Wolfovitz, "Combinatorial model and bounds for target set selection," *Theoretical Computer Science*, vol. 411, no. 44, pp. 4017–4022, 2010.

[5] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '03, 2003, pp. 137–146.

[6] A. Swaminathan, "An Algorithm for Influence Maximization and Target Set Selection for the Deterministic Linear Threshold Model," Master's thesis, Virginia Polytechnic Institute and State University, USA, 2014.

[7] S. Kochemazov and A. Semenov, "Using synchronous boolean networks to model several phenomena of collective behavior," *PLOS ONE*, vol. 9, no. 12, pp. 1–28, 12 2014.

[8] S. Kochemazov, "Comparative study of combinatorial algorithms for solving the influence maximization problem in networks under a deterministic linear threshold model," *Procedia Comput. Sci.*, vol. 136, pp. 190 – 199, 2018.

[9] S. Kochemazov and A. Semenov, "Computational study of time constrained influence maximization problem under deterministic linear threshold model for networks with nonuniform thresholds," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2019, pp. 1121–1125.

[10] E. Boros and P. L. Hammer, "Pseudo-boolean optimization," *Discrete Applied Mathematics*, vol. 123, no. 1, pp. 155 – 225, 2002.

[11] S. Luke, *Essentials of Metaheuristics*, 2nd ed. Lulu, 2013.

[12] H. Mühlenbein, "How genetic algorithms really work: Mutation and hillclimbing," in *PPSN*, 1992, pp. 15–26.

[13] S. Droste, T. Jansen, and I. Wegener, "On the analysis of the (1+1) evolutionary algorithm," *Theor. Comput. Sci.*, vol. 276, no. 1–2, p. 51–81, 2002.

[14] S. Russell and P. Norvig, *Artificial Intelligence – A Modern Approach*, 3rd ed. Pearson Education, 2010.

[15] B. Doerr, H. P. Le, R. Makhmara, and T. D. Nguyen, "Fast genetic algorithms," in *GECCO*, 2017, p. 777–784.

[16] A. Semenov, I. Otpuschennikov, and K. Antonov, "On some variants of the merging variables based (1+1)-evolutionary algorithm with application to maxsat problem," in *Mathematical Optimization Theory and Operations Research*, ser. LNCS, vol. 12755, 2021, pp. 111–124.

[17] A. Biere, M. Heule, H. van Maaren, and T. Walsh, Eds., *Handbook of Satisfiability - Second Edition*, ser. Frontiers in Artificial Intelligence and Applications. IOS Press, 2021, vol. 336. [Online]. Available: https://doi.org/10.3233/FAIA336

[18] A. Ignatiev, A. Morgado, and J. Marques-Silva, "PySAT: A Python toolkit for prototyping with SAT oracles," in *SAT*, 2018, pp. 428–437. [Online]. Available: https://doi.org/10.1007/978-3-319-94144-8_26

Neuroevolution for the Sustainable Evolution of Neural Networks

E. Otović[1,3], J. Lerga[1,3], D. Kalafatovic[2,3] and G. Mauša[1,3]
1 University of Rijeka, Faculty of Engineering, Department of Computer Engineering, Rijeka, Croatia
2 University of Rijeka, Department of Biotechnology, Rijeka, Croatia
3 University of Rijeka, Center for Artificial Intelligence and Cybersecurity, Rijeka, Croatia
eotovic@riteh.hr

***Abstract* - The predictive performance of a neural network depends on its weights and architecture. Optimizers based on gradient descent are most commonly used to optimize the weights, and grid search is utilized to find the most suitable architecture from the list of predefined architectures. On the other hand, neuroevolution offers a solution for the simultaneous growth of neural network architecture and the evolution of its weights. Thus, it is not limited by the user-defined list of possible architectures and can find configurations optimal for a specific task. Both approaches can be effectively parallelized and take advantage of modern multi-process systems. In this research, we compare neuroevolution and backpropagation in terms of the time consumed by the algorithm, the predictive performance of the neural network, and the complexity of the neural network. The total time for each algorithm is measured along with the times for each section of the algorithm and the time spent on synchronization due to the multi-process setting. The neural networks are compared by their predictive performance in terms of Matthews correlation coefficient score and their complexity as the number of nodes and connections. The case study is based on two synthetic and two real-world datasets for classification tasks.**

Keywords - machine learning; neuroevolution; neural networks; genetic algorithm; sustainability; classification

I. Introduction

Machine learning is a rapidly growing branch of computer science and is considered to be one of the enabling factors for industry 4.0. However, the powerful hardware required to run machine learning algorithms and its energy consumption have become a concern in terms of sustainability [1, 2]. Neuroevolution, a form of machine learning that evolves neural networks specific to a given task through an evolutionary process, may provide a more sustainable alternative to traditional gradient-based machine learning methods.

Gradient-based optimization algorithms have become a *de facto* standard for training neural networks. These algorithms are generally computationally efficient and converge relatively fast, especially when datasets with a small number of features are provided. However, the design of neural network architecture and the choice of adequate hyperparameters poses an obstacle, especially when it comes to the application of machine learning to a problem that has not been previously studied [3]. Suggested configurations and practices from literature can serve as a good starting point, but ultimately, the near-optimal configuration is found through experimentation. Grid search is an automated way of searching for the best-performing hyperparameters configuration from the predefined set of possible hyperparameter values. It is a widely used method because it can be easily implemented and understood, but is also easily scaled to multiple processes for faster execution [3]. The search space for grid search is defined by the user and it quickly becomes inefficient when the search space is *too big*, as it grows rapidly with the number of dimensions and possible values. It may also end up with suboptimal solutions since it considers only the predefined values for each hyperparameter.

On the other hand, neuroevolution is a form of genetic algorithm which allows for the exploration of a large solution space which can lead to the discovery of novel and creative solutions and its search space is not bounded by user-defined configurations. Since it performs a global search, it is less likely to get stuck in local optima and is less dependent on initial conditions [4]. Finally, it yields neural networks that have the architecture as well as the weights adapted for the given task. Genetic algorithms are known for their high scalability and neuroevolution is not an exception. Being a gradient-free method, it allows the optimization of the neural network directly on the non-differentiable objective function. Because of this, it has gained popularity in reinforcement learning settings as it can directly aim to maximize the reward signal [5].

Neuroevolution removes the burden of user-defined search spaces and finds novel topologies, however, it is considered to converge slower than backpropagation-based neural network training. Optimizing the neural network topology using the backpropagation-based methods requires training and validation for every considered topology, while neuroevolution speeds up this process by optimizing weights and topology simultaneously. Hence, the goal of this study was to evaluate and compare neuroevolution and grid search in terms of execution time, model's predictive performance and complexity.

II. Materials and Methods

A. NeuroEvolution of Augmenting Topologies

Neuroevolution of augmenting topologies (NEAT) is a method first proposed by Kenneth Stanley in 2002 for evolving artificial neural networks through the use of genetic algorithms [6]. Its main distinction from traditional neural network training is that it allows for the topology to change during the evolution process. This can

lead to the discovery of novel topologies that are better suited for a specific task.

The algorithm flow diagram is shown in Fig. 1. It starts by initializing the random population of neural networks (solutions). The initial neural networks contain only the input and output nodes that are fully connected and the weights of connections are randomly initialized. Internally, it represents a neural network with a list of nodes and connections among nodes. Then, the population of initial solutions is evaluated. In the speciation step, the population is separated into multiple species based on genetic similarity with the aim to prevent premature convergence and maintain diversity. The main loop then starts and consists of three steps: reproduction, fitness evaluation and speciation. More fit individuals get to reproduce more often through natural selection and lead the population towards solutions that maximize the fitness function. The generated offspring form a population for the next generation. The algorithm runs until a stopping criterion is met, then exits the loop and returns the fittest individual as a solution to the problem.

The next generation is formed through elitism, crossover and mutation operators. In this paper, we employ species-wise elitism which preserves two best-performing individuals from each species into the next generation and thus ensures that best-performing individuals are not lost during the evolution process. A crossover operator in neuroevolution involves combining genes (nodes, connections and weights) from two individuals to create an offspring. This allows for the exploration of new combinations of genes, potentially leading to improved solutions. We used three mutations like the ones introduced in the original NEAT implementation. *Add node mutation* disables and replaces a single existing connection in a network with a node and two connections; one going from the source to the newly inserted node and the other going from the inserted node to the destination. The weight of the incoming connection is 1, and the weight of the outgoing connection is set to the weight of the disabled connection. The activation function for a node is randomly sampled from a list of available activation functions. *Add connection mutation* randomly inserts a single connection between two previously unconnected nodes and randomly initializes its weight. Since this study was focused only on feedforward neural networks, only connections that do not form a cycle in a network are allowed. *Weight mutation* randomly changes a fraction of connection weights in a network, either by adding a random value to them or by replacing them with a random value. The probability for each mutation is configured separately. The probabilities for *add node mutation* and *add connection mutation* control the rate at which new topologies are explored. If they are too low, the exploration will be slow and it will slow down the convergence of the algorithm. On the other hand, if they are too high, individuals may not have enough time to get their weights optimized through crossover and *weight mutation*. The fraction of offspring generated only by mutation is controlled by the *crossover probability* parameter.

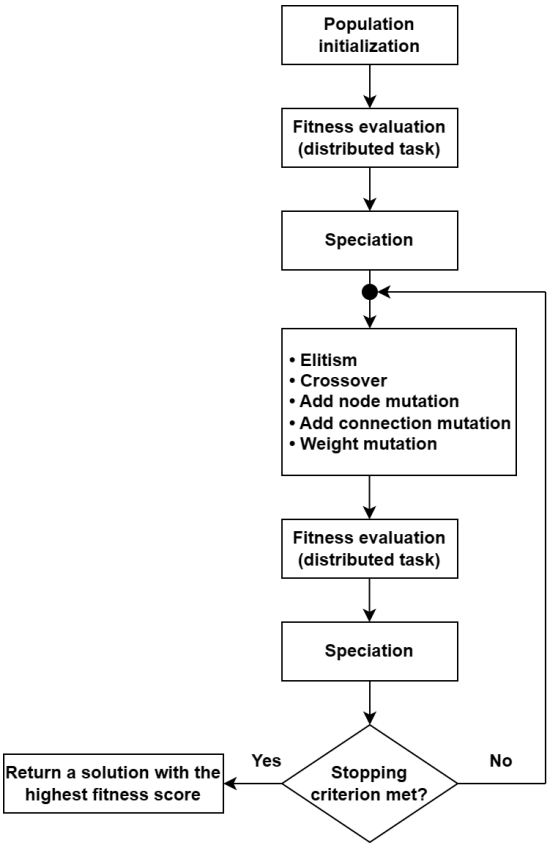


Figure 1. Program flow of NEAT.

NEAT aims to prevent premature convergence and preserve diversity in a population through speciation. Individuals from the population are grouped into species based on their genetic similarity and the best-performing individual from each species is chosen as its representative. When a newly formed individual has to be assigned to a species, it is sequentially compared against the representative from each species. The unnormalized compatibility distance **δ** shown in (1) is used in the comparison and if it is below a certain threshold $\boldsymbol{\delta}_t$ then the individual is assigned to that species.

$$\delta = c_{disjoint} * D + c_{weights} * \overline{W} \qquad (1)$$

The compatibility distance computes the genetic similarity between individual and species representatives. Its lowest value is zero when two genomes are equal, while higher values indicate greater difference. It considers the number of disjoint genes D (nodes and connections), which represents the number of genes that are present only in one of the individuals, but not both. It also considers the average difference in connection weights $\overline{W}$ for connections that are common to both individuals. The contribution of genes and weights to the compatibility distance can be regulated with $c_{disjoint}$ and $c_{weights}$ parameters. NEAT effectively tracks genes by assigning them historical markings that make it easy to determine genes that are shared by two individuals. Speciation allows the algorithm to explore solutions around multiple maxima at the same time. Members of

TABLE I. OVERVIEW OF USED DATASETS AND THEIR PROPERTIES.

Dataset name	Type	Number of features	Number of instances per class	Total size
Moons	Synthetic	2	10000, 10000	20000
Circles	Synthetic	2	5000, 5000	10000
Iris	Real-world	4	50, 50, 50	150
Breast cancer	Real-world	30	212, 357	569

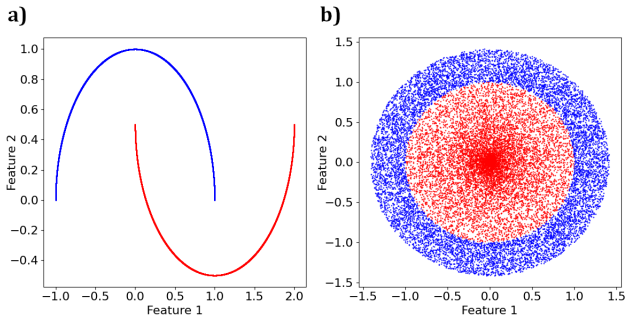


Figure 2. Visualization of moons (a) and circles (b) datasets.

the same species reproduce among themselves and on rare occasions interspecies mating occurs. The species that have not increased their fitness for a certain number of generations become stagnant and are not allowed to reproduce.

Fitness sharing is also introduced to prevent a single species from taking over the entire population. Equation (2) demonstrates the calculation of the shared (modified) fitness *f'* for an individual with fitness *f* that belongs to the species of size *n*.

$$f' = \frac{f}{n} \tag{2}$$

B. Algorithm and Model Evaluation

In this study, the algorithms and models are evaluated in terms of execution time, predictive performance of the model and model's complexity.

The total execution time is measured for both algorithms. In the NEAT algorithm, we as well measure the time required to complete the fitness function evaluation and the time spent on speciation and reproduction.

We used the Matthews correlation coefficient (MCC) to assess the predictive performance of the models. For the binary classification problems, it is defined as in (3):

$$MCC = \frac{tp \times tn - fp \times fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}} \tag{3}$$

Variables *tp*, *tn*, *fp* and *fn* denote the number of true positives, true negatives, false positives and false negatives, respectively. Even though a confusion matrix cannot be described perfectly with a single number, the MCC score is regarded as one of the best options since it considers all four cells of the confusion matrix in the case of binary classification. This also makes it suitable for datasets with imbalances among the classes. MCC score has been generalized and can also be used in multi-class and multi-label classification problems [7].

The Wilcoxon signed-rank test is a non-parametric statistical test used to compare paired samples. In this case, it was used to compare the MCC score of neuroevolution models and grid search models in order to determine whether one of two algorithms performs significantly better than the other one. We used a significance level of 0.05.

The complexity of a neural network can be expressed as the number of nodes and connections between them. For the purpose of this paper, this count does not include the nodes in the input and output layers because we are interested only in the number of hidden nodes that are required to solve the task. We count each bias term as a single connection.

III. CASE STUDY

A. Datasets

In this case study we employ two synthetic and two real-world classification datasets. Both synthetic datasets have two features and two equally distributed classes. These datasets contain 2D points belonging to two objects; in the case of the *moons* dataset it is two curves resembling the shape of the moons (Fig. 2a), and in the case of the *circles* dataset it is a circle with a ring around it (Fig. 2b). Two input features represent *(x, y)* coordinates of the point while the task of the model is to predict which object the point belongs to.

The *iris* dataset contains instances of three different species of Iris flowers (Iris setosa, Iris virginica and Iris versicolor) with four features representing various aspects of the flowers (sepal length, sepal width, petal length and petal width). Breast Cancer Wisconsin (Diagnostic) dataset contains malign and benign instances of breast cancer. Thirty provided features describe the characteristics of the cell nuclei present in the digitized image of a fine needle aspirate (FNA) of a breast mass. *Moons*, *Iris* and *Breast cancer* datasets were used from the *scikit-learn* library, while *circles* dataset was manually created [8, 9]. The overview of the used datasets is given in Table I.

Features used to describe instances in a dataset can have different units and scales which poses a problem for neural networks because the ones with a greater range of values may have a greater impact on the prediction outcome. To combat this, features were standardized for NEAT and backpropagation algorithms by centering the data points around zero and scaling to the unit variance [10]. If observed values for feature *i* are denoted as x_i, then standardization of that feature can be expressed as in (4).

$$x_i' = \frac{x_i - mean(x_i)}{std(x_i)} \tag{4}$$

To accurately estimate an algorithm's ability to adapt and model's ability to predict new data, we use stratified 10-fold cross-validation. The stratification guarantees that the distribution among classes will be preserved and thus prevents the under-representation of the minority class. In each iteration, nine folds are designated for training and one fold is held out for testing. In this way, each instance will be used exactly once for testing and thus minimizing the impact of random sampling on the performance estimation.

B. Algorithm Parameters

We used the NEAT algorithm with a population size of 250 and it was run for a maximum of 1000 generations or was stopped earlier if the maximal fitness did not increase for 300 generations. Output neurons used the softmax activation function while ReLU was used for hidden nodes. Weights were randomly sampled from a normal distribution with a mean of 0 and a standard deviation of 0.4. The crossover probability was set to 0.8, and the probability for interspecies mating was set to 0.01. *Add node mutation* probability was set to 0.1, *add connection mutation* was set to 0.3 and *weight mutation* was set to mutate 40% of all weights. A dynamic compatibility threshold was employed to preserve the number of species around 10 and the species would become stagnant after 20 generations without improvement. Compatibility parameters $c_{disjoint}$ and $c_{weights}$ were set to 2 and 1, respectively.

$$f = MCC + 1.001 \quad (5)$$

Equation (5) shows the fitness function that was used for neuroevolution. The original NEAT implementation cannot handle negative fitness scores due to its implementation of fitness sharing which is a problem since the lowest possible value of MCC is -1. Therefore, we had to increment it by 1.001 which allowed us to use it as a strictly positive fitness function.

Grid search tested three values for ADAM optimizer learning rate (0.01, 0.001, 0.0001) while other parameters were kept at Keras default values. We also varied the number of hidden layers (1, 2, 3), the number of neurons in each of them (10, 30, 50, 70, 90, 110) and dropout (0, 0.1, 0.2, 0.3, 0.4, 0.5). In total, this yields 4644 configurations that had to be tested. All the hidden layers used ReLU as the activation function. The output layer used the softmax activation function and the number of neurons corresponded to the number of classes in a given dataset. We used the Adam optimizer with a batch size of 32. Early stopping and dropout on each hidden layer were used to prevent overfitting. The training proceeded for the maximum of 20 epochs for *moons* and *circles* datasets due to their relatively large size, while the maximal number of epochs for *iris* and *breast cancer* datasets was set to 300 which provided enough time for convergence. 30% of the training was kept aside as the validation set and was used to evaluate and pick the best performing neural network configuration after training. The one with the highest MCC was evaluated on a held-out test set and that score is reported in Table II.

TABLE II. AVERAGE VALUES OF PREDICTIVE PERFORMANCE AND COMPLEXITY OF NEURAL NETWORKS FOUND BY NEAT AND GRID SEARCH. THE NUMBER OF EVALUATED NETWORKS, EXECUTION TIME AND THE NUMBER OF GENERATIONS ARE ALSO SHOWN. THE ASTERISK INDICATES STATISTICAL SIGNIFICANCE AT THE CHOSEN LEVEL OF SIGNIFICANCE (α=0.05).

	NEAT				**Grid search**			
	Moons	**Circles**	**Iris**	**Breast cancer**	**Moons**	**Circles**	**Iris**	**Breast cancer**
MCC score (test set)	1.0	0.913*	0.972	0.974*	1.0	0.981*	0.922	0.941*
Number of hidden nodes	8.8	14.9	7.5	8.2	10.0	229.0	69.0	89.0
Number of connections	39.3	69.9	39.6	87.6	52.0	12 717.0	2 578.0	3 597.0
Generations count	506	962	487	617	/	/	/	/
Evaluations count	99 038	192 412	95 378	122 940	4 644	4 644	4 644	4 644
Total time (per fold)	2 153s	4 135s	345s	610s	7 197s	6 447s	2 088s	2 261s
Evaluations per second	46.0	46.53	276.46	201.54	0.65	0.72	2.22	2.05

C. Experimental Setup

The experiments were carried out on a single computer with two Xeon E5 processors (24 physical cores; 48 threads) with 64 GB of RAM. Both algorithms were run in parallel with 12 workers. In the case of NEAT, multiple workers were used for parallel evaluation of genomes and in the case if grid search, it was used for parallel evaluation of neural network configurations. Due to the relatively large memory footprint of Tensorflow, it was not possible to use more workers on the same machine. Parallel execution of Python code was achieved with the *Dask* library which enables parallel execution on a single or across multiple computers. The dataset was distributed to all workers prior to the execution of the experiment. During each fitness evaluation in NEAT, genomes from the population had to be scattered among the workers, while in the case of the grid search, only numerical values describing neural network configuration and hyperparameters had to be sent to the workers. The fitness function or specific neural network configuration is evaluated in a worker process and the result is returned to the main process.

IV. Results and Discussion

Results from Table II show that NEAT completed on average 3.7 times faster than grid search. Even though our experiment was set up in a way that would allow both algorithms to find relatively good solutions in a reasonable time, it is impossible to fairly compare their performances in terms of total execution time. This is because the duration of the execution of each algorithm primarily depends on the configuration set by the user, e.g. stopping criteria for NEAT and search space size for a grid search. Grid search evaluated 4 644 pre-set neural network configurations for each dataset, while the number of neural networks evaluated by NEAT depends on many (stochastic) factors and cannot be known in advance. Therefore, we considered the number of neural networks evaluated per second which reflects the exploration rate of each algorithm. We found that NEAT is capable of performing exploration 90 times faster than the grid search. Grid search spends most of its time in the training phase tuning the weights of neural networks while the evaluation phase is relatively cheap time-wise. On the other hand, NEAT is capable of creating a new candidate neural network and evaluating it much more quickly. If the topology is predefined and fixed, backpropagation is an efficient algorithm to fine-tune the weights. Some researchers have proposed their combination in which NEAT would evolve the topology while backpropagation is used to fully or partially tune the weights [11, 12]. However, such methodology may not be applicable in some cases, e.g. if a non-derivable activation function is used.

Results from Table III show that NEAT spends most of the time evaluating the fitness function (on average 73.5%) even with parallelization introduced. This is expected because it involves the construction of a neural network from genes in a genome and performing forward pass to compute the output. The former depends on the size of the neural network while the latter depends on the size of the neural network and the size of the dataset. It is noticeable that larger datasets (*moons* and *circles*) required more time for forward pass than smaller datasets (*iris* and *breast cancer*). The second most intensive step is reproduction which consists of elitism, crossover and mutation and takes 19% of the time. The most expensive operations from this step are the identification of joint and disjoint genes during the crossover, verification if a new connection can be inserted without creating a cycle during *add connection mutation* and copying large genomes during elitism and crossover. The third most time-consuming step is speciation which computes the genetic distance between the genome and the representative of each species and takes 6.8% of the time. This involves finding the number of joint and disjoint genes as well as the average distance in connection weights which makes this computation relatively expensive. All other steps such as the computation of shared fitness are relatively fast and take on average 0.8% of the time. Therefore, fitness evaluation, reproduction and speciation are all dependent on the size of the genome and as evolution progresses and genomes get bigger, these steps require more time to complete.

Comparing the performance of the model evolved by neuroevolution and the model found by grid search, it is evident that both of them have achieved high MCC scores for all datasets. Both algorithms have achieved the highest possible score on the *moons* dataset which can be attributed to its relatively simple *shape* and the fact that classes are well separated (Fig. 2a). Statistical test revealed that grid search outperformed NEAT on *circles* dataset by a statistically significant margin of 7.4%. NEAT outperformed grid search by 5.4% and 3.5% on the *iris* and *breast cancer* datasets, respectively, which are relatively small in comparison to the other two datasets. While a statistically significant difference was found only in the case of the *breast cancer* dataset, it is worth noting that NEAT was significantly faster than grid search in every case, highlighting its potential as a more efficient approach. These results may suggest that NEAT is more suitable for smaller datasets.

In this research, we employed early stopping and dropout to control the overfitting. Early stopping mechanism evaluates model performance (generalization power) on a portion of the training set that was not used for training and stops training if validation loss stagnates

Table III. Measurements of Time NEAT Spent in Each Part of the Algorithm During a Single Generation. Measurements are Reported in Milliseconds.

Function	**Moons**	**Circles**	**Iris**	**Breast cancer**
Fitness evaluation	4 020 **(92%)**	4 003 **(91%)**	457 **(55%)**	582 **(56%)**
Speciation	77 **(2%)**	91 **(2%)**	87 **(10%)**	133 **(13%)**
Reproduction time	284 **(6%)**	298 **(7%)**	282 **(34%)**	298 **(29%)**
Other	8 **(<1%)**	12 **(<1%)**	9 **(1%)**	21 **(2%)**
Total	**4 389**	**4 404**	**835**	**1 034**

for 10 epochs. We tested multiple values for dropout which increased the search space by a factor of 6. When early stopping criteria is met, training is stopped and weights are restored to the ones with the lowest validation loss. Hyperparameters that are commonly used for addressing known problems can be found in the literature. However, for new or unfamiliar problems, these parameters need to be determined either manually or through an automated approach such as grid search. This also introduces the problem of choosing the amount of data that should be kept aside for early stopping. The more data is kept aside, the better the estimate of the model's performance will be, but at the same time less data will be available for training. On the other hand, setting too little data aside will result with high variance in the estimation of generalization by early stopping and may stop training too soon or too late. Therefore, it is a trade-off between the precision of the generalization power estimate and the amount of information that can be extracted from the training set.

The results from Table II reveal that neural networks evolved by NEAT are consistently less complex with regard to the number of hidden nodes and connections and this can be attributed to two factors. The limiting factor for traditional machine learning is that layers are stacked upon each other and fully connected. Even though this increases the model capacity, it introduces unnecessary connections and makes the neural network more prone to overfitting. On the other hand, NEAT starts with the population of minimal fully connected neural networks and the complexity of candidate topologies gradually increases through trial and error exploration. The evolutionary process of NEAT prevents the addition of unnecessary connections by favouring the solutions with higher fitness. Therefore, if a newly added node or a connection does not contribute to the fitness, eventually it will cease to exist. Additionally, using the historical markings, NEAT keeps track of previously inserted nodes and connections and utilizes it to prevent duplicate nodes and connections to be inserted. Neuroevolution operates on a finer level, e.g. activation function is separately set for each neuron and custom connections between nodes. Grid-search can be modified to consider more specialized architectures to some extent, e.g. by trying multiple activation functions for each layer, but such configurations rapidly increase the search space and are still inferior to NEAT in terms of customizability.

V. Conclusion

In this paper, we compared the NEAT neuroevolution algorithm and grid search in terms of execution time, predictive performance and neural network complexity. The measurements of execution time have shown that NEAT is capable of exploring the search space 90 times faster than grid search. However, in some cases, it is possible to combine neuroevolution and backpropagation to preserve the exploration of topologies and efficient weight-tuning mechanisms. With the addition of more parameters to the grid search, better results may be achieved at the cost of execution time. We demonstrated the ability of neuroevolution to more effectively search a much larger search space of possible configurations that are specialized for the given task. This resulted in NEAT consistently evolving neural networks with fewer nodes and connections than the ones that were found by grid search. In future studies, it would be valuable to compare NEAT against other commonly used hyperparameter optimization techniques, such as random search or Bayesian optimization, in order to gain a deeper understanding of the relative strengths and weaknesses of these methods and their applicability to different neural network optimization problems. We also plan to explore the potential of NEAT to improve the accuracy and efficiency of predictive models in peptide chemistry where a high number of available representation schemes together with high dimensionality and imbalance of data are an important challenge that requires smart optimization procedures. Our research findings suggest that NEAT is a promising approach for improving peptide activity prediction, particularly in cases where datasets are limited in size. Simpler neural networks are more suitable for deployment on battery-powered embedded devices in order to reduce computational demands and energy consumption while improving the responsiveness of the system to the input. This can lead to more sustainable and eco-friendly applications of artificial intelligence in emerging IoT devices.

Acknowledgment

This project was conducted using the resources of the supercomputer Bura at the University of Rijeka, Center for Advanced Computing and Modelling and supported by the Croatian Science Foundation/Hrvatska zaklada za znanost (grant no: UIP-2019-04-7999 and DOK-2020-01-4659).

This paper acknowledges the support of the Erasmus+ Key Action 2 (Strategic partnership for higher education) project No. 2020-1-PT01-KA203-078646: "SusTrainable - Promoting Sustainability as a Fundamental Driver in Software Development Training and Education". The information and views set out in this paper are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

References

[1] Peres, Ricardo Silva, et al. "Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook." IEEE Access 8 (2020): 220121-220139.

[2] "The Fourth Industrial Revolution: what it means, how to respond", World Economic Forum, Jan. 2016, https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

[3] Yang, Li, and Abdallah Shami. "On hyperparameter optimization of machine learning algorithms: Theory and practice." Neurocomputing 415 (2020): 295-316.

[4] Kumar, Manoj, et al. "Genetic algorithm: Review and application." Available at SSRN 3529843 (2010).

[5] Igel, Christian. "Neuroevolution for reinforcement learning using evolution strategies." The 2003 Congress on Evolutionary Computation, 2003. CEC'03.. Vol. 4. IEEE, 2003.

[6] Stanley, Kenneth O., and Risto Miikkulainen. "Evolving neural networks through augmenting topologies." Evolutionary computation 10.2 (2002): 99-127.

[7] Gorodkin, Jan. "Comparing two K-category assignments by a K-category correlation coefficient." Computational biology and chemistry 28.5-6 (2004): 367-374.

[8] Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.

[9] Newman, David J., et al. "UCI repository of machine learning databases, 1998." (1998).

[10] Kotsiantis, Sotiris B., Dimitris Kanellopoulos, and Panagiotis E. Pintelas. "Data preprocessing for supervised learning." International journal of computer science 1.2 (2006): 111-117.

[11] Chandra, Rohitash, and Christian W. Omlin. "The Comparison and Combination of Genetic and Gradient Descent Learning in Recurrent Neural Networks: An Application to Speech Phoneme Classification." Artificial Intelligence and Pattern Recognition. 2007.

[12] Cui, Xiaodong, et al. "Evolutionary stochastic gradient descent for optimization of deep neural networks." Advances in neural information processing systems 31 (2018).

IMAGE AND VIDEO PROCESSING AND ANALYSIS

Influence of quality of pixel level annotations on text detection performance in natural images

I. Dorkić*, M. Brisinello*, R.Grbić** and M. Herceg**
* TTTech Auto Central and Eastern Europe, Osijek, Croatia
** Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Osijek, Croatia
ivan.dorkic@tttech-auto.com, matteo.brisinello@tttech-auto.com, ratko.grbic@ferit.hr, marijan.herceg@ferit.hr

***Abstract* - Text detection in natural images is a task that arises in many computer vision applications. State-of-the-art text detection methods are mainly based on deep neural networks designed for instance segmentation task. However, most of the available datasets for text detection do not have fine annotations at the pixel level which are required during supervised learning of such networks. Usually, a whole or reduced text bounding box is used as a segmentation mask. In this paper, a method that generates a synthetic dataset with precise annotations at the pixel level is proposed. The method is based on the available Synthtext script for generating synthetic datasets with text instances. By creating synthetic datasets with precise and coarse annotations at the pixel level we explore the efficiency of the state-of-the-art text detector TextFuseNet.**

Keywords* - *object detection; text detection; instance segmentation; Synthtext; TextFuseNet

I. INTRODUCTION

The text represents a set of words that conveys a certain meaning to the reader. In our daily life, we often encounter text while reading newspapers, emails, information and traffic signs, street names, and so on. Text detection and recognition are for an adult person relatively simple task. Reading a text with a computer is successfully solved within domains where the structured text appears like in digitalized books, official documents, license plates, and so on. Such text is usually on a legible background, with a standard font and line spacing and horizontally aligned with no occlusions. However, text can appear in so-called unstructured form - we find it in random places in natural images. Typical examples are scenes of the streets or shops where text appears in posters, advertisements, cafe names, products, and so on. Such text instances can have non-standardized font and spacing, can be in mixed colors, multi-oriented, occluded, curved, and with a colorful background. Detection of such text in natural images is still quite challenging for computer algorithms.

Nowadays, text detection in natural images is based on deep Artificial Neural Networks (ANNs). The most effective methods [1] are based on deep Convolutional Neural Networks (CNNs) that perform instance segmentation task. To successfully learn such a network in a supervised manner, an appropriately annotated dataset is required which means that each text instance in the image is annotated at a pixel level. However, most of the available datasets provide only a bounding region of a text instance in form of a rectangle or polygon, and usually, every pixel within such bounding region is considered a part of the text by the applied method for text detection [2]. In the rest of this paper, we call such pixel-level annotations coarse. We believe that more precise text instance annotations on a pixel level of the training dataset can potentially boost text detector efficiency.

To explore whether more precise annotations at the pixel level of the training dataset can boost text detector efficiency, we propose an approach for obtaining precisely annotated synthetic datasets for text detection. The available code from [3,4] is used to generate synthetic images where the text appears in natural images. The code is further modified to produce precise annotations on a pixel level in form of a segmentation mask or polygon for each character and whole text instance. Such datasets are then used to train state-of-the-art text detector TextFuseNet [1] to quantify the importance of precise annotating of text instances in the training dataset.

The paper is structured as follows. In Section II, a problem of text detection and text annotation is presented, followed by an overview of existing datasets for text detection. After that, an overview of the most recent solutions for text detection is given. A process of creating datasets with precise and coarse annotations on a pixel level is given in Section III along with the process of text detector training. The obtained results on the created datasets and the accompanying discussion are given in Section IV. At the very end of the paper, a conclusion is given.

II. RELATED WORK

Most text detection methods, at least the most effective ones, are based on deep neural networks for instance segmentation task like the well-known Mask RCNN [5]. These networks are further modified to take into account text appearance in natural images and are trained in a supervised manner based on available text detection datasets. Most of the datasets do not provide the fine annotations of each character or the whole text instance at the pixel level (in form of a mask or a polygon). In theory, the most precise way of specifying a text mask would result in a mask consisting of pixels within the bounding box that exactly belong to the text in the image. An example of an image with several text instances is shown in Fig. 1 a) while the difference between the commonly

used coarse approach (Fig.1 b)) and the precise labeling approach (Fig.1 c)) is clearly visible. Most of the instance segmentation methods based on Mask R-CNN solve this problem by treating all pixels within the bounding box as a segmentation mask during training.

A. A brief overview of available datasets for text detection

The ICDAR 2013 dataset [2] consists of 229 training images and 233 testing images. All images are of different resolutions. It is a set in which all the images are photographed with a camera and the text on them is natural, i.e. it is not in focus. It is a standard reference dataset for horizontal text detection.

Synthtext in the wild [3] is a synthetically generated dataset in which words are placed in images of natural scenes. The text instances are placed in an image by considering the outline of the scene and the depth map of the scene. In this way, the text was added to sufficiently large areas. It consists of 800,000 images with approximately 8 million words artificially added to the images. The images are of different resolutions. Each text instance is annotated with word-level and character-level bounding boxes.

The COCO_Text dataset [6] contains 63,686 images with 145,859 text instances, where words are annotated by using polygons. It is the largest real-world dataset of images with text appearance in a natural form. It consists mostly of horizontal and multi-oriented text instances which are rotated by a certain angle. A small part of the images also contains curved text.

The Multi-Lingual scene Text (MLT-2019) dataset [7] contains 20,000 images. The images are of different resolutions. The text in the images is displayed in its natural form. Annotation of the dataset contains word level and character level text bounding boxes along with the corresponding transcription and language class. It contains text in 10 different languages.

The Incidental Scene Text (ICDAR 2015) dataset [8] contains 1,670 images and 17,548 annotated regions. This dataset contains images in which the text is not in the focus. Therefore, most of the text instances in the images are out of focus and blurry. Text instances are annotated with word-level bounding boxes in form of quadrilaterals.

B. A brief overview of text detection methods

In [9], a method called Mask TextSpotter is proposed, which is an end-to-end neural network for spotting text with arbitrary shapes. Mask TextSpotter uses a simple end-to-end learning procedure that can achieve the detection and recognition of text directly from two-dimensional space using semantic segmentation. This method is efficient in detecting instances of irregular shape text, such as curved text. It is evaluated on four English datasets and one multi-language dataset, achieved results are much better than other similar methods.

Most existing solutions for text detection in images are based on neural networks primarily intended for instance segmentation. In [1], a method called TextFuseNet is proposed, which is mainly based on the architecture of Mask R-CNN and TextSpotter, which uses richer fused features to detect text. TextFuseNet collects and fuses text features from different levels using a multi-path fusion architecture that can efficiently match and fuse different representations. TextFuseNet detects text at the word and character level. During training, TextFuseNet uses a more precise polygon segmentation mask. The text instance annotation is a "narrowed" bounding box compared to the approach of most other methods, where everything inside the bounding box is considered a segmentation mask. The detection success measured with F1 score on the ICDAR2013 dataset is 94.3%, and on the ICDAR2015 dataset it is 92.1%

It appears that one problem in building efficient object detectors is the lack of large data sets that have precisely annotated text at the pixel level. In [10], a method for annotating text on images at the level of pixels is proposed. The previously mentioned COCO_Text data set was used. From the images in that set, 1,000,000 text images per border of the rectangular bounding box are cropped. These images were then used to train a deep neural network that is used for text segmentation, that is, the background is separated from the text itself using this network. After the training phase, two thresholds were determined. If the probability is less than 0.3, that part of the image is considered the background, if the probability is greater than 0.7, the image contains text, and if the probability is between these two thresholds, the network is unsure of its decision. Using this method, 14,690 images were generated where text is precisely annotated at the pixel level.



a) b) c)

Figure 1. Examples of annotated images: a) input image with text bounding boxes, (b) coarse segmentation mask, (c) precise segmentation mask.

III. CREATING PRECISELY ANNOTATED SYNTHETIC DATASETS

This section describes the proposed approach for the precise annotating of text instances in synthetic natural images. For this purpose, we used an existing script available at [4] which was used for developing a famous Synthtext in the wild dataset. Based on created synthetic datasets we trained and tested TextFusenet text detector.

The code from [4] produces natural images where text instances are placed at suitable locations in the image. The script for each text instance provides coordinates of the text bounding box and information which image pixels correspond to the characters of the text. The example of generated text instance is shown in Fig. 2 where annotation at a pixel level provided by the script [4] is overlaid in red color. It can be noticed that generated annotation is quite imprecise. Certain pixels that belong to the background are annotated as text, so the actual shape of each character is not clearly visible from the generated mask.

To annotate the text instance more precisely at the pixel level we first extract the text instance from the image based on the provided text bounding box. Then *k*-means clustering algorithm [11] is applied to the extracted image to perform color segmentation. By doing so the segmentation of the image is performed based on the color similarity of the pixels. *K*-means clustering is performed in RGB color space. The resulting centers of *k*-means algorithm represent the dominant colors in the image. Since every character of each generated text instance is always in the same color, the assumption is that the pixels that belong to the text will be assigned to a single group and all other pixels (e.g. background) to other groups. In this paper, the optimal number of clusters was determined experimentally, and it is four. If less than four clusters are used, the text wouldn't always be segmented properly. Segmentation of the image shown in Fig. 2 in four clusters by *k*-means algorithm is shown in Fig. 3. In this case the first cluster represents pixels belonging to text (Fig. 3 a)) while the remaining clusters represent the background and text border. Obviously, cluster representing text can be used as a more precise text mask than the annotation produced by the script [4] which is shown in Fig. 2. Furthermore, since script [4] provides rectangle bounding boxes on a character level also, we can easily extract precise mask for each character of a text instance from the obtained cluster representing text.

Since the result of the k-means algorithm depends on the initial initialization of the centers and the appearance of the image itself, it is not possible to know in advance in which group the pixels belonging to the text will be assigned. For this purpose, we developed CNN that performs binary classification whether the image contains text or not. In that way, we can determine which cluster actually contains pixels belonging to the text, i.e. required precise mask of the text. The CNN structure is shown in Table I. It is based on convolutional and max pooling layers. An input image is scaled to the size of 64x64 pixels. Dropout layer is used to prevent overfitting. To train such CNN, 812 different images containing text instances (like Fig. 4) were segmented with *k* equal to four



Figure 2. Example of generated text instance with overlaid annotation in red color produced by script [4].



Figure 3. Segmentation of the image shown in Fig. 2 using *k*-means algorithm in four clusters.

resulting in 3,248 images (similar to Fig. 3). The resulting images were manually labeled to class 1 or 0. It is necessary to do so because images containing text are considered positive examples, while images without text are considered negative examples. Several examples of such images are shown in Fig. 4. The images were divided into a training set, a validation set, and a test set. The training dataset contains 512 images with text and 1536 images without text. The validation set contains 50 images with text and 150 without text. The test set contains 250 images with text and 750 without text. The number of training epochs was 40 and the batch size was 32. During learning online augmentation was used such as rotation, horizontal flip, vertical flip, zoom, etc.

The evaluation of the built CNN is shown in Table II using the confusion matrix on the test dataset with 0.5 threshold applied to network output. The built CNN shows accuracy on the test of over 99%, with precision equal to 97.6% and recall equal to 98.4%.

The built CNN was implemented along with *k*-means algorithm in the script [4] to produce precise annotation at the pixel level for each text instance during the generation process (on-the-fly). Based on the resulting mask, a polygon or RLE record is determined for each text instance and saved to the JSON file containing annotations in COCO style [12]. Although the RLE notation is somewhat more precise, in this paper we used the polygon annotations for pixel-level segmentation since the default implementation of TextFuseNet requires segmentation annotation in form of polygons. Since the script from [4] provides a bounding box for each character of the text instance, we simply annotated each character at pixel level by extracting only part of the text mask that intersects with the character bounding box.

The code from [4] script uses a database of 8000 images but during a single run over 7000 images with

TABLE I. CNN STRUCTURE FOR BINARY CLASSIFICATION OF IMAGES CONTAINING TEXT OR NOT .

Layer	Input size	Output size
InputLayer	(64,64,3)	(64,64,3)
Conv2D	(64,64,3)	(64,64,32)
Conv2D	(64,64,32)	(62,62,32)
MaxPooling2D	(62,62,32)	(31,31,32)
Droput	(31,31,32)	(31,31,32)
Conv2D	(31,31,32)	(31,31,64)
Conv2D	(31,31,64)	(29,29,64)
MaxPooling2D	(29,29,64)	(14,14,64)
Droput	(14,14,64)	(14,14,64)
Flatten	(14,14,64)	(12544)
Dense	(12544)	(512)
Droput	(512)	(512)
Dense	(512)	(2)

TABLE II. CONFUSION MATRIX OBTAINED ON TEST DATASET

		Predicted	
		Text	*Not text*
True	*Text*	244	6
	Not text	4	746



a) b) c) d)

Figure 4. Examples of training images obtained by *k*-means clustering, a) - b) images containing text, c) - d) images without text.

random text instances are generated because some images do not always provide a convenient place to add text instance of a certain size and length. A typical example of the generated image is visible in Fig. 5 and a single text instance with precise and coarse annotations at the pixel level is shown in Fig. 6 a) and b).

In the end, six different training datasets were generated with varying numbers of training images (around 1000, 3000, and 5000) and with different types of annotations at the pixel level (precise or coarse). The generated datasets are summarized in Table III.

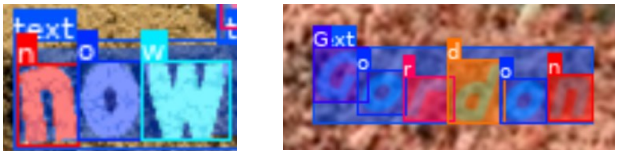
IV. RESULTS AND DISCUSSION

This section presents the results of TextFuseNet detectors training and testing on the generated datasets. A PC with a Linux operating system (Ubuntu 20.04. LTS), NVIDIA GeForce RTX 3080Ti graphics card, and Intel Core i9-11900F processor with 32 GB of RAM was used for training and testing. The training comprised of 60 epochs and the model with the lowest validation loss was selected for further evaluation.

All built detectors were evaluated on the same test data set with respect to text instances detection. COCO evaluation metrics are used for built detectors evaluation: average precision over 10 IoU thresholds from 0.5 to 0.95



Figure 5. Examples of the generated image with precise annotations using modified *Syntext* script.



a) b)

Figure 6. Examples of text instance with a) precise annotation, b) coarse annotation.

TABLE III. GENERATED SYNTHETIC DATASETS

Dataset name	Number of images in each dataset
Precise_1k	1029
Precise_3k	3265
Precise_5k	5099
Coarse_1k	1029
Coarse_3k	3265
Coarse_5k	5099
Validation	1038
Test	1032

with a step size of 0.05 - **AP**, average precision at IoU equal to 0.5 – $\mathbf{AP^{50}}$, average precision at IoU equal to 0.75 – $\mathbf{AP^{75}}$, average precision for small objects with an area less than 32^2 – $\mathbf{AP^S}$, average precision for medium objects with an area between 32^2 and 96^2 – $\mathbf{AP^M}$, and average precision for large objects with an area greater than 96^2 – $\mathbf{AP^L}$.

The evaluation results of TextFuseNet detectors which were trained on precisely and coarsely annotated datasets with a varying number of training images are presented in Table IV and Table V. In both tables, it can be observed that metrics slightly increase with the increase of the number of training images which was expected. AP^{50} values are only slightly lower than AP^{75} indicating good localization properties of all detectors. All models have high values of average precision for medium and large objects but struggle with small ones. If we compare AP^{50} value of detectors trained on precisely and coarsely annotated datasets, then it can be noticed that the obtained values are only slightly in favor of more precise

TABLE IV. TEST RESULTS OF TEXTFUSENET DETECTORS TRAINED ON PRECISELY ANNOTATED DATASETS

Training dataset	AP	AP^{50}	AP^{75}	AP^{S}	AP^{M}	AP^{L}
Precise_1k	73.02	87.05	82.77	66.92	81.65	86.82
Precise_3k	74.75	88.04	83.85	68.58	83.18	88.23
Precise_5k	75.49	88.03	84.86	69.40	83.92	89.75

TABLE V. TEST RESULTS OF TEXTFUSENET DETECTORS TRAINED ON COARSELY ANNOTATED DATASETS

Training dataset	AP	AP^{50}	AP^{75}	AP^{S}	AP^{M}	AP^{L}
Coarse_1k	73.32	87.04	82.71	63.55	81.12	89.04
Coarse_3k	74.25	87.08	83.79	64.49	82.43	89.22
Coarse_5k	74.96	87.08	83.92	65.09	83.04	89.79



Figure 7. TextFuseNet output for a single test image.

annotating process. However, further analysis of average precision for different size of text instances shows certain benefits when annotating instances more precisely in the case of small and medium-sized objects. It is especially pronounced for small objects where AP^{S} is around 6% higher for detectors trained on precisely annotated datasets. Interestingly, AP^{L} is slightly decreased when the detectors are trained on precisely annotated datasets. This suggests that different strategies can be used when data annotating concerning the object size for the best TextFuseNet detector performance. In Fig. 7. the detector output for one test image is displayed.

V. CONCLUSION

In this paper, text detection in natural images is studied in more detail regarding the quality of annotations in the training dataset. A method that produces fine annotations on a pixel level is proposed and implemented within available Synthext in the wild script to obtain precisely annotated synthetic datasets for text detection. Precisely and imprecisely datasets are used to train text detectors based on TextFuseNet. The obtained results suggest that average precision on IoU equal to 0.75 does not significantly increase if precise annotations at pixel level are used. However, average precision for small and medium-sized objects increases, while average precision for large objects decreases.

Future work will include pretraining of TextFuseNet with precisely annotated synthetic datasets and then training on available real-world datasets. Additionally, modifications to TextFuseNet will be performed to use masks as annotations rather than polygons annotations. Furthermore, we believe that the recognition process that should be performed after text detection can benefit when detectors are trained with precisely annotated text instances, and this will be explored in more detail.

REFERENCES

[1] J. Ye, Z. Chen, J. Liu, and B. Du, "TextFuseNet: Scene Text Detection with Richer Fused Features," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, Yokohama, Japan, Jul. 2020, pp. 516–522. doi: 10.24963/ijcai.2020/72.

[2] D. Karatzas *et al.*, "ICDAR 2013 Robust Reading Competition," in *2013 12th International Conference on Document Analysis and Recognition*, Washington, DC, USA, Aug. 2013, pp. 1484–1493. doi: 10.1109/ICDAR.2013.221.

[3] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic Data for Text Localisation in Natural Images," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2315–2324. doi: 10.1109/CVPR.2016.254.

[4] A. Gupta, "SynthText." Aug. 23, 2022. Accessed: Aug. 23, 2022. [Online]. Available: https://github.com/ankush-me/SynthText

[5] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Oct. 2017, pp. 2980–2988. doi: 10.1109/ICCV.2017.322.

[6] A. Veit, T. Matera, L. Neumann, J. Matas, and S. Belongie, "COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images." arXiv, Jun. 19, 2016. Accessed: Aug. 23, 2022. [Online]. Available: http://arxiv.org/abs/1601.07140

[7] N. Nayef *et al.*, "ICDAR2019 Robust Reading Challenge on Multi-lingual Scene Text Detection and Recognition -- RRC-MLT-2019." arXiv, Jrul. 01, 2019. Accessed: Nov. 10, 2022. [Online]. Available: http://arxiv.org/abs/1907.00945

[8] D. Karatzas *et al.*, "ICDAR 2015 competition on Robust Reading," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, Aug. 2015, pp. 1156–1160. doi: 10.1109/ICDAR.2015.7333942.

[9] M. Liao, P. Lyu, M. He, C. Yao, W. Wu, and X. Bai, "Mask TextSpotter: An End-to-End Trainable Neural Network for Spotting Text with Arbitrary Shapes." arXiv, Aug. 22, 2019. Accessed: Dec. 13, 2022. [Online]. Available: http://arxiv.org/abs/1908.08207

[10] S. Bonechi, M. Bianchini, F. Scarselli, and P. Andreini, "Weak supervision for generating pixel–level annotations in scene text segmentation," *Pattern Recognit. Lett.*, vol. 138, pp. 1–7, Oct. 2020, doi: 10.1016/j.patrec.2020.06.023.

[11] "What is K-Means Clustering? - Definition from Techopedia," *Techopedia.com*. http://www.techopedia.com/definition/32057/k-means-clustering (accessed Aug. 23, 2022).

[12] S. Bonechi, P. Andreini, M. Bianchini, and F. Scarselli, "COCO_TS Dataset: Pixel-level Annotations Based on Weak Supervision for Scene Text Segmentation," vol. 11730, 2019, pp. 313–325. doi: 10.1007/978-3-030-30490-4_26.

Effects of Applying Identified Road Lane Lines on Vehicle Autopilot Model Driving Performance

Vjekoslav Diklić*, Franjo Matković†, Krešimir Pardon‡
*, ‡ dSPACE Engineering d.o.o., Zagreb, Croatia
† Fakultet elektrotehnike i računarstva, Zagreb, Croatia
e-mail address: vdiklic@dspace.hr franjo.matkovic@fer.hr kpardon@dspace.hr

***Abstract*—Autonomous Vehicle driving is one of the most actively developed systems in the automotive industry. Producing a reliable autonomous driving system is a complex multistage problem. The road lane lines are traffic signs with high level of semantics that determine how a vehicle should move, what actions are allowed to be performed, etc. Comparing different modifications of the PilotNet neural network model, the usage of detected lane lines, and their implications on vehicle autopilot driving performance is the topic of this paper. The lane lines are identified by neural network that has been trained by processing images with previously marked lane lines from CULane dataset. Vehicle autopilot driving performance is analyzed using autopilot model running on Euro Truck Simulator 2. Proposed models are analyzed and compared in multiple driving and weather environments including highway, off-road and urban driving during day, night and rain conditions. The lane line recognition (classified by a U-Net neural network model from images) is compared with additional control and path planning to determine whether it leads to better autonomous driving performance and smaller driving system model.**

Keywords*— *Autonomous Driving, Autopilot, Convolutional Neural Networks, Euro Truck Simulator 2, Image Processing, Virtual Controller

I. INTRODUCTION

Autonomous vehicle driving is one of the most actively developed systems in the automotive industry. The initiative comes from both industry and science as high-performance computing and artificial intelligence technology become cheaper, widely available and more advanced. Such systems offer increased comfort, driving experience and safety, while reducing road accidents. Based on the degree of automation, there are six levels of autonomous driving, which are: level zero (no automation at all), level one (very light automation; cruise control, etc.), level two (some automation but requires continuous human attention), level three (self driving with required intervention in severe conditions), level four (highly autonomous), level five (completely autonomous). Road lane lines are traffic signs with high level of semantics, since they determine vehicles movement and actions that are allowed to be performed. This makes lane lines critical component in autonomous driving as they describe the path for self driving cars and ensure the vehicle will not unpredictably change lanes. Comparing different modifications of PilotNet neural network model [1], usage of detected lane lines based on LaneUNet model [2], [3], and its implications on vehicle autopilot driving performance is the topic of this paper. The remainder of this paper is organized as follows: section II gives a brief overview of previous and current work on lane detection and autonomous driving pilots. Section III provides description of the problem being addressed, along with the algorithms used and the formation of the dataset. The training process, evaluation and results are displayed in Section IV. Section V concludes this paper.

II. RELATED WORK

Producing a reliable autonomous vehicle driving system is complex multistage problem, and critical part of it is lane detection. The following papers are dealing with different approaches to the problem of visual road lane lines detection.

Zheng et al. [4] proposed a Cross Laxer Refinement Network (CLRNet) for detecting lines by utilizing both high- and low-level semantic features in the images which correlate to the drive lanes.

Dong et al. [5] proposed a hybrid spatio-temporal deep learning network for continuous lane detection based on multiple sequential frames. The described working theory gives that lane lines can be better detected if past knowledge is used by sending multiple frames in the sequence, and the lane is detected only on the last frame. The authors have shown through experiments that their model outperforms other state-of-the-art models by utilizing single frame feature extraction module combined with spatio-temporal recurrent network model feature integrator and coder-decoder structure, which enabled it to be end-to-end learned.

Liu et al. [6] proposed a new approach to tackle lane detection in scenarios with complex topologies. Proposed approach is a top-to-bottom lane detection method that first detects line instances, then predicts lane shape for every detected instance. A conditional lane detection strategy based on conditional convolution and row-wise formulation is introduced to resolve lane instance-level discrimination problem.

Qiu et al. [7] introduced a transformer network to the road lane detection problem. The authors proposed a novel framework called PriorLane, which enhances segmentation

of the fully vision transformer by introducing a low-cost prior knowledge. Results on their Zjlab dataset showed that the proposed method outperforms other SOTA methods.

Jin et al. [8] used a novel method to detect road lanes in eigenlane space. First, authors have introduced Eigenlanes, which are data-driven lane descriptors for structurally diverse lines (like curves and straight lines). Then, by obtaining best candidate lanes through approximation, authors have used a detection neural network to find optimal lanes.

In [9], He et al. proposed a generative adversarial network to find an enhanced feature space where lane features are distinctive while maintaining a similar distribution of road lanes in the wild. Novel Repainting and Imitating Learning (RIL) framework experiments proved effectiveness both on CULane and TuSimple for road lane detection methods.

Borjanski et al. [10] proposed a convolutional neural network (CNN) for learning autonomous driving from raw pixels of a single front-facing camera by mapping them into steering commands.

Chitta et al. [11] proposed a mechanism for integrating image and LIDAR representations using self-attention. This approach uses transformers at multiple resolutions to fuse perspective view with the bird's eye view feature maps. The motivation for the approach was under-performance of the end-to-end learning in complex driving scenarios with heavy traffic and dynamic agents high-density.

Chekroun et al. [12] presented a novel method called General Reinforced Imitation (GRI), which combines benefits from exploration and expert data. The authors proposed a simplifying hypothesis, that expert demonstrations can be seen as perfect data whose underlying policy receives a constant high reward. GRI combines offline demonstration agents and online RL exploration agents.

Hao et al. [13] introduced a novel safety-enhanced autonomous driving framework based on transformers called the Interpretable Sensor Fusion Transformer (InterFuser). The framework fuses information from multi-modal multi-view sensors for comprehensive scene recognition and adversarial event detection. The framework also generates intermediate interpretable features that provide more semantics that are exploited to better constrain actions.

Wu et al. [14] introduced a novel two-branch method for autonomous driving. Branches are based on trajectory planning and direct control of vehicle. While trajectory planning branch predicts the future trajectory, the control branch receives the corresponding guidance from the trajectory branch. The control branch makes a multi-step prediction scheme in the way the relationship between current action and future states can be reasoned. Outputs from the branches are fused together to an achieve advantage from each branch.

Toromanoff et al. [15] presented a novel technique that used a reinforcement learning (RL) approach to autonomous driving. It contains coined implicit affordances to leverage RL for urban driving, which include lane keeping, pedestrian and vehicles avoidance and traffic light detection. The authors have successfully managed to present an RL agent which is capable of handling traffic light detection.

III. The proposed approach

This paper compares how pre-marked road lane lines impact autonomous driving system when the system is trained with and without pre-marked lines. The original authors of PilotNet [16], [17] developed a deep learning neural network for autonomous driving and stated that the deep learning network will perform better due to internal components self-optimizing to maximize overall system performance if it were given no human-selected intermediate optimizing criteria, such as lane line detection. Such criteria are understandably selected for ease of human interpretation, but this doesn't automatically guarantee maximum system performance. This paper examines the influence of human interpretable data as additional information provided to deep neural networks. The Europilot [18] project is an interface between Euro Truck Simulator 2 (ETS2) [19] and Python code that was used to investigate the impact of the pre-marked lane lines (lane detection) on a driving performance of an PilotNet model. The autopilot model uses only raw pixels from simulator as input, extended with pre-marked lane lines generated by LaneUNet.

A. Self-driving algorithms

Authors [1] argue that providing the neural network with raw data is the best way to create a robust self-driving system which can navigate its way even on unmarked lanes (unpaved roads, parking lots):

> "The system automatically learns internal representations of the necessary processing steps such as detecting useful road features with only the human steering angle as the training signal. We never explicitly trained it to detect, for example, the outline of roads. Compared to explicit decomposition of the problem, such as lane marking detection, path planning, and control, our end-to-end system optimizes all processing steps simultaneously. We argue that this will eventually lead to better performance and smaller systems. Better performance will result because the internal components self-optimize to maximize overall system performance, instead of optimizing human-selected intermediate criteria, e.g., lane detection. Such criteria understandably are selected for ease of human interpretation which doesn't automatically guarantee maximum system performance."

PilotNet CNN model was used in [1], [18] for autonomous driving system only on captured frames of Euro Truck

TABLE I: PilotNet architecture

Layer type	Layer description
Input planes	200 × 200 RGB
2D CNN layer	24 filters, kernel_size=(5, 5), strides=(2, 2) Normalization layer
2D CNN layer	36 filters, kernel_size=(5, 5), strides=(2, 2) Normalization layer
2D CNN layer	48 filters, kernel_size=(5, 5), strides=(2, 2) Normalization layer
2D CNN layer	64 filters, kernel_size=(3, 3), strides=(1, 1) Normalization layer
2D CNN layer	64 filters, kernel_size=(3, 3), strides=(1, 1) Normalization layer
FC layer	100 neurons Normalization layer
FC layer	50 neurons Normalization layer
FC layer	10 neurons Normalization layer
FC layer	1 neuron Normalization layer

Simulator 2. The model architecture is composed of following layers as shown in Table I.

B. Lane Lines Detection

A road lane is a portion of roadway allocated to a single line of vehicles and their movement to control and guide drivers in order to reduce traffic confusion. It is indicated by painted longitudinal lines or embedded markings on roadway pavement, along with road surface markings such as lane direction indicators. Lane lines detection is one of the critical detection problems in the vision navigation systems of the intelligent vehicles [4]. It is a traffic sign with high-level semantics which determines how a vehicle should move and what actions it should perform in traffic (such as switching lanes if the line is dashed or not switching/overtaking if the lane is solid or double solid, etc.).

1) Lane Lines Detection Model: Lane line classification from an image was done by a U-Net-like model [2], [3], called LaneUNet. LaneUNet consists of 4 different stages: downscale, encoder, decoder, and output generator, as shown in Figure 1. The model is configured to work with images with a resolution of (400×400) pixels, with two classes: Lane and Not Lane. The chosen training optimizer for this model was RMSprop, with the loss function defined as categorical cross-entropy.

2) PilotNet model and modifications: The PilotNet [1] model that was inherited from Europilot project [18], has five convolutional layers followed by three fully connected layers. The optimizer used for learning PilotNet is stochastic gradient descent (SGD), while the loss function is mean square error (MSE) in reference to the angle of the recorded steering wheel data The LaneUNet model, with frozen internal layer weights, was added to input layer of PilotNet in three different ways:

- Direct input (1 Channel PilotNet), was made for control purposes as proof that lane lines are not reference data for driving and has smallest size of input data $400 \times 400 \times 1$.
- Overlayed in green channel (3 Channel PilotNet), was introduced as a way to include lane lines without increasing size of input data.
- Additional channel of input (4 Channel PilotNet), adds lane lines as additional channel and has biggest size of input data $400 \times 400 \times 4$.

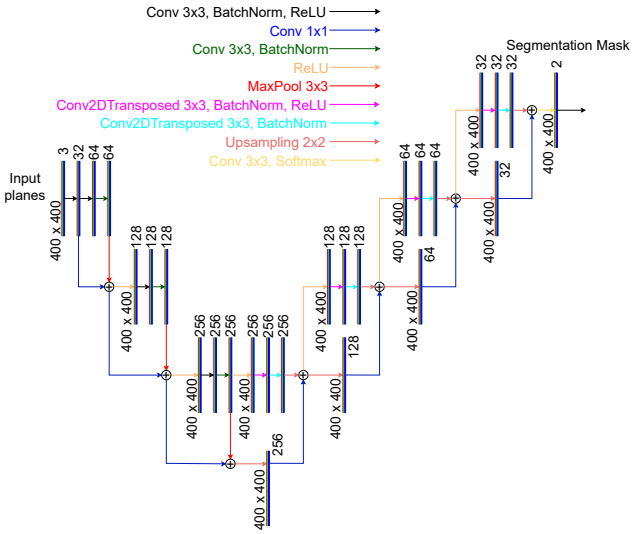


Fig. 1: LaneUNet structure

The original PilotNet model was included for comparison as reference with an input data size as $400 \times 400 \times 3$. A complete overview of all aforementioned modifications done to PilotNet model can be seen in Figure 2.

C. Dataset generation

Both the LaneUNet and PilotNet models requires data for training, since pretrained models were not used. For LaneUNet, a modified and prepared dataset was used, while PilotNet training dataset were recorded.

1) LaneUNet Training Dataset: LaneUNet training data is based on CULane [20] dataset and extended with data augmentation. CULane lane line was plotted with a line width of 13 pixels. Data augmentation was done to increase the size of dataset inspired by the work of Mamun et al. [21] and to increase robustness of the model. Every image in the dataset was kept as original, mirrored randomly by up/down or left/right, randomly cropped, and randomly perspective cropped, also brightness was changed by a random value. The complete number of frames after combining original CULane dataset and augmented frames was 668021 frames. Learning and validation subsets were divided from this combined number of frames in an $(80 : 20)$ ratio.

2) PilotNet Training Dataset: Data for training PilotNet models was recorded using ETS2 and a steering wheel Controller (Thrustmaster F430) by utilizing Europilot [18] data generation script. While recording data, the ETS2 internal resolution was set to 1152×864 and the graphical

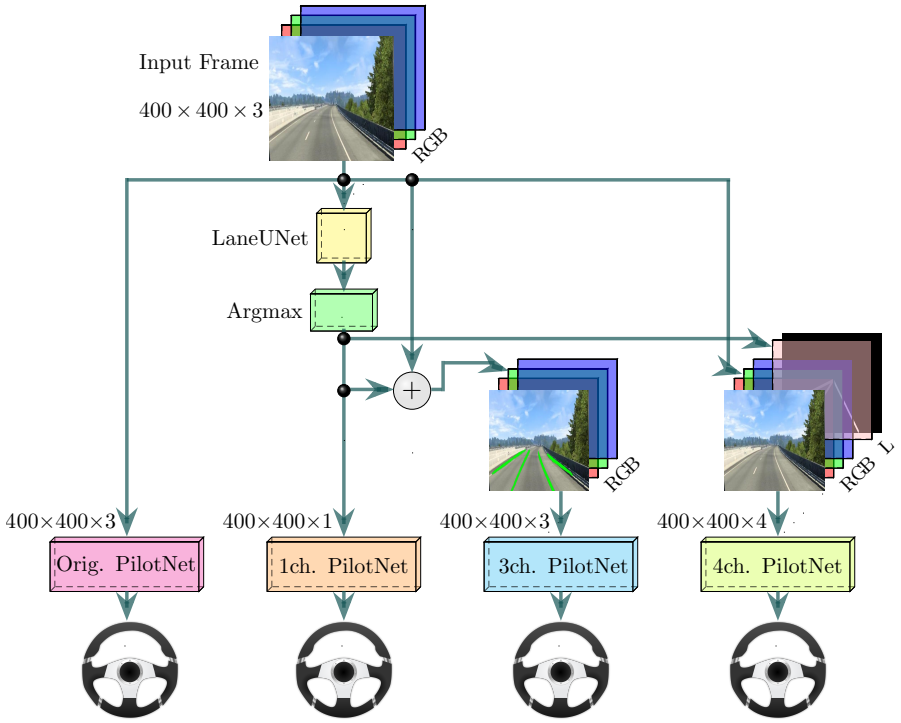


Fig. 2: PilotNet Modifications Overview

TABLE II: Rough estimate: recorded data locations and conditions

Location \ Condition	Day& Clear	Day& Rain	Night& Clear
Urban	7.86%	0%	5.96%
Offroad	3.46%	0%	1.41%
Contry road	1.12%	0%	0.91%
Highway	34.48&	6.21%	38.64%

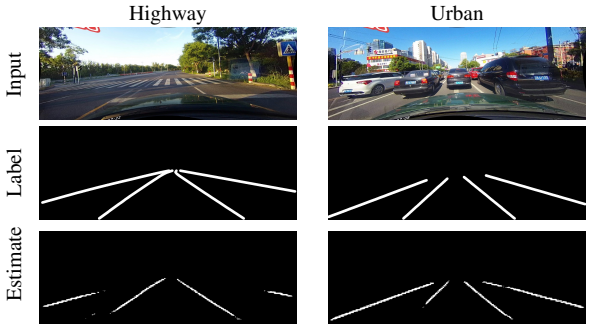
level of details was set to "High" in order to make recorded frames close to real-life frames of CULane dataset. The recorded data consists of 249190 frames captured at rate of 25 frames per second (fps) which equals of 2 hours and 46 minutes. Dataset was divided into training and validation subsets by 80 : 20 ration. Recorded data consist of multiple driving conditions as shown in Table II. After recording, frames were cropped to include only road view out of the vehicle cabin as suggested in the Europilot [18] project post processing script. After recording, frames were cropped to include only road view out of vehicle cabin as suggested in Europilot [18] project post processing script . Captured steering wheel data contains three values: steering angle, brake and acceleration pedal position. For the purpose of this paper only steering angle values were used in order to simplify the autonomous driving problem.

IV. Experiments

A. LaneUNet Evaluation

Examples for the evaluation of final LaneUNet model after 8935 epochs with an evaluation loss of 0.0598 and a validation accuracy of 97.94229% on a validation part of dataset is shown in Table III. Additional evaluation results of the model on ETS2 frames are shown in Table IV. It is obvious that model was not able to detect lane lines in off-road scenarios. This comes from the fact that CULane does not include frames that are similar to off-road scenarios from ETS2.

TABLE III: Evaluation Example: LaneUNet on CULane



B. PilotNet evaluation

PilotNet and its modifications were trained for 320 epochs. Neural network models for evaluation in the simulator were chosen based on the best results of validation loss and accuracy value shown in Table V, and additionally for each model epochs 100, 200 and 320 were also included. These models were chosen as a uniform distribution through learning epochs. This totals 20 models for live evaluation tests in ETS2. Due to limitations of ETS2, evaluation tests of each neural model were done

TABLE IV: Evaluation Example: LaneUNet on ETS2

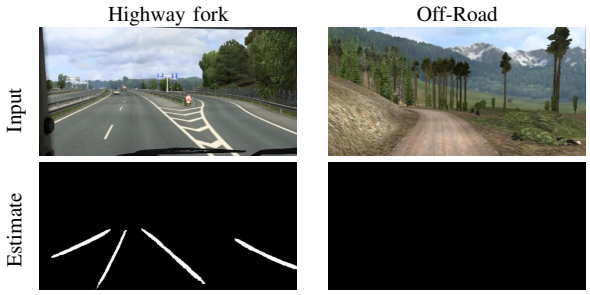


TABLE V: PilotNet Models Training results

Model	Value	Loss	Accuracy	Validation Loss	Validation Accuracy
Original PilotNet	epoch	305	305	315*	307**
	value	1.13104	0.21225	0.75563	0.26586
1 Channel PilotNet	epoch	170	1	183*	1**
	value	59.70996	0.01470	58.38287	0.0
3 Channel PilotNet	epoch	320	300	243*	308*
	value	0.21720	0.27705	0.54382	0.27416
4 Channel PilotNet	epoch	320	315	276*	284**
	value	0.29463	0.27160	0.57606	0.27568

by accelerating the vehicle to specific velocity, engaging the built in ETS2 speed cruise control and then activating the neural model for steering input through a virtual steering wheel controller. The evaluation test is manually stopped either when the vehicle hits roadside or it drives off road. There are 4 evaluation locations: Urban, Off-Road, Country Road, Highway. Due to the nature of the simulation, each test was run only once in identical initial conditions and yielded consistent results when repeated. Results of evaluation are shown in Tables VI, VII, VIII and IX. The displayed results show PilotNet models that have LaneUNet pre-marked lane lines in preprocessing have better driving performances. Time value in the table represents the time that model under test is able to drive before hitting a road barrier or making a dangerous action. The result in tables shows the best performing model on Urban location for Day&Rain is 3 Channel PilotNet. For Night&Clear shows that 4 Channel PilotNet is the best. For Night&Rain the best performing network is 4 Channel PilotNet, even though this condition is not present in learning dataset. Original PilotNet is the best performing in Day&Clear conditions.

The bast performing model for Off-Road location is 4 Channel PilotNet, except for Night&Rain condition where 3 Channel PilotNet is slightly better. Original PilotNet has performance values close to the best model in Day&Rain and Night&Rain conditions. 4 Channel PilotNet model is the best in Country Road location at Day&Rain and Night&Clear conditions, while 3 Channel PilotNet is better at Day&Clear and Night&Rain conditions.

The highway location is dominated by 4 Channel PilotNet model except for the Day&Rain condition where the Original PilotNet model is slightly better.

TABLE VI: PilotNet Models Evaluation Results, Location: Urban

Location: **Urban**	Conditions:	Day & Clear	Day & Rain	Night & Clear	Night & Rain
Model	Epoch	Time [s]			
Original PilotNet	100	3.83834	2.71345	10.96311	14.28656
	200	11.23782	2.96604	11.00926	14.27493
	320	4.09575	2.71534	10.96144	26.98413
	307**	**16.23458**	16.55962	14.14848	14.38183
	315*	4.35174	16.51747	14.09681	14.34698
3 Channel PilotNet	100	7.13	3.43547	10.19311	11.28226
	200	9.24653	3.65635	11.05906	11.22218
	320	7.46921	3.26165	14.05149	10.92544
	243*	9.2222	**17.83614**	10.09567	10.61436
	308**	9.008	12.51868	14.15266	11.36733
4 Channel PilotNet	100	10.88346	4.5254	11.0694	**45.60521**
	200	10.67793	12.44566	**18.45405**	30.41055
	320	3.12718	7.26811	11.49798	17.51122
	276*	3.89543	10.67699	11.26646	17.54752
	284**	2.98151	9.71796	11.06521	17.72489

TABLE VII: PilotNet Models Evaluation Results, Location: Off-Road

Location: **Off-Road**	Conditions:	Day & Clear	Day & Rain	Night & Clear	Night & Rain
Model	Epoch	Time [s]			
Original PilotNet	100	54.20072	53.19633	6.55412	7.84882
	200	57.95325	92.22533	9.05661	21.15158
	320	53.32286	12.84984	6.95825	20.03792
	307**	53.11098	53.8781	3.49425	15.174
	315*	52.54838	54.28565	3.39048	12.439
3 Channel PilotNet	100	52.34927	53.25124	12.75766	19.81261
	200	16.60616	52.3291	21.84162	**21.41379**
	320	19.10798	47.94633	19.51536	20.50395
	243*	50.51435	45.65567	18.86555	21.32946
	308**	19.47457	53.00088	21.68742	21.01616
4 Channel PilotNet	100	7.75637	53.66708	**21.86947**	17.89866
	200	**133.53286**	38.02302	25.09455	3.60717
	320	52.75734	91.98988	21.74667	3.98164
	276*	57.83613	**92.43867**	6.53626	4.50465
	284**	14.25077	3.34132	21.65484	3.69087

V. Conclusion

Comparing different modifications of PilotNet neural network model [1], usage of detected lane lines based on LaneUNet model [2], [3], and its implications on vehicle autopilot driving performance is the topic of this paper.

The first modification is based on sending the results of LaneUNet directly to the PilotNet model, replacing its original input. The second modification overlays lane lines detection results with the green channel of raw input frames from ETS2. The third modification adds extra channel to the input data of PilotNet which consists of LaneUNet detection results on ETS2 raw frames. The Original PilotNet model is used as reference during training and evaluation. Overall, the best result is obtained by the third modification of PilotNet, followed by the second modification, the Original PilotNet and lastly the first modification. ETS2 is used to train and validate PilotNet

TABLE VIII: Evaluation Results, Location: Country Road

Location: **Country Road**	Conditions:	Day & Clear	Day & Rain	Night & Clear	Night & Rain
Model	Epoch	Time [s]			
Original PilotNet	100	93.59565	93.66441	35.67297	9.7689
	200	94.48444	91.17316	47.53247	9.88838
	320	95.95082	86.93309	47.17973	38.77904
	307**	96.15499	54.81467	47.33901	48.01871
	315*	49.66787	95.80688	47.79666	48.64261
3 Channel PilotNet	100	**111.42274**	93.6237	38.54933	12.17085
	200	93.86836	98.7547	38.62285	87.21712
	320	16.3591	98.97377	39.31017	**88.75275**
	243*	92.71974	98.42113	38.49544	18.61541
	308**	99.97491	98.56087	37.84902	85.84725
4 Channel PilotNet	100	53.6726	**103.0861**	**68.14147**	75.26335
	200	60.62461	79.25365	67.89574	67.955
	320	55.4819	92.57005	14.27094	67.53863
	276*	60.73591	85.66176	69.02957	31.79299
	284**	60.68326	95.33997	67.81089	68.2345

TABLE IX: PilotNet Models Evaluation Results, Location: Highway

Location: **Highway**	Conditions:	Day & Clear	Day & Rain	Night & Clear	Night & Rain
Model	Epoch	Time [s]			
Original PilotNet	100	22.89169	2.33809	41.45732	22.26164
	200	23.93349	2.86061	41.88455	30.73024
	320	24.81868	**5.97896**	41.3217	31.16824
	307**	7.82739	2.56142	41.61631	9.56341
	315*	24.05545	2.43687	41.03164	30.54046
3 Channel PilotNet	100	2.69504	2.56999	23.25907	25.09734
	200	2.71988	2.59072	29.00045	24.95291
	320	2.00677	2.36893	59.41756	25.29419
	243*	4.23775	2.22972	59.00256	23.57244
	308**	2.25181	2.46994	59.54084	24.61422
4 Channel PilotNet	100	2.48298	2.87298	62.71282	**31.1772**
	200	14.85202	5.44053	41.97563	4.46748
	320	11.49844	3.03276	26.73725	3.47817
	276*	14.99146	3.66295	**89.29299**	5.14221
	284**	**60.05419**	3.14683	10.62356	6.1917

models on four different locations. The LaneUNet model, used for lane line detection, is trained on augmented CULane dataset.

In future work, both models for lane line detection and PilotNet can be improved by including a temporal component such as LSTM or GRU. The learning dataset for the PilotNet model can be further extended by diversifying locations and conditions. Also, instructions for lane line changes and pedal controls (acceleration, braking) can be added.

REFERENCES

[1] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, "End to end learning for self-driving cars," 2016. [Online]. Available: https://arxiv.org/abs/1604.07316

[2] F. Chollet, Deep Learning with Python, Second Edition. Manning, 2021. [Online]. Available: https://books.google.hr/books?id=mjVKEAAAQBAJ

[3] ——, Apr 2020. [Online]. Available: https://keras.io/examples/vision/oxford_pets_image_segmentation/

[4] T. Zheng, Y. Huang, Y. Liu, W. Tang, Z. Yang, D. Cai, and X. He, "Clrnet: Cross layer refinement network for lane detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2022, pp. 898–907.

[5] Y. Dong, S. Patil, B. van Arem, and H. Farah, "A hybrid spatial–temporal deep learning architecture for lane detection," Computer-Aided Civil and Infrastructure Engineering, vol. 38, no. 1, pp. 67–86, feb 2022. [Online]. Available: https://doi.org/10.1111%2Fmice.12829

[6] L. Liu, X. Chen, S. Zhu, and P. Tan, "Condlanenet: a top-to-down lane detection framework based on conditional convolution," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 3753–3762, 2021.

[7] Q. Qiu, H. Gao, W. Hua, G. Huang, and X. He, "Priorlane: A prior knowledge enhanced lane detection approach based on transformer," 2022. [Online]. Available: https://arxiv.org/abs/2209.06994

[8] D. Jin, W. Park, S.-G. Jeong, H. Kwon, and C.-S. Kim, "Eigenlanes: Data-driven lane descriptors for structurally diverse lanes," 2022. [Online]. Available: https://arxiv.org/abs/2203.15302

[9] Y. He, M. Jiang, X. Ye, L. Du, Z. Zou, W. Zhang, X. Tan, and E. Ding, "Repainting and imitating learning for lane detection," 2022. [Online]. Available: https://arxiv.org/abs/2210.05097

[10] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, "End to end learning for self-driving cars," 2016. [Online]. Available: https://arxiv.org/abs/1604.07316

[11] K. Chitta, A. Prakash, B. Jaeger, Z. Yu, K. Renz, and A. Geiger, "Transfuser: Imitation with transformer-based sensor fusion for autonomous driving," 2022. [Online]. Available: https://arxiv.org/abs/2205.15997

[12] R. Chekroun, M. Toromanoff, S. Hornauer, and F. Moutarde, "Gri: General reinforced imitation and its application to vision-based autonomous driving," 2021. [Online]. Available: https://arxiv.org/abs/2111.08575

[13] H. Shao, L. Wang, R. Chen, H. Li, and Y. Liu, "Safety-enhanced autonomous driving using interpretable sensor fusion transformer," 2022. [Online]. Available: https://arxiv.org/abs/2207.14024

[14] P. Wu, X. Jia, L. Chen, J. Yan, H. Li, and Y. Qiao, "Trajectory-guided control prediction for end-to-end autonomous driving: A simple yet strong baseline," 2022. [Online]. Available: https://arxiv.org/abs/2206.08129

[15] M. Toromanoff, E. Wirbel, and F. Moutarde, "End-to-end model-free reinforcement learning for urban driving using implicit affordances," 2019. [Online]. Available: https://arxiv.org/abs/1911.10868

[16] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang et al., "End to end learning for self-driving cars," arXiv preprint arXiv:1604.07316, 2016.

[17] M. Bojarski, P. Yeres, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, and U. Muller, "Explaining how a deep neural network trained with end-to-end learning steers a car," arXiv preprint arXiv:1704.07911, 2017.

[18] MARSAUTO, "Europilot." [Online]. Available: https://github.com/marsauto/europilot

[19] SCS Software, "Euro truck simulator 2," DVD,Steam, 2012.

[20] P. Xingang, S. Jianping, L. Ping, W. Xiaogang, and X. Tang, "Spatial as deep: Spatial cnn for traffic scene understanding," in AAAI Conference on Artificial Intelligence (AAAI), February 2018.

[21] A. A. Mamun, E. P. Ping, J. Hossen, A. Tahabilder, and B. Jahan, "A comprehensive review on lane marking detection using deep neural networks," Sensors, vol. 22, no. 19, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/19/7682

Estimating a nonradial vignetting shape

Dorotea Potoč, Davor Petrinović
University of Zagreb
Faculty of Electrical Engineering and Computing
dorotea.potoc@fer.hr, davor.petrinovic@fer.hr

***Abstract*—Vignetting is a phenomenon characterized by a decrease in illumination towards the edges of an image. This effect is typically represented by a radially symmetrical model, however, this paper aims to demonstrate a non-radial model of vignetting and estimate its shape. To accomplish this, a synthetic image was created and the angular vignetting shape has been modeled as a sum of harmonics. The magnitudes and amplitudes of these harmonics were obtained and used to construct the desired angular vignetting shape. Once the synthetic image with the modeled vignetting shape and added noise was created, it was used as input into a function for vignetting estimation. Also, the inputs have been a fixed vignetting center and different initial values of harmonics' magnitudes and phases. With that inputs, despite the level of the noise, we have successfully estimated vignetting function by non-linear optimization. The function has attempted to determine the original harmonics' used to create the vignetting angular shape. When the vignetting model is calculated, we removed it in order to get a homogeneous image. While it may be difficult to obtain the exact original values of the harmonics', the shape can be estimated with a high level of accuracy. The paper shows that highly accurate models can be estimated for a lower number of angular harmonics, with a residual gain error standard deviation of less than 0.03%. Even in the presence of 5dB noise in the images, the gain error standard deviation remains below 3%, as long as proper parameter initialization is performed prior to optimization.**

***Keywords*—vignetting, non-radial model, vignetting correction, optimization**

I. Introduction

Vignetting is a photometric phenomenon in which the brightness of an image decreases towards its periphery, relative to its center. The cause of vignetting can be due to limitations in camera and lens design, such as the aperture value, or the geometric distortion of oblique light beams. It can introduce inaccuracies in computer vision algorithms that rely on precise intensity data. Vignetting correction techniques range from commercial to non-commercial solutions, but the majority assume radial distribution [1], [2], leading to potential inaccuracies if the vignetting is actually non-radial. The degree of vignetting is influenced by factors such as lens, aperture, and exposure, making automatic correction challenging [3].

The characteristics of the lens and its aperture determine the degree of vignetting. Natural vignetting occurs due to the cosine fourth law of illumination fall-off, which states the decrease in light is proportional to the fourth power of the cosine of the angle between peripheral lens rays and the optical axis. The aperture, or diaphragm, regulates the amount of light reaching the sensor and controls the f-stop, a logarithmic measure of the ratio of maximum to minimum intensity [4].

Vignetting has four different types: optical vignetting is a lens-based phenomenon of gradual darkening in images that are more prominent at wider apertures. It can be corrected by reducing the aperture size [5]. Natural vignetting is caused by the light reaching the sensor at different angles, particularly with wide-angle lenses, and is a result of the cosine fourth law of light fall-off [6]. It cannot be corrected with aperture adjustment but can be compensated through filters or image processing. Mechanical vignetting results from physical obstruction of light by filter rings, hoods, or other objects and is less noticeable with wider apertures and zoom lenses. Pixel vignetting is due to lower illumination of corner pixels compared to center pixels, which is a result of the flat construction of image sensors and the angle of the light impinging on the sensor. It can be reduced by using micro-lenses on the sensor and perpendicular angles, though it cannot be eliminated entirely [7], [8].

The vignetting can be reduced using radial graduated neutral density filters or by post-processing methods. Some cameras have a built-in vignetting correction for JPEG images, but not for RAW images. Image processing can be used to enhance images and correct vignetting, but various lens-related factors such as angle of incidence, chief ray angle (CRA) misalignment, and natural vignetting must be considered. Correction methods for RAW images require specialized software. Image processing methods can also enhance the quality of the image. Most causes of vignetting can be reduced by decreasing the lens aperture by 2 f-stops, or through the use of a telecentric lens that produces uniform illumination of the image plane. Mechanical vignetting can be reduced by using longer focal lengths, using a telecentric lens, or using flat field correction.

Any optical system consisting of multiple lenses can cause non-radial vignetting. This occurs because lenses typically have around 20 elements, and their misalignment with the optical axis can lead to non-radial vignetting. Specifically, due to the mechanical and lens imperfections and for lenses with variable focus. Even high-end lenses may not have a perfect internal construction, which can

This work was supported in part by the Croatian Science Foundation under Project IP-2019-04-6703 and DOK-2021-02-6215 and in part by the European Regional Development Fund under Grant KK.01.1.1.01.0009 (DATACROSS). Any opinions, findings, and conclusions to recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Croatian Science Foundation.

lead to this phenomenon. For example, there is one lens that has 17 elements divided into 12 groups [9], it is a very high-rated lens, but can still cause this phenomenon.

Non-radial vignetting can also occur in scenarios where the illumination is not sufficiently homogenous.

II. Related works

Vignetting correction methods aim to eliminate the darkening of image corners caused by the vignette effect. These techniques can be broadly divided into two groups: those that use a reference image and those that do not. The former group relies on a reference image to determine a vignetting function, which is obtained by approximating the function with parametric models.

Physically-based models within this group require detailed knowledge of camera lens parameters and can be challenging to implement [10], [11]. Single-image [12], [13], [14] and image sequence methods [15], [16], [17], [18] within this group estimate vignetting by minimizing an objective function with the assumption that vignetting is a radial function. These methods often require additional image processing techniques, such as image segmentation [1].

The effectiveness of methods that use a single image to estimate a vignetting function depends on the precision of localization of corresponding pixels and usually use additional image processing methods, eg. image segmentation. Most single-image correction methods do not require prior knowledge of the vignette model or optical system and can be achieved through: radial and tangential gradient analysis [19], Gaussian quadric fitting [20], and estimation of correction factors at each pixel position [21]. Some of these methods also take into consideration pupil aberration and the symmetric distribution of the radial gradient [22], [23]. In most cases, all compared images require an acquisition in the same scene conditions and any change in the scene may influence the outcome vignetting function. The effectiveness of these methods strongly depends on the uniformity of scene illumination.

In the image sequence method, they use a set of not entirely overlapping images of the same reference scene to calculate vignetting function. It is done by minimizing the objective function which depends on the differences between values of corresponding pixels in different images, which represent the same scene point. Multiple-image correction techniques may require prior knowledge of the optical system or vignetting model and use information from multiple images to estimate a vignetting model. These methods can be model-based or otherwise utilize techniques such as wavelet decomposition and hyperbolic function approximation

Methods that do not use reference images are based on flat-field correction and assume a reference vignetting image that represents a uniformly illuminated surface. There can be used different types of vignetting models, such as 2D polynomial [15], [24], [25], exponential 2D polynomial [24], smooth non-iterative local polynomial [26], radial polynomial [27], hyperbolic cosine [28], and Gaussian function [29]. Most models represent a radial model, but there are only few methods that cover non-radial models, and are easy to use.

Type of Methods		Reference
With reference image		
Physically based model		
	Natural, optical, pixel	[10], [11]
Single image:		[12] - [14]
	Radial and tangential gradient analysis	[19]
	Gaussian quadric fitting	[20]
	Estimation of correction factors	[21],
	Non-radial methods	[3], [30], [31]
Image mosaic:		
	Model-based	[17] , [18]
	Wavelet decomposition	[2]
	Hyperbolic function approximation	[15]
Without reference image		
Flat field:		
	2D polynomial	[24], [25]
	Exponential 2D polynomial	[24]
	Smooth non-iterative local polynomial	[26]
	Radial polynomial	[27]
	Hyperbolic function	[28]
	Gaussian function	[29]

TABLE I: Type of methods for vignetting correction

A. Non-radial vignetting correction methods

Non-radial vignetting correction methods are mostly done as single-image vignetting correction. There are a few methods that deal with this type of vignetting.

Non-radial vignetting correction described in [3] is based on a local parabolic model of vignetting. This algorithm includes compensation for non-uniform scene luminance. The method shows better results on artificial images than on natural ones. The usual quality metrics for luminance compensation accuracy are MAE and RMSE measures.

Another approach, Deformable Radial Polynomial (DRP) model combines the simplicity of the commonly used radial polynomial(RP) model with the universality of more complex methods. This model uses a distance transformation and minimization method to match the radial vignetting model to the non-radial vignetting of the analyzed lens-camera system and can give better results than the RP model [30].

An alternative approach of local fitting of the vignetting model to the measure data [31] is based on the local polynomial model in which the order of the polynomial is a parameter of the model and allows to better fit the model to the real vignetting of camera-lens systems.

Our preliminary results are not yet comparable to the results obtained by existing methods. Further research and optimization are necessary to fully evaluate the potential of the proposed approach.

III. Proposed non-radial model of vignetting

Radial vignetting shown in Figure 1 is the most widely recognized form of vignetting and has been well studied, leading to the development of numerous correction algorithms that take its radial shape into account. However, non-radial vignetting, shown in Figure 2, is a more complex form of vignetting that is more difficult to quantify and

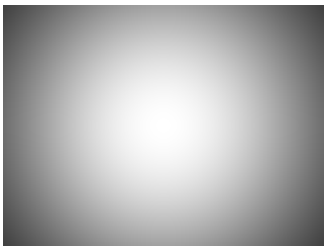


Fig. 1: Radial model of vignetting [32]

correct. This is due to the lack of a uniform and systematic pattern in the reduction of brightness or saturation across the image plane. Non-radial vignetting correction methods are still an active area of research in the field of image processing.



Fig. 2: Example of a non-radial model of vignetting

The general form of vignetting can be expressed mathematically as a function of the normalized radius, as shown in equation 1.

$$V = f(R') \tag{1}$$

In the proposed model, in the equation 2, the normalized radius is defined by multiplying the actual radius by an angular dependent factor k(θ) that accounts for the desired non-radial shape of the vignetting.

$$R' = \frac{R}{maxR} \cdot k(\theta) \tag{2}$$

The factor k(θ) is a function of the polar angle, θ, and allows for the creation of a non-uniform reduction of image brightness or saturation across the image plane. The non-radial form of vignetting can be used to model more complex forms of vignetting beyond the traditional radial model, allowing for greater control over the shape and intensity of the vignetting effect.

For a radially symmetric vignetting model, the factor k(θ) is a constant, meaning that the reduction in image brightness or saturation is uniform in all directions. In contrast, for a non-radially symmetric vignetting model, the factor k(θ) varies as a function of the polar angle, θ, which allows for the creation of non-uniform vignetting patterns that fall off at different rates in different directions. The larger the value of k(θ), the more rapid the reduction of brightness or saturation in that direction. By modeling vignetting in this manner, it is possible to create more complex and nuanced vignetting effects.

$$k(\theta) = 1 + \sum_{n=1}^{Nh} (ma_i \cdot \cos(\arctan(\frac{y - y_0}{x - x_0}) \cdot i + an_i) \tag{3}$$

The proposed angular term k(θ) of the non-radial vignetting model is determined as a simple sum of the harmonics, as expressed in equation 3 where (x, y) is the pixel position in the frame, (x_0, y_0) is the center of the shifted vignetting model, and R is the euclidean distance of these two points. Such a definition ensures natural periodicity with the period of 2π, thus achieving the angular smoothness of the model. In this equation, ma_i represents the magnitude and an_i represents the angle of each harmonic term. The sum of these harmonics results in the final expression for angular dependent normalized radius R'. The proposed approach allows for a flexible parametric representation of the vignetting effect and the ability to control the shape and intensity of the vignetting by adjusting the magnitude and angle of each harmonic term. By combining multiple harmonics, it is possible to create complex and non-uniform vignetting patterns that cannot be represented by simple radial models.

IV. Proposed method of non-radial vignetting correction

A. Synthetic image of non-radial model

Since such a model has not been studied in the literature so far, before comparing it to other models and before its validation on real-world images it was necessary to establish the feasibility of model estimation from synthetic images with given model parameters. This represents the main research objective presented in this paper.

To validate the vignetting correction function, a synthetic image, shown in Figure 3 was created as follows: 1) a homogeneous gray image was established as the background to simulate the uniform part of the night sky glow, 2) white pixels simulating stars were added to the constant value of the image, 3) the described non-radial vignetting model was then applied to the image, and 4) Poisson noise was applied to the image to simulate the realistic noise present in images captured by cameras. The presence of noise can make it difficult to accurately estimate the form of vignetting, potentially leading to discrepancies between the estimated and actual vignetting models, especially when the sensor's response to sky glow intensity is of a similar magnitude to the expected noise level. Therefore, it is essential to consider the impact of noise when estimating the vignetting correction, as it can further exacerbate the vignetting correction problem.

B. Function for vignetting correction

The function operates based on the principle of pixel sensing and attempts to identify the shape of vignetting by searching for areas where the function assumes the same intensity values. In simple radial models, search is performed along contour lines of equal intensity which are

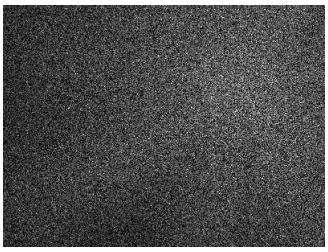


Fig. 3: Example of synthetic image with non-radial vignetting, Nh = 9

represented by concentric circles. Based on the proposed angular model from equation 3 with assumed harmonics values, the contour lines become non-circular as shown in Figure 4 for the same example from Figure 3.

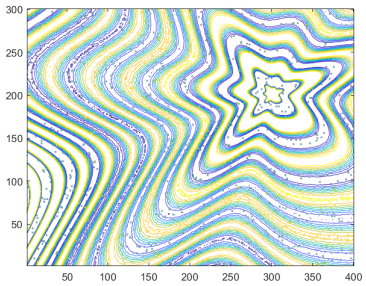


Fig. 4: Contour lines of non-circular model, Nh = 9

The radial cut of the normalized vignetting function f(R') is then determined by averaging across all angles θ. Introducing such angular dependent scaling of radial distance to the vignetting model's center is equivalent to the image frame spatial distortion in which the vignetting of the original undistorted image would be a simple radial model of the normalized radius R'. The original and distorted frames are shown in Figure 5.

Ideally, the vignetting estimation function shall result with the optimal set of parameters: model center (x_0,y_0), angular model harmonics (ma_i and an_i), and the normalized radial cut function V(R'), using the non-linear optimization for the chosen loss function. Since this is a non-trivial problem we have experimented with a different kind of parameter initialization which affect the final solution.

The inputs to the optimization function are thus initial magnitudes and phases of harmonics and the vignetting center point which is either assumed as a free optimization variable as well or is fixed to the actual known position to simplify the problem. The radial cut function is also modeled as a harmonic function of the normalized radius R' with the chosen number of radial harmonics (20 in all our experiments). The parameters of this averaged radial harmonic model are always determined as an optimal solution for the current parameters of the assumed angular vignetting model k(θ).

In this paper, we will present the estimation results for the fixed center position in order to validate the possibility of estimating an accurate angular model for such a simplified case.

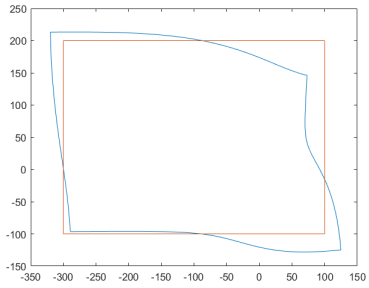


Fig. 5: Distorted shape of the frame due to angular dependent radial scaling, Nh=2

The estimation process is inherently nonlinear, rendering the task of identifying the optimal solution intractable through analytical methods. A reliable estimate of the model will be obtained if the application of the reciprocal model to the image results in a uniform brightness across the entire frame, with minimal variations. To minimize these variations in consideration of the model's parameters, numerical optimization was performed in Matlab utilizing the fminunc function, as well as a similar function developed within our department.

V. Experiments and Results

Experimental verification of vignetting estimation and compensation was done on the synthetic image (300x400 pix) with x_0 = 200, y_0 = 150. In order to evaluate our estimation methodology, the number of harmonics of the angular model was varied from the set Nh = {1, 2, 9}, so we can validate the estimation function behavior for a small number of harmonics, as well as a larger value. To verify the robustness of the final estimation result to the proper parameter initialization, optimization was initialized with the scaled values of the ideal harmonics magnitudes and phases which were used for the image synthesis. The same scalar scaling factor α was used for all initial parameters, which was chosen from the set α = {1, 0.7, 0.4, 0.2, 0, -1}. Every experiment was performed for the noise-free synthetic image as well as for the Poisson corrupted image with an SNR of 20dB and 5 dB. The normalized radial cut function of the synthetic image was chosen as V(R') = tan($\frac{\pi}{4}$ - $\frac{\pi}{8}$ $\frac{R'}{maxR'}$), such that V(0) = 1, and V(maxR') = 0.4142.

For the trivial scale factor α = 1, initial optimization parameters were identical to the parameter set used for synthesis which truly minimizes the objective function so this case represents the baseline for comparison.

A. Evaluation metrics

The success of the vignetting estimation and compensation was measured by two factors: standard deviation (std) of the residual gain variation to the unit value and the percentage of valid pixels (valid) whose gain variation is within +/-3std (to account for inevitable outliers). The robust estimate of the std was used to establish the

validity threshold to account for heavy-tailed gain error distribution. Therefore, both values must be considered when interpreting and comparing the experimental results.

B. Experiments with noise-free images

The estimation results of compensating noise-free images are presented in Figures 6, 7, and 8 for scaling factor α = 0.4. The correction is demonstrated to be very successful with minimal deviation from the unit value of less than 1 permille relative gain variation for Nh = 1. Residual radial fringes are caused by the mismatch between the radial cut model used for the synthesis (tan) and the non-parametric harmonic model which is used in the estimation to account for the arbitrary radial cut shape. The angular model error can be observed for Nh = 9 as angular zoning combined with radial fringes, but the error is still within 1% of the ideal flatness.

The numerical simulation results for all three choices of Nh are presented in Tables II to IV. It can be observed from Table II that for an angular model with a single harmonic, the true model parameters (ma_1, an_1) can be easily resolved for all positive values of scaling factor α. In table III, for a model with 2 harmonics the true parameter values are restored for scaling factor $\alpha \geq 0.7$, while for the most complex model with Nh = 9, most of the parameters are restored to the true value for scaling factor $\alpha \geq 0.7$ (Table IV). However, for smaller or negative scaling of the initial parameters, in all 3 cases, the minimization function becomes trapped in an alternative minimum and approaches its closest values which are not the global minimum.

The reconstruction accuracy of the compensated synthetic image is also presented in Tables II, III and IV in the last three rows through the percentage of valid pixels and corresponding standard deviation of the residual gain variation. The last row displays the required number of optimization iterations, showing that a more complex model as well as poor initialization increases its value. Similar behavior is observed for the compensation accuracy demonstrating the importance of appropriate initialization.

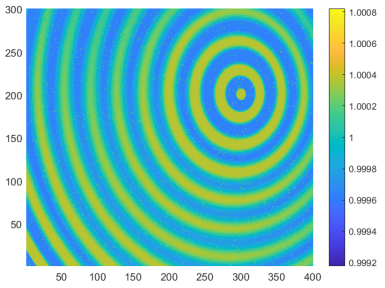


Fig. 6: Residual gain variations for noise-free image, Nh=1, α = 0.4

C. Testing images with SNR = 20dB

The estimation results of the vignetting correction in the image with the signal-to-noise ratio (SNR) of 20 dB is illustrated in Figure 9 for α = 0.7. The compensation error

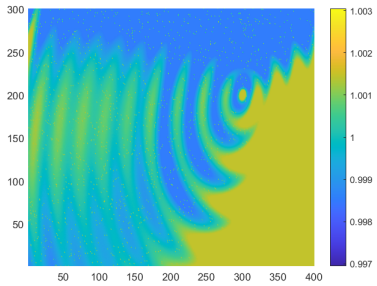


Fig. 7: Residual gain variations for noise-free image, Nh=2, α = 0.4

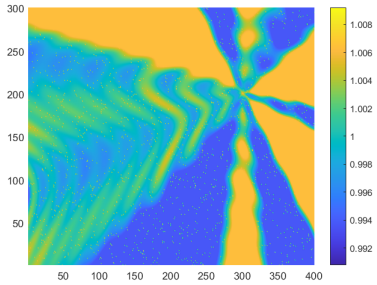


Fig. 8: Residual gain variations for noise-free image, Nh=9, α = 0.4

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0799	0.0799	0.0798	0.0798	0.0798	0.0791	-0.0798
an_1	-0.0666	-0.0666	-0.0680	-0.0676	-0.0687	-0.0689	0.4330
valid %	-	99.200	99.099	99.153	99.151	99.114	99.164
std %	-	0.0306	0.0328	0.0316	0.0350	0.0357	0.0306
Iter num	-	4	32	48	43	63	77

TABLE II: Values of the estimated parameters for scaled initialization for noise-less image, Nh = 1

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0986	0.0986	0.0690	0.1007	0.0795	0.1023	-0.2977
ma_2	0.1819	0.1819	0.1281	0.1971	0.2020	0.1680	-0.0327
an_1	0.1011	0.1011	0.0710	0.0567	0.0448	0.1433	-0.1894
an_2	-0.0703	-0.0703	-0.0503	-0.0728	-0.0779	-0.0672	0.2558
valid %	-	99.095	98.360	80.483	80.713	80.443	72.894
std %	-	0.0297	2.0874	0.1503	0.1690	0.1390	1.1579
Iter num	-	6	18	86	55	146	64

TABLE III: Values of the estimated parameters for scaled initialization for noise-less image, Nh = 2

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0466	0.0466	0.0340	0.0340	-0.0431	-0.0460	-0.0111
ma_2	0.1014	0.1014	0.0945	0.0641	0.0408	-0.0200	-0.0287
ma_3	0.0360	0.0360	0.0164	-0.0141	-0.0067	0.0510	0.0029
ma_4	0.1408	0.1408	0.1318	0.1154	0.1107	0.1673	-0.0404
ma_5	0.0895	0.0895	0.0700	0.0572	0.0046	-0.0206	-0.1819
ma_6	0.0176	0.0176	-0.0098	-0.0951	-0.0234	-0.0570	0.0644
ma_7	0.1566	0.1566	0.1570	0.1288	0.1703	0.1584	-0.0029
ma_8	0.0218	0.0218	0.0066	0.0425	0.0326	-0.0230	0.0970
ma_9	0.0366	0.0366	0.0524	0.0369	0.0652	0.0300	0.0191
an_1	0.4492	0.4492	0.3070	0.1696	0.1059	0.0706	-0.4483
an_2	0.2998	0.2998	0.2399	0.1314	0.1048	-0.0430	-0.2563
an_3	-0.4346	-0.4346	-0.3194	-0.1710	-0.0987	-0.1294	0.4460
an_4	-0.0787	-0.0787	-0.0655	-0.0658	0.0682	0.0265	0.1312
an_5	-0.2185	-0.2185	-0.1720	-0.1298	-0.0842	0.0072	0.1818
an_6	-0.4015	-0.4015	-0.2829	-0.1572	-0.0938	-0.0136	0.3269
an_7	-0.0884	-0.0884	-0.0911	-0.0400	-0.0771	-0.1028	0.1899
an_8	0.4547	0.4547	0.3243	0.1795	0.1059	0.0196	-0.4922
an_9	-0.0004	-0.0004	0.0294	0.0423	0.0244	-0.0105	-0.0235
valid %	-	98.301	80.421	76.681	68.514	69.574	69.040
std %	-	0.0388	0.2673	0.6437	0.2115	0.1508	0.2545
Iter num	-	20	360	303	451	534	565

TABLE IV: Values of the estimated parameters for scaled initialization for noise-less image, Nh = 9

is shown for Nh = 2, but similar results were obtained for

Nh = 1 and Nh = 9 as well. The correction is successful with the maximal deviation from the unit value of slightly above 1% gain error. Detailed results presented in Tables V to VII indicate that the correction is slightly worse than for the noise-free images. The percentage of compensated pixels drops rapidly for $\alpha \leq 0.7$, which gives even higher significance to proper parameter initialization. We can also observe that the higher number of angular harmonics requires a significantly larger number of iterations (e.g. Nh = 9).

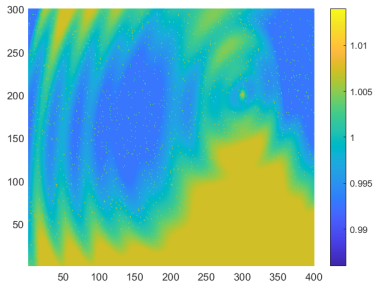


Fig. 9: Residual gain variations for image with SNR = 20dB, Nh=2, α = 0.7

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0799	0.0806	0.0559	0.0319	0.0169	0.0630	-0.0790
an_1	-0.0666	-0.0678	-0.0466	-0.0266	-0.0129	-0.0006	0.0683
valid %	-	97.356	99.962	99.940	99.824	99.838	98.691
std %	-	0.2433	0.6300	1.0944	1.3533	0.7765	2.0206
Iter num	-	9	4	4	12	21	17

TABLE V: Values of the estimated parameters for scaled initialization for image with SNR = 20dB, Nh = 1

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0986	0.1014	0.0704	0.0030	0.0262	0.0125	-0.1579
ma_2	0.1819	0.1828	0.1632	0.1899	0.0937	0.0628	-0.1050
an_1	0.1011	0.1005	0.0915	0.0650	0.0374	0.0002	-0.1451
an_2	-0.0703	-0.0677	-0.0995	-0.0596	0.0076	-0.0009	0.2053
valid %	-	99.164	99.319	100.000	99.946	99.988	100.000
std %	-	0.1641	0.7355	1.4962	3.8717	4.2820	2.6032
Iter num	-	27	37	47	22	24	44

TABLE VI: Values of the estimated parameters for scaled initialization for image with SNR = 20dB, Nh = 2

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0466	0.0472	0.0322	-0.0066	0.0154	0.0191	-0.0974
ma_2	0.1014	0.1008	0.0822	0.0153	0.0284	0.0567	-0.0727
ma_3	0.0360	0.0366	0.0286	-0.0184	-0.0146	0.0080	0.0617
ma_4	0.1408	0.1402	0.1103	0.0535	0.0548	0.0839	-0.0788
ma_5	0.0895	0.0894	0.0597	0.0686	0.0184	-0.0342	-0.1614
ma_6	0.0176	0.0175	-0.0049	-0.0446	-0.0224	-0.0768	0.0536
ma_7	0.1566	0.1562	0.1333	0.0207	0.0397	0.0992	-0.0030
ma_8	0.0218	0.0218	0.0563	0.0423	-0.0009	-0.0781	0.1066
ma_9	0.0366	0.0370	0.0470	0.0611	0.0372	0.0529	0.0238
an_1	0.4492	0.4490	0.3256	0.1566	0.0939	0.0032	-0.4687
an_2	0.2998	0.2996	0.2110	0.1821	0.0432	0.0210	-0.3730
an_3	-0.4346	-0.4345	-0.3093	-0.1735	-0.0892	0.0075	0.4193
an_4	-0.0787	-0.0777	-0.0630	-0.0323	-0.0222	0.0000	0.0955
an_5	-0.2185	-0.2181	-0.1700	-0.0872	-0.0577	0.0335	0.1729
an_6	-0.4015	-0.4015	-0.2704	-0.1629	-0.0944	-0.0348	0.3874
an_7	-0.0884	-0.0882	-0.0652	-0.0586	-0.0371	-0.0269	0.0878
an_8	0.4547	0.4552	0.3213	0.2018	0.0806	0.0102	-0.4813
an_9	-0.0004	-0.0004	0.0075	-0.0063	-0.0041	0.0115	-0.0391
valid %	-	99.141	91.550	98.643	99.827	97.293	95.299
std %	-	0.2548	0.9074	3.2596	4.1844	3.0358	1.7616
Iter num	-	81	105	111	85	184	295

TABLE VII: Values of the estimated parameters for scaled initialization for image with SNR = 20dB, Nh = 9

D. Correction of low SNR images

For the same initial scaling of α = 0.7, the optimization result of the vignetting correction of the low SNR image (SNR = 5 dB) is illustrated in Figure 10 for Nh = 2. The values shown in Tables VIII to X demonstrate that the results are worse than those obtained from images with a higher SNR, or those of the noise-free images, yet the vignetting function can still be approximately reconstructed with magnitudes and angles affected by the high noise level. It is observed that the smaller the number of harmonics, the easier it is to attain desired results, as the model depends on fewer free parameters. Additionally, it can be seen that even for such a high noise level, the percentage of the correctly compensated pixels is in the range of 97-99% with the residual std of around 2% gain error for $\alpha \geq 0.7$. The error clearly increases with poorer initialization.

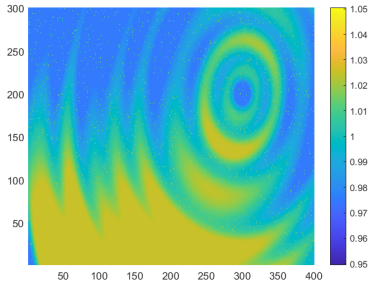


Fig. 10: Residual gain variations for image with SNR = 5dB, Nh=2, α = 0.7

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0799	0.0799	0.0559	0.0329	0.0160	0.0023	-0.0800
an_1	-0.0666	-0.0666	-0.0466	-0.0264	-0.0133	0.0004	0.0751
valid %	-	96.567	97.386	97.969	98.568	99.030	99.596
std %	-	1.2126	1.3729	1.6930	2.0619	2.2066	2.4473
Iter num	-	4	4	12	4	17	20

TABLE VIII: Values of the estimated parameters for scaled initialization for image with SNR = 5dB, Nh = 1

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0986	0.1006	0.0683	0.0394	0.0204	-0.0021	-0.0994
ma_2	0.1819	0.1825	0.1270	0.0728	0.0369	0.0010	-0.1819
an_1	0.1011	0.1011	0.0712	0.0404	0.0203	0	-0.1010
an_2	-0.0703	-0.0669	-0.0474	-0.0281	-0.0135	0	0.0698
valid %	-	96.961	99.742	99.857	99.944	100.000	99.998
std %	-	0.8132	2.3364	3.9619	4.7796	5.3943	6.6510
Iter num	-	20	19	6	18	19	25

TABLE IX: Values of the estimated parameters for scaled initialization for image with SNR = 5dB, Nh = 2

VI. Conclusion

The new non-radial vignetting model was introduced in the paper based on a low-order harmonic representation of the angular-dependent vignetting shape. The feasibility of model estimation through non-linear optimization was investigated and the influence of the proper parameter initialization on the estimated model accuracy was evaluated by comparing the estimated model to the synthetic one. The effect of image noise level on reconstruction accuracy was also considered. The presented results are

param	true\α	1	0.7	0.4	0.2	0	-1
ma_1	0.0466	0.0459	0.0330	0.0175	0.0093	0.0001	-0.0465
ma_2	0.1014	0.1005	0.0709	0.0400	0.0203	0.0002	-0.1018
ma_3	0.0360	0.0371	0.0254	0.0142	0.0072	0	-0.0366
ma_4	0.1408	0.1399	0.0984	0.0559	0.0282	0.0008	-0.1409
ma_5	0.0895	0.0901	0.0626	0.0355	0.0179	0.0002	-0.0896
ma_6	0.0176	0.0144	0.0121	0.0071	0.0035	-0.0004	-0.0180
ma_7	0.1566	0.1560	0.1097	0.0625	0.0313	-0.0005	-0.1564
ma_8	0.0218	0.0219	0.0158	0.0088	0.0044	-0.0008	-0.0219
ma_9	0.0366	0.0368	0.0257	0.0145	0.0073	0.0005	-0.0363
an_1	0.4492	0.4496	0.3148	0.1800	0.0898	0	-0.4484
an_2	0.2998	0.2984	0.2103	0.1191	0.0600	0	-0.2992
an_3	-0.4346	-0.4351	-0.3041	-0.1744	-0.0869	0	0.4350
an_4	-0.0787	-0.0781	-0.0549	-0.0322	-0.0157	0	0.0781
an_5	-0.2185	-0.2181	-0.1530	-0.0876	-0.0437	0	0.2191
an_6	-0.4015	-0.4022	-0.2811	-0.1607	-0.0803	0	0.4019
an_7	-0.0884	-0.0895	-0.0620	-0.0359	-0.0177	0	0.0883
an_8	0.4547	0.4541	0.3184	0.1817	0.0909	0	-0.4545
an_9	-0.0004	-0.0002	-0.0003	-0.0003	-0.0001	0	0.0007
valid %	-	97.850	97.970	98.882	98.677	98.366	98.635
std %	-	1.5375	2.6741	4.3433	5.4530	6.2868	9.4832
Iter num	-	125	60	61	20	60	81

TABLE X: Values of the estimated parameters for scaled initialization for image with SNR = 5dB, Nh = 9

so far limited to the simplified case with a priori known origin of the model. The paper shows that highly accurate models can be estimated especially for a lower number of angular harmonics with residual gain error std of less than 0.03%. Even for images corrupted with 5dB noise the gain error std is still below 3% with proper parameter initialization prior to optimization. Such results are very encouraging and in our future work, the model origin estimation will also be considered, as well as appropriate analytical methods for parameter initialization. The higher-order models with a larger number of angular harmonics are more challenging for estimation as was illustrated for the case with 9 harmonics. The study of non-radial vignetting and its correction provides valuable insights into the nature of vignetting and the potential for new and improved correction methods.

REFERENCES

[1] Y. Zheng, S. Lin, C. Kambhamettu, J. Yu, and S. B. Kang, "Single-image vignetting correction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2243–2256, 2009.

[2] S. J. Kim and M. Pollefeys, "Robust radiometric calibration and vignetting correction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 562–576, 2008.

[3] A. Kordecki, H. Palus, and A. Bal, "Practical vignetting correction method for digital camera with measurement of surface luminance distribution," *Signal, Image and Video Processing*, vol. 10, pp. 1417 – 1424, 2016.

[4] S. Cushing, "Optical vignetting," https://u42.co/Tech-Blog/Vignetting/Vignetting.html, [Online; accessed 21.12.2022.].

[5] RED-101, "Understanding lens vignetting," https://www.red.com/red-101/lens-vignetting, [Online; accessed 21.12.2022.].

[6] P. Kumar, "What is lens vignetting - and how to eliminate it in embedded cameras?" https://www.e-consystems.com/blog/camera/technology/what-is-lens-vignetting-and-how-to-eliminate-it-in-embedded-cameras/, 2022, [Online; accessed 21.12.2022.].

[7] N. Mansurov, "What is Vignetting," https://photographylife.com/what-is-vignetting, 2020, [Online; accessed 21.12.2022.].

[8] W. Emery and A. Camps, "Optical imaging systems," in *Introduction to Satellite Remote Sensing*. Elsevier, 2017, pp. 85–130. [Online]. Available: https://doi.org/10.1016/b978-0-12-809254-5.00003-8

[9] B. Carnathan, "Sigma 105mm f/1.4 DG HSM Art Lens Review," https://www.the-digital-picture.com/Reviews/Sigma-105mm-f-1.4-DG-HSM-Art-Lens.aspx, 2018, [Online; accessed 20.03.2023.].

[10] N. Asada, A. Amano, and M. Baba, "Photometric calibration of zoom lens systems," in *Proceedings of the 1996 International Conference on Pattern Recognition (ICPR '96) Volume I - Volume 7270*, ser. ICPR '96. USA: IEEE Computer Society, 1996, p. 186.

[11] P. B. Catrysse, X. Liu, and A. E. Gamal, "Qe reduction due to pixel vignetting in cmos image sensors," in *Electronic imaging*, 2000.

[12] T. Rohlfing, "Single-image vignetting correction by constrained minimization of log-intensity entropy," Sep 2012.

[13] Y. Zheng, C. Kambhamettu, and S. Lin, "Single-image optical center estimation from vignetting and tangential gradient symmetry," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2058–2065, 2009.

[14] H. Cho, H. Lee, and S. Lee, "Radial bright channel prior for single image vignetting correction," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 189–202.

[15] D. B. Goldman and J.-H. Chen, "Vignette and exposure calibration and compensation." USA: IEEE Computer Society, 2005. [Online]. Available: https://doi.org/10.1109/ICCV.2005.249

[16] H. T. Lin, Z. Lu, S. J. Kim, and M. S. Brown, "Nonuniform lattice regression for modeling the camera imaging pipeline," in *Computer Vision – ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 556–568.

[17] D. B. Goldman, "Vignette and exposure calibration and compensation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2276–2288, 2010.

[18] C. Doutre and P. Nasiopoulos, "Fast vignetting correction and color matching for panoramic image stitching," *2009 16th IEEE International Conference on Image Processing (ICIP)*, pp. 709–712, 2009.

[19] S. B. Kang and R. Weiss, "Can we calibrate a camera using an image of a flat, textureless lambertian surface?" in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2000, pp. 640–653. [Online]. Available: https://doi.org/10.1007/3-540-45053-x_41

[20] K. He, P.-F. Tang, and R. Liang, "Vignetting image correction based on gaussian quadrics fitting," in *2009 Fifth International Conference on Natural Computation*, vol. 5, 2009, pp. 158–161.

[21] D. Zhang, Q. Y. Yang, and T. Chen, "Vignetting correction for a single star-sky observation image," *Applied Optics*, vol. 58, no. 16, p. 4337, May 2019. [Online]. Available: https://doi.org/10.1364/ao.58.004337

[22] M. Aggarwal, H. Hua, and N. Ahuja, "On cosine-fourth and vignetting effects in real lenses," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 1, 2001, pp. 472–479 vol.1.

[23] Y. Zheng, J. Yu, S. B. Kang, S. Lin, and C. Kambhamettu, "Single-image vignetting correction using radial gradient symmetry," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[24] A. A. Sawchuk, "Real-time correction of intensity nonlinearities in imaging systems," *IEEE Transactions on Computers*, vol. C-26, no. 1, pp. 34–39, 1977.

[25] M. Brady and G. Legge, "Camera calibration for natural image studies and vision research," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, vol. 26, no. 1, pp. 30–42, Jan. 2009.

[26] A. Bal and H. Palus, "A smooth non-iterative local polynomial (snilp) model of image vignetting," *Sensors*, vol. 21, no. 21, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/21/7086

[27] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics (TOG)*, vol. 2, no. 4, pp. 217–236, 1983.

[28] W. Yu, "Practical anti-vignetting methods for digital cameras," *IEEE Trans. Consumer Electron.*, vol. 50, pp. 975–983, 2004.

[29] F. J. W.-M. Leong, M. Brady, and J. O. McGee, "Correction of uneven illumination (vignetting) in digital microscopy images," *Journal of Clinical Pathology*, vol. 56, no. 8, pp. 619–621, 2003. [Online]. Available: https://jcp.bmj.com/content/56/8/619

[30] A. Bal and H. Palus, "Image vignetting correction using a deformable radial polynomial model," *Sensors*, vol. 23, no. 3, p. 1157, 2023.

[31] A. Kordecki, A. Bal, and H. Palus, "Local polynomial model: A new approach to vignetting correction," in *Ninth International Conference on Machine Vision (ICMV 2016)*, vol. 10341. SPIE, 2017, pp. 463–467.

[32] D. Potoč; and D. Petrinović;, "Creating a synthetic image for evaluation of vignetting modeling and estimation," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2022, pp. 918–923.

The Impact of Image Processing on Perceptual Hash Values

Matej Ferenčak*, Petra Grd† Igor Tomičić†
* Erste&Steiermärkische Bank, Zagreb, Croatia
† Faculty of Organization and Informatics, Varazdin, Croatia
mferencak@foi.unizg.hr; petra.grd@foi.unizg.hr: igor.tomicic@foi.unizg.hr

***Abstract*—In the last few decades image processing has become the focus of research in different fields. It can be used to extract information from images, improve their quality and make it easier for computers to understand them. Some of the challenges that pose a great problem for working with images are detecting modified images or detecting similar images. In order to address those challenges, image hashing can be used. Hashing is calculating a digest value from images and perceptual hash algorithms are a type of hash algorithms with the main idea that similar data has similar hash values, which means that the hash values remain approximately the same if the content is not significantly modified. This paper will compare different perceptual hash algorithms, apply them to differently modified images and analyse the impact of those modifications on perceptual hash values provided by different algorithms. The obtained values will be compared and advantages and disadvantages of different algorithms will be discussed.**

Keywords—image processing, perceptual hash, A-Hash, P-Hash, D-Hash, W-Hash

I. INTRODUCTION

Nowadays, more than ever, due to the widespread usage of image processing applications and availability of image editing software, establishing the legitimacy of image data has become an important challenge and has emerged as a major research topic. Recognizing modified photographs and locating similar images are two obstacles that, when combined, pose a substantial barrier to advancement when working with images. There are different techniques to overcome these obstacles, but these approaches fall mostly into two categories [1]: (1) image watermarking techniques and (2) image hashing techniques. In some applications, using an image watermark is not suitable due to capacity, security aspects and transparency [2], so image hashing has become a prevalent technique. Application areas of perceptual hashing vary across different fields, some of them are [3]–[5]: (1) biometrics, (2) security and (3) forensics. In biometrics, perceptual hashing can be used to find similar biometric traits, such as fingerprints of the same person. In security, perceptual hashing can be used for spam filtering, where hash values of an email digest can be used to determine whether an email is a spam, or for digital rights management where similar video, audio or image files can be found based on their perceptual hashes. Perceptual hash in forensics can be used for detection and removal of known terrorist content or child sexual abuse material shared online.

Hashing is the process of calculating a digest value from images, and perceptual hash algorithms are a type of hash algorithms. Whereas cryptographic hash functions operate on the byte-level and small changes in an image lead to significant changes in cryptographic hash, perceptual image hashing functions are based on the extraction of certain robust or invariant features from the image to produce a hash with the property that two completely different images produce uncorrelated hashes, whereas two visually similar images produce highly correlated hashes [3]. This suggests that the hash values remain relatively unchanged if the content is not substantially altered. This is especially important in the image processing domain where minor changes like compression or scaling should not impact the hash value as much.

There are different properties of perceptual hash algorithms and this study will analyse perceptual robustness to examine the similarities and differences between various perceptual hash algorithms, apply those algorithms to images that have been altered in a variety of ways, and analyze how those image modifications affect the perceptual hash values produced by different algorithms. The image modifications can be categorized into content changing and content preserving. Dittmann et al. [6] and Du et al. [7] state that content-preserving manipulations are: (1) transmission errors/noise/data storage errors, (2) compression, quantization/brightness reduction, (3) resolution reduction/scaling, (4) color conversions (from one color space to another, where the goal is to have the colors in different color spaces be as similar as possible), (5) y-distortion and (6) hue and saturation changes; and content modifying changes are: (1) removing image objects, (2) changing image elements positions, (3) adding new objects, (4) changes of image characteristics: color, textures, structure, impression, (5) changes of the image background and (6) changes of light conditions (shadow manipulations). In this study we will focus mostly on content-preserving changes.

The remainder of this study is organized as follows. Section II will describe a state of the art research effort within the field. Section III will describe the dataset used for the analysis and experiments conducted, after which, the obtained results will be discussed. In the end, the conclusion and future research directions will be given.

II. RELATED WORK

In recent years a number of different perceptual hash algorithms have been developed. This section will focus on the state of the art algorithms and comparison of robustness of those algorithms to image manipulations.

In general, perceptual hash methods include the following steps (Figure 1): preprocessing, feature extraction, feature quantification and hash production. Preprocessing removes superfluous information from an image, making it easier to extract features later. Feature extraction step finds the primary visual features using an extraction method and is the most challenging step of perceptual hashing algorithms. Feature compression blurs the extracted features to improve their robustness and hash creation makes quantitative qualities more abstract [8].

The authors of [9], one of the earliest studies on image hashing utilizing image features, suggested an image authentication system that is tolerant to image compression with loss. They used visually prominent image characteristics and Mexican-Hat wavelets to extract these features from photos. The strategy they proposed was able to determine when portions of an image were eliminated, but they also highlighted areas for improvement in order to apply their method to other image modifications.

The paper by Venkatesan et al. [10] presented a novel image-hashing technique that employs a wavelet representation of images and new random processing strategies for hashing. Error-correcting code-based constructions lower the length of the hash value while maintaining a low collision probability. With this approach, two images can be compared by comparing two bit strings for exact equality, rather than attempting to compare changeable image data, which is a significantly more complex challenge. In their experiments, image hashes proved resistant to a variety of attacks, such as ordinary image processing and malicious distortions.

In [11] a novel and robust hashing paradigm that uses iterative geometric techniques and relies on observations that main geometric features within an image would approximately stay invariant under small perturbations was proposed. The authors use Discrete Wavelet Transformation (DWT) to capture significant image features via time and frequency localization and select relevant regions using a threshold. Experiments revealed robustness to Stirmark attacks, with the exception of extreme rotation and cropping, and that the distances between the hash values of perceptually comparable images are distinct from the distances between images.

Monga and Evans [12] propose an iterative feature detector that uses the first derivative of Gaussian and Morlet wavelets to extract relevant geometry-preserving feature points. To increase perceptual resilience, the authors apply probabilistic quantization to the obtained features. The suggested hash algorithm is resistant to standard benchmark attacks, such as compression, geometric distortions of scaling and small-angle rotation, as well as common signal processing operations. Moreover, content-changing alterations of image data are accurately recognized.

The authors of [13] offer a new pseudo-random (PR) signal representation technique for images, in which they regard images as a succession of linear operator representatives (i.e. matrices). The authors employ Singular Value Decomposition (SVD) to extract images' semi-global, robust PR features. In contrast to DCT/DWT-type fixed basis transforms, SVD selects the optimal basis vectors, and their approach is robust against extreme geometric changes.

Swaminathan et al. [14] created a novel approach for producing an image hash using Fourier-Mellin transform features and controlled randomization. The suggested algorithm is resistant to modest levels of filtering, compression, and basic geometric operations up to 10°rotation and 20% cropping. The proposed hashing system is also capable of identifying malicious operations, such as copy-and-paste editing.

Lv and Wang [15] offer a novel shape-contexts-based approach to image hashing utilizing robust local feature points. Using the robust SIFT-Harris detector, they chose the most stable SIFT keypoints under various content-preserving distortions and generated compact and resilient image hashes by embedding the detected local features into shape-contexts-based descriptors. The suggested technique delivers improved identification performance against geometric attacks and comparable performance against classical distortions.

The previously described algorithms all have one notable problem and that is that they are not robust to rotation changes. In [16], an image hashing algorithm with a ring partition and a Nonnegative Matrix Factorization (NMF) that is both rotation-robust and discriminative was designed. The most significant contribution is a novel construction of rotation-invariant secondary image, which is employed for the first time in image hashing and contributes to making image hash rotation-resistant. In addition, content-preserving modifications modify NMF coefficients almost linearly in order to evaluate hash similarity with correlation coefficient. Experiments have demonstrated that the suggested hashing technique outperforms all other hashing algorithms in terms of robustness and discrimination. The continuation of this research can be found in [17], where authors add ring partition and invariant vector distance to image hashing algorithms to improve rotation robustness and discriminative capabilities. As ring partition is rotation-invariant, this approach to hashing is resistant to image rotation at any angle.

Ding et al. [8] suggested a new perceptual hash technique for multispectral (MS) remote sensing image authentication. In order to compactly represent the perceptual characteristics of the multispectral image, the authors implemented an affinity propagation method to classify the MS images into several clusters based on the information of the bands of these images. By dividing each band into a grid, the features of the grid cell at the same place within

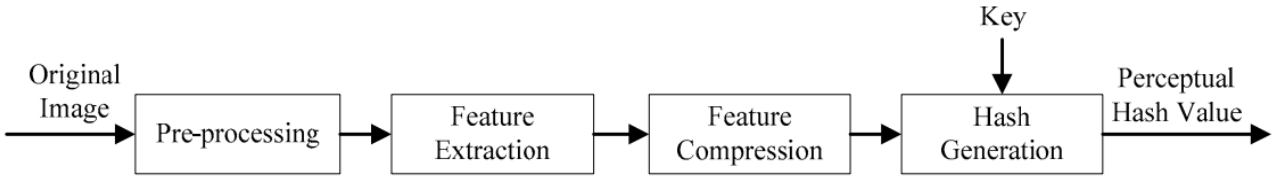


Fig. 1: Steps of the perceptual hash algorithm [8]

the cluster are retrieved and fused using DWT, while PCA-based data compression on the fused feature reduces the effect of noise.

In [18] a scheme is described for efficiently producing image hashing by combining local texture and color angle features. Weber's law is utilized to extract features, and the Weber Local Binary Pattern (WLBP) value of each low-frequency wavelet coefficient is generated prior to a histogram count and PCA dimensionality reduction operation. For color feature extraction, the color angular pattern is applied to each pixel in order to determine the correlations between the various color channels. The color angle values are then dimensionally reduced using a DCT and significant coefficients are chosen prior to cascading the two features to produce an integrated hash.

Qi and Zhao [19] presented a Color Opponent Component (COC) and Quadtree-based perceptual image hashing algorithm. Combining color features with structural features improves the image classification performance of the suggested technique. The input image is subjected to image normalization and Gaussian low-pass filtering to generate a secondary image from which COC is extracted. The color change information is derived from the COC as a color feature, and Quadtree decomposition (QD) is applied to the intensity image of the secondary image to get Quadtree structure features.

Singh et al. [20] created an efficient image hashing technique for authenticating image content based on local and global characteristics. Local information is derived using KAZE features that employ a non-linear diffusion filter, while global features are estimated using a reference image formulation.

Taking into account all of the mentioned papers, there is only a small number of papers dealing with the impact of image processing and manipulation on perceptual hash values and those that analyse the impact, do so on a small set of image data or on images of a particular subjects. This study will analyse a large set of images of different categories with different image manipulations and different perceptual hash algorithms to reach the general conclusion.

III. EXPERIMENTS

In this section different experiments have been conducted to analyse the impact of image processing on most often used perceptual hash algorithms in practice, as identified in literature [3]. Perceptual hash values of original images obtained with Average Hash (A-Hash),

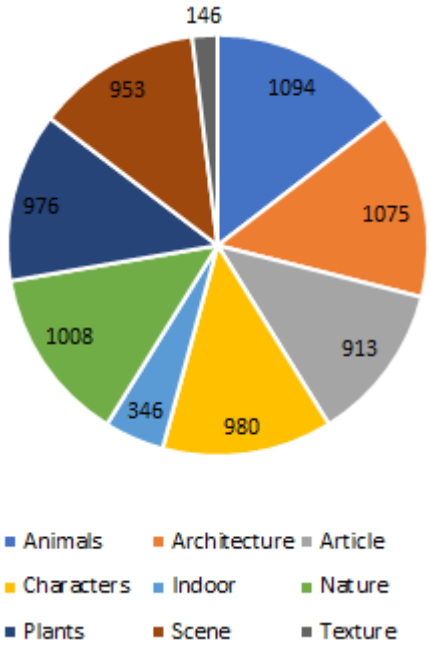


Fig. 2: Distribution of images per class in CASIA 2.0 authentic dataset

Perceptive Hash (P-Hash), Difference Hash (D-Hash) and Wavelet Hash (W-Hash) algorithms are compared with perceptual hash values of images processed with image scaling, rotation, images converted to grayscale, brightness modification, image sharpening, gaussian blurring, median filter and noise adding.

A. Dataset

In order to conduct the experiments and achieve the objective of this study, authentic images from CASIA 2.0 image tampering detection evaluation dataset [21] are used. The authentic part of the dataset consists of 7 491 images of animals, architecture, articles, characters, indoor images, nature, plants, scene and texture. The dataset was selected because of the many different categories of images which was important for reaching general conclusions. All images have the dimensions of 256x384 px or 384x256 px and the number of images in each category can be seen in Figure 2. The dataset is slightly imbalanced, with texture and indoor class having less images (146 and 346 respectively) than the other classes.

All modifications analysed in this study were applied to all 7 491 images, which resulted with the final dataset of 59 928 images.

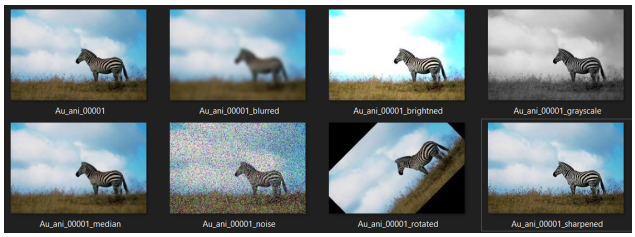


Fig. 3: Example of image modifications

B. Testing and Results

The goal of this study, as mentioned earlier, is to analyse how different image modifications impact the perceptual hash values of different algorithms. To this end, we applied eight modifications to each image from the CASIA 2.0 authentic dataset. Table I shows the description and parameters for each image modification and an example of an image and its modification can be seen in Figure 3.

TABLE I: Image processing parameters

Modification	Parameter
Image Scaling	Factor=0.75
Image Rotation	Angle=45°
Grayscale Conversion	Intensity range=[0,255]
Brightness Modification	Factor=1.5
Image Sharpening	Factor=4
Gaussian Blurring	Radius=3
Median Filter	Filter size=3
Salt & Pepper Noise	Amount=0.3

In order to make necessary image modifications, we created a Python script using Python Imaging Library (PIL) which resulted with a new dataset consisting of 59 928 images with different modifications and original images.

Next step was to calculate perceptual hash values for the original image and each modified image. The values were calculated with A-Hash, P-Hash, D-Hash and W-Hash algorithms of the ImageHash Python library. The first step in all of these methods is converting an image into grayscale and scaling an image into an 8x8 image. After that, the computations for each of these 64 pixels are done depending on the algorithm used, and it gives each one of them a binary value of either 1 or 0. A-Hash calculates the average value of pixels and gives an output of 1 if the pixel is more than or equal to the average, otherwise the output is 0. P-Hash performs the same operation as A-Hash, but it performs a Discrete Cosine Transformation (DCT) and operates in the frequency domain first. D-Hash is a gradient hash and calculates the difference between every pixel and compares it to the average difference. W-Hash is wavelet hashing, which operates in the frequency domain like P-Hash, but employs Discrete Wavelet Transform (DWT) rather than DCT [22].

After we obtained all hash values, the differences between those values need to be calculated. Different distance measures between hash values can be utilized, such as Hamming distance, Euclidean distance, Manhattan distance and Cosine distance [23], but most papers use Hamming distance or Normalised Hamming distance to compare hash values. Hamming distance is the number of symbols or positions of two strings at which their corresponding characters are different (Eq. 1) [24]. Normalized Hamming distance is the ratio of the Hamming distance to the length of the string, or in this case 16.

$$D_H = \sum_{i=1}^{k} |x_i - y_i|, \qquad \begin{array}{l} x = y \Rightarrow D = 0 \\ x \neq y \Rightarrow D = 1 \end{array} \tag{1}$$

In this study, we will use Normalized Hamming distance metric to evaluate the similarity between hash values. The distance was calculated between each hash value (A-Hash, P-Hash, D-Hash, W-Hash) of each original image and all of its modifications.

To discuss the results and draw the conclusions, min, max and mean values of perceptual hash differences have been calculated and are shown in Table II. It is important to note that the closer the Normalized Hamming distance is to zero, the more similar the images are. The results obtained in the experiments show that the A-Hash method achieves the lowest mean Normalized Hamming distances between reference hashes and their respective adjusted counterparts in all cases except for brightness modification and grayscale conversion where W-Hash has the lowest mean Normalized Hamming distance and outperforms the A-Hash. In most cases (Median Filtering, Salt & Pepper noise, Image Scaling, Image Rotation and Image Sharpening), W-Hash has been found to be in the second-place position. In total, D-Hash has been shown to be the least robust to changes in most cases. Also, while image rotation could be classified as a content modifying manipulation, it was interesting to see how different algorithm hash values behave during image rotation. The results have shown that P-Hash has the highest mean Normalized Hamming distance between original image and rotated image.

Given that the dataset contains nine distinct image categories, it was intriguing to determine whether algorithmic behavior varies depending on image category. If we take a look at the hash values per image category (Table III), we can conclude that there are no significant differences in hash value changes depending on image subject. However, in many categories and image manipulations, A-Hash and W-Hash have the same performance.

IV. CONCLUSION

A robust perceptual hashing technique should be one that yields the same or similar hashes for visually identical photos even if their digital representations are no longer the same. The objective of this study was to analyse the changes in perceptual hash values during different image manipulations. We carried out a series of experiments to evaluate the perceptual resilience of the four algorithms and for each of the algorithms, we calculated the hashes from the original 7 491 images and their corresponding changed versions according to the eight modifications

TABLE II: Normalized Hamming distances for different image manipulations

	A-Hash			P-Hash			D-Hash			W-Hash		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Brightness Modification	0.000	0.875	0.112	0.000	0.875	0.198	0.000	0.938	0.142	0.000	1.000	**0.085**
Gaussian Blurring	0.000	0.250	**0.016**	0.000	0.375	0.022	0.000	0.313	0.044	0.000	0.500	0.043
Grayscale Conversion	0.000	0.250	**0.004**	0.000	0.313	0.010	0.000	0.438	0.014	0.000	0.625	0.032
Median Filtering	0.000	0.625	**0.019**	0.000	0.625	0.032	0.000	0.438	0.037	0.000	0.750	0.021
Salt & Pepper noise	0.000	0.750	**0.051**	0.000	0.750	0.120	0.000	0.813	0.169	0.000	0.813	0.076
Image Scaling	0.000	0.250	**0.006**	0.000	0.438	0.012	0.000	0.500	0.017	0.000	0.438	0.006
Image Rotation	0.188	1.000	**0.818**	0.563	1.000	0.914	0.375	1.000	0.904	0.125	1.000	0.828
Image Sharpening	0.000	0.625	**0.036**	0.000	0.563	0.062	0.000	0.688	0.055	0.000	0.625	0.045

TABLE III: Mean Normalized Hamming distances for each modification per image category

	Animals				Architecture				Art			
	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash
Brightness Modification	0.107	0.180	0.135	**0.082**	0.103	0.202	0.155	**0.083**	0.127	0.211	0.158	**0.099**
Gaussian Blurring	**0.016**	0.021	0.041	0.041	**0.015**	0.022	0.051	0.044	**0.015**	0.026	0.042	0.045
Grayscale Conversion	0.005	0.010	0.014	**0.004**	**0.004**	0.009	0.020	**0.004**	0.004	0.009	0.013	**0.003**
Median Filtering	**0.018**	0.029	0.031	**0.018**	**0.019**	0.041	0.046	0.023	**0.020**	0.031	0.040	0.021
Salt & Pepper noise	**0.060**	0.119	0.169	0.090	**0.045**	0.124	0.198	0.069	**0.051**	0.121	0.162	0.074
Image Scaling	**0.005**	0.012	0.016	0.006	**0.005**	0.012	0.019	0.007	**0.006**	0.013	0.015	0.007
Image Rotation	**0.830**	0.908	0.897	0.836	**0.807**	0.920	0.903	0.823	**0.804**	0.907	0.897	0.810
Image Sharpening	**0.034**	0.060	0.049	0.044	**0.038**	0.074	0.067	0.050	**0.036**	0.065	0.057	0.042

	Characters				Indoor				Nature			
	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash
Brightness Modification	0.127	0.207	0.151	**0.107**	0.109	0.192	0.113	**0.075**	0.110	0.239	0.163	**0.058**
Gaussian Blurring	**0.012**	0.018	0.040	0.039	0.019	**0.018**	0.036	0.041	**0.013**	0.022	0.049	0.033
Grayscale Conversion	0.003	0.006	0.010	**0.002**	0.004	0.008	0.011	**0.003**	0.004	0.013	0.018	**0.003**
Median Filtering	**0.014**	0.025	0.028	0.016	**0.021**	0.031	0.029	0.022	**0.017**	0.032	0.039	0.019
Salt & Pepper noise	**0.042**	0.094	0.136	0.068	**0.050**	0.118	0.137	0.072	**0.039**	0.142	0.209	0.069
Image Scaling	**0.004**	0.008	0.013	0.005	**0.006**	0.009	0.012	**0.006**	**0.004**	0.014	0.021	0.006
Image Rotation	**0.828**	0.920	0.914	0.836	**0.803**	0.907	0.899	0.809	**0.808**	0.919	0.896	0.825
Image Sharpening	**0.029**	0.046	0.042	0.036	**0.038**	0.062	0.048	0.041	**0.031**	0.068	0.058	0.041

	Plants				Scene				Texture			
	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash	aHash	pHash	dHash	wHash
Brightness Modification	0.096	0.161	0.105	**0.073**	0.116	0.199	0.139	**0.091**	0.117	0.143	**0.115**	0.117
Gaussian Blurring	**0.019**	0.020	0.042	0.047	**0.016**	0.022	0.047	0.044	**0.028**	0.048	0.039	0.083
Grayscale Conversion	0.005	0.008	0.011	**0.003**	0.004	0.009	0.013	**0.002**	0.012	0.020	0.017	**0.009**
Median Filtering	**0.021**	0.026	0.032	0.023	**0.020**	0.037	0.041	0.022	**0.045**	0.057	0.058	0.055
Salt & Pepper noise	**0.062**	0.119	0.144	0.088	**0.046**	0.113	0.168	0.067	**0.116**	0.188	0.191	0.141
Image Scaling	0.008	0.010	0.015	**0.007**	0.005	0.010	0.017	**0.004**	**0.017**	0.024	0.030	**0.017**
Image Rotation	**0.810**	0.907	0.909	0.821	**0.836**	0.918	0.911	0.845	**0.855**	0.916	0.910	0.860
Image Sharpening	**0.040**	0.052	0.048	0.046	**0.040**	0.068	0.061	0.048	**0.078**	0.092	0.086	0.089

described in Table I. After that, the Normalized Hamming distance between each original hash and its associated altered hash is calculated. The experiment results have shown that A-Hash method achieves the lowest mean Normalized Hamming distances between reference hashes and their respective adjusted counterparts in all cases except for brightness modification and grayscale conversion where W-Hash has the lowest mean Normalized Hamming distance. The future work will focus on the analysis and comparison of additional algorithms and additional manipulations, as well as the comparison of hash values for different content modifying manipulations.

ACKNOWLEDGMENT

This research was funded by the project "Development of CSTI platform for retrieval and analysis of structured and unstructured data". The project received funding from the European Regional Development Fund through OP Competitiveness and Cohesion 2014–2020 within the Call for Proposals "Development of the products and services arising from research and development activities—Phase II", under grant number KK.01.2.1.02.0310.

REFERENCES

[1] S. P. Singh, G. Bhatnagar, and A. K. Singh, "A new robust reference image hashing system," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 4, pp. 2211–2225, 2022.

[2] R. Chandramouli, N. Memon, and M. Rabbani, "Digital watermarking," *Encyclopedia of Imaging Science and Technology*, vol. 10, p. 0471443395, 2002.

[3] M. Hamadouche, K. Zebbiche, M. Guerroumi, H. Tebbi, and Y. Zafoune, "A comparative study of perceptual hashing algorithms: Application on fingerprint images," 2021.

[4] M. Rickard. Perceptual hashing. [Online]. Available: https://matt-rickard.com/perceptual-hashing

[5] Ofcom. Perceptual hashing technology. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0036/247977/Perceptual-hashing-technology.pdf

[6] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," in *Proceedings IEEE International Conference on Multimedia Computing and Systems*, vol. 2, 1999, pp. 209–213 vol.2.

[7] L. Du, A. T. Ho, and R. Cong, "Perceptual hashing for image authentication: A survey," *Signal Processing: Image Communication*, vol. 81, p. 115713, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0923596519301286

[8] K. Ding, S. Chen, and F. Meng, "A novel perceptual hash algorithm for multispectral image authentication," *Algorithms*, vol. 11, no. 1, 2018.

[9] S. Bhattacharjee and M. Kutter, "Compression tolerant image authentication," in *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)*, vol. 1, 1998, pp. 435–439 vol.1.

[10] R. Venkatesan, S.-M. Koon, M. Jakubowski, and P. Moulin, "Robust image hashing," in *Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101)*, vol. 3, 2000, pp. 664–666 vol.3.

[11] M. K. Mıhçak and R. Venkatesan, "New iterative geometric methods for robust perceptual image hashing," in *Security and Privacy in Digital Rights Management*, T. Sander, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 13–21.

[12] V. Monga and B. Evans, "Robust perceptual image hashing using feature points," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, vol. 1, 2004, pp. 677–680 Vol. 1.

[13] S. Kozat, R. Venkatesan, and M. Mihcak, "Robust perceptual image hashing via matrix invariants," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, vol. 5, 2004, pp. 3443–3446 Vol. 5.

[14] A. Swaminathan, Y. Mao, and M. Wu, "Robust and secure image hashing," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 215–230, 2006.

[15] X. Lv and Z. J. Wang, "Perceptual image hashing based on shape contexts and local feature points," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1081–1093, 2012.

[16] Z. Tang, X. Zhang, and S. Zhang, "Robust perceptual image hashing based on ring partition and nmf," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 3, pp. 711–724, 2014.

[17] Z. Tang, X. Zhang, X. Li, and S. Zhang, "Robust image hashing with ring partition and invariant vector distance," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 200–214, 2016.

[18] C. Qin, Y. Hu, H. Yao, X. Duan, and L. Gao, "Perceptual image hashing based on weber local binary pattern and color angle representation," *IEEE Access*, vol. 7, pp. 45 460–45 471, 2019.

[19] Q. Shen and Y. Zhao, "Perceptual hashing for color image based on color opponent component and quadtree structure," *Signal Processing*, vol. 166, p. 107244, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165168419302907

[20] S. P. Singh, G. Bhatnagar, and A. K. Singh, "A new robust reference image hashing system," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 4, pp. 2211–2225, 2022.

[21] J. Dong, W. Wang, and T. Tan, "Casia image tampering detection evaluation database," in *2013 IEEE China Summit and International Conference on Signal and Information Processing*, 2013, pp. 422–426.

[22] D. Petrov. Wavelet image hash in python. [Online]. Available: https://fullstackml.com/wavelet-image-hash-in-python-3504fdd282b5

[23] S. Huang, J. Qin, T. Tao, L. Yang, and X. Liu, "A novel method to perceive self-vehicle state based on vehicle video by image similarity calculation," *IEEE Open Journal of Instrumentation and Measurement*, vol. 1, pp. 1–11, 2022.

[24] X.-S. Yang, "Chapter 6 - genetic algorithms," in *Nature-Inspired Optimization Algorithms (Second Edition)*, second edition ed., X.-S. Yang, Ed. Academic Press, 2021, pp. 91–100. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128219867000135

Usporedba funkcija gubitka za semantičku segmentaciju objekata u prometu

G. Oreški*, S. Aničić*
* Sveučilište Jurja Dobrile u Puli/Fakultet informatike, Pula, Hrvatska
goreski@unipu.hr

***Sažetak*—Autonomna vožnja postaje sve važniji čimbenik svakodnevnog života, čime raste značaj dubokog učenja na kojem se takvi sustavi temelje. Posljednjih godina predložene su mnoge funkcije gubitka za problem semantičke segmentacije, čiji odabir ima direktan utjecaj na sposobnost modela da uči i uspješno generira predikcije. Ovo istraživanje analizira utjecaj različitih funkcija gubitka na performanse semantičke segmentacije na 23 različita objekta prisutna u prometu. Istraživanje je provedeno na skupu podataka prikupljenom na CARLA simulatoru prometa u urbanom okruženju. Rad obuhvaća sve najčešće korištene funkcije gubitka: *Binary Cross entropy*, *Dice*, *Focal*, *Jaccard* te *Tversky*, kako bi analizirali njihov utjecaj na konvergenciju modela. Usporedba je izvršena na tri različite arhitekture mreža: U-net, DeepLabV3Plus i FPN. Rad donosi preporuke prikladnosti pojedinih funkcija na problemu semantičke segmentacije objekata u prometu.**

Ključne riječi*—*semantička segmentacija, duboko učenje, autonomna vozila, funkcija gubitka

I. Uvod

Danas sve više proizvođača automobila stavlja fokus na autonomnu vožnju, koja se brzo razvija i obećava revoluciju u životima pojedinaca. Stvaranje autonomnog sustava vožnje uključuje ne samo kontrolu i upravljanje vozilom, nego zahtijeva da vozilo i njegov autonomni sustav vožnje precizno percipiraju okolinu, očekuju i predviđaju različite pojave i događaje na cesti te donose brze i točne odluke u stvarnom vremenu, kako bi putnicima i drugim sudionicima u prometu omogućili sigurno putovanje.

U rješavanju izazova autonomne vožnje trenutno se u svijetu nameću modeli dubokog učenja. Takvi modeli su sposobni predvidjeti različita ponašanja vozila i okoline s iznimnom točnošću. Duboko učenje i računalni vid zahtijevaju korištenje skupa podataka koji je dovoljno velik i sadrži različite scenarije, tako da model može naučiti razlikovati različite obrasce i pojave na cesti. Stvaranje skupa podataka je dugotrajan i skup proces. To se pogotovo odnosi na označavanje podataka za treniranje modela s nadziranim učenjem [1]. CARLA simulator za gradsku vožnju predstavlja alternativu klasičnom načinu prikupljanju podataka. Simulator nudi niz različitih elemenata u svakodnevnom prometu, kao što su: pješaci, vozila, zgrade, ulični znakovi te drugi elementi. Također nudi mogućnost promjene doba dana, vremena, senzora, brzine, ubrzanja i drugih varijabilnih elemenata [2]. Najveća prednost takvih simulatora je mogućnost automatskog generiranja labela, čime se značajno ubrzava proces stvaranja skupa podataka.

U ovom istraživanju smo iskoristili CARLA simulator da bi istražili performanse različitih funkcija gubitka na problemu semantičke segmentacije u prometu. Funkcija gubitka je osnovni element svakog modela. Ključna je u procesu treninga te stoga njezin odabir predstavlja bitnu odluku.

II. Pregled literature

Iako je problem autonomne vožnje danas vrlo popularan postoji malo literature vezane uz funkcije gubitka, pogotovo za semantičku segmentaciju. Jedan od razloga može biti ranije spomenuta cijena prikupljanja takvih skupova podataka. U nastavku donosimo istraživanja koja su povezana s navedenom temom.

S. Malec [3] je istražio semantičku segmentaciju za treniranje modela autonomne vožnje proširenu s podacima generiranim iz CARLA simulatora. U radu je korištena isključivo *Cross entropy* funkcija gubitka. M. Boudissa et al. [4] su korištenjem U-net modela nastojali stvoriti model za korištenje u prometu, odnosno prepoznavanje objekata te su koristili dvije funkcije gubitka: *Dice* i *Focal*. *Dice* se pokazao kao bolji izbor zbog performansi detekcije rubova objekata. O. Natan i J. Miura [5] su uz pomoć CARLA simulatora razvili model za semantičku segmentaciju i procjenu dubine. Tijekom treninga semantičke segmentacije koristili su *Dice* funkciju kombiniranu s *pixel-wise* binarnom *Cross-entropy* funkcijom u svrhu postizanja veće raznolikosti pogreške. F. Nesti et al. [6] su usporedili standardnu *Cross entropy* funkciju s vlastitom, nadograđenom funkcijom koja bolje sprječava *Adversarial Patch* napade tijekom semantičke segmentacije.

J. Ma et al. [7] su analizirali veliki broj funkcija gubitka za semantičku segmentaciju (više od 10) na medicinskim slikama. U radu se zaključuje da su najrobustnije složene funkcije gubitka između kojih je i *Dice* funkcija. M. Yeung et al. [8] su predložili novu funkciju, *Unified Focal*, koja generalizira dvije najčešće funkcije gubitka za zadatak segementacije: *Dice* i *Cross-entropy*. Istraživanje je također provedeno na medicinskim slikama. H Ghandorh et al. [9] su istražili način kako uz pomoć semantičke segmentacije prepoznati cestu na visokokvalitetnim satelitskim slikama, kao funkciju gubitka koristili su težinsku *Cross entropy* te *Focal Tversky* funkciju gubitka. J. Su et al. [10] su usporedili različite funkcije gubitaka za nebalansiranu semantičku segmentaciju te su predložili *Cross Dropout Focal* funkciju gubitka. Koristeći istu ostvareno je 2.5%

veća točnost u usporedbi sa standardnom *Cross-entropy* funkcijom. H.X. Cheng et al. [11] su predložili novi model koji smanjuje računsku i vremensku kompleksnost semantičke segmentacije. Koristili su kombinaciju funkcija gubitka: težinski *Cross-entropy*, *Lovasz-Softmax* te *Boundary*. Kombinacijom su nastojali riješiti izazove: nebalansiranih podataka, optimizacije Intersection over Union (IoU) mjere i nejasnih granica segmentacije. S. Jadon [12] je napravio iscrpnu analizu nekoliko funkcija gubitaka te ih usporedio na primjeru semantičke segmentacije. Identificirao je u kojim situacijama je jedna prikladnija od druge. Istraživanje je proveo na medicinskim slikama.

Ovim radom donosimo sustavnu usporedbu različitih funkcija gubitka za semantičku segmentaciju na tri najpopularnije arhitekture modela. Na podacima iz prometa smo ispitali njihovu prikladnost i donijeli preporuke za korištenje. U analizu smo uključili najpopularnije funkcije spomenute u pregledu literature.

III. Funkcije gubitka

Algoritmi dubokog učenja počivaju na mjeri koja određuje kvalitetu njihove izvedbe, odnosno koja računa veličinu pogreške između predviđenih i stvarnih (ciljnih) podataka. Funkcija gubitka predstavlja taj izračun te se koristi kao polazište za prilagođavanje parametara modela kako bi se smanjila pogreška te poboljšale performanse modela. Izbor funkcije gubitka može uvelike utjecati na performanse algoritma, samim time neprikladna funkcija će negativno utjecati na performanse i konvergenciju modela. U nastavku donosimo kratki opis funkcija obuhvaćenih ovim istraživanjem. Uz svaku funkciju dana je i njezina formula, čije detaljno objašnjenje se može pronaći u referenciranim radovima.

A. Cross Entropy

Cross entropy funkcija gubitka je definirana kao negativni logaritam predviđene vjerojatnosti za ispravnu klasu [13] te je najčešće korištena funkcija gubitka za treniranje klasifikacijskih modela. Funkcija se pokazala uspješnom na brojnim problemima te se često koristi za semantičku segmentaciju. Njezin nedostatak je što sve piksele tretira podjednako te veliki objekti dominiraju mrežom stoga su performanse na malim objektima loše. Formula kategorijske *Cross entropy* funkcije za C klasa je dana jednadžbom 1 gdje je t_i oznaka istinite klase a $f(s_i)$ *Softmax* vjerojatnost za i-tu klasu.

$$CE = -\sum_{i}^{C} t_i \log(f(s_i))) \tag{1}$$

Binary Cross entropy je varijanta navedene funkcije koja se koristi u binarnoj klasifikaciji.

$$BCE = -\sum_{i}^{C=2} t_i log(f(s_i)) \tag{2}$$

B. Dice

Dice funkcija gubitka bazirana je na Sørensen–Dice koeficijentu ili Tversky indeksu [14]. Lažno pozitivni i lažno negativni primjeri su jednake važnosti te je ona imunija na problem nebalansiranosti [15]. Postoji nekoliko nedostataka, odnosno ograničenja kod treniranja neuronskih mreža s ovom funkcijom [16]: neosjetljivost na udaljenost regija koje se ne preklapaju, [17] neosjetljivost na granice objekata [18] te osjetljivost na male difuzne lezije [19]–[23]. Formula *Dice* funkcije dana je jednadžbom 3 i 4 gdje A predstavlja istinitu masku, a B predviđenu masku. Presjek maski se računa kao umnožak po elementima (engl. element-wise) te se potom sumira rezultirajuća matrica.

$$Dice = 2 * \frac{|A \cap B|}{|A| + |B|} \tag{3}$$

$$DiceLoss = 1 - Dice \tag{4}$$

C. Jaccard/Intersection over union (IoU)

Jaccard indeks, odnosno IoU gubitak daje rezultat u rasponu [0,1] za presjek između istinite i predviđene maske preko njihove unije s pretpostavkom da je 0/0 jednako 1 [24]. IoU rezultat želimo maksimizirati dok IoU gubitak želimo minimizirati tijekom treniranja. Nedostatak funkcije su loše performanse na malim objektima segmentacije. U tim slučajevima nema varijacije indeksa i ti objekti ostaju nezapaženi tijekom treninga [25]. Formula za Jaccard/IoU funkciju gubitka dana je jednadžbom 5. Oznake A i B imaju isto značenje kao u formuli 3.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{5}$$

D. Focal

Focal je varijacija *Cross entropy* funkcije. Autori su predložili dodavanje faktora $(1 - p_t)^\gamma$ standardnoj funkciji gubitka *Cross entropy* te se postavljanjem $\gamma > 0$ može smanjiti gubitak na ispravno predviđenom skupu primjera [26]. Funkcija dijeli primjere u lagane i teške te pomoću faktora povećava vrijednost gubitka za teške, odnosno smanjuje ga za lagane primjere. Funkcija gubitka klasificira svaki piksel pojedinačno što rezultira time da pogrešna anotacija značajno utječe na performanse mreže [27]. Navedeni problem nije značajan kada se koriste podaci generirani iz simulatora. Formula Focal funkcije dana je jednadžbom 6 gdje je p_t predviđena vjerojatnost ispravne klase, γ parametar koji kontrolira stupanj za koji će gubitak biti smanjen za ispravno klasificirane primjere a α težinski faktor za svaki primjer.

$$FL(p_t) = -\alpha(1 - p_t)^\gamma \log(p_t) \tag{6}$$

E. Tversky

Tversky proširuje *Dice* funkciju uvođenjem dvaju koeficijenata koji kontroliraju osjetljivost gubitka na lažno pozitivne i lažno negativne rezultate. Koeficijenti α i β

omogućuju određivanje kompromisa između preciznosti i odziva u izračunu gubitka. To rezultira fleksibilnijom funkcijom gubitka u usporedbi s gubitkom *Dice* funkcije budući da se može bolje nositi s neuravnoteženim skupovima podataka i smanjiti utjecaj klasa s velikim brojem piksela [28]. Tversky funkcija gubitka je dana jednadžbom 7 gdje TP označava istinito pozitivne piksele, FP lažno pozitivne, FN lažno negativne a α i β su ranije spomenuti koeficijenti.

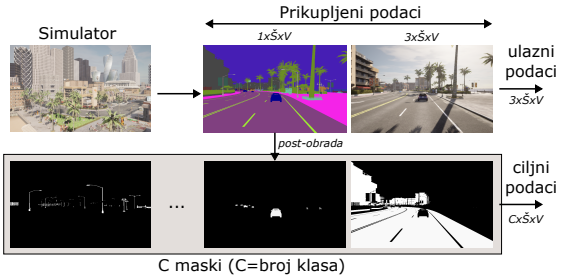
$$T_{\alpha,\beta}\left(y,\hat{y}\right) = \frac{TP}{TP + \alpha FP + \beta FN} \tag{7}$$

IV. Eksperiment i rezultati analize

Da bi dali sustavan prikaz utjecaja funkcija gubitka na semantičku segmentaciju prometa, u istraživanje su uključene tri arhitekture najčešće korištene u literaturi što ukupno daje 15 različitih modela koje smo vrednovali koristeći četiri mjere uspješnosti predikcije: IoU, F mjera, preciznost i odziv. Detaljan opis eksperimenta zajedno s rezultatima je dan u nastavku.

A. Skup podataka iz simulatora CARLA

Skup podataka je prikupljen koristeći CARLA simulator autonomne vožnje. Za generiranje podataka kreirana su dva različita senzora, jedan za generiranje RGB slike, a drugi za generiranje semantičke segmentacije. Trenutačna verzija CARLA-e sadržava 23 različite klase (C=23). Senzor za semantičku segmentaciju označava klase svakog piksela oznakama od 1 do 23 u crvenom (R) kanalu slike, dok su ostala dva kanala (G i B) prazna, odnosno sadržavaju vrijednost 0. Koristeći funkciju *cityScapesPalette* ta slika se može transformirati iz slike s vrijednosti samo u crvenom kanalu u sliku s vrijednostima u svim kanalima te se grafički lijepo prikazati. Za generiranje prometa koristila se skripta *generate_traffic.py*, a za postavljanje Ego vozila (vozila s kamerom), pokretanje senzora te autopilot, odvojena, vlastita skripta.



Slika 1: Generiranje skupa podataka

Slika 1 prikazuje dohvaćanje podataka iz simulatora. Podaci sa senzora se dohvaćaju u paru te sadržavaju RGB sliku i masku u jednom kanalu (koja se iscrtava kao RGB sa spomenutom *cityScapesPalette* funkcijom). Slika je veličine 3 x Širina(Š) x Visina(V), a maska 1xŠxV. U post-obradi tj. pripremi ciljnih podataka, maska se tehnikom *one-hot encoding* dijeli na C zasebnih kanala u kojima se nalaze odvojene maske svake klase. Na Slici 1 su odvojene maske vidljive kao crno-bijele slike gdje bijeli pikseli prikazuju prisutnost pojedine klase. Modelu se tijekom treniranja kao ulazni podatak daje slika i maska veličine CxŠxV. Model predviđa masku za svaku pojedinu klasu na zasebnom kanalu. Da bi se nakon predikcije došlo do inicijalne dimenzije maske (1xŠxV) izlaz se agregira koristeći *argmax* funkciju po prvoj dimenziji. Prikupljeno je 11000 slika, koje su raspodijeljene u omjeru 70/20/10 u trening, validacijski i testni skup respektivno.

B. Modeli

Ovaj rad analizira pet različitih funkcija gubitaka koje su korištene za trening u kombinaciji s tri različite arhitekture mreže. Kao osnovna mreža (*engl. backbone*) odabran je *resnext50*, a korištene težine su pred-trenirane na *imagenet* skupu podataka. Za arhitekture smo koristili: FPN (*Feature Pyramid Network*) [29], DeepLabV3+ [30] te U-net [31].

FPN arhitektura se koristi za izdvajanje značajki slike (engl. feature extractor). Kao ulaz prima sliku proizvoljne veličine i generira mape značajki proporcionalne veličine na više razina. DeepLabV3Plus koristi modul *atrous spatial pyramid pooling* (ASPP) za izdvajanje kontekstualnih informacija na više razina (engl. multi-scale) iz mapa značajki konvolucijske neuronske mreže. U-net je enkoder-dekoder tip arhitekture koji prvo uz pomoć enkodera izdvaja značajke te potom dekoderom rekonstruira segmentacijsku mapu.

C. Evaluacijske mjere

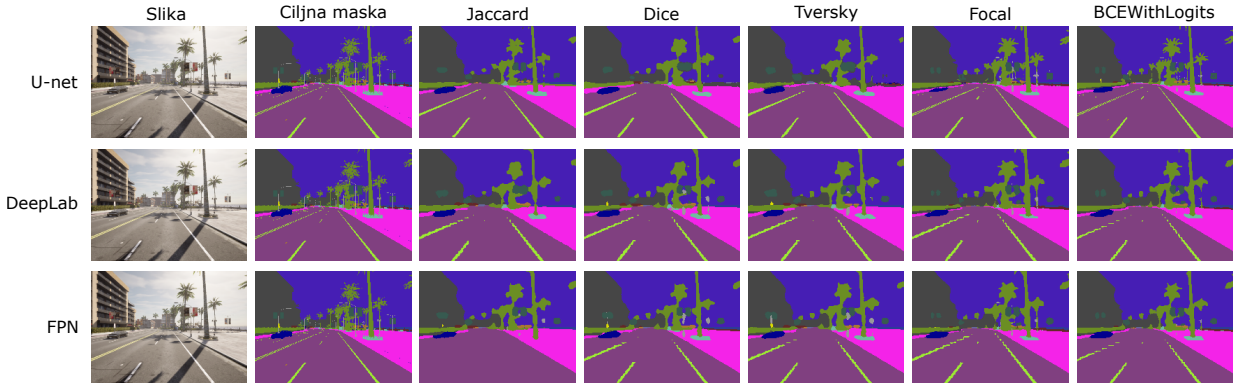
Evaluacijske mjere se koriste za procjenu uspješnosti modela, a mi ih koristimo za usporedbu prikladnosti odabranih funkcija gubitka. Mnoge mjere imaju polazište u matrici konfuzije (engl. confusion matrix). U matrici postoje 4 moguća ishoda predikcije: istinito pozitivna, istinito negativna, lažno pozitivna te lažno negativna. Različite mjere predstavljaju različite odnose tih ishoda. U nastavku su opisane korištene evaluacijske mjere u radu.

Preciznost je udio točno klasificiranih piksela u skupu pozitivno klasificiranih piksela. Odziv je udio točno klasificiranih piksela u skupu svih pozitivnih piksela. F-mjera je harmonijska sredina preciznosti i odziva. IoU mjera je udio ispravno klasificiranih piksela klase u skupu svih piksela koji toj klasi pripadaju ili su svrstani u nju.

D. Rezultati

Trening modela je zaustavljan kada je vrijednost funkcije gubitka prestala padati na validacijskom skupu podataka. Konačni rezultati uspješnosti modela na testnom skupu podataka, iskazani ranije spomenutim mjerama, su vidljivi u Tablici I. Na Slici 2 je dan usporedni prikaz predikcije modela za isti ulazni podatak. Analizirane su performanse pojedinih modela kroz funkcije gubitka, a potom i korištenu arhitekturu.

Iz Tablice I je vidljivo da je *Binary Cross entropy* funkcija ostvarila najbolje rezultate mjereno F mjerom, IoU mjerom i preciznosti predikcije. U Poglavlju III je



Slika 2: Prikaz predikcija svih modela na istoj slici

istaknuto da je to najčešće korištena funkcija te da ostvaruje dobre rezultate na semantičkoj segmentaciji što ovo istraživanje potvrđuje. Štoviše, iz Slike 2 je vidljivo da su i manji objekti prepoznati, što se često navodi kao glavni nedostatak te funkcije. Rezultati pokazuju da je *Binary Cross* entropy robusna funkcija te da se može koristiti s bilo kojom arhitekturom.

Focal funkcija gubitka ima vrlo slične rezultate kao i *Binary Cross entropy*, što dokazuje da se radi o njezinoj varijanti koja dodaje jedan termin. Obzirom da je *Focal* funkcija parametrizirana s γ, kojim se određuje koliko će biti smanjen utjecaj laganih primjera na vrijednost funkcije, njegovim mijenjanjem se može utjecati na konačan ishod. U eksperimentu smo zadržali zadanu vrijednost γ=2 koju autori navode kao najbolji odabir. *Focal* funkcija je konzistentno kroz sve arhitekture imala nešto bolju preciznost i manji odziv od *Cross entropy* funkcije. Ako usporedimo dvije funkcije na Slici 2, vidimo da udaljene linije između prometnih traka (manji objekti) bolje prepoznaje *Focal* funkcija, dok automobil bolje prepoznaje *Cross entropy*.

Dice funkcija i njezino proširenje *Tversky* ostvaruju slične rezultate, koji su lošiji od ranije spomenutih tehnika. Ukoliko se koristi *U-net* ili *FPN* tada *Dice* postiže bolje rezultate, dok s *DeepLabV3Plus* je obrnuto. Najveća razlika je vidljiva na U-net mreži gdje *Dice* ostvaruje značajno bolji rezultat. Za razliku od prošle dvije tehnike, ove obje imaju veći odziv nego preciznost, bez obzira na korištenu mrežu. Vizualnom analizom predikcija sa Slike 2 je vidljivo da obje funkcije postižu lošije rezultate od *Cross entropy* i *Focal* funkcije, iako na nekim elementima kao što je stup rasvjete s lijeve strane ceste (žuta boja na maski) imaju bolju detekciju.

Jaccard funkcija postiže dobre rezultate ako se ne koristi s FPN mrežom u kojem slučaj rezultira s najlošijim modelom. Iznimno je dobra u kombinaciji s *U-net* arhitekturom jer postiže rezultate tek nešto lošije od *Focal* funkcije. Međutim, za razliku od nje ima bolje vrijednosti odziva. Na slici je vidljivo da *Jaccard* funkcija najlošije prepoznaje male i udaljene objekte kao što su znakovi, stupovi i linije prometnih traka.

Tablica I: Rezultati usporedbe funkcija gubitka

Arhitektura mreže	Funkcija gubitka	Mjera$_{test}$			
		F mjera	IOU	Preciznost	Odziv
U-net	Jaccard	0.8836	0.7943	0.8352	**0.9385**
	Dice	0.8805	0.7889	0.8387	0.9272
	Tversky	0.849	0.74	0.7781	0.935
	Focal	0.9078	0.8351	**0.9552**	0.8653
	BCEWithLogits	**0.9179**	**0.8524**	0.9386	0.8983
DeepLabV3Plus	Jaccard	0.8205	0.6974	0.7192	0.9562
	Dice	0.7558	0.6086	0.6224	**0.9637**
	Tversky	0.762	0.6172	0.6319	0.9619
	Focal	0.8914	0.8076	**0.9591**	0.8334
	BCEWithLogits	**0.908**	**0.835**	0.9371	0.8809
FPN	Jaccard	0.5535	0.3833	0.3907	0.9504
	Dice	0.6142	0.4437	0.4479	**0.9784**
	Tversky	0.6021	0.4313	0.4352	**0.9784**
	Focal	0.8903	0.8056	**0.9617**	0.8296
	BCEWithLogits	**0.9081**	**0.8351**	0.9389	0.8795

Analizom rezultata iz Tablice I jasno se vidi da je vrlo bitno voditi brigu koju funkciju gubitka koristiti s kojom arhitekturom jer rezultati mogu biti značajno različiti. *U-net* mreža je ostvarila najbolje rezultate. Sve funkcije u kombinaciji s njom postižu najbolje rezultate F i IoU mjere. Slike koje prikazuju kombinaciju *U-net* i *Cross entropy* ili *Focal* funkcije su iznimno dobre te se na njima vide detalji kao što su grane i lišće na drvećima, što nije prisutno na slikama ostalih modela. S druge strane *FPN* nije dobro kombinirati s *Jaccard*, *Dice* i *Tversky* funkcijama jer generirani modeli imaju značajno lošije performanse.

V. Zaključak

Cilj ovog rada je bio usporediti najpopularnije funkcije gubitka za semantičku segmentaciju na skupu slika iz prometa. U tu svrhu smo generirali slike i njihove anotacije pomoću CARLA simulatora. U eksperiment smo uključili tri različite arhitekture mreža koje su najzastupljenije u dostupnoj literaturi.

Nakon provedenog istraživanja možemo zaključiti da je *Binary Cross entropy* funkcija postigla najbolje rezultate bez obzira na odabranu arhitekturu. Dovoljno je robusna da bi ju svakako trebalo razmotriti u inicijalnom koraku stvaranja modela semantičke segmentacije. Najbolja se pokazala njezina kombinacija s U-net arhitekturom koja je ostvarila vrlo uvjerljive rezultate i impresivne slike maski. Osim *Binary Cross entropy* funkcije i *Focal* funkcija je dobar odabir jer otvara prostor da pomoću validacije γ parametra dođemo do još boljih rezultata. Ukoliko je odziv

bitna mjera tada je preporučena kombinacija *U-net* mreže i *Jaccard* funkcije.

LITERATURA

[1] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, "A comprehensive survey of loss functions in machine learning," *Annals of Data Science*, pp. 1–26, 2020.

[2] W. Yuan and W. Xu, "Gaploss: A loss function for semantic segmentation of roads in remote sensing images," *Remote Sensing*, vol. 14, no. 10, p. 2422, 2022.

[3] S. Malec, "Semantic segmentation with carla simulator," 2021.

[4] M. Boudissa, H. Kawanaka, and T. Wakabayashi, "semantic segmentation of traffic landmarks using classical computer vision and u-net model," in *Journal of Physics: Conference Series*, vol. 2319, no. 1. IOP Publishing, 2022, p. 012031.

[5] O. Natan and J. Miura, "Semantic segmentation and depth estimation with rgb and dvs sensor fusion for multi-view driving perception," in *Pattern Recognition: 6th Asian Conference, ACPR 2021, Jeju Island, South Korea, November 9–12, 2021, Revised Selected Papers, Part I*. Springer, 2022, pp. 352–365.

[6] F. Nesti, G. Rossolini, S. Nair, A. Biondi, and G. Buttazzo, "Evaluating the robustness of semantic segmentation for autonomous driving against real-world adversarial patch attacks," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 2280–2289.

[7] J. Ma, J. Chen, M. Ng, R. Huang, Y. Li, C. Li, X. Yang, and A. L. Martel, "Loss odyssey in medical image segmentation," *Medical Image Analysis*, vol. 71, p. 102035, 2021.

[8] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, 2022.

[9] H. Ghandorh, W. Boulila, S. Masood, A. Koubaa, F. Ahmed, and J. Ahmad, "Semantic segmentation and edge detection—approach to road detection in very high resolution satellite images," *Remote Sensing*, vol. 14, no. 3, p. 613, 2022.

[10] J. Su, S. Anderson, and L. S. Mihaylova, "A deep learning method with cross dropout focal loss function for imbalanced semantic segmentation," in *2022 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2022, pp. 1–6.

[11] H.-X. Cheng, X.-F. Han, and G.-Q. Xiao, "Cenet: Toward concise and efficient lidar semantic segmentation for autonomous driving," in *2022 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2022, pp. 01–06.

[12] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*. IEEE, 2020, pp. 1–7.

[13] L. Li, M. Doroslovački, and M. H. Loew, "Approximating the gradient of cross-entropy loss function," *IEEE Access*, vol. 8, pp. 111 626–111 635, 2020.

[14] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, and I. B. Ayed, "Boundary loss for highly unbalanced segmentation," in *International conference on medical imaging with deep learning*. PMLR, 2019, pp. 285–296.

[15] X. Li, X. Sun, Y. Meng, J. Liang, F. Wu, and J. Li, "Dice loss for data-imbalanced nlp tasks," *arXiv preprint arXiv:1911.02855*, 2019.

[16] Y. Zhang, S. Liu, C. Li, and J. Wang, "Rethinking the dice loss for deep learning lesion segmentation in medical images," *Journal of Shanghai Jiaotong University (Science)*, vol. 26, pp. 93–102, 2021.

[17] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 658–666.

[18] J. Dolz, C. Desrosiers, and I. B. Ayed, "3d fully convolutional networks for subcortical segmentation in mri: A large-scale study," *NeuroImage*, vol. 170, pp. 456–470, 2018.

[19] N. Guizard, P. Coupé, V. S. Fonov, J. V. Manjón, D. L. Arnold, and D. L. Collins, "Rotation-invariant multi-contrast non-local means for ms lesion segmentation," *NeuroImage: Clinical*, vol. 8, pp. 376–389, 2015.

[20] R. Harmouche, N. K. Subbanna, D. L. Collins, D. L. Arnold, and T. Arbel, "Probabilistic multiple sclerosis lesion classification based on modeling regional intensity variability and local neighborhood information," *IEEE transactions on Biomedical Engineering*, vol. 62, no. 5, pp. 1281–1292, 2014.

[21] M. Styner, J. Lee, B. Chin, M. Chin, O. Commowick, H. Tran, S. Markovic-Plese, V. Jewells, and S. Warfield, "3d segmentation in the clinic: A grand challenge ii: Ms lesion segmentation," *MIDAS journal*, vol. 2008, pp. 1–6, 2008.

[22] K. C. Wong, M. Moradi, H. Tang, and T. Syeda-Mahmood, "3d segmentation with exponential logarithmic loss for highly unbalanced object sizes," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part III 11*. Springer, 2018, pp. 612–619.

[23] C. Lucas, A. Kemmling, A. M. Mamlouk, and M. P. Heinrich, "Multi-scale neural network for automatic segmentation of ischemic strokes on acute perfusion images," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 1118–1121.

[24] M. Berman, A. R. Triki, and M. B. Blaschko, "The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4413–4421.

[25] U. Kamal, T. I. Tonmoy, S. Das, and M. K. Hasan, "Automatic traffic sign detection and recognition using segu-net and a modified tversky loss function with l1-constraint," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1467–1479, 2019.

[26] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[27] L. Wang, C. Wang, Z. Sun, and S. Chen, "An improved dice loss for pneumothorax segmentation by mining the information of negative areas," *IEEE Access*, vol. 8, pp. 167 939–167 949, 2020.

[28] H. Gábor, "The tversky loss function and its modifications for medical image segmentation," 2022.

[29] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.

[30] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.

[31] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

NATURAL LANGUAGE PROCESSING

The Role of Knowledge Management in Transition to Industry 5.0

S. Lovrenčić
University of Zagreb Faculty of Organization and Informatics, Varaždin, Croatia
sandra.lovrencic@foi.unizg.hr

***Abstract* - Fifth industrial revolution, or Industry 5.0 is gaining a momentum in last several years, additionally enhanced with many disruptive events in global economy. Its main keywords are sustainable, human-centric and resilient, and it should bring these benefits to industry and society. Knowledge management enables the cycle of knowledge through organization and is a human oriented dicipline that continually embraces new technologies, helping organizations to achieve their goals. This paper explains that sustainability, human-centricity and resilience are naturally supported by knowledge management. Investigation of connection of elements of Industry 5.0 action plan with knowledge management processes and technologies shows that knowledge management has an important role in achieving Industry 5.0, especially through processes of knowledge creation/acquisition and storing and supporting artificial intelligence technologies.**

Keywords – knowledge management; Industry 5.0; technology

I. INTRODUCTION

The society is continually and inevitably changing, rushing forward (some would say backward) with big strides. The advancement of technology is ever faster, with new possibilites it offers causing both thrill and fear. Of course, it is inevitable to mention large language model ChatGPT [1], that is causing a lot of controversy and polemics, also about its impact on future of certain professions.

Humans have embraced almost every new technology for personal and business benefits. Organizations carefully consider positive and negative impacts that new technology can bring them for competitive advantage and profit. Many information systems that help organization in everyday business, including knowledge management systems (KMS), already rely on a wide range of latest technology, including artificial intelligence (AI).

Recent disruptions in various forms have shown that businesses and society in general lack resilience, despite all available technology. Covid-19 pandemic and, recently war in Ukraine are most warning examples on global level. Brexit also caused disruptions, and more localy, earthquakes in Croatia and Turkey. A single container ship that was blocking the Suez Canal for six days in 2021 caused global logistic problems.

Negative impact of humans on climate change and biodiversity collapse is making headlines for a long time, with grim warnings for future generations. Technology use and production are seen very negative in this context, with urge for lower energy consumption, waste reduction, recycling and re-using natural resources. For that reason, the concept of sustainability requires of current generation not to compromise the needs of future generations with satisfying their own needs.

All those issues caused the move of current industry trends to Industry 5.0, that tries to take into account needs of humans without geopardizing the planet, and to the consideration of means how to achieve it . This paper is organized as follows. In Section II the concept of Industry 5.0 and its goals are explained. Section III gives review of knowledge management and Section IV discusses how knowledge management can help in realization of Industry 5.0 concepts and action plan. Conclusion remarks are given in Section V.

II. INDUSTRY 5.0

Industry 4.0 (coined in 2011) was, or is, considering how technology advancements can help economy, and is focused to a great extent on interconnection of physical and virtual world, humans and devices, concentrating on Internet of Things (IoT) and cyber-physical objects, although it also considers ecological aspects [2,3]. It fosters enhanced efficiency through the technology for the purpose of minimisation of costs and maximisation of profit and, since it is concentrated more on AI-driven technologies and digitalisation than on sustainability and social fairness, it is not sufficient to tackle climate, environmental or social issues [3, 4].

Raised concerns about future human and planet well-being caused deliberations of how the abundance of technology that are developed and used every day can be used to avoid those issues, hence leading to the idea and the concept of Industry 5.0. European Union (EU) puts much effort into Industry 5.0 in the scope of its support to research and innovation. Fig 3. [5] depicts three core elements of the concept [3,5]:

- Human-centricity – Human needs and interests should be put first with ensuring that technology is used to serve people and adapt production process to workers, enabling them to develop new skills and roles and ensuring that their rights are respected. Workers are investment (not cost) and must have opportunity to continually learn to advance their careers. Technology should be adaptive to make

workplaces safer and more inclusive, including remote work, thus increasing job satisfaction and taking care of physical and mental health and well-being. General Data Protection Regulation, White Paper on Artificial Inteligence and Digital Education Action Plan (2021-2027) are examples of steps towards human-centricity in EU.

- Sustainability – Processes should be circular and re-use natural resources, as well as reduce waste and energy consumption, in order to minimize environmental impact, respect the boundaries of the planet and consequently enable safe future for living beings. Circular economy should ensure efficient and effective resource use. Reflection paper Towards Sustainable Europe by 2030 and European Green Deal show EU determination to achieve climate-neutrality.
- Resilience – Production should be more robust and resilient to disruptions, including resilient value chains, flexible business processes and support to critical infrastructure. Resilient industry should cope with various vulnerabilities on global and local level, caused by various sources, such as political or natural. As a key instrument to help in achieving more resilient economy and society, EU statrted Recovery and Resilience Facility.

Industry 5.0 implementation requires new approach to business, as well as new design for business models and decentralisation, meaning that organizations must change, ensuring that sustainability is part of the strategy and that progress can be measured [4]. The concept can also be connected with the concept of Society 5.0 that is characterized by balance between economic development and environmental and societal issues and is a society where technology is used for the benefit of all people [3].

There are many technologes that should enable Industry 5.0, as identified at a virtual workshop with Europe's technology leaders [2]:

- Human-centric solutions and human-machine interaction (robotics, augmented reality, enhancement of physical or cognitive human capabilites…)
- Bio-inspired technologies and smart materials (self-repairing, recyclable, biosensors, adaptive ergonomics…)
- Real time based digital twins and simulation (digital twins of products, processes or systems, simulation of environmental impact…)
- Cyber safe data transmision, storage, and analysis technologies (networked sensors, big data management, edge computing…)
- Artificial intelligence (brain-machine interfaces, energy-efficient AI, causality-based AI…)
- Technoloiges for energy efficiency and trustworthy autonomy (smart dust, low energy data transmission…)

Both chalenges and enablers of Industry 5.0 are social, economic, governmental and political dimension, as well as scalability, interdisciplinarity and transdisciplinarity [2]. Therefore, the success of the new industry paradigm depends also on government support in the form of creating new policies and active participation in the transformation [4].



Figure 1. Industry 5.0 [5]

III. KNOWLEDGE MANAGEMENT

A. *Development of Knowledge Management*

Interest in knowledge management, as well as a number of research papers about the discipline started to grow in 1990s [6], especially after the first knowledge management related scientific conference in 1993 [7]. Development of the field was driven by widespread use of information and communication technologies (ICT) and Internet [8]. Of course, research in topics of knowledge management started even earlier than book „The Practice of Management“ by Peter L. Drucker, the most prominent figure in the field, was published in 1954 [9].

Since then, a large number of defintions of knowledge management were created, in general or from various prespectives. According to one research [10], there are over one hundred definitions, from which authors, regarding most used terms, created two similar definitions, one of them being: „Knowledge Management is the process of creating, sharing, using and managing the knowledge and information of an organization“ [10, p. 14].

Researchers usually divide knowledge management processes into three to six or more [11], but all proposals reflect the flow (or cycle) of knowledge through organization, from its creation or acquisition, to storage, sharing and using. It is obvious that knowledge creation and acquisition are very important processes in knowledge management strategy [12], but knowledge sharing is usually considered to be a critical process within an organization for achieving better performance, competitive advantage and innovativeness [13,14,15,16]. Final goal is, of course, the use of knowledge for achieving organizational goals and strategy.

Stages of knowledge management development identified by McInerney and Koenig [8] show that, although largely impacted with information technology at the beginning, knowledge management over time added

important elements needed for a long-term success and positive impact on organization:

- Stage 1: Information technology – The initial driver of knowledge management was information technology that could enable faster and more effective knowledge sharing with the use of both Internet and intranet. Knowledge management became new product to preserve intellectual capital and enable storage of valuable experience (in the form of „lessons learned“).
- Stage 2: Human and cultural dimension – This stage marks the realization that knowledge management system has little value if people don't use it, or, better to say, participate in it. The purpose of the knowledge management is to help people in making decisions and achieving organizational goals. On the other side, the system also needs employees' knowledge, especially tact expert knowledge. To ensure this, knowledge management needs to be part of organizational culture that foster the learning organization. In this stage communities of practice became important.
- Stage 3: Content and retrievability – With large ammount of information available (today known as big data), it was important for it to be structured and described, so that specific information can be easily retrieved when needed. Enterprise content management became important, as well as structuring knowledge in taxonomies that evolved into ontologies as a form of a knowledge bases.
- Stage 4: External knowledge and information – External knowledge was allways important to organizations, but it is a two way process. Organizations need external knowedge, and they are also providing information to to their customers or suppliers. Opening communities of practice to a wider group of participants can enable more knowledge flow into organization, but it also has some risks of knowledge leak. Usability of the available knowledge depends on its context and relation to other knowledge and informaiton.

McInerney and Koenig [8] marked the start of stage four in the beginning of 2010s, but also noticed that this may be merely the maturing of knowledge management.

B. Knowledge Management Processes

Knowledge management processes are supported with various mechanisms and technologies. Mechanisms are structural organizational means that foster knowledge processes, such as meetings, conferences, tutoring, presentations, projects, models, prototypes, lessons learned, organizational policies or standards [17]. All those mechanisms don't have to but can be supported with various technologies [11, 17, 18]:

- Knowledge creation/acquisition – data and text mining, big data analytics, natural language processing (NLP), machine learning (ML), Internet of Things, augmented reality (AR), content management systems, business intelligence, videoconferences, social media, communites of practice…
- Knowledge storing – databases, knowledge bases, warehouses, blockchain, document management systems, repositories…
- Knowedge sharing – visualisation tools, simulations, webinars, videoconferences, socal media, communities of practice, blogs, wikis, NLP, chatbots, virtual assistants…
- Knowledge use – knowledge-based systems, enterprise resource planning systems, management information systems, cognitive computing, robotics…

One technology and/or mechanism can support several processes. For example, at a videoconference one person shares knowledge and the other acquires it. All processes are connected and successful knowledge management system depends on their smooth performance, ensuring uninterrupted flow of knowledge through organization, although it can be developed to support all of some of them.

C. Artificial Intelligence Supported Knowledge Management Systems

Knowledge management systems can include various information technologies mentioned in previous subsection and many more, depending on the needs of the organization. They can also incorporate AI technologies, also listed in previous subsection, to support some or all knowledge management processes. Most of the systems on the market that include AI technologies are aimed at improving employee decision making and customer service [19,20,21,22,23] with supporting processes of knowledge creation/acquisition, storing an sharing, and consequently knowledge using. Therefore, examples of those systems will be presented.

Document360 [19] is an AI supported knowledge base platform that uses natural language processing, AI algorithms and text analytics to foster knowledge creation, knowledge storing in public or private knowledge base and knowledge sharing. AI algorithms improve content creation and collaboration and enable employees to quickly find and use relevant knowledge. They also enable performance analytics of documents and thus help creators to improve the knowledge base content. Document360 also has various extensions and allows integration with other tools, such as Zendesk or Google Analytics.

Lucy [20] is an AI-powered enterprise knowledge management system and answer engine (chatbot) that also supports knowledge creation, storing, sharing and use. Information in knowledge base can be received in various formats (documents, audio, video) and Lucy uses NLP and machine learning to give relevant answers and knowledge for better decision making.

eGain Knowledge Hub knowledge management software [21] provides content and workflow management, search and knowledge analytics, giving support to all knowledge processes. With the use of NLP, machine learning and case-based reasoning, it enables knowledge access through virtual assistant and supports decision making. New eGain 21 [22] is organised as three hubs: Knowledge Hub, Conversation Hub and Analytics Hub. eGain enables integration with call tracking systems

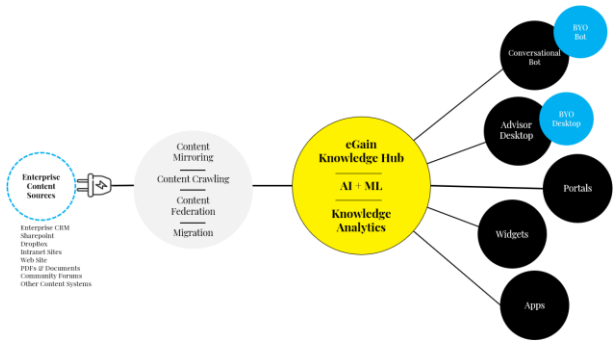


Figure 2. eGain Knowledge Hub with connections [23]

and third party tools, such as Lotus Notes or Microsoft Share Point [21]. Fig. 2 [23] shows eGain Knowledge Hub and connection to various third party information sources as well as with eGain developed tools.

IV. KNOWLEDGE MANAGEMENT SUPPORT TO INDUSTRY 5.0

A. *Industry 5.0 Core Elements and Knowledge Management*

During last ten years in the focus of research came various issues related to the rising concerns or possibilites for future workers because of advancement of the technology: from low-skill tasks workers displacement and less work for information technology specialists due to outosurcing, over possibility for creativity and new human capacities, and lastly to human-computer collaboration with emergence of Industry 5.0, titled Employment 5.0 [24].

Knowledge management benefited from many Industry 4.0 technologies, such as Internet of Things, cloud computing, Blockchain, simulation and modelling, visualisation or automation, especially for knowledge sharing, but also for other processes [25]. These technologies comprise also Industry 5.0 and continue to support knowledge management. On the other hand, knowledge management recognizes the importance of people from its second development stage [8] and acknowledges that all technology that is used to help knowledge management processes is actually used for the purpose of ensuring that workers have the knowledge they need when they need it to make decisions and develop new products and services, thus enabling innovation and progress. Therefore, knowledge management naturally suports humans as the important element of Industry 5.0 and Employment 5.0, where technology is used to serve people.

Sustainability issues are already considered by many organizations [26,27,28], where knowledge management plays an important role in providing critical knowledge for making the organization and subsequently its processes and products more sustainable, especially through innovation [29,30,31]. It also fosters resilience - many organizations that possesed required knowledge and technology at the beginning of Covid-19 pandemic made easy transition to making products needed to overcome this crisis [32,33]. Another example is a supply chain – suply chain risks lead to more focus on all knowledge management processes in organizations, which improves risk management culture and consequently supply chain resilience [34].

B. *Industry 5.0 Action Plan and Knowledge Management Processes*

ESIR, expert group on the economic and societal impact of research and innovation that advises European Commission, has proposed action plan [4] that should lead to Industry 5.0 implementation. This plan is in large part oriented towards sustainability, meaning that other two aspects, human-centricity and resilience should follow from required transformations of business. To achieve sustainability through technology and production, knowledge of new technologies and knowledge how to implement and use them to achieve it is necessary, as well as knowledge that can lead to innovations in sustainable technologies, processes and products.

The role of knowledge management and its processes in supporting Industry 5.0 can be discussed through support in achieving specific six elements of the Industry 5.0 action plan [4]. Active participants in implementation of the plan are government, industry, financial sector, as well as civil society and NGOs. The role of knowledge management in this paper is observed from the industry side and knowledge management processes are connected to four of action plan six elements, which is supported with literature. Last two elements are dependent on the government support.

First element, „The triple Imperative and the role of industry“ [4] requires from industry transformation of business models towards sustainablity, changing the mind-set and redesign of value chains to include new technologies, circular economy and well-being, as well as to adopt metrics that will measure the progress. For all major changes in organizations, such as change of business model and value chain, knowledge needed for making optimal decisions must be available. Innovation of green processes and green technology is important for the transition to sustainability, where knowledge management plays a major role [29,30,31]. Although all processes can impact green innovation [29], knowledge creation and acquisition process is crucial in gaining knowledge that can result in innovation.

„New Economic Orientation and New Approaches to Industry Performance“ [4], as a second element, asks industry to observe return on investment with the purpose to optimise a return on invested energy or natural assets, to support „resource efficiency first“ principle and to review all value chains that result in products on EU market. It is well known that innovations in energy and resource efficiency require knowledge acquisition from various stakeholders and other sources [35] and that knowledge sharing and transfer is critical for decision making and their succesful implementation [36]. This emphasizes knowledge storing after acquisition and before sharing as important process that enables formating and representing gathered knowledge for transfer and use.

Reshoring (backshoring) of economic activities to lessen carbon and material footprint is required from industry in third element, „New Design for Business

Models, Value Chains and Supply Chains“ [4]. When returning production to Europe or any home country, it is important that all relevant knowledge is also transfered and reintegrated [37,38], and therefore process of knowledge sharing is the one that plays important role in achieving this element. With the change of employee base, part of tacit knowledge can be lost [37], which emphaiszes knowledge creation and acquisition process during and after reshoring for continuation of business with less disruption.

According to fourth element, „New Purpose for Digital Transformation, Achieve Life within Planetary Boundaries“ [4], digital technologies should be used in industry to make positive impact on climate. This again stresses connection of knowledge management and sustainability [29,30,31], and need for green innovations in technology with the help of knowledge creation process. If environmental aspects are included in knowledge management, all its processes positively impact on green technology innovation [29], especially knowledge acquisition that includes big data analytics [39].

Fifth element of Industry 5.0, „New Approaches to Policy-Making“ [4], includes development of green and social industrial strategy and policies and interaction with public sector during transformation (where the role of public sector in transition to Industry 5.0 should be revised). Final, sixth element, „New Capabilites and Approaches to Research and Innovation“ [4] requires the change of regulations and research and innovation policies, as well as increase of funding connected to sustainable solutions. Both elements are dependent on the goverment, which should interact with and consult other stakeholders. Therefore it is clear that industry role in those two elements is only to share its knowledge to provide input for new strategies and policies for easier implementation of sustainability in business.

C. Industry 5.0 Action Plan and Knowledge Management Technologies

Industry 5.0 action plan calls for sustainability with new business models, redesigned value chains, circular economy, resource efficiency, reshoring and green digital technologies supported with new government strategy, policies and regulations. Various knowledge management technologies mentioned in Section III B and many of them in the scope of Industry 4.0 and belonging to AI are used to achieve these specific goals.

Circular economy itself implies changes in business model and value chains, and demands resource efficiency through reuse of materials and products or their parts or through products durability [40]. The smart circular economy framework, shown in Fig. 3 [41] explains connection of AI technologies with data transformation into knowledge and how they can help in resource optimization. These technologies, that are listed on the right side of the Fig. 3 [41], support processes of knowledge creation/acquisition as well as knowledge storing in desired forms and subsequently its use - for example, various sensors can acquire data about product condition in phases of production and use, and analysis of

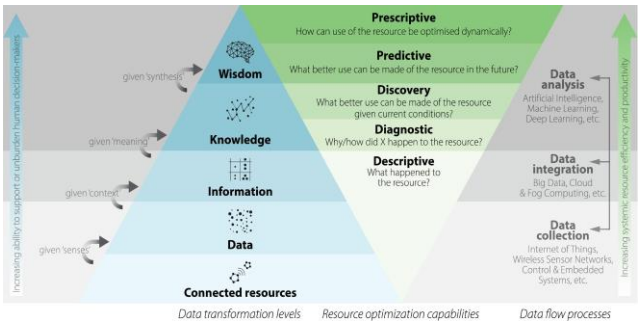


Figure 3. The smart circular economy framework [41]

these data can result in knowledge that can be applied for improvements in product quality and durability.

Internet of Things in general supports knowledge creation from data embedded into objects, and augmented reality also creates, stores and transforms vast ammounts of data into knowledge that can be shared to enable easier learning and then the use of learned knowledge [42]. Machine learning can be used to design and test new sustainable materials [43] through knowledge acquisition and storage (and analysis) processes. This can lower the need for raw natural materials, ensure material reuse and durability, thus enabling resource efficiency. New digital technologies, such as IoT, sensors for real-time data collection or data analytics, can be of support to reshoring [44], and they foster knowledge creation and acquisition process that is important, as already mentioned, because of need for recovery and reintegration of relevant and tacit knowledge [37,38].

Need for technology support, especially for knowledge creation/acquisition and then storing (for subsequent sharing and use) to achieve elements of action plan is obvious. All AI supported KMS solutions in subsection III C and similar systems can therefore help in achieving Industry 5.0 goals. Action plan asks for those digital technologies to have positive impact on climate [4]. From technologies mentioned in this subsection and examples of AI supported KMS in subsection III C it is clear that many AI technologies for support to knowledge management processes in the scope of Industry 5.0 are related to data collection and analysis. However, there is currently a lot of discussion about large ammounts of energy that large ammounts of data, information and knowledge consume [45]. Controversial topic is also a large energy consumption that an AI model training requires [46], thus making a large carbon footprint.

The shift to renewable energy resources is invitable, where many AI technologies have an important role, for example smart grids or virtual power plants, making knowledge sharing process also important to enable succesful technology implementation [47]. For instance, Google claims that it operates their data centres on carbon-free and renewable energy, continually lowers power consumption, uses wastewater for cooling and new solutions to lower water consumption (including ML), and that companies moving to their cloud solutions from local servers also add to energy efficiency [48,49,50]. 2023 Best Practice Guidelines for the EU Code of Conduct on Data Centre Energy Efficiency [51] list many factors that can make data centres more sustainable, such as

proportion of energy used from renewable sources, optimal extended operating temeperature and humidity range, air flow and liquid management for cooling, equipment containing high efficiency AC/DC power converters, deployment of virtualisation technologies, selection and/or development of energy efficient software, or data management policy (to lower energy consumption with deleting unnecessary data), to name a few. Carbon footprint of machine learning models can be evaluated by tracking electric energy consumption allowing for energy efficiency improvement with different training methods [52]. Selection of efficient ML model architectures, processors optimized for ML training and cloud computing using carbon neutral cloud can significantly lower the energy consumption [53]. In this way digital transformation can get closer to support Industry 5.0 action plan.

V. CONCLUSION

Industry 5.0 in a way represents new hope for the future of human society, trying to overcome issues that are concerning it for a long time. It represents the idea to divert the technological advancement into more greener track, because it can not be stoped. And it also requires more technological innovations that can enable climate-neutral production based in large part on recycling and reuse (and other R's of circular economy), with less energy and resources consumption. Besides this, they must also support the well-being of workers, of those that will create them and of those that will use them.

It is obvius that those goals first require new perspective, and then research that will lead to creation of new knowledge that will then lead to innovations – known as knowledge and innovation cycle, which is part of knowledge management. Investigation of connection of KM processes and technologies with Industry 5.0 action plan clearly shows the role that knowledge creation/acquisition and storing, and associated supporting AI technologies have in achieving Industry 5.0 action plan. Therefore knowledge management should be considered as an important part of transition to Industry 5.0.

Besides the problem of AI technology sustainability, which is addressed in previous section, an important issue is also AI ethics, especially conisdering human element of Industry 5.0. This problem is already tackled, for example with EU Ethics Guidelines for Trustworthy AI [54], but is still gaining momentum because of ChatGPT [1] and other large language models. Government support with new regulations is crucial for dealing with this issue, as is in acheiving elements of Industry 5.0 action plan.

To successfuly implement action plan, technologies that support knowledge management processes must themselves be aligned with principles of Industry 5.0. Only continual access to needed knowledge and carefull implementation of innovative technologies, taking into account specific needs of organizations can foster more human-centric, sustainable and resilient industry and consequently society.

REFERENCES

[1] OpenAI, „ChatGPT: optimizing language models for dialogue,“ https://openai.com/blog/chatgpt/

[2] J. Müller, „Enabling technologies for Industry 5.0,“ European Commission, Brussels, Belgium, Ind. Exp. Rep., Sept. 2020.

[3] M. Breque, L. De Nul, and A. Petridis, „Industry 5.0 – Towards a sustainable, human-centric and resilient European industry,“ European Commission, Brussels, Belgium, Policy Brief, Jan. 2021.

[4] S. Dixon-Declève et al., „Industry 5.0: A transformative vision for Europe,“ European Commission, Brussels, Belgium, ESIR Policy Brief No. 3, Dec. 2021.

[5] X. Xun et al., „Industry 4.0 and Industry 5.0—Inception, conception and perception,“ Journal of Manufacturing Systems, vol. 61, pp. 530-535, October 2021

[6] R. De Hoog, „Knowledge Management: From idea to a discipline,“ in Knowledge Management, Organizational Intelligence and Learning, and Complexity, Vol. III, D. L. Kiel, Ed. Oxford, UK: Eolss Publishers Co. Ltd., 2009, pp. 163-178.

[7] L. Prusak, „Where did knowledge management come from?,“ IBM Systems Journal, vol. 40, no. 4, pp. 1002-1007, October 2001.

[8] C. R. McInerney and M. E. D. Koenig, Knowledge Management (KM) Processes in Organizations,“ San Rafael, CA: Morgan & Claypool, 2011.

[9] The Editors of Encyclopeedia Britannica, „Peter F. Drucker,“ Britannica, OpenAI, "ChatGPT: optimizing language models for dialogue," https://www.britannica.com/biography/Peter-F-Drucker

[10] J. Girard, and J. Girard, „Defining knowledge management: toward an applied compendium,“ Online Journal of Applied Knowledge Management, vol. 3, no. 1, pp. 1-20, 2015.

[11] K. Dalkir, Knowledge Management in Theory and Practice, Third Edition, Cambridege, MA: The MIT Press, 2017.

[12] B. Choi and H. Lee, „Knowledge management strategy and its link to knowledge creation process,“ Expert Systems with applications, vol. 23, no. 3., pp. 173-187, October 2002.

[13] P. Hendriks, „Why share knowledge? The influence of ICT on the motivation for knowledge sharing,“ Knowledge and Process Management, vol. 6, no. 2, pp. 91-100, June 1999.

[14] S. Wang and R. A. Noe, „Knowledge sharing: A review and directions for future research,“ Human Resource Management Review, vol. 20, no. 2, pp. 115-131, June 2010.

[15] D. Mesner Andolšek and S. Andolšek, „Knowledge sharing in an organization from the perspective of the individual,“ International Journal of Cognitive Research in Science, Engineering and Education, vol. 3, no. 2, pp. 65-76, December 2015.

[16] M. Asrar-ul-Haq and S. Anwar, „A systematic review of knowledge management and knowledge sharing: trends, issues, and challenges,“ Cogent Business & Management, vol. 3, no. 1, 1127744, 2016.

[17] I. Becerra-Fernandez and R. Sabherwal, Knowledge management: systems and processes. New York, NY: Routledge, 2014.

[18] S. Lovrenčić and V. Sekovanić, „Knowledge management in disruptive times,“ in Proc. of Third International Scientific Conference on Economics and Management, pp. 373-381, March 2019.

[19] S. Murugesan, „Leveraging Artificial Intelligence in knowledge management,“ Document360, https://document360.com/blog/ai-in-knowledge-management/ (accessed Mar. 22, 2023)

[20] Equals 3, „AI-Powered Enterprise Knowledge Management,“ Equals 3, https://www.lucy.ai/ai-powered-enterprise-knowledge-management (accessed Mar. 22, 2023)

[21] eGain, „#1 Knowledge Management Platform - eGain Knowledge Hub,“ eGain 21, https://www.egain.com/products/knowledge-hub/ (accessed Mar. 22, 2023)

[22] eGain, „What's New in eGain 21,“ eGain 21, https://ebrain.egain.com/kb/newin21/home (accessed Mar. 22, 2023)

[23] eGain, „New in Knowedge Hub,“ eGain 21, https://ebrain.egain.com/kb/newin21/browse/easy-1622/New-in-Knowledge-Hub (accessed Mar. 22, 2023)

[24] O. Kolade and A. Owoseni, „Employment 5.0: the work of the future and the future of work,“ Technology in Society, vol. 71, 102086, November 2022.

[25] A. P. Lista, G. L. Tortorella, and M. Bouzon, „Knowledge management benefited by Industry 4.0 integrationa – a scoping review,“ in XI International Congress on Knowledge and Innovation, 2021.

[26] S. Todd-Ryan, „Who Are The 100 Most Sustainable Companies Of 2020?,“ Forbes, https://www.forbes.com/sites/samanthatodd/2020/01/21/who-are-the-100-most-sustainable-companies-of-2020/ (accessed Mar. 21, 2023)

[27] C. Ames, „The 15 Most Environmentally Friendly & Sustainable Companies (2022),“ GROW ensemble, https://growensemble.com/environmentally-friendly-companies/ (accessed Mar. 21, 2023)

[28] American Sustainable Business Network, „Business Members,“, ASBM, https://www.asbnetwork.org/asbn-members (accessed Mar. 21, 2023)

[29] A. Jawad, and M. Sağsan, „Impact of knowledge management practices on green innovation and corporate sustainable development: A structural analysis,“ Journal of cleaner production vol. 229, pp. 611-620, August 2019.

[30] A. Mwanzu, R. W. Dickens, and F. Kibet, „Knowledge management for innovativeness and sustainable organizational development: a case study of selected corporate institutions in Kenya,“ Library Philosophy and Practice (e-journal), 5086, 2021.

[31] S. Wang et al., „Achieving green innovation and sustainable development goals through green knowledge management: Moderating role of organizational green culture,“ Journal of Innovation & Knowledge, vol. 7, no. 4, 100272, October 2022.

[32] World Economic Forum, „From perfume to hand sanitiser, TVs to face masks: how companies are changing track to fight COVID-19,“ WEF, https://www.weforum.org/agenda/2020/03/from-perfume-to-hand-sanitiser-tvs-to-face-masks-how-companies-are-changing-track-to-fight-covid-19/ (accessed Mar. 21, 2023)

[33] T. Šimunić-Bendić, „Kaštelanin se bacio u izradu zaštitnih pregrada i sada se ne može okrenuti od posla: Još da proizvodim pleksiglas, postao bih Rockefeller,“, Slobodna Dalmacija, https://slobodnadalmacija.hr/vijesti/biznis/kastelanin-se-bacio-u-proizvodnju-zastitnih-pregrada-i-sada-se-ne-moze-okrenuti-od-posla-jos-da-proizvodim-pleksiglas-postao-bih-rockefeller-1019121 (accessed Mar. 21, 2023)

[34] I. Ali, I. Golgeci, and A. Arslan, „Achieving resilience through knowledge management practices and risk management culture in agri-food supply chains,“ Supply Chain Management: An International Journal, vol. 28, no. 2, pp. 284-299, March 2023.

[35] K.Rennings, and C. Rammer, „Increasing Energy and Resource Efficiency Through Innovation – An Explorative Analysis Using Innovation Survey Data,“ Discussion Paper No. 09-056, Centre for European Economic Research, 2009.

[36] M- Adams, and S. Comber, „Knowledge transfer for sustainable innovation: a model for academic-industry interaction to improve resource efficiency within SME manufacturers,“ Journal of Innovation Management in Small & Medium Enterprises, vol. 2013, 999612, 2013.

[37] B. B. Nujen, and R. Damm, „The need for knowledge management when backsourcing is embraced,“ in Proc. of IFIP International conference on Advances in Production Management Systems (APMS), pp. 748-755, September 2016.

[38] D. Fjellstrom et al., „Manufacturing Relocation ambiguity model: a prerequisite for knowledge management,“ British Journal of Management, Special Issue Article, Early View, 2023.

[39] S. Sahoo, A. K. Saumyaranjan, and A. Upadhyay, „How do green knowledge management and green technology innovation impact corporate environmental performance? Understanding the role of green knowledge acquisition,“ Business Strategy and the Environment, vol. 32, no. 1, pp. 551-569, January 2023.

[40] M. Geissdoerfer et al., „Circular business models: A review,“ Journal of cleaner production, vol. 277, 123741, December 2020.

[41] E. Kristoffersen et al., „The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies,“ Journal of business research, vol. 120, pp. 241-261, November 2020.

[42] T. Miandar, A. Galeazzo, and A. Furlan, „Coordinating knowledge creation: a systematic literature review on the interplay between operational excellence and Industry 4.0 technologies,“ in Knowledge Management and Industry 4.0: New Paradigms for Value Creation, M. Bettiol, E. Di Maria, and S. Micelli, Eds. Cham: Springer Nature Switzerland, 2020, pp. 137-162.

[43] R. Hardian, et al., „Artificial intelligence: The silver bullet for sustainable materials development,“ Green Chemistry, vol. 21, pp. 7521-7528, November 2020.

[44] S. Cosimato, and R. Vona, „Digital innovation for the sustainability of reshoring strategies: A literature review,“ Sustainability, vol. 13, 7601, 2021.

[45] F. Lucivero, „Big data, big waste? A reflection on the environmental sustainability of big data initiatives,“ Science and engineering ethics, vol. 26, no.2, pp. 1009-1030, 2020.

[46] J. Saul, and D. Bass, „Artificial Intelligence is Booming – So Is Its Carbon Footprint, Bloomberg,“ https://www.bloomberg.com/news/articles/2023-03-09/how-much-energy-do-ai-and-chatgpt-use-no-one-knows-for-sure? (accessed Mar. 25, 2023)

[47] A. Tanveer, et al., „Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities,“ Journal of Cleaner Production, vol. 289, 125834, March 2021.

[48] Google, „Realizing a carbon-free future - Google’s Third Decade of Climate Action,“, gstatic.com, https://www.gstatic.com/gumdrop/sustainability/carbon-free-by-2030.pdf (accessed Mar. 25, 2023)

[49] DeepMind, „DeepMind AI Reduces Google Data Centre Cooling Bill by 40%“, Deepmind, https://www.deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40 (accessed Mar. 25, 2023)

[50] Google, „Google Data Centers - Efficiency“, Google, https://www.google.com/about/datacenters/efficiency/ (accessed Mar. 25, 2023)

[51] M. Action, P. Bertoldi, and J. Booth, Eds., „2023 Best Practice Guidelines for the EU Code of Conduct on Data Centre Energy Efficiency,“ European Comission, Ispra, Italy, JRC Tech. Rep. 132576, 2023.

[52] S. A. Budennyy et al., „Eco2ai: carbon emissions tracking of machine learning models as the first step towards sustainable AI,“ Doklady Mathematics, vol. 106, no. 1., pp. 118-128, December 2022.

[53] D. Patterson et al., „The carbon footprint of machine learning training will plateau, then shrink,“ Computer, vol. 55, no.7, pp. 18-28, July 2022.

[54] High-Level Expert Group on Artificial Intelligence, „Ethics Guidelines for Thrustworthy AI,“ European Comission, Brussels, Belgium, April 2019.

Learning Translation Model to Translate Croatian Dialects to Modern Croatian Language

Blagica Penkova[1], Maja Mitreska[1], Kiril Ristov[1], Kostadin Mishev[1,2] and Monika Simjanoska[1,2]
[1]iReason, LLC, Skopje, N. Macedonia
[2]Ss. Cyril and Methodius University,
Faculty of Computer Science and Engineering,
Skopje, N. Macedonia
{blagica.penkova, maja.mitreska, kiril.ristov, kostadin.mishev, monika.simjanoska}@ireason.mk,
{kostadin.mishev, monika.simjanoska}@finki.ukim.mk

***Abstract*—The task of translating dialects into modern language is a challenging task since it requires enormous parallel data. Such data is hard to find especially when it comes to low-resource languages. Among them is the Croatian language which has very few datasets in the standardized version, let alone enough resources for its dialects. In the Croatian language, there are three main groups depending on the geographical position Shtokavian, Kajkavian, and Chakavian which also include other more specific local dialects. For solving these kinds of problems, unsupervised neural machine translation (UNMT) models are considered a good solution since they can be trained on monolingual data. In this paper, we propose an application of a modified version of the state-of-the-art UNMT model for dialect translation on monolingual data of the standardized Croatian language and its dialects. We experimented with several types of cross-lingual embeddings of the input data to determine the best approach that can leverage the similarities and differences between the language and the dialect. All techniques are evaluated on a small parallel dataset using the BLEU metric. Translating these dialects to the modern Croatian language helps in improving communication and access to information for all speakers.**

***Keywords*—Unsupervised Neural Machine Translation, Dialect Translation, Low-resource languages, Cross-lingual embeddings, Croatian Language, Croatian Dialects.**

I. Introduction

Machine translation (MT) is an automatic approach to converting written text from one language to another. The translating task is widely explored in the field of Natural Language Processing (NLP) and many researchers are introducing novel solutions as the development of the technology continues. The translation task can be roughly divided into two groups: statistical MT and neural MT, which took a rise with the introduction of sequence neural networks [1]. The automatic MT replaces the need to use human resources to translate textual forms, hence accelerating the process of their generation. The translation task enables more content to be available to larger groups of audiences and therefore all the information is publicly accessible.

There is extensive research in this field and many different translation architectures have been covered throughout the years. Beginning with more traditional approaches from the statistical MT branch and continuing with state-of-the-art neural networks. The traditional approaches require extensive calculations and are more time and resources-consuming. On the other hand, the unsupervised approach of MT is gaining momentum since the need for parallel sentences cannot be always satisfied. This approach tries to find the underlying hidden meaning and characteristics of an unlabeled dataset. Since the unsupervised approach finds the internal structure of the dataset it solves the problems which emerge from the disadvantages of the supervised approach.

Most of the papers that propose using the unsupervised approach for machine translation are basing their architectures on using monolingual datasets in the two languages and construing the network around them. The establisher of the usage of monolingual data in an unsupervised approach is [2] where they provide several decipherment approaches. In [3], the authors map the two monolingual datasets into the same latent space and learn to reconstruct the sentences from the shared features. The state-of-the-art model [4] is the most widely used and explored model when experimenting with translation between two languages.

With the advancement of unsupervised approaches, not only the translation of two mostly standardized languages is enabled, but a number of possibilities emerge for re-purposing this approach for dialect translation. Dialect translation is an even more challenging task since many dialects can be found in spoken form rather than written form, especially when it comes to dialects spoken by a small number of people. Moreover, the acquisition of sentences in a dialect that comes from a low-resource language is particularly complex. Same to regular translation, dialect translation can be solved using supervised or unsupervised approaches. In the supervised approach, first, a number of parallel sentences are gathered and then phrase-based architectures [5] are used for learning the mappings of the words.

Some papers exploring this field focus on constructing the parallel dataset and then proposing appropriate architecture for solving the translation problem. In [6] a grapheme-to-phoneme (g2p) model is proposed for learning the Malay dialect translation. The vast majority of the papers that propose parallel datasets for supervised learning implement more methods from the statistical approach [7]–[10], or combine them with a language model [11]. The unsupervised approach also

developed popularity in the field of dialect translation since here even fewer resources are available. The pioneer in this field is [12] where they propose two systems, the first one is a standard attention sequence-to-sequence model with cosine similarity to capture the similarities between the language and the dialect, and a second system based on the Google NMT (GNMT) model. The paper explored the methods of the Modern Standard Arabic (MSA) language. In [13] and [14] Transformer based architectures are implemented like in [14] for Mandarian-Cantonese language translation.

Since most of the papers do not include specific language characteristics and leave it to the model to extract the hidden relationships, can be repurposed for other languages. The usage of monolingual data is immensely helpful when working with low-resource languages where the resources for standardized languages are scarce, and the resources in dialects are even scarcer. Such a language is the Croatian language, the standard language in Croatia. The Croatian language consists of three primary dialects: Shtokavian, Kajkavian and Chakavian. The Shtokavian dialect is the one that the standard Croatian language is based on. The dialects are named after the interrogatory pronouns "što", "kaj" and "ča". Since there is no modern literature written in these languages we combined them to create one model that can learn to distinguish the dialects from the standard Croatian language, and also learn the similarities and the differences between the dialects and the modern version.

In this paper, we propose developing a dialect translation model that can translate a dialect sentence into a sentence from the standard Croatian language. The model uses an unsupervised approach on two monolingual datasets. The first one consists only of Croatian sentences in the modern version and the other only dialect sentences. It is necessary to be mentioned that these sentences are not parallel and are acquired from various resources. So, the first contribution of this paper is constructing three datasets, one consisting only of Shtokavian sentences, one of the Kajkavian sentences and the last one consisting of sentences in Chakavian. All combined form the dialect dataset with 53K sentences. The second contribution of this work is that, to our knowledge, the dialect translation model is the first one in Croatian. There is no prior work in the dialect translation tasks for this kind of low-resource language, especially in Croatian, or any of the Slavic languages.

In the next few paragraphs, we talk about recent advancements in the field of UNMT architectures for dialect translation. In section III, we discuss the introduced dataset and the implemented translation model. In IV we present the results obtained through the experiment as well as their fluency with the help of a human expert in the Croatian language. Our work is concluded in the final section V where we present our findings. The presented work is made publicly available at [1].

II. Related Work

A major problem when dealing with low-resource languages is the lack of available data. In the course of our research, we found related work that allows translation to be performed without a parallel corpus [4]. This paper suggested a way to be able to translate using only monolingual data from the source and the target language. The paper proposes two model variants, a neural and a phrase-based model. Long Short-Term Memory [15] and Transformer cells [16] are the basis for the NMT models. For the Transformer, they use 4 layers both in the encoder and in the decoder. During the first generation, Moses' smoothed n-gram language model is used by default in the PBSMT. In the paper, five language pairs are considered: English-French, English-German, English-Romanian, English-Russian and English-Urdu. They use a dataset of around 2 million sentences of the source and the target language. For tokenization, they use Moses scripts.

The NMT is trained with a 60,000 Byte-Pair Encoding (BPE) tokenizer in the pre-processing phase so the vector embedding is kept in a reasonable dimension while PBSMT is trained with true-casing. Both models required either n-gram embeddings or cross-lingual BPE embeddings. With an embedding dimension of 512 and a context window of size 5 and 10 negative samples, the embeddings are made using fastText [17]. For NMT, fastText is used after the source and target corpora have been concatenated. For PBSMT, n-gram embeddings are first created independently on the source and target corpora, and they are then aligned using the MUSE library. The results in this paper demonstrate that the unsupervised NMT and PBSMT systems are significantly superior to previous unsupervised baselines. For the translation task, the Urdu-English result that the model achieves 12.3 BLEU using only a validation set of 1800 sentences and on the Russian-English translation task the model receives a BLEU score of 16.6.

Even though some languages are widely spoken, there is no known natural language processing (NLP) work on these languages, such as Pidgin English, which is the most widely spoken language in West Africa [18]. The goal of this paper is to translate from Pidgin English to English only by using a monolingual dataset. The model that is used in this paper is based on the previous paper, UNMT [4]. They used a Transformer with 10 attention heads. There are 4 encoder and 4 decoder layers with 3 encoder and decoder layers shared across both languages. They obtained the data from newspapers and collected a corpus of 56695 Pidgin sentences. They trained cross-lingual embedding via monolingual mapping where a linear mapping is learned between already trained monolingual word embedding. They got the English embeddings by using already pre-trained Glove English embeddings. And the Pidgin embeddings are done by training new word embedding on Pidgin (since they share common words with English) meaning first, they got the English embeddings and fine-tuned them with new Pidgin words. And the Unsupervised cross-lingual embeddings are made using MUSE. The training with an Unsupervised Neural Machine Translation model between Pidgin and English achieves BLEU scores of 7.93 from Pidgin to English and 5.18 from English to Pidgin.

Related research on dialect translation has been done in many languages. For instance, the multitask learning model proposed in this article [19]is for converting Arabic dialects

[1]https://github.com/Blagica88/CroatianDialectTranslation.git

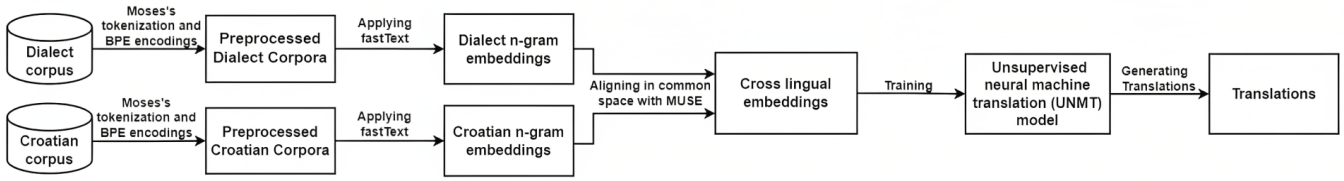


Fig. 1. The structure of our proposed methodology

into Modern Standard Arabic (MSA). In this study, an integrated neural machine translation model was developed and trained with decoders shared across all language pairs and each source language with its own encoder. Consequently, 13805 sentence pairs were trained for Levantine dialects (LD) and 17736 sentence pairs for Maghrebi dialects (MD) that are collected from TV shows, movies and social media. Arabic tokens are space-separated, and modern standard Arabic and English languages are tokenized using the Python tokenizer with default English settings. The findings in this publication demonstrate that, in comparison to the individually trained model, their suggested MTL model may guarantee superior translation quality.

This paper [14] is most similar to our problem because it attempts to exploit the commonality and diversity between dialects to create an unsupervised translation model that accesses only monolingual data. The translation task is between Mandarin and Cantonese, which are official languages and the most widely used dialect in China. Each of the MAN and CAN monolingual training corpora consist of 20M sentences and they also use parallel datasets for evaluation. There are two stages to the training procedure. 1) A commonality model that learns to identify common features in all dialects 2) Diversity modelling that creates connections between various expressions. The commonality modelling is done by training one model for both two dialects and diversity modelling is done with back translation. The model architecture is the same as in the previous paper [4]. They also use UNMT. The main difference is in the embeddings because they propose pivot-private embedding and layer coordination to jointly balance commonality and diversity. The Pivot learns to share a part of the features while the Private captures the word-level characteristics in different dialects. The results are that their model outperforms rule-based simplified and traditional Chinese conversion and conventional unsupervised translation models over 12 BLEU scores.

III. Methodology

In this section, we explain our proposed methodology demonstrated in Figure 1. We present the dataset and its creation, data pre-processing, the cross-lingual embeddings, and the utilization of the UNMT architecture to build our translation model.

A. Dataset

The biggest challenge was gathering enough data for the language model that would be used to translate dialects of Croatian into standard Croatian language. The data used to generate the Croatian dataset was collected from a "24sata" online news portal. The 24sata news portal covers a variety of news articles. More than 650,000 Croatian articles from 2007 to 2019 are included in the collection, along with tags [20]. Our dataset consists of 53K sentences extracted from the news articles from "24sata".

All the resources that we managed to find in Kaykavian, Shtokavian and Chakavian are before the standardization of the Croatian language hence all of them are old poems and/or dramas. We found the following books on the Shtokavian language: "Pesme Meneti I Ora Dria" collected by Vatroslav Jagic [21], "Novela od Stanca" by Marin Držić [22], "Pjesni ljubene" by Ore Dria [23],and "Razgovor ugodni naroda slovinskoga" by Andrije Kaia Mioia [24], "Satir Iliti Divji Čovik" by Matija Antun Relković [25], "Priče iz davnine" by Ivana Brlić Mažuranić [26], "Kvas bez kruha" by Antun Nemčić [27] and "Izabrane štokavske pjesme" by Fran Galović [28]. The following books are in Chakavian: "Skladanja izvarsnih pisan razliih" by Hanibala Lucia [29], "Ribanje i ribarsko prigovaranje" by Petra Hektorovia [30], "Jeđupka" and "Pelegrin Sabu Mietiu" by Mike Pelegrinovia [31], [32] and "Vazetje Sigeta Grada" by Brne Kar [33]. Additionally, we were only able to find one book, "Matija, Grabancija" by Tituš Brezovačk [34] on the Kajkavian dialect. The discovery of TV series on Croatian dialects had the largest influence on our dataset. We were able to find two written books for the series "Projsaci i Sinovi" by Ivan Raos [35] and "Velo Misto" by Miljenko Smoje [36] which are written in the Shtokavian and Kajkavian dialects, respectively. From all of the collected data, we extracted each sentence as a separate input sample. After cleaning, we were able to gather 53K sentences.

Even though we acquired monolingual data for training, parallel phrases were necessary for testing and validating our model. Since parallel sentences were required for the translation of Croatian dialects into the standardized form of the language, we collaborated with a human Croatian specialist to create them. As a result, our validation and test datasets include 30 sentences each in a dialect as well as 30 sentences in standard Croatian.

Moreover, we were able to find parallel word dictionaries for both Shtokavian and Chakavian and utilized them in the cross-lingual phase.

B. Cross-lingual Embedding

For the purposes of the paper and the model, before training the model on the monolingual data, we first need to embed the

words of the sentences to capture their context as well as the context of the words. The architecture described in III-C requires monolingual embeddings for both of the languages and shared cross-lingual embeddings learnt over the concatenated datasets.

1) Dialect Monolingual Embedding: The dialect dataset is first tokenized using Moses tokenizer and encoded using Byte-Pair Encoding (BPE). Furthermore, the dataset is used to train a fastText model to learn the representations of the word.

2) Croatian Monolingual Embedding: Similarly to the dialect monolingual embeddings, the Croatian monolingual embeddings are learnt using the fastText skip-gram model. The learnt embeddings are further used in the creation of cross-lingual embeddings.

3) Cross-lingual Embedding: For the cross-lingual embeddings, we experimented with two approaches. The first approach is based on [14] where besides the monolingual embeddings, we train another fastText model on the concatenated and shuffled dataset created from the two monolingual datasets. The second approach is based on MUSE [37] where the monolingual fastText embeddings are aligned in a common space to obtain multilingual word embeddings. The alignment is done using the unsupervised approach and only utilizes a bilingual dictionary of pairs of dialect words and standard words for the evaluation. The MUSE-aligned centred embeddings between the two languages are learnt using adversarial training and (iterative) Procrustes refinement. For the purposes of our work, we experiment with both approaches and obtained better results with the MUSE-aligned approach.

C. Architecture

We used [14] approach as a reference and built our model on top of their source code. The used architecture is a Transformer based. The number of layers in the encoder and decoder and the number of layers to share between the encoder and decoder were all assigned to 4. Adam is the optimizer utilized for the encoder and decoder, and the learning rate is 0.0001. The pre-trained language embeddings described above are employed in the training process. A larger embedding dimension can potentially capture more complex relationships between words and their context, and as a result, the dimensionality of our embeddings is set to 1024 and on shared embedding layers to 512. The number of training examples per epoch is 200 and the number of examples in each training batch is 32. Unlike [14] which uses cross-lingual embeddings consisting of pivot-private embeddings, we employ MUSE-aligned embedding on both of the datasets. That way we obtained higher results when evaluating the model.

D. Evaluation Metrics

The evaluation of the model is done using the BLEU metric which is standard in machine learning problems, especially in translating tasks. It measures the quality of the translated text to the original text. The standard definition of the BLEU metric is that it compares the machine-generated sentence to a set of reference sentences. The obtained score indicated how similar the candidate text is to the reference texts [38].

IV. Results

In this section, a comparison is made of the experiments done over the architecture with the two different embeddings. The cross-lingual shared embeddings learned with MUSE are in fact aligned embeddings from the two separate fastText models trained on separate language-specific data. For the second experiment, the fastText embeddings are trained on the concatenated dataset of the sentences in Croatian and in dialect. Moreover, the embeddings utilized in the training process of the model are in the form of shared-private, meaning that the final word embedding consists of shared embeddings learned from the concatenated dataset, and private embeddings learned from the language-specific datasets.

TABLE I
Results of the models during evaluation

Metric	Valid	Test
Dialects-Croatian	9.83	12.8
Croatian-Dialects	7.86	9.71
Dialects-Croatian-Dialects	42.27	29.07

Table I presents the best-obtained results from the first experiment with MUSE embeddings. We evaluated the experiments on a manually gathered and translated dataset with the help of a human expert. The models were trained to translate the sentences in both directions, dialect-to-standard language translation and standard-to-dialect language translation. Moreover, we included the dialect-standard-dialect back translation to check the quality of the translated sentences.

In Figure 2, four different BLEU scores are reported where each BLEU score indicates the degree of overlap between the predicted translations and reference translations in terms of n-gram precision. The notation is in the form of score metric, source language, target language, and type of dataset, where the source language and the target language can receive a 'di' value for the dialect sentences or a 'cr' value for the sentences in the standard Croatian language and the type of dataset can be validation or test set.

- **bleu_di_cr_valid:** Model's performance in translating from Croatian dialect to standard Croatian on the validation set.
- **bleu_cr_di_valid:** Model's performance in translating from standard Croatian to Croatian dialect on the validation set.
- **bleu_di_cr_test:** Model's performance in translating from Croatian dialect to standard Croatian on the test set.
- **bleu_cr_di_test:** Model's performance in translating from standard Croatian to Croatian dialect on the test set.

In Figure 3, two different BLEU scores are reported that measure the quality of the back-to-back translations:

- **bleu_di_cr_di_valid:** Model's performance in translating from dialect to standard Croatian and back to dialect on the validation set.
- **bleu_di_cr_di_test:** Model's performance in translating from dialect to standard Croatian and back to dialect on the test set.



Fig. 2. Comparison of BLEU CR-DI and DI-CR scores across epochs

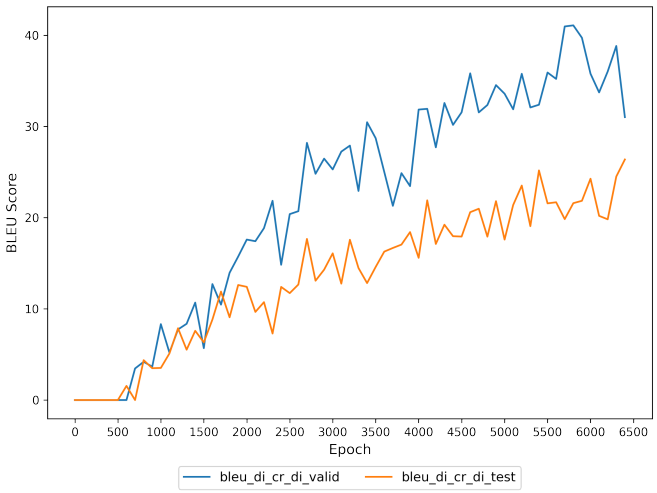


Fig. 3. Comparison of BLEU scores from the back translation across epochs

The flow of how the number of epochs affects our model's ability to learn more efficiently is depicted in Figure 2. For instance, when we tested the model to translate from dialect to standard Croatian language, the BLEU score in the 600 epoch is only 6, but as the number of epochs increases, so does the BLEU score. We currently have the best BLEU score in the 5000 epoch, which is 12.8. The lowest score was received in the validation phase when translating from Croatian to dialect languages. The comparison of BLEU scores from the back translation across epochs is shown in Figure 3. As we can see, we received the highest score of 42 BLEU in the validation phase when translating a back-to-back dialect-Croatian-dialect.

An effort was made to create a language model for translating Croatian dialects to the country's official language, and initial success was achieved. The model did in fact learn some interesting cases and captured the differences between the dialects and the standard language. In the examples below, we present the reference sentences written in the standard language and in dialect, and also the outputs of the model when the different versions of translations are done over the input sentences.

Example 1:

Reference sentence in standard language: E, moje dijete, tako je onda bilo.

Reference sentence in dialect language: E, moje dite, tako je unda bilo.

Dialect-Standard Translation: E, moje, dijete ako je unda bilo.

Standard-Dialect Translation: E moje dite drugačije je unda bilo

Dialect-Standard-Dialect Translation: E moje dite drugačije je bilo unda ništa.

Example 2:

Reference sentence in standard language: Ustao samo rano jer samo morao ići na tržnicu po rajčice.

Reference sentence in dialect language: Usto sam rano jer sam moro ić na tržnicu po rajčice.

Dialect-Standard Translation: Usto sam jer sam rano moro na tržnicu po suzama .

Standard-Dialect Translation: Ustao samo rano jer morao na tržnicu po kući.

Dialect-Standard-Dialect Translation: Usto sam jer sam ja moro na tržnicu po zdravlje rajčice.

From the outputs, it can be concluded that the model is able to translate some different words and make changes. As we can see from Example 1, the model was able to change the word "dite" which is in the Shtokavian dialect to "dijete" in the standard Croatian language. That is a sign that the model indeed learned the difference and has the ability to evolve and upgrade itself. However, the lack of sufficient data in the dialect dataset prevents the model from learning more effectively.

V. Conclusion

This paper represents the first-ever work done in the field of Croatian dialect translation to the standardized version of the language. Since it is the first one of its kind it has its deficiencies due to the lack of available digital data in all of the dialects. The unsupervised approach is proven to work in various types of settings and languages, but it requires a vast amount of monolingual data. Despite the need for no parallel data, it needs a lot of sentences to learn the underlying characteristics, similarities and differences between the languages so it can learn and align the words in a shared latent space. This sort of model can be used in speech-to-text systems in the Croatian language to serve as many customers as possible from different backgrounds and from different locations all over Croatia. At the moment we are trying to acquire as many dialect resources as possible to improve the results and create a system that can translate from Shtokavian, Kajkavian and Chakvian dialects to the standardized version of the Croatian language.

REFERENCES

[1] P. Koehn, "Neural machine translation," *arXiv preprint arXiv:1709.07809*, 2017.

[2] S. Ravi and K. Knight, "Deciphering foreign language," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, Jun. 2011, pp. 12–21. [Online]. Available: https://aclanthology.org/P11-1002

[3] G. Lample, A. Conneau, L. Denoyer, and M. Ranzato, "Unsupervised machine translation using monolingual corpora only," *arXiv preprint arXiv:1711.00043*, 2017.

[4] G. Lample, M. Ott, A. Conneau, L. Denoyer, and M. Ranzato, "Phrase-based & neural unsupervised machine translation," *arXiv preprint arXiv:1804.07755*, 2018.

[5] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst, "Moses: Open source toolkit for statistical machine translation," in *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*. Prague, Czech Republic: Association for Computational Linguistics, Jun. 2007, pp. 177–180. [Online]. Available: https://aclanthology.org/P07-2045

[6] T.-P. Tan, S.-S. Goh, and Y.-M. Khaw, "A malay dialect translation and synthesis system: Proposal and preliminary system," in *2012 International Conference on Asian Language Processing*. IEEE, 2012, pp. 109–112.

[7] S. Kchaou, R. Boujelbane, and L. H. Belguith, "Parallel resources for tunisian arabic dialect translation," in *Proceedings of the Fifth Arabic Natural Language Processing Workshop*, 2020, pp. 200–206.

[8] K. Meftouh, S. Harrat, S. Jamoussi, M. Abbas, and K. Smaili, "Machine translation experiments on padic: A parallel arabic dialect corpus," in *Proceedings of the 29th Pacific Asia conference on language, information and computation*, 2015, pp. 26–34.

[9] F. Huang, "Improved arabic dialect classification with social media data," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015, pp. 2118–2126.

[10] S. Chakraborty, A. Sinha, and S. Nath, "A bengali-sylheti rule-based dialect translation system: Proposal and preliminary system," in *Proceedings of the International Conference on Computing and Communication Systems: I3CS 2016, NEHU, Shillong, India*. Springer, 2018, pp. 451–460.

[11] W. Salloum and N. Habash, "Elissa: A dialectal to standard arabic machine translation system," in *Proceedings of COLING 2012: Demonstration Papers*, 2012, pp. 385–392.

[12] W. Farhan, B. Talafha, A. Abuammar, R. Jaikat, M. Al-Ayyoub, A. B. Tarakji, and A. Toma, "Unsupervised dialectal neural machine translation," *Information Processing & Management*, vol. 57, no. 3, p. 102181, 2020.

[13] M. Dare, V. F. Diaz, A. H. Z. So, Y. Wang, and S. Zhang, "Unsupervised mandarin-cantonese machine translation," *arXiv preprint arXiv:2301.03971*, 2023.

[14] Y. Wan, B. Yang, D. F. Wong, L. S. Chao, H. Du, and B. C. Ao, "Unsupervised neural dialect translation with commonality and diversity modeling," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 9130–9137.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[17] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017.

[18] K. Ogueji and O. Ahia, "Pidginunmt: Unsupervised neural machine translation from west african pidgin to english," *arXiv preprint arXiv:1912.03444*, 2019.

[19] L. H. Baniata, S. Park, S.-B. Park *et al.*, "A neural machine translation model for arabic dialects that utilizes multitask learning (mtl)," *Computational intelligence and neuroscience*, vol. 2018, 2018.

[20] M. Purver, R. Shekhar, M. Pranjić, S. Pollak, and M. Martinc, "24sata news article archive 1.0," 2021, slovenian language resource repository CLARIN.SI. [Online]. Available: http://hdl.handle.net/11356/1410

[21] V. Jagić, *Pjesme Šiška Menčetića Vlahovića i Gjore Držića*, ser. Stari pisci hrvatski. Zagreb: JAZU, 1870, vol. 2.

[22] M. Držić, *Novela od Stanca*. Zagreb: Školska knjiga, 2006.

[23] D. Držić, *Pjesni ljuvene*, ser. Stari pisci hrvatski. Zagreb: JAZU, 1965.

[24] A. Kačić Miošić, *Razgovor ugodni naroda slovinskoga*. pò A. Czesaru, 1801.

[25] M. A. Relković, *Satir ili Divji čovik*. Zagreb, Croatia: Tiskom Josipa Kossicha, 1783.

[26] I. Brlić-Mažuranić, *Priče iz davnine*. Zagreb: Nakladni zavod Hrvatske, 1916.

[27] A. Nemčić, *Kvas bez kruha*. Zagreb: Profil knjiga, 2017.

[28] F. Galović, *Izabrane štokavske pjesme*. Zagreb: Mladost, 1984.

[29] H. Lucić, *Skladanja izvrsnih pisan različitih*. Venice: Tisak Franje pl. Kesterčaneka, 1556.

[30] P. Hektorović, *Ribanje i ribarsko prigovaranje*. Zagreb: Nakladni zavod Matice hrvatske, 1968.

[31] M. Pelegrinović, *Jeđupka*. Zagreb: Matica Hrvatska, 1584.

[32] ——, *Pelegrin Sabu Mišetiću*. Zagreb: Matica Hrvatska, 1578.

[33] B. Karnarutića, *Vazetje Sigeta Grada*. Zagreb: Mletci, 1584.

[34] T. Brezovački, *Matijaš Grabancijaš dijak*. Zagreb, Croatia: Nakladom Stjepana Kugli, 1899.

[35] I. Raos, *Prosjaci i sinovi*. Mladost, 1971.

[36] M. Smoje, *Velo misto*. Zagreb: Školska knjiga, 1981.

[37] A. Conneau, G. Lample, M. Ranzato, L. Denoyer, and H. Jégou, "Word translation without parallel data," *arXiv preprint arXiv:1710.04087*, 2017.

[38] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 2002, pp. 311–318.

Sentiment of the Tweets on Russo-Ukrainian War: the Social Network Analysis

Andrija Poleksić[1], Sanda Martinčić-Ipšić[1,2]
[1]Faculty of Informatics and Digital Technologies, [2]Center for Artificial Intelligence and Cybersecurity,
University of Rijeka,
Radmile Matejčić 2, 51000 Rijeka, Croatia;
Email: andrija.poleksic@student.uniri.hr, smarti@uniri.hr

***Abstract*—This paper presents the analysis of social network posts using standard natural language processing (NLP) methods and undirected graph representations. The Twitter data used in this work is based on two keywords: Ukraine and Russia in May, October, November and December 2022. After standard pre-processing of the raw data and sentiment classification using "Valence Aware Dictionary and sEntiment Reasoner" (VADER), we proceed to the construction of a weighted and undirected network with hashtags as nodes and hashtag co-occurrences as weighted edges. This representation enables topic extraction based on Louvain community detection. For the quantification of the polarity in the network–community, we propose the "stanciness" metric based on "pnlogratio". The results show that combining sentiment with community structures gives us a deeper insight into the polarity of public opinion on the Russian-Ukrainian war.**

***Keywords*—Sentiment analysis, Undirected graphs, Community detection, Natural language processing (NLP)**

I. INTRODUCTION

Social networks are a valuable source of information and can serve as an important communication platform during global crises such as the COVID-19 pandemic [1], [2], wars [3]–[6], climate change [7], economic crises [6], [8], migration crises, etc. On one side social platforms can serve as the key communication platform that enables sharing of valuable information [1], but on the other, they can propel multiple negative effects: (i) foster negative emotions [5], [6], [9], [10] (ii) generation of insulting or hate speech [5], [11], (iii) spreading of fake news and misinformation [7], etc. Still, understanding how people post comments in online communication may shed light on the fundamental mechanisms by which collective thinking emerges in a social group [12].

Social networks analysis (SNA) and natural language processing (NLP) provide relevant methods for studying user-posted content and user behaviour in social networks, hence we engage both. Sentiment analysis (or opinion mining) is the field of NLP study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [13]. Sentiment analysis has been applied to identify attitudes, opinions and public perceptions on many topics, usually identified with topic modelling methods [10], [14]. Social network analysis provides a set of methods for analysing relationships and interactions between actors, explaining its underlying social structure, such as the communities of users, network hubs and central nodes, spreading patterns, etc. [15]–[17].

It is well-known that Twitter data can be used to study many phenomena like the growth mechanisms of social interactions, assessing user influence, monitoring trends, or sentiment analysis [18]–[21]. The goal of this work is to obtain insights into Twitter communication on the topic of the major global crisis caused by the ongoing war. We aim to provide insight into public opinion on the Russian-Ukrainian war by combining the polarity of tweets sentiment with analysis of hashtags graph structures.

Hence, we collect the dataset based on two keywords: Ukraine and Russia during May, October, November and December 2022. For the detection of the polarity of tweets (positive, negative and neutral) we use "Valence Aware Dictionary and sEntiment Reasoner" (VADER). Next, we construct an undirected weighted network with hashtags as nodes and hashtag co-occurrences as weights on the edges. This representation enables community detection for topic extrapolation. Finally, we propose a metric that quantifies the "stanciness" of the tweet based on "pnlogratio", which calculates the overall sentiment of the node based on the compound score of incident edges. The results show that "pnlogratio" can be used to obtain insights into the polarity of public opinion on the topic captured with hashtags.

Section II contains a short overview of the research of Twitter posts and hashtags analysis, specifically for the ongoing Ukrainian war domain. In Section III, we provide a brief overview of the methods for sentiment classification and graph analysis and define a novel "stanciness" measure. Section IV elaborates upon the data preparation, sentiment detection and network construction. The results are presented and discussed in Section V. In the last Section, we give some concluding remarks and plans for future work.

II. RELATED WORK

Twitter is a popular social network that enables publishing of short (i.e. up to 280 characters) publicly visible messages called tweets. Tweets typically consist of content (i.e. text), links (i.e. URLs), user mentions (with sign), retweet information (RT) and hashtags. Hashtags

are marked with the # sign and are used for meta-tagging, which enables users to retrieve specific content. Hashtags do not follow a predefined structure, yet they can capture the essence of messages, much like keywords or keyphrases do [22].

The thorough report on the tweets hashtag networks construction principles and network analysis, along with the study of the growth mechanism via link prediction is reported in our previous work [22]. Therefore, here we focus on the overview of the related work centred around the sentiment analysis of tweets for the selected domain of the Russo-Ukrainian war.

Chen and Ferrara in [3] collect a large dataset of 63M tweets in many languages and report English for 72.58%, Spanish for 5.6%, and French for 3,77% tweets as the top frequent languages. Based on the simple occurrence frequency they report the top 15 hashtags: ukraine, russia, putin, standwithukraine, kyiv, ukrainerussiawar, stopputin, ukrainerussianwar, russian, ukraineunderattack, nato, stoprussia, kiev, ucrania, ukrainian. Finally, they also study the region of the tweet origin.

Garcia and Yabut in [4] perform lexicon-based tweet polarity detection with a report on the most frequently used hashtags: ukraine war, worldwar3, ukrainian, putin, ukraineRussaconflict, Russia, ukraineinvasion, ukraine, putin. The negative polarity and emotion of sadness prevail over fear and anger.

Sazzed in [5] goes a step forward and employs a methodologically similar setup to our work. The sentiment of 80,000 tweets is determined according to the VADER scores followed by LDA (Latent Dirichlet Allocation) and BERT (Bidirectional Encoder Representations from Transformers) topic modelling. The results suggest that although both positive and negative sentiments are present the negative prevails and topic modelling reveals a myriad spectrum of topics.

Polyzos in [6] conducts the meticulous sentiment analysis on 42M million tweets and calculates impulse response for 15 economic and financial indicators (stock markets, commodities, interest rates, currencies and cryptocurrencies). He shows that sentiment analysis determined with pre-trained deep learning model on an extended sample of 42M tweets can proxy the public's perception of the crisis situation, and influence the financial sector. The paper uses Vector AutoRegressive (VAR) model to quantify the response of stock markets, oil, gold and bitcoin prices, to Russo-Ukrainian War.

Finally, kindly note that we limit our study to isolated time periods and the English language, hence the generalization can be achieved only with extension to a broader set of languages (i.e. geographical regions) and the whole duration of the war. This remains a future challenge.

III. METHODOLOGY

This section gives a brief overview of the methods used for sentiment analysis and network construction and characterization.

A. Sentiment Analysis

The central part of the methodology focuses on the 3-dimensional sentiment analysis, which consists of a positive, neutral and negative dimension. To determine the sentiment, we use the Valence Aware Dictionary and sEntiment Reasoner (VADER) [23] due to its proven good performance when applied to Twitter data. VADER uses valence scores from -4 to 4 (extremely negative to extremely positive) to enrich a lexicon of words with their intensity. Compared to related work, VADER relies on 5 heuristics and an optimized human scoring method using the wisdom-of-the-crowd (WotC) approach [24]. The use of the 5 heuristics: punctuation, capitalization, degree modifiers, constructive conjunction "but", and a tri-gram strategy to resolve the effects of negation on polarity, as elaborated in [23]. VADER requires (works best with) raw, unprocessed microblog-like (i.e. short posts on social networks) data.

For the given input, (e.g. a tweet text), VADER computes 4 scores based on the valence scores of each word in the lexicon. The first three refer to the positive, neutral, and negative scores ("pos", "neu", "neg") as a ratio for the proportion of the text in each category. The compound score is calculated as the sum of the valence scores of each word in the lexicon (where unknown words have a valence score of 0). The sum is then normalized to fit the $[-1, 1]$ interval. The compound score is calculated with

$$\frac{s}{\sqrt{s^2 + \alpha}}$$

where "s" is the sum of the valence scores of all words and alpha, as used by the author, is equal to 15. Using a recommended compound score threshold of 0.05, the data is sorted into three categories: "positive" for scores greater than or equal to 0.05, "neutral" between -0.05 and 0.05, and "negative" for less than or equal to -0.05.

B. Network analysis and construction

Next, we overview used network construction and characterization principles. Note that, we interchange the terminology between graph and network, both are accurate but stem from different scientific disciplines i.e. sociology and mathematics. The graph consists of nodes (i.e. vertices) representing the actor. We construct the weighted and undirected network: hashtags are nodes; undirected edges (i.e. links) represent the co-occurrence of hashtags in the tweet. The weights on the edges are derived from the co-occurrence frequency of the hashtags. With this setup, it is possible to calculate network characterization measures: centrality measures (e.g. degree, betweenness, eigenvector and closeness centralities) - metrics that quantify the importance of nodes - and community detection methods (e.g., Louvain clustering) - relevant for the revealing of the structure of communities or clusters. The details about complex network analysis can be found in [25], still, in

continuation, we provide only the essential descriptions to ensure the integrity of the paper.

Degree centrality is computed as the total number of edges incident with a node normalized by the maximum degree in a network [25]. Betweenness centrality measures the extent to which a node lies along the shortest paths between other nodes. A node with high betweenness centrality acts as a bridge between other nodes, and information flow is more likely to pass through it [25]. Eigenvector centrality is similar to degree centrality and attributes the value of a node higher when connected to nodes with a high value. Hence, a node with a high eigenvector centrality value is incident to high-valued nodes [25]. Closeness centrality measures the average distance from a node to all other nodes, or a node with high closeness is close to all other nodes in the network [25]. The idea behind this measure is that the closer a node is to all other nodes, the more central it is, and the faster it can access and disseminate information to others.

Communities (i.e clusters of nodes) are groups of nodes that share properties and/or play similar roles within the network [26]. Detecting communities in complex networks is of particular interest when identifying nodes that share properties and dynamics [27]. In this research, we apply the Louvain algorithm [28] for community detection. As stated in [29], for analyzing network data where links represent not flows but rather pairwise relationships, it may be useful to detect the structural components - communities [26], [27]. The Louvain algorithm [28] is an unsupervised community detection method that consists of two basic steps: modularity optimization and community aggregation. In the first step, the modularity measure is iteratively maximized by comparing the modularity gain for combinations of different node partitionings. In the second step, each community discovered in the first step is abstracted into a single node. The algorithm is repeated until no further improvements are possible.

We propose a novel metric "stanciness" of tweets based on the "pnlogratio". "Stanciness" explains the amount or level of attitude present in a piece of text, which can be either positively or negatively polarized. Therefore, we use "stanciness" to approximate the total quantity of sentiment (both positive and negative) in texts. "Pnlogratio" quantifies the overall sentiment of nodes based on the compound score of each edge with which that node incidents. The formula is based on the ratio of the precomputed attributes *posc* and *negc* as follows:

$$pnlogratio = log_{10}(\frac{\sum posc}{\sum abs(negc)}),$$

where *posc* is a sum of all positive (>0) and *negc* is a sum of all negative (<0) compound scores of weights on incident links. The "pnlogratio" expresses the ratio between positive and negative nodes advantageously and centres the results around the value of zero to enable the best visualisation interpretability of the calculated values.

IV. DATA AND EXPERIMENTAL SETUP

A. Data collection

Aiming to analyze the content on tweeter posts on the crises caused by the Russo-Ukrainian war, collecting relevant data is a prerequisite. Therefore, we use the open-source alternative to Twitter's official API, the Twitter intelligence tool Twint [30]. Twint simulates Twitter search as a guest user obtaining tweets fitting the preset time span. For comparison purposes, we decide to gather 20000 tweets for each of the two selected keywords, Russia and Ukraine, on a daily basis. The selection of these two keywords is grounded on the frequency obtained on a large Twitter dataset [3]. Initial data collection conducted in May 2022 gathered a total of 134,297 and 120,038 tweets for Ukraine and Russia keywords respectively. During October, November and December, the data collection totalled 1,952,474 tweets (942,616 collected on the keyword Ukraine, 1,009,858 on Russia). The final dataset consists of two sets of 60 comma-separated values (csv) files (corresponding to a total of 60 days of data collection) containing 36 columns of various tweet metadata.

B. Data preprocessing

The data preparation resulted in two csv files (each per one keyword, and 60 days) containing eight filtered columns (out of 36): id, conversation_id, user_id, created_at, tweet, urls, hashtags, and language. Additionally, we filter the data by language to remove all non-English tweets. Note, due to the selected sentiment analysis method (see Section III), no further NLP preprocessing steps like stopwords removal or stemming are required.

Next, we proceed with assessing the polarity of tweets with VADER as defined in Section III. Tab. I shows two examples of Twitter data evaluated by VADER, where the first column is the text of the tweet and the second column is the evaluation score in the form of a 5-tuple. The tuple consists of the "pos", "neu", and "neg" mentioned above, a compound score, and the class to which it belongs (e.g. "negative").

Fig. 1 shows the overall distribution in two datasets where negative, neutral, and positive are red, blue, and green, respectively. As expected, higher overall negativity in the dataset is associated with the keyword Russia. And both parts of the dataset are leaned negatively (51% and 47% of negative tweets for Russia and Ukraine respectively).

TABLE I: Polarity of tweets with VADER scores

Tweet text	VADER Score
@JJcycles Trumpers in tears...now back to Russia and Ukraine war. Still think Putin will do something horrible soon.	(0.316, 0.684, 0.0, -0.8126, "Negative")
@Flingan67 Ukraine you are winning. There is no need to do that to POWs. Yes the ruzzians are terrible to Ukrainian POWs, but they are not Ukrainians. Be the better people.	(0.097, 0.692, 0.21, 0.6428, 'Positive')

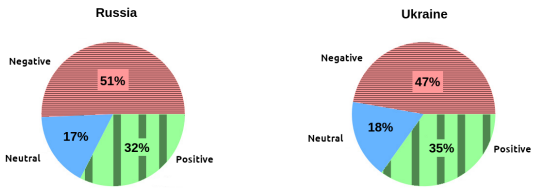


Fig. 1: Ratio of positive, neutral and negative classes in the dataset for Russia (left) and Ukraine (right)

C. Graph construction

Next, we proceed with the construction of hashtags network by merging two datasets (Ukraine and Russia) into one. By their nature, two sets are expected to overlap (e.g. contain both keywords), so we deduplicate the data, which resulted in removing a total of 197,499 rows and retaining 1,754,975 rows for the network construction. The data is then filtered by the number of hashtags to ensure that there are at least two in a tweet. Additionally, we normalize the case-sensitive variations of hashtags and remove emojis and unrecognized characters.

Graph (Network) construction is implemented with a Python package for creating, manipulating and studying the structure, dynamics and functions of complex networks, the NetworkX [31]. The first version of the network of hashtags consists of 50,522 vertices and 486,612 edges with "weight" and 4 more attributes:

- "posc": Sum of all positive (>0) compound scores of tweets in which the given "combination of hashtags" (from now on, edge) occurs,
- "poscount": Count of all positive (>0) compound scores of tweets in which the edge occurs,
- "negc": Sum of all negative (<0) compound scores of tweets in which the given edge occurs,
- "negcount": Count of all positive (<0) compound scores of tweets in which the edge occurs.

These attributes enable different calculations including averages, positive-negative ratios and weighted metrics. In the second version of the network, we filter out unconnected nodes, retaining only the largest connected component (LCC - i.e. the largest component of connected nodes in the network) with 47,395 nodes and 483,023 edges.

Fig. 2 shows a preview of the graph created from approximately 0.009% of the data, enabling the visualization. The colors of the nodes represent the overall sentiment of the incident edges: red, grey, and green indicating negative, neutral, and positive sentiment, respectively. Color is determined by comparing the values of "a" and "b": red (-1) if $a < b$, grey (0) if $a = b$, and green (1) if $a > b$. where

$$a = \sum posc, b = \sum negc$$

and *posc* and *negc* correspond to all edges incident to the particular node. The edge weights are visualised by the thickness of the edges.

D. Network analysis

Enriching nodes with numerical attributes such as centrality measures and sentiment averages allows us to gain a deeper understanding of public opinion on specific topics. Using the constructed graph, we can identify important nodes, their "stanciness", and the context in which they are mentioned. Therefore, we apply the proposed metric for measuring "stanciness" based on "pnlogratio" and add it to the node attributes.

In the second step, we compute four centrality measures: degree, eigenvector, betweenness, and closeness, to detect the most important nodes - prominent hashtags. The centrality measures can provide different insights as we have elaborated in [17], [32]. Therefore, here we examine several centrality values to detect the most prominent nodes, assuming that the majority of the important nodes will be closely related to the issues of the Russo-Ukrainian war.

The final step of analysis involves the detection of communities using Louvain clustering with a granularity of 0.75 while also removing communities containing less than 20 nodes. In the context of this analysis, we interpret the detected communities as a rough estimate of the topics related to the Twitter dataset we collected, which explains the underlying themes discussed in the context of the dataset. Still, whether detected communities really represent topics remain to be confirmed with topic modelling in future research.

V. RESULTS

In this section, we report the results of the combined network and sentiment analysis. First, we report the results of the centrality measures. Tab. II lists the top 15 nodes (hashtags), excluding the obvious russia and ukraine hashtags, based on each of the degree, eigenvector, betweenness and closeness centralities.

Different centrality measures detect different hashtags and derive different rankings, although as expected there

TABLE II: Top 15 hashtags according to the centrality value

Degree	Eigenvector	Betweenness	Closeness
putin	ukrainewar	urw	putin
usa	kherson	ukrainewar	usa
war	urw	ukrainewillwin	war
russian	putin	standwithukraine	russian
nato	nato	kherson	nato
ukrainewar	usa	crimea	ukrainewar
urw	war	etsyseller	urw
china	russian	ukrainianarmy	china
kherson	news	kharkiv	news
ukrainian	urwr	stopputin	urw
news	kyiv	dekorstyle	ukrainian
ruw	unitedstates	lgbtq	biden
urw	ruw	armukrainenow	ruw
biden	germany	nft	europe
iran	china	stoprussia	kherson

abbreviations: urw - ukrainerussi(an)awar; ruw - russiaukrainewar.

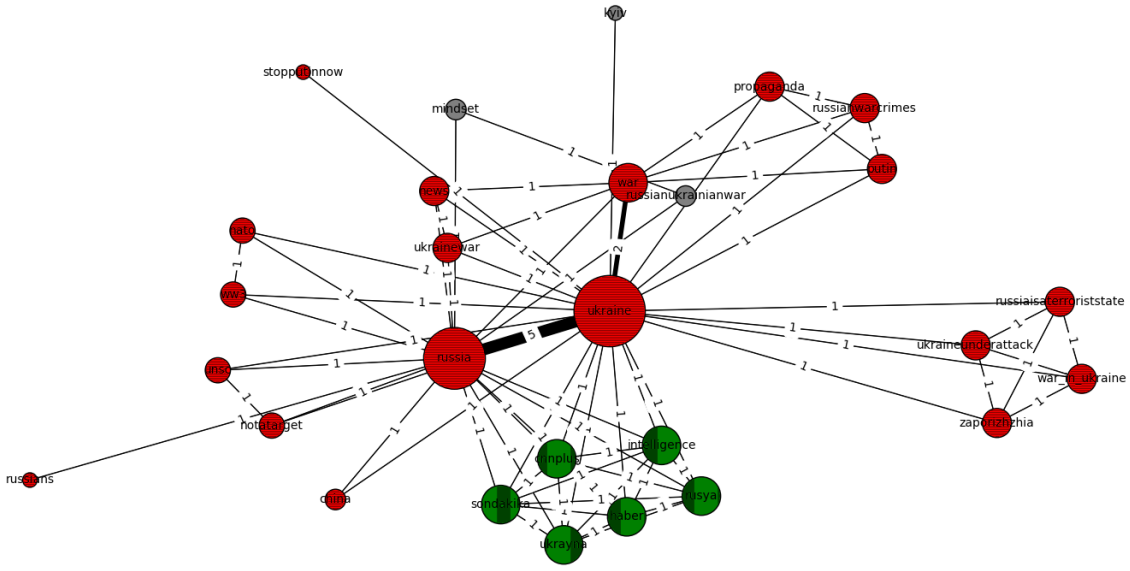


Fig. 2: Weighted graph example: negative hashtags-red-(horizontal lines), positive-green-(vertical stripes), neutral-grey-(no pattern)

is an overlap between ranked keywords (e.g. ukrainewar, putin, etc.). As might be expected, most of the detected keywords refer to or are directly related to recent events in the Russo-Ukrainian war. Betweenness centrality ranks the most out-of-the-domain hashtags in the top 15 (e.g. etsyseller, dekorstyle, lgbtq). Therefore degree, eigenvector and closeness are preferred centrality measures in this context. Moreover, it seems that centrality measures, regardless of which applied (with the exception of the betweenness) contribute to better extraction of representative hashtags in the domain (i.e keywords that capture meaning) compared to simple frequency calculation as also reported in [3], [4].

Second, we move on to the results of the sentiment analysis using the proposed "pnlogratio" metric. Fig. 3 shows the kernel density estimation (KDE) for estimating the probability density function of "pnlogratio" using Matplotlib [33] and Seaborn [34] Python libraries. The graph provides insight into the distribution of negative, neutral, and positive sentiments of the hashtags at the level of the largest connected component. It is possible to notice a slight tilt towards negative sentiment with the central peak around 0. The two opposing peaks (the first around -2 and the second around 2) indicate a large concentration of extremely negative and extremely positive hashtag contexts in the domain. Next, we compare the KDE density at the global (LCC) level with the KDE density at the level of identified communities.

The weighted Louvain clustering algorithm detected 39 communities of varying sizes. Fig. 4 and 5 illustrate the distribution of "pnlogratio" within a community. In the left top corner of the figures, we list the top-ranked nodes (i.e. hashtags) within the community, ordered by the value of degree centrality. The same can be reported for other

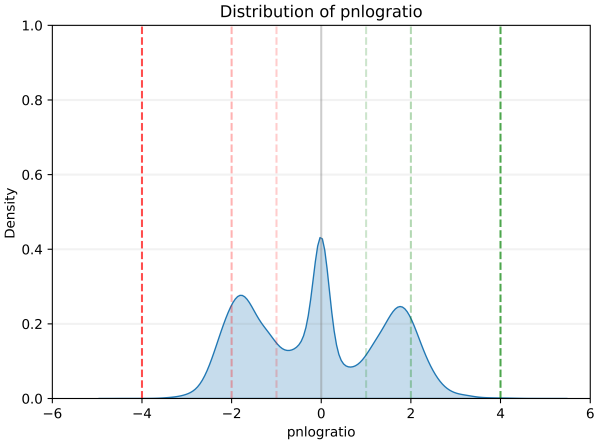


Fig. 3: The distribution of "pnlogratio" attribute at the level of the largest connected component

centrality measures, as well.

Examining the KDE curve in Fig. 4 compared to Fig. 3, one can notice a significant increase in the density of negative hashtags (nodes) and an overall decrease in positive and neutral "pnlogratio", indicating that the topic in Fig. 4 is connected to a more negative opinion and topic can be semantically defined as general war-related posts. The same holds for the community (or cluster) 5, which is related to energy as seen in Fig. 5. A similar shift towards positivity or negativity can be detected for all clusters, so in the Appendix, we provide additional examples of one positive (Fig. 6) and one negative (Fig. 7) cluster.

Please note, that it is possible to refine the proposed "pnlogratio" metric by including the weight attribute in the calculation and recalculating the metric at the local level,

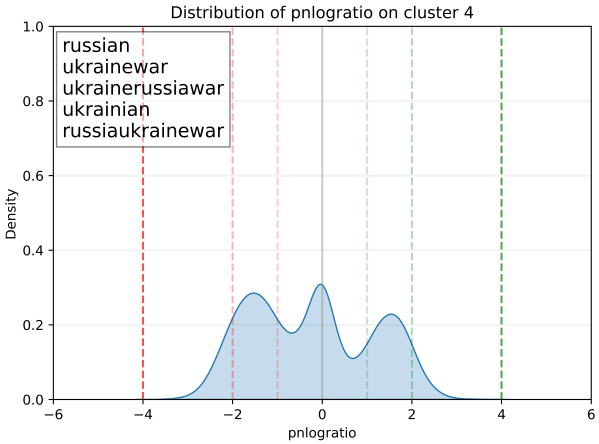


Fig. 4: The distribution of "pnlogratio" attribute in cluster 4: War related

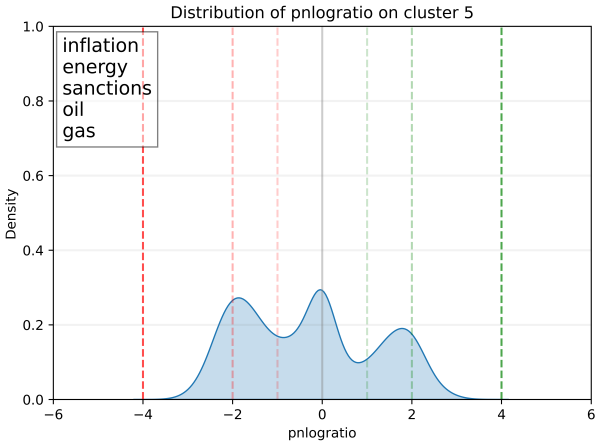


Fig. 5: The distribution of "pnlogratio" attribute in cluster 5: Energy related

e.g., within a community or cluster. In this way, the metric can assess the more or less context and possibly contribute to more or less fine-grained detection of the polarity. This mechanism still needs to be examined, but we assume that the proposed solution can be of help to detect the public-opinion polarity at different scales of granularity of topics, and subtopics extrapolated with communities of hashtags.

VI. Conclusion

In this paper, we combine methods of natural language processing, sentiment analysis, and network analysis to identify and explore the polarity of themes related to the recent event of the Russo-Ukrainian war. Initially, we collect and analyse Twitter data collected by two keywords, Ukraine and Russia. After preprocessing the data and determining sentiment polarity using the VADER dictionary, and constructing the undirected and weighted network of co-occurring hashtags, based on the value of centrality measures - degree, eigenvector, betweenness, closeness - we rank the top hashtags. Then, the Louvain algorithm detects communities that roughly extrapolate topics. Finally, we calculate the "stanciness" of topics based on "pnlogratio". "Stanciness" approximates the total quantity of both positive and negative sentiments in hashtags communities.

The obtained results confirm that cross-fertilization of NLP and social network analysis is of benefit to the gained insights. To conclude as expected and already reported in [4]–[6], we detect both positive and negative clusters of hashtags, but the negative polarity prevails. We show that proposed "stanciness" can quantify the overall sentiment of the hashtags community, being the alternative methodological solution to standardly applied topic modelling. Still, this remains to be further investigated and compared, especially in terms of the adequacy of obtained results and the potential limitations and biases of the measure. Still, the proposed measure of "stanciness", reveals many potentials and seems promising for further study.

Future work will include fine-graining of the Louvain clustering and/or using similar topic extraction methods for more accurate and thematically narrowed results at different scales of granularity. Next, it is also possible to extend the spectre of emotions extracted from the text and reach beyond the limited positive-neutral-negative perspective. Finally, detected communities need to be confirmed with topics determined with standard NLP - LDA techniques.

Acknowledgment

This work has been fully supported by the University of Rijeka under project number uniri-drustv-18-20.

References

[1] C. Cuello-Garcia, G. Pérez-Gaxiola, and L. van Amelsvoort, "Social media can have an impact on how we manage and investigate the covid-19 pandemic," *Journal of clinical epidemiology*, vol. 127, pp. 198–201, 2020.

[2] D. C. Glik, "Risk communication for public health emergencies," *Annual Review of Public Health*, vol. 28, no. 1, pp. 33–54, 2007, pMID: 17222081. [Online]. Available: https://doi.org/10.1146/annurev.publhealth.28.021406.144123

[3] E. Chen and E. Ferrara, "Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between ukraine and russia," 2022. [Online]. Available: https://arxiv.org/abs/2203.07488

[4] M. B. Garcia and A. Cunanan-Yabut, "Public sentiment and emotion analyses of twitter data on the 2022 russian invasion of ukraine," in *2022 9th ICITACEE*, 2022, pp. 242–247.

[5] S. Sazzed, "The dynamics of ukraine-russian conflict through the lens of demographically diverse twitter data," in *2022 IEEE International Conference on Big Data*, 2022, pp. 6018–6024.

[6] E. Polyzos, "Escalating tension and the war in ukraine: Evidence using impulse response functions on economic indicators and twitter sentiment," *SSRN Electronic Journal*, 3 2022. [Online]. Available: https://ssrn.com/abstract=4058364

[7] A. Samantray and P. Pin, "Credibility of climate change denial in social media," *palgrave communications*, vol. 5, pp. 1–8, 2019.

[8] S. C. Long, B. Lucey, Y. Xie, and L. Yarovaya, ""I just like the stock": The role of Reddit sentiment in the GameStop share rally," *Financial Review*, vol. 58, no. 1, pp. 19–37, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/fire.12328

[9] J. Xue, J. Chen, R. Hu, C. Chen, C. Zheng, Y. Su, and T. Zhu, "Twitter discussions and emotions about the covid-19 pandemic: Machine learning approach," *Journal of medical Internet research*, vol. 22, no. 11, p. e20550, 2020.

[10] K. Babić, M. Petrović, S. Beliga, S. Martinčić-Ipšić, M. Matešić, and A. Meštrović, "Characterisation of COVID-19-Related Tweets in the Croatian Language: Framework Based on the Cro-CoV-cseBERT Model," *Applied Sciences*, vol. 11, no. 21, p. 10442, Nov 2021. [Online]. Available: http://dx.doi.org/10.3390/app112110442
[11] J. Uyheng and K. M. Carley, "Characterizing network dynamics of online hate communities around the covid-19 pandemic," *Applied Network Science*, vol. 6, no. 1, pp. 1–21, 2021.
[12] A. N. Medvedev, R. Lambiotte, and J. Delvenne, "The anatomy of Reddit: An overview of academic research," *CoRR*, vol. abs/1810.10881, 2018. [Online]. Available: http://arxiv.org/abs/1810.10881
[13] B. Liu, "Sentiment Analysis and Subjectivity," in *Handbook of Natural Language Processing*, 2010.
[14] P. K. Bogović, A. Meštrović, and S. Martinčić-Ipšić, "Topic Modeling for Tracking COVID-19 Communication on Twitter," in *Information and Software Technologies*, A. Lopata, D. Gudonienė, and R. Butkienė, Eds. Springer International Publishing, 2022, pp. 248–258.
[15] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.
[16] D. S. Martinčić-Ipšić, D. Margan, and A. Meštrović, "Multilayer network of language: a unified framework for structural analysis of linguistic subsystems," *Physica A: Statistical Mechanics and its Applications*, vol. 457, pp. 117–128, 2016.
[17] N. Matas, S. Martinčić-Ipšić, and A. Meštrović, "Comparing network centrality measures as tools for identifying key concepts in complex networks: A case of wikipedia," *Journal of Digital Information Management (JDIM)*, vol. 15, pp. 203–213, 2017.
[18] M. Mathioudakis and N. Koudas, "Twittermonitor: Trend detection over the twitter stream," in *Proc. 2010 ACM SIGMOD Int. Conf. on Management of Data*. ACM, 2010, pp. 1155–1158.
[19] X. Wang, F. Wei, X. Liu, M. Zhou, and M. Zhang, "Topic sentiment analysis in twitter: A graph-based hashtag sentiment classification approach," in *20th ACM Int. Conf. on Infor. and Know. Management*, B. Berendt, A. de Vries, W. Fan, C. Macdonald, and I. Ounis, Eds. ACM, 2011, pp. 1031–1040.
[20] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *4th Int. AAAI Conf. on Weblogs and Social Media*. AAAI Press, 2010, pp. 178–185.
[21] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña-López, and A. Montejo-Ráez, "Sentiment analysis in twitter," *Natural Language Engineering*, vol. 20, pp. 1–28, 2012.
[22] S. Martinčić-Ipšić, E. Močibob, and M. Perc, "Link prediction on twitter," *PloS ONE*, vol. 12, no. 7, p. e0181079, 2017, q1, IF 2.806.
[23] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14550
[24] J. Surowiecki, *The Wisdom of Crowds*. Anchor Books, August 2005. [Online]. Available: http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20\&path=ASIN/0385721706
[25] M. Newman, *Networks: An Introduction*. Oxford University Press, 03 2010. [Online]. Available: https://doi.org/10.1093/acprof:oso/9780199206650.001.0001
[26] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, feb 2010. [Online]. Available: https://doi.org/10.1016\%2Fj.physrep.2009.11.002
[27] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, "Characterization of complex networks: A survey of measurements," *Advances in Physics*, vol. 56, no. 1, pp. 167–242, 2007. [Online]. Available: https://doi.org/10.1080/00018730601170527
[28] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, oct 2008. [Online]. Available: https://doi.org/10.1088\%2F1742-5468\%2F2008\%2F10\%2Fp10008
[29] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0706851105
[30] F. Poldi and C. Zacharias, "twint," 2017, gitHub repository, https://github.com/twintproject/twint.
[31] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using networkx," in *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, and J. Millman, Eds., 2008, pp. 11 – 15.
[32] S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, "Selectivity-based keyword extraction method," *International Journal on Semantic Web and Information Systems*, vol. 12, no. 3, pp. 1–26, 2016.
[33] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
[34] M. L. Waskom, "seaborn: statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021. [Online]. Available: https://doi.org/10.21105/joss.03021

VII. Appendix

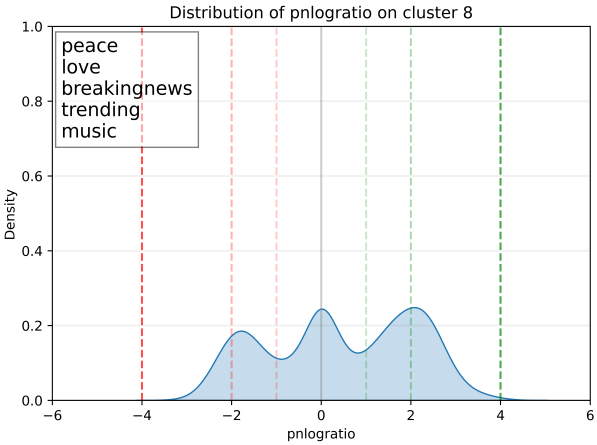


Fig. 6: The distribution of "pnlogratio" attribute in cluster 8: peace related

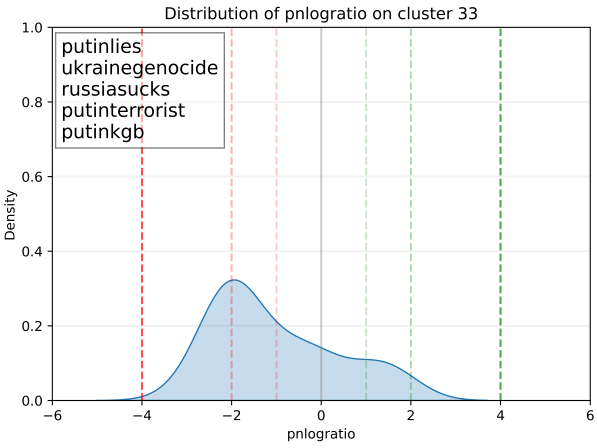


Fig. 7: The distribution of "pnlogratio" attribute in cluster 33: genocide related

Compressing Sentence Representation with Maximum Coding Rate Reduction

Domagoj Ševerdija*, Tomislav Prusina*, Antonio Jovanović*, Luka Borozan*,
Jurica Maltar*, Domagoj Matijević*
* Department of Mathematics, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia
dseverdi@mathos.hr

***Abstract*—In most natural language inference problems, sentence representation is needed for semantic retrieval tasks. In recent years, pre-trained large language models have been quite effective for computing such representations. These models produce high-dimensional sentence embeddings. An evident performance gap between large and small models exists in practice. Hence, due to space and time hardware limitations, there is a need to attain comparable results when using the smaller model, which is usually a distilled version of the large language model. In this paper, we assess the model distillation of the sentence representation model Sentence-BERT by augmenting the pre-trained distilled model with a projection layer additionally learned on the Maximum Coding Rate Reduction (MCR^2) objective, a novel approach developed for general purpose manifold clustering.**

We demonstrate that the new language model with reduced complexity and sentence embedding size can achieve comparable results on semantic retrieval benchmarks.

***Keywords*—Sentence embeddings, model distillation, Maximum Coding Rate Reduction, semantic retrieval**

I. Introduction

Dense vector representations of words, or word embeddings, form the backbone of most NLP applications and can be constructed using context-free (see [1], [2], [3]) or contextualized methods (see [4], [5] for more details). In practice, few NLP applications often benefit from having sentence or document representations in addition to word embeddings. In most cases, one can use the weighted average (aka pooling) over some or all of the word embeddings from a sentence or document. Although it disregards word order while pooling, this approach has shown to be reasonably performant [6]. Pre-trained language models like BERT have shown success on many NLP tasks through fine-tuning. Unfortunately, using contextualized word vectors from these models as a sentence representation is significantly inferior in terms of semantic textual similarity compared to approaches when one uses non-contextualized word vectors, which are trained with a much simpler model (see [7] for more details). Therefore, more sophisticated methods were derived to find efficient and performant universal sentence encoders. Reimers et al. in [7] developed the Sentence-BERT model by fine-tuning pre-trained BERT architecture on sentence pair scoring tasks using a Siamese architecture to learn better sentence representations, showing much improvement in downstream NLP tasks. Their approach ended up with a relatively large model size (hundreds of millions to billions of parameters) and sentence embedding dimension 768, a relatively large number for efficient search and retrieval operations over databases. In this paper, we focus on reducing the dimensionality of sentence embeddings up to 50%-70% while still achieving comparable results across the board of NLP benchmarks. This opens up a possibility of deploying AI models on smaller-scale computer systems like embedded systems.

A. Related Work

Following the distributional hypothesis, Mikolov et al. in [2] showed that computing dense vectors of lower dimension for word embeddings give interesting mathematical properties of words. Inspired by the same idea, Kiros et al. [8] and Lee et al. [9] tried to derive a model which predicts surrounding sentences. Sent2Vec [10] generates context-free sentence embeddings as averages of word vectors and n-gram vectors (similar to FastText [11] for words). Conneau et al. [12] computed contextualized sentence embeddings using a BiLSTM Siamese network that was fine-tuned on pairs of semantically similar sentences. This approach was extended to fine-tuning pre-trained language models like BERT in [7]. Recently, Gao et al. [13] improved this approach by suggesting a contrastive learning method and achieved state-of-the-art results. Projecting sentence embeddings to lower dimensions was motivated by projecting word vectors. In most cases, PCA methods gave surprisingly good results and even retrofitted the word vectors in such a way that it made vectors more isotropic which had a good impact on NLP benchmarks. Li et al. [14] showed that this anomaly is also apparent in sentence vectors and gave a normalizing flow method to retrofit such vectors. Recent work of [15] introduced Maximum Coding Rate Reduction (MCR^2), a novel learning objective that enables for learning a subspace representation given the clustering[1]. They also demonstrated how to extend the approach to the problem of unsupervised clustering.

B. Our contribution

We use a pre-trained sentence embedding model like Sentence-BERT (SBERT) as a sentence encoder and train a non-linear mapper atop the encoder using a Maximal Coding Rate Reduction as a training objective for

[1]The formal definition of MCR^2 will be given in Section II.

learning discriminative low-dimensional structures that preserve all the essential information encoded into the high-dimensional data. This approach allows for more robust training than standard training objectives like cross-entropy and produces clusters in the embedding space. The main contribution of our paper is a sentence embedding compression technique that achieves comparable results with smaller sentence embedding sizes on semantic NLP benchmarks compared to the baseline sentence encoder.

The paper is organized as follows. In Section II we describe Maximum Rate Coding Reduction training objective for computing subspace embedding space. Furthermore, SBERT architecture is described as a sentence encoder followed by a definition of the projection layer. In Section III we experimentally evaluate our method and conclude with a results discussion.

II. Method

For a given set of sentences S and for each sentence

$$(word_1, word_2, \ldots, word_{n_i}) \in S$$

our task is to construct a lower dimensional embedding $z_i \in \mathbb{R}^d$ that contains important semantic information characteristic for that sentence. Our idea is to extend SBERT and from it's embedding compute a small projector to reduce the dimension, i.e. given the set of SBERT's embeddings $Z \in \mathbb{R}^{d\times n}$ of the dataset S, find a $\hat{Z} \in \mathbb{R}^{\hat{d}\times n}$ that retains semantic information extracted by SBERT.

A. Learning a subspace representation with MCR^2

Using the idea from Li et al. [16] we aim to minimize the angle between similar sentences and maximize the entropy of the whole dataset. For two representations $\hat{z}_1, \hat{z}_2 \in \mathbb{R}^d$ of two sentences we measure how similar they are by cosine similarity

$$D\left(\hat{z}_1, \hat{z}_2\right) = \frac{\cos\left(\hat{z}_1^\top \hat{z}_2\right)}{\|\hat{z}_1\|_2 \|\hat{z}_2\|_2}.$$

For two sets $\hat{Z}_1, \hat{Z}_2 \in \mathbb{R}^{d\times b}$ we define this function as

$$D(\hat{Z}_1, \hat{Z}_2) = \frac{1}{b}\sum_{i=1}^{b} D\left(\hat{z}_{1,i}, \hat{z}_{2,i}\right) \tag{1}$$

where $\hat{z}_{1,i}$ is the i-th element of $\hat{Z}_1$ and $\hat{z}_{2,i}$ is the i-th element of $\hat{Z}_2$. Given pairs of similar sentences we want them to have the D score as large as possible.

For a set of representations $\hat{Z} \in \mathbb{R}^{d\times n}$ with n elements, its entropy is defined as

$$R_\varepsilon(\hat{Z}) = \frac{1}{2}\log\det\left(I + \frac{d}{n\varepsilon^2}\hat{Z}\hat{Z}^\top\right) \tag{2}$$

for a given parameter ε and identity matrix I. This function is approximately the Shannon coding rate function for multivariate Gaussian distribution given average distortion ε [17]. Maximizing (2) we maximize the volume of the ball in which the embeddings are packed. The theory behind this is well over the scope of this paper. It is given in the paper by Ma et al. [18] where they explore rate distortion, ε-ball packing and lossy encoding with normally distributed data. By optimizing it in parallel with (1) we try to distance each sentence from others, except for the similar pairs that we try to keep close. Additionally, given cluster assignments, we can measure the entropy of each cluster with

$$R_\varepsilon\left(\hat{Z}, \Pi_k\right) = \frac{n_k}{2n}\log\det\left(I + \frac{d}{n_k\varepsilon^2}\hat{Z}\Pi_k\hat{Z}^\top\right) \tag{3}$$

where Π_k is a diagonal matrix with i-th entry being 1 if the i-th sentence belongs to cluster k, otherwise 0, and $n_k = \mathrm{tr}\left(\Pi_k\right)$, trace of matrix Π_k, i.e. number of points in this cluster. Combining functions (1), (2) and (3) into one we get the MRC^2 loss function defined with

$$L(\hat{Z}, \Pi) = -R_\varepsilon(\hat{Z}) + \sum_{i=1}^{k} R_\varepsilon(\hat{Z}, \Pi_i) - \lambda D(\hat{Z}_1, \hat{Z}_2) \tag{4}$$

for some hyperparameter λ and pairs of similar sentences respectively divided into two sets $\hat{Z}_1, \hat{Z}_2$. Π denoted in $L(\hat{Z}, \Pi)$ is the clustering of data given by the user or learned by the architecture. The choice of λ depends on how close we want to keep similar sentences in our projection. For larger values of λ the network focuses on collapsing similar pairs into the same vector which, if one is not careful enough, can lead to collapsing all vectors into one. For smaller values of λ the network has more freedom to decide which vector embeddings to keep close. This, on the other hand, can lead to an unwanted vector representation that tends to maximally distance vectors from each other. By minimizing (4) we

- maximize the volume of all embeddings, $R_\varepsilon(\hat{Z})$,
- minimize the volume of each cluster, $\sum_{i=1}^{k} R_\varepsilon(\hat{Z}, \Pi_i)$,
- maximize the cosine similarity of pairs of similar sentences, $\lambda D(\hat{Z}_1, \hat{Z}_2)$.

The consequence of this is that after the minimization we have an embedding in which different clusters are orthogonal to each other (see [15] for more details), i.e.

$$i \neq j \implies \hat{Z}_i\hat{Z}_j^\top = 0. \tag{5}$$

B. Architecture

Our model receives as input a batch of sentences S, encodes a sentence representations Z and outputs projected sentence representations $\hat{Z}$ together with cluster assignments Π for S. The overall architecture is shown in Fig. 1.

1) Sentence encoder: BERT [4] and its variants has set a new state-of-the art performance on sentence-pair regression and classification tasks. Unfortunately, it requires that both sentences are fed into network causing a computation overhead which renders simple tasks like finding similar sentence pairs in large datasets a costly procedure. Therefore, SBERT [7] is a modification of the BERT network which uses siamese network that is able to derive semantically meaningful sentence representations.

The model consists of BERT as a pre-trained encoder, a pooling layer that computes sentence representation as an average of hidden states from the last layer of BERT. SBERT is trained on the combination of the SNLI [19] and MultiNLI [20] datasets.

2) Projection layer: Following Li et al. [16] we use above mentioned SBERT as a backbone and two last linear heads used to produce features and cluster logits. Features given by the first head are additionally normalized to unit sphere and the clusters are learned from the given pairs of similar sentences. The whole architecture is described in Fig. 1 where blue denotes the SBERT model and gray denotes a feed forward neural network that we call a projection layer. In this projection layer we have two heads. The first head colored in red is a single linear layer that collects information about the clusters and applies Gumbel-Softmax [21]. The second head colored in green is again a single linear layer that outputs features which are in turn normalized to zero mean and unit variance. The ELU activation function is used due to its good properties [22].

III. Experiments

We trained our model on StackExchange duplicate questions as title/title pairs, used from CQADupStack [23]. The pipeline from Sentence Transformers[2] for SBERT and projection layer was used with default settings, 256 batch size and 50 epochs. The backbone *all-mpnet-base-v2* and distilled model *all-MiniLM-L6-v2* [24] as pre-trained SBERT were frozen and the only trained part was the projection layer. We refer to the former as MPNET and the latter as MiniLM. Hyperparameter λ from equation (4) was set to 2000 for dimensions 50 and 100, and to 4000 for all other dimensions. Our model is evaluated on several downstream NLP tasks. First of all, we test our model on those benchmarks that can include clustering, namely, semantic retrieval tasks. We also show that computed low-dimension sentence representations behave reasonably well on other semantic benchmarks. The sizes of these dimensions are motivated by experimental observation of suitable word vector sizes from [25] and [14] in which a connection between word vectors and sentence embeddings is established. For downstream NLP tasks such as standard textual similarity, sentiment analysis and question-type classification tasks we use available datasets from SentEval evaluation toolkit [26] for sentence embeddings. See [26] and references therein for dataset descriptions.

All our experiments were evaluated on AMD Ryzen Threadripper 3990X 64-Core Processor @ 4.3GHz, Nvidia GeForce RTX 3090 GPU, CUDA 11.6 with PyTorch implementation 1.9.1.

A. Semantic Retrieval (SR) Task

The semantic retrieval (SR) task is to find all sentences in the retrieval corpus that are semantically similar to the query sentence. The basic framework is to compute sentence embeddings for the retrieval corpus and the query sentence. The goal is to find closest points in retrieval corpus embedding space to the query. Sometimes, to speed up the process [27], one can cluster sentences in the retrieval corpus embedding space into k clusters and use query sentence to find the closest cluster of sentences. The Quora Duplicate Question Dataset[3] is used to evaluate our method. This dataset consists of 500k sentences with over 400k annotated question pairs if they are duplicates or not.

B. Semantic Textual Similarity (STS) Task

One of the baseline benchmarks in natural language processing is the semantic textual similarity (STS) task that qualitatively assesses the semantic similarity between two sentences (i.e., text snippets). Our model is evaluated by computing cosine similarity between sentence pair embeddings on standard STS tasks: STS 2012-2016 and STS Benchmark available in SentEval. These datasets were labeled between 0 and 5 scores indicating the semantic relatedness of sentence pairs. Evaluation on these datasets is conducted using Spearman rank correlation which measures the correlation quality between calculated and human labeled similarity. It is valued from -1 and 1 which will be high if the ranks of predicted similarities and human labels are similar.

C. Sentence classification (SC) Task

Sentiment classification tasks involve assigning a score for a sentiment of a snippet of text. It is formulated as a classification of text into two or more sentiment classes, namely negative, positive or neutral, or something in-between. Datasets SST, SUBJ, CR, MR are typical benchmarks for sentiment analysis. Moreover, another example of a sentence classification task is to assign a question type for a question, like in the TREC task. In the paraphrase detection problem (like MRPC), one must classify if one sentence is a paraphrase of the other. MPQA dataset is an example of opinion classification task. The performance metric for these benchmarks is given as accuracy. All these datasets are available in SentEval toolkit.

IV. Results

This section compares our method as a clustering and compression algorithm, respectively. In Table I, we compare how our clustering competes with k-means[4] algorithm and report time performance (needed time for encoding vectors, clustering, and total time) In the second part, we test our compression against semantic relatedness tasks. The results are reported in Tables I, II and III. Model names in these tables are structured as follows: the sBERT pretrained model name, the abrreviation MCR indicates that MCR[2] is used as a projection, the number followed is a projection dimension, and optionally if k-means is used.

[2]https://github.com/UKPLab/sentence-transformers

[3]https://www.kaggle.com/datasets/sambit7/first-quora-dataset

[4]implemented in *scikit-learn* Python package

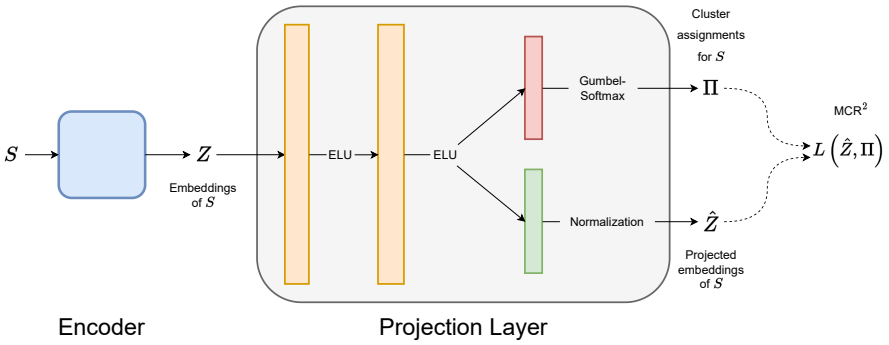


Fig. 1: The overall architecture

TABLE I: For Semantic Retrieval (SR) tasks the `all-mpnet-base-v2` SBERT model with MCR^2 projection to dimension 200 achieves best accuracy without no additional time for clustering like in the same setup with k-means (denoted with *). Clustering backbone sentence embeddings from SBERT with k-means (denoted with **) took almost an hour.

model	accuracy	TIME encoding	clustering	total
all-mpnet-base-v2 + MCR50	0.562	00:04:15	-	00:04:15
all-mpnet-base-v2 + MCR100	0.545	00:04:15	-	00:04:15
all-mpnet-base-v2 + MCR200	0.645	00:04:16	-	00:04:16
all-mpnet-base-v2 + MCR300	0.632	00:04:17	-	00:04:17
all-mpnet-base-v2 + MCR50 + kmeans	0.671	00:04:15	00:07:08	00:11:23
all-mpnet-base-v2 + MCR100 + kmeans	0.650	00:04:15	00:07:22	00:11:37
all-mpnet-base-v2 + MCR200 + kmeans*	0.635	00:04:16	00:09:08	00:13:24
all-mpnet-base-v2 + MCR300 + kmeans	0.631	00:04:17	00:11:09	00:15:26
all-mpnet-base-v2 + kmeans (768)**	0.648	00:04:17	00:59:57	01:04:12
all-mpnet-base-v2 + MCR768 + kmeans	0.630	00:04:17	00:18:15	00:22:32

TABLE II: For Semantic Textual Similarity (STS) tasks backbone model achieves the best results (bolded) for Spearman rank correlation coefficient on multiple benchmarks, although we observe comparable results of our method compared to the backbone and distilled model `all-MiniLM-L6-v2`.

model	STSb	STS12	STS13	STS14	STS15	STS16
all-mpnet-base-v2 + MCR50	0.749	0.666	0.739	0.713	0.754	0.768
all-mpnet-base-v2 + MCR100	0.788	0.696	0.782	0.753	0.791	0.793
all-mpnet-base-v2 + MCR200	0.818	0.712	0.812	0.779	0.819	0.816
all-mpnet-base-v2 + MCR300	0.821	0.718	0.817	0.783	0.827	0.823
all-mpnet-base-v2 (768)	0.836	0.722	0.821	0.790	0.838	0.831
all-MiniLM-L6-v2 + MCR50	0.752	0.654	0.690	0.682	0.741	0.737
all-MiniLM-L6-v2 + MCR100	0.778	0.685	0.742	0.721	0.780	0.777
all-MiniLM-L6-v2 + MCR200	0.810	0.705	0.773	0.751	0.813	0.792
all-MiniLM-L6-v2 + MCR300	0.813	0.710	0.780	0.759	0.826	0.800
all-MiniLM-L6-v2 (384)	0.824	0.711	0.790	0.772	0.838	0.812

TABLE III: For Sentence Classification (SC) tasks backbone model achieves the best results (bolded) for accuracy on multiple benchmarks, although we observe comparable results of our method compared to the backbone and distilled model.

model	SST2	SST5	MR	CR	SUBJ	MPQA	TREC
all-mpnet-base-v2 + MCR50	75.45	36.43	69.67	63.76	79.16	68.84	51.6
all-mpnet-base-v2 + MCR100	82.54	39.19	75.85	63.76	81.86	68.77	60.0
all-mpnet-base-v2 + MCR200	86.55	42.76	80.62	72.77	88.28	82.27	71.0
all-mpnet-base-v2 + MCR300	87.59	44.66	82.33	79.71	90.73	85.76	79.8
all-mpnet-base-v2 (768)	88.74	49.00	85.05	86.84	93.97	89.32	94.0
all-MiniLM-L6-v2 + MCR50	65.95	31.76	61.61	63.76	79.19	68.77	41.0
all-MiniLM-L6-v2 + MCR100	72.27	33.94	66.38	63.82	83.97	76.58	64.0
all-MiniLM-L6-v2 + MCR200	77.54	37.10	70.06	69.17	86.87	81.83	69.2
all-MiniLM-L6-v2 + MCR300	79.35	39.50	72.95	75.07	88.47	84.13	72.0
all-MiniLM-L6-v2 (384)	81.44	42.99	75.98	80.56	91.80	87.38	90.0

The number given in parenthesis is the default embedding size.

a) Results on SR tasks: We evaluate our projection layer capacity of clustering against k-means clustering in retrieval space of sentence embeddings. The query sentence is assigned to a cluster of semantically related sentences, and we compare whether the ground truth duplicate belongs to that cluster, reported as an accuracy score. In all our experiments, the number of clusters was up to 128 (chosen empirically). In Table I and Fig. 2, accuracy scores and overall time for computation (i.e., encoding of sentence embeddings and clustering)

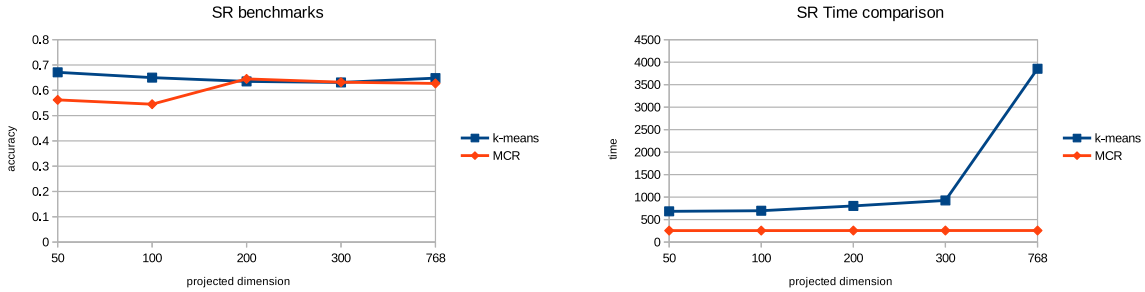


Fig. 2: Performance comparison on SR task

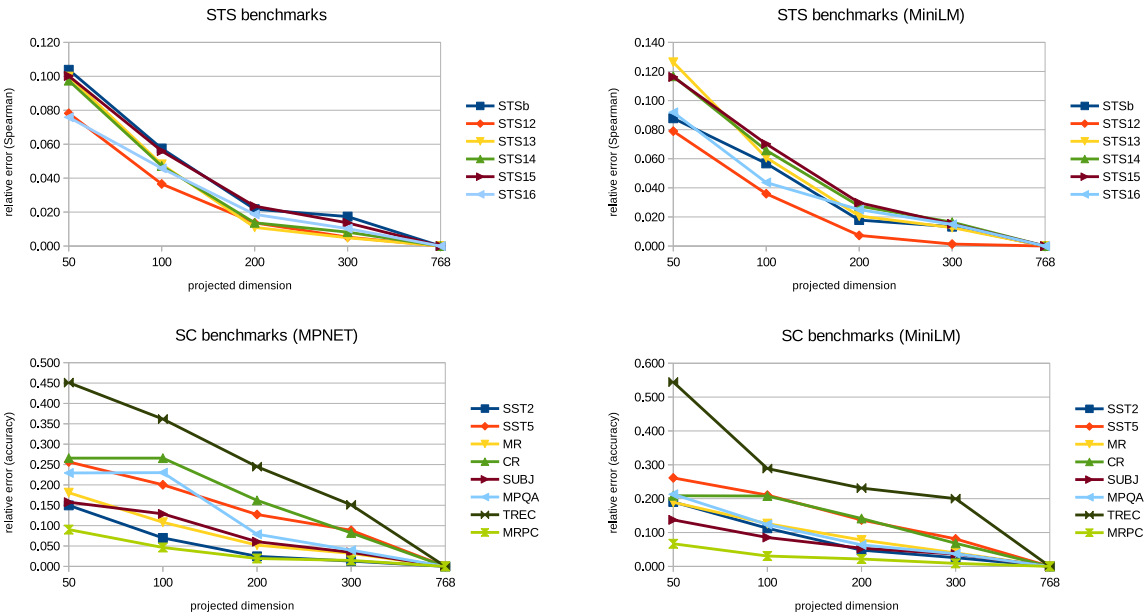


Fig. 3: Relative error in STS and SC benchmarks

depending on embedding size and type of model (MCR^2 with implicit clustering or k-means) are presented. Our method is comparable to k-means algorithm down to a certain dimension. On dimensions less than 200 k-means performed slightly better due to the fact that we did not put much effort into finding suitable λ values (suggested values of λ are from [16]). Our method computes clusters during inference which is much faster comparing to using k-means. Also worth noting, our projection layer used as a non-linear mapper in the original space (i.e., without any dimensionality reduction) retrofitted the sentence embeddings, enabling faster convergence of k-means algorithm (shown in the last row of Table I).

b) Results on STS tasks: Table II presents results for baseline (MPNET) and distilled model (MiniLM) coupled with the projection layer (MCR) with various embedding sizes. As one can see, a relative error of up to 13% in Spearman rank correlation is incurred if the sentence embedding dimension is as low as 6% of the original sentence embedding size. We conclude that due to the projection layer's ability to preserve cosine distance in lower-dimensional space, the neighborhood of points is preserved, resulting in less performance degradation. This trend is visible on all STS benchmarks with both models. The relative error in Spearman rank correlation coefficient with respect to projection dimension is shown in the first row of Fig. 3 for both the baseline and distilled model.

c) Results on SC tasks: As seen in Table III, it is observable that per-sentence classification problems like SST2 and MRPC have less performance degradation than per-token sentence classification problems like MPQA and per-sentence multi-classification problems like TREC, respectively. This is because fine-grained semantics for such tasks could not be preserved as much during projection. In the worst case, the performance degradation went up to 45% for the baseline model and up to 60% for the distilled model, respectively, at 6% of the original embedding size. The relative error in accuracy with respect to projected dimension is shown in the second row of Fig. 3 for both the baseline and distilled model.

V. Conclusion

In this paper, we demonstrated how MCR^2 technique could be used to obtain lower-dimension embeddings of

sentence representation for fast semantic retrieval tasks up to 70% of its original size. Also, we argued that these embeddings are comparable with SBERT results on standard semantic NLP benchmarks. Due to the projection layer's ability to cluster data, we were able to cluster our sentences without any extra time cost and further reduce the representation of sentences to a reasonable dimension size without significant loss of the important semantic features. We hope our approach gives new insights for possible applications in deploying AI models in smaller-scale computer systems.

REFERENCES

[1] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A neural probabilistic language model," *J. Mach. Learn. Res.*, vol. 3, no. null, p. 1137–1155, mar 2003.

[2] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," 2013. [Online]. Available: https://arxiv.org/abs/1310.4546

[3] J. Pennington, R. Socher, and C. Manning, "GloVe: Global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1532–1543. [Online]. Available: https://aclanthology.org/D14-1162

[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[5] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. New Orleans, Louisiana: Association for Computational Linguistics, Jun. 2018, pp. 2227–2237. [Online]. Available: https://aclanthology.org/N18-1202

[6] H. Aldarmaki and M. Diab, "Evaluation of unsupervised compositional representations," 2018. [Online]. Available: https://arxiv.org/abs/1806.04713

[7] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," Aug. 2019, arXiv:1908.10084 [cs]. [Online]. Available: http://arxiv.org/abs/1908.10084

[8] R. Kiros, Y. Zhu, R. Salakhutdinov, R. S. Zemel, A. Torralba, R. Urtasun, and S. Fidler, "Skip-thought vectors," in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'15. Cambridge, MA, USA: MIT Press, 2015, p. 3294–3302.

[9] L. Logeswaran and H. Lee, "An efficient framework for learning sentence representations," *CoRR*, vol. abs/1803.02893, 2018. [Online]. Available: http://arxiv.org/abs/1803.02893

[10] M. Pagliardini, P. Gupta, and M. Jaggi, "Unsupervised learning of sentence embeddings using compositional n-gram features," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. Association for Computational Linguistics, 2018.

[11] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *arXiv preprint arXiv:1607.04606*, 2016.

[12] A. Conneau, D. Kiela, H. Schwenk, L. Barrault, and A. Bordes, "Supervised learning of universal sentence representations from natural language inference data," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics, Sep. 2017, pp. 670–680. [Online]. Available: https://aclanthology.org/D17-1070

[13] T. Gao, X. Yao, and D. Chen, "SimCSE: Simple contrastive learning of sentence embeddings," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 6894–6910. [Online]. Available: https://aclanthology.org/2021.emnlp-main.552

[14] B. Li, H. Zhou, J. He, M. Wang, Y. Yang, and L. Li, "On the sentence embeddings from pre-trained language models," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics, Nov. 2020, pp. 9119–9130. [Online]. Available: https://aclanthology.org/2020.emnlp-main.733

[15] Y. Yu, K. H. R. Chan, C. You, C. Song, and Y. Ma, "Learning Diverse and Discriminative Representations via the Principle of Maximal Coding Rate Reduction," Jun. 2020, arXiv:2006.08558 [cs, math, stat]. [Online]. Available: http://arxiv.org/abs/2006.08558

[16] Z. Li, Y. Chen, Y. LeCun, and F. T. Sommer, "Neural Manifold Clustering and Embedding," Jan. 2022, arXiv:2201.10000 [cs]. [Online]. Available: http://arxiv.org/abs/2201.10000

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.

[18] Y. Ma, H. Derksen, W. Hong, and J. Wright, "Segmentation of multivariate mixed data via lossy data coding and compression," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, pp. 1546–1562, 2007.

[19] S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning, "A large annotated corpus for learning natural language inference," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal: Association for Computational Linguistics, Sep. 2015, pp. 632–642. [Online]. Available: https://aclanthology.org/D15-1075

[20] A. Williams, N. Nangia, and S. Bowman, "A broad-coverage challenge corpus for sentence understanding through inference," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. New Orleans, Louisiana: Association for Computational Linguistics, Jun. 2018, pp. 1112–1122. [Online]. Available: https://aclanthology.org/N18-1101

[21] I. A. M. Huijben, W. Kool, M. B. Paulus, and R. J. G. van Sloun, "A review of the gumbel-max trick and its extensions for discrete stochasticity in machine learning," 2021. [Online]. Available: https://arxiv.org/abs/2110.01515

[22] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," 2015. [Online]. Available: https://arxiv.org/abs/1511.07289

[23] D. Hoogeveen, K. M. Verspoor, and T. Baldwin, "Cqadupstack: A benchmark data set for community question-answering research," in *Proceedings of the 20th Australasian Document Computing Symposium (ADCS)*, ser. ADCS '15. New York, NY, USA: ACM, 2015, pp. 3:1–3:8. [Online]. Available: http://doi.acm.org/10.1145/2838931.2838934

[24] N. Reimers and I. Gurevych, "Making monolingual sentence embeddings multilingual using knowledge distillation," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2020. [Online]. Available: https://arxiv.org/abs/2004.09813

[25] K. Patel and P. Bhattacharyya, "Towards lower bounds on number of dimensions for word embeddings," in *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Taipei, Taiwan: Asian Federation of Natural Language Processing, Nov. 2017, pp. 31–36. [Online]. Available: https://aclanthology.org/I17-2006

[26] A. Conneau and D. Kiela, "SentEval: An evaluation toolkit for universal sentence representations," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan: European Language Resources Association (ELRA), May 2018. [Online]. Available: https://aclanthology.org/L18-1269

[27] J. Johnson, M. Douze, and H. Jégou, "Billion-scale similarity search with GPUs," *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535–547, 2019.

DistilBERT and RoBERTa Models for Identification of Fake News

A. Kitanovski, M. Toshevska and G. Mirceva
Faculty of computer science and engineering, Ss. Cyril and Methodius University in Skopje, Skopje, Macedonia
aleksandar.kitanovski@students.finki.ukim.mk, martina.toshevska@finki.ukim.mk, georgina.mirceva@finki.ukim.mk

***Abstract* - The proliferation of fake news has become a significant issue in today's society, affecting the public's perception of current events and causing harm to individuals and organizations. Therefore, the need for automated systems that can identify and flag fake news is critical. This paper presents a study on the effectiveness of DistilBERT and RoBERTa, two state-of-the-art language models, for detecting fake news. In this study, we trained both models on a dataset of labelled news articles and evaluated them on two different datasets, comparing their performance in terms of accuracy, precision, recall and F1-score. The results of our experiments show that both models perform well in detecting fake news, with RoBERTa model achieving slightly better results in overall. Our study highlights the ability of these models to effectively identify fake news and help combat misinformation.**

Keywords - fake news; deep learning; transformer models; DistilBERT; RoBERTa

I. INTRODUCTION

The rise of the internet has brought about a wealth of information at our fingertips, but it has also led to the spread of false information, commonly referred to as fake news. Fake news can have serious consequences, including influencing public opinion, spreading misinformation, and even disrupting elections. As a result, detecting fake news has become an important task in the field of natural language processing (NLP). Recent advancements in NLP and deep learning have led to the development of various techniques for fake news detection. One such approach involves the use of pre-trained language models. Two such models are DistilBERT [1] and RoBERTa [2], these models have achieved astonishing results in various NLP tasks. The current study aims to investigate the performance of these models in detecting fake news. DistilBERT [1] is a distilled version of BERT [3], which is a pre-trained transformer model, while RoBERTa [2] is an improved version of BERT.

The research will be conducted by fine-tuning these models on a fake news detection dataset and comparing their performance on two different fake news datasets. This study is significant in light of the increasing spread of fake news and the need for effective methods to detect and curb its impact. With the growing reliance on social media and the internet for news and information, it has become more important than ever to ensure the accuracy and authenticity of the information we consume. The findings of this study will be of interest to researchers and practitioners in the field of NLP, as well as those concerned with the spread of fake news and its impact on society.

The rest of this paper is organized in the following way. Section 2 presents some related work for solving this task. Section 3 gives detailed description about the approach that is used containing details regarding the datasets that are used for training the prediction models, the preprocessing steps that are made, how the deep learning models, i.e. DistilBERT and RoBERTa are created, as well as the evaluation measures that are used to estimated their predictive performance. The experimental results are presented in Section 4 and are further discussed in Section 5, while Section 6 concludes the paper and gives some directions for further research.

II. RELATED WORK

In recent years, the detection of fake news has been an active area of research. Many studies have used different techniques to address this problem, such as using traditional machine learning algorithms, deep learning algorithms, and natural language processing techniques.

A. Traditional Machine Learning Approaches

Traditional machine learning algorithms, such as Naive Bayes and Support Vector Machines, have been widely used for fake news detection. These algorithms rely on hand-crafted features, such as the frequency of certain words, to classify news articles as true or fake. Ahmed et al. [4] evaluated the performance of six machine learning algorithms including Support Vector Machines (SVM), Linear Support Vector Machines (LSVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), Stochastic Gradient Descent (SGD) and Decision Trees (DT). These classification algorithms were applied after extracting TF and TF-IDF feature vectors. The authors in [4] achieved good results with each of the models, with Linear Support Vector Machines achieving the best results.

B. Deep Learning Models

Recently, deep learning algorithms, such as Convolutional Neural Networks and Recurrent Neural Networks, have been used for fake news detection. These algorithms are able to learn complex representations of the articles, and they are better able to handle the context of the words in the articles. However, these methods can be computationally expensive and require large amount of data to train effectively. Saleh et al. [5] compared the performance of several different machine learning

algorithms with three deep learning algorithms, i.e. Recurrent Neural Networks (RNN), Long-Short Term Memory Networks (LSTM) and Convolutional Neural Networks (CNN) optimized for fake news detection. All of their models achieved good results, but CNNs outperformed the rest of the models.

C. Transformers Models

Transformer models have also been used for the task of fake news detection. Mateusz Szczepański et al. [6] have suggested a technique for enhancing the explainability of BERT-based models used in fake news detection. Their approach involves adding an explainability module onto the existing model architecture, thereby avoiding the need for a redesign. A different team of researchers has presented an approach to identify fake news by leveraging the RoBERTa model to detect emotions, which are subsequently employed as features in a Random Forest Classifier to improve the detection accuracy [7]. In [8], another research team has introduced and evaluated an approach utilizing overlapping window strides with multiple transformer models, specifically DeBERTa, RoBERTa, XLM-RoBERTa and BigBird.

This paper focuses on comparing the performance of two specific transformer models, namely DistilBERT and RoBERTa, in detecting fake news. Therefore, our research adds to the existing literature by providing insights into the relative strengths and weaknesses of these models for solving the task at hand.

We want to note that in [9] we already used the RoBERTa [2] model, as well as BERT [3], for solving fake news detection task focused solely on data regarding COVID-19. Namely, in [9], another dataset was used that contains data about tweets related to COVID-19, while in this study we aim to create model for general purpose covering various topics that is a more challenging task.

III. Methodology

A. Datasets

We utilized two datasets in the analysis made in this study. The first dataset was split into three subsets so that 10% of the samples were used for training the models, 20% were used for validation purposes, and the remaining 70% being reserved for testing. We refer to this first dataset as ds-1 and its three subsets are referred to as ds-1 (train), ds-1 (validation), and ds-1 (test) respectively. The reason ds-1 (train) and ds-1 (validation) are only 10% and 20% respectively of the ds-1 dataset is because of the hardware limitations. The ds-1 dataset was sourced from Hadeer Ahmed et al. [10].

The second dataset, referred to as ds-2, was exclusively used for testing purposes. The idea for using this dataset was to being able to test how well the models generalize and if they could be used for new data. The ds-2 dataset was obtained from Kaggle [11].

The class distribution of the classes for all subsets of the ds-1 dataset can be seen in Table 1. It can be seen that the dataset is balanced, which is also a case with its three subsets.

TABLE I. The Distribution of The Classes in the Ds-1 Dataset

Dataset	True	Fake
ds-1 (train)	2044	2321
ds-1 (validation)	4133	4597
ds-1 (test)	14625	15932

TABLE II. The Distribution of The Classes in the Ds-2 Dataset

Dataset	True	Fake
ds-2 (original)	3171	3164
ds-2 (filtered)	3155	3123

The ds-2 dataset has 3164 fake news samples and 3171 true news samples, but some of them were removed due to being very similar to news samples from ds-1, as it will be described in the next subsection. The distribution of the classes in ds-2 before and after this filtering can be seen in Table 2. From Table 2 it can be noted that both the original and the filtered datasets are balanced.

Both DistilBERT and RoBERTa models were fine-tuned on ds-1 (train) using ds-1 (validation) for validation purposes. After fine-tuning the models, they were evaluated on ds-1 (test) and ds-2 (filtered).

B. Preprocessing

To preprocess the data we followed these steps: mapping all letters to lowercase, removing any non-whitespace or non-alphanumeric characters, and removing stopwords from the text. Initially, these preprocessing steps resulted in good results when evaluating on ds-1 (test), but poor results when evaluating on ds-2. Upon further analysis, we discovered that all news articles labelled as "true" in ds-1 began with the same text and that all of them were sourced from Reuters. To address this issue, we added an additional preprocessing step to remove this text after converting all letters to lowercase. This significantly improved the results on ds-2.

To ensure that ds-2 was different enough from ds-1 and to test the models' ability to generalize, we removed all articles from ds-2 (original) that were similar to the articles in ds-1. This was done by computing the tf-idf embeddings of the articles and calculating the pairwise cosine similarity between the samples from the two datasets. Higher value for the cosine similarity indicates dissimilar samples, while low values correspond to high similarity. Through empirical evaluation, we determined that articles with a cosine similarity greater than 0.65 were sufficiently similar and were removed thus obtaining the dataset denoted as ds-2 (filtered). As we can see from Table 2, only a few articles were removed and the class distribution in ds-2 (filtered) remained balanced.

C. DistilBERT and RoBERTa Models

As previously mentioned, in this study we evaluate the performance of two state-of-the-art language models (i.e. DistilBERT [1] and RoBERTa [2]) for solving the fake news detection task.

DistilBERT [1] learns an approximate version of BERT using a knowledge distillation technique [12], [13]. It has only one half of the layers of the original BERT model, thus reducing the number of parameters by 40%. DistilBERT is designed to be smaller and faster than BERT [3], while still retaining much of its accuracy.

RoBERTa [2] is also built upon BERT [3] but has been trained on a larger corpus of text and with a different training procedure. It is trained with dynamic masking, where the masking pattern is generated every time when a sequence is fed to the model, as opposed to static masking in the original BERT implementation where the same training mask was used. RoBERTa has been trained without next sentence prediction objective, with bigger batches over more data and longer sequences.

We have used DistilBERT and RoBERTa implementations from the huggingface library [14]. We fine-tuned both models for 3 epochs using the Adam optimizer. We also used padding and truncation for the tokenizers of both models.

D. Evaluation Measures

We evaluated the models' performance using classification accuracy, precision, recall, and F1-score as evaluation measures. The classification accuracy is an appropriate measure in our case because the datasets are well balanced. These metrics are defined as presented in Eq. (1), Eq. (2), Eq. (3) and Eq. (4).

$$\text{accuracy}=\frac{\text{TP}+\text{TN}}{\text{N}} \tag{1}$$

$$\text{precision}=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{2}$$

$$\text{recall}=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{3}$$

$$\text{F1-score}=\frac{2\times\text{precision}\times\text{recall}}{\text{precision}+\text{recall}} \tag{4}$$

In the preceding definitions, TP denotes the number of true positives (the samples that are correctly classified as positive), TN is the number of true negatives (the samples correctly classified as negative), FP denotes the number of false positives (the samples that are misclassified as positive), FN is the number of false negatives (the samples that are misclassified as negative), while N is the total number of samples.

TABLE III. THE RESULTS FOR DISTILBERT ON DS-1 (TEST)

Evaluation measure	True	Fake
Accuracy	0.99	0.99
Precision	0.99	0.98
Recall	0.98	0.99
F1-score	0.99	0.99

TABLE IV. THE RESULTS FOR DISTILBERT ON DS-2 (FILTERED)

Evaluation measure	True	Fake
Accuracy	0.68	0.68
Precision	0.70	0.67
Recall	0.64	0.72
F1-score	0.67	0.69

TABLE V. THE RESULTS FOR ROBERTA ON DS-1 (TEST)

Evaluation measure	True	Fake
Accuracy	0.99	0.99
Precision	0.99	0.98
Recall	0.98	0.99
F1-score	0.99	0.99

TABLE VI. THE RESULTS FOR ROBERTA ON DS-2 (FILTERED)

Evaluation measure	True	Fake
Accuracy	0.71	0.71
Precision	0.77	0.67
Recall	0.61	0.81
F1-score	0.68	0.74

IV. EXPERIMENTAL RESULTS

In this section, we present the results of our experiments with the DistilBERT and RoBERTa models on the two datasets, ds-1 (test) and ds-2 (filtered). The results for DistilBERT on ds-1 (test) and ds-2 (filtered) are shown in Table 3 and Table 4, respectively. Similarly, the results for RoBERTa on ds-1 (test) and ds-2 (filtered) are shown in Table 5 and Table 6, respectively.

Table 3 and Table 4 show that DistilBERT achieved good performance on ds-1 (test), with high accuracy, precision, recall, and F1-score. However, its performance was not as good on ds-2 (filtered), with lower precision, recall, and F1-score.

Table 5 and Table 6 show that RoBERTa performed similarly to DistilBERT on ds-1 (test), with high accuracy, precision, recall, and F1-score. On ds-2 (filtered),

RoBERTa outperformed DistilBERT, with higher accuracy and F1-score compared to DistilBERT.

Overall, the results suggest that RoBERTa is a better choice for fake news detection compared to DistilBERT, especially considering the fact that the recall is a very important evaluation measure for fake news detection since we want to detect as much of the fake news as possible, namely RoBERTa has a much higher recall for fake news class than DistilBERT on ds-2 (filtered).

We also compared our models with the models obtained in [4], [15], where TF and TF-IDF feature vectors were used in combination with six machine learning algorithms, i.e. Support Vector Machines (SVM), Linear Support Vector Machines (LSVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), Stochastic Gradient Descent (SGD) and Decision Trees (DT). We want to note that in [4], [15] the division of the samples into training and test sets was not the same as in our paper. In Table 7 the results from the comparison are provided. From the results it is evident that the DistilBERT and RoBERTa models obtained in this paper are significantly more accurate than the models obtained in [4], [15], which are based on several well-known machine algorithms.

V. DISCUSSION

RoBERTa showed better performance compared to DistilBERT on ds-2 (filtered), especially when considering the recall for the fake news class where RoBERTa was much better than DistilBERT. This is very important since higher recall for the fake news class indicates that the model is more able to detect the fake news. This is especially important in solving fake news detection task in order to reduce the number of unidentified samples that are fake in order to prevent such news to be presented to the end users.

The lower predictive performance obtained on ds-2 (filtered) compared to the case when ds-1 (test) was used could be due to using samples that are focused on other topics for which is harder to determine whether are fake or real based on the data used for training the model. Namely the two datasets could cover texts for very distinct topics.

TABLE VII. THE RESULTS FROM THE COMPARISION WITH SEVERAL EXISTING MODELS

Model	**Accuracy (%)**
Our DistilBERT model	99
Our RoBERTa model	99
SVM	86
LSVM	92
KNN	83
DT	89
SGD	89
LR	89

Our analysis showed that the additional preprocessing step of removing the specific text present at the beginning in the "true" news in ds-1 improved the performance of both models on ds-2 (filtered). These findings suggest that careful preprocessing is crucial for the effective use of these models in detecting fake news. Before applying this preprocessing step, the model identified that specific text that was present in the “true” class and the revealed models were mostly based on that irrelevant knowledge, which was not relevant for ds-2 (filtered) where this text was not present. This way, we became aware that this preprocessing step is required in order to build relevant model, which was later confirmed with the results obtained when adding this preprocessing step. Due to this, we want to note that it is better to assess the model on additional dataset in order to become aware if such patterns are present in the data that should be eliminated in order to build relevant models.

VI. CONCLUSION

In this paper, we focused on solving the task for fake news detection. For that purpose, we utilized the DistilBERT and RoBERTa transformer models in order to create prediction models for solving the task at hand. Two datasets were used in this study. The first one was divided into subsets for training, validation and testing, while the second dataset was solely used for testing purposes.

We compared the performance of the DistilBERT and RoBERTa models for detecting fake news over the two datasets. The results indicate that both models performed well in terms of accuracy, precision, recall, and F1-score. However, RoBERTa outperforms DistilBERT on ds-2 (filtered) regarding the recall for the fake class, thus it better identifies fake news. Additionally, the obtained DistilBERT and RoBERTa models were compared with several models from the literature that are based on several well-known machine learning algorithms, and our models obtained significantly better results.

In conclusion, our study highlights the potential of using transformers such as DistilBERT and RoBERTa in detecting fake news, and underscores the importance of carefully preprocessing the data for improved performance. Further research is needed to explore the use of these models and to evaluate their performance in real-world applications. Besides application of these models in various contexts, also as future work we plan to utilize some other transformer models, which are very popular nowadays since they offer very accurate predictions.

ACKNOWLEDGMENT

This work was partially financed by the Faculty of computer science and engineering at the “Ss. Cyril and Methodius University in Skopje”, Skopje, Macedonia.

REFERENCES

[1] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter,” arXiv preprint arXiv:1910.01108, 2019.

[2] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “Roberta: A robustly

optimized bert pretraining approach," arXiv preprint arXiv:1907.11692, 2019.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

[4] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," Security and Privacy, vol. 1, no. 1, p. e9, 2018.

[5] H. Saleh, A. Alharbi, and S. H. Alsamhi, "Opcnn-fake: Optimized convolutional neural network for fake news detection," IEEE Access, vol. 9, pp. 129471–129489, 2021.

[6] M. Szczepański, M. Pawlicki, R. Kozik, and Michał Choraś, "New explainability method for BERT-based model in fake news detection," Scientific Reports, vol. 11, 23705, 2021.

[7] V. Kolev, G. Weiss, and G. Spanakis, "FOREAL: RoBERTa Model for Fake News Detection based on Emotions," 14th International Conference on Agents and Artificial Intelligence (ICAART 2022), vol. 2, pp. 429-440, 2022.

[8] H. R. LekshmiAmmal and A. K. Madasamy, "NITK-IT NLP at CheckThat! 2022: Window based approach for Fake News Detection using transformers," Conference and Labs of the Evaluation Forum, 2022.

[9] T. Pavlov and G. Mirceva, "COVID-19 Fake News Detection by Using BERT and RoBERTa models," 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO 2022), pp. 312-316, Opatija, Croatia, 2022.

[10] S. S. Hadeer Ahmed and I. Traore, "Kaggle," 2020, accessed: 2022-01-15. [Online]. Available: https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset.

[11] A. Motefaker, "Kaggle," 2022, accessed: 2022-01-15. [Online]. Available: https://www.kaggle.com/datasets/amirmotefaker/detecting-fake-news-dataset.

[12] C. Bucilua, R. Caruana, and A. Niculescu-Mizil, "Model compression," 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 535–541, 2006.

[13] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," arXiv preprint arXiv:1503.02531, 2015.

[14] Huggingface library, accessed: 2022-01-15. [Online]. Available: https://huggingface.co/.

[15] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," In: I. Traore, I. Woungang, A. Awad (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, ISDDC 2017, Lecture Notes in Computer Science, vol. 10618, Springer, Cham, pp. 127-138, 2017.

Analysis of Pro-Russian Tweets during Russian Invasion of Ukraine

Josip Katalinić
Department for information and communication sciences,
Faculty of Humanities and Social Sciences, University of
Zagreb, Zagreb, Croatia
josipkatalinicandroid@gmail.com

***Abstract* – With the Russian invasion of Ukraine on February 24, 2022 a new aspect of the conflict opens up, the fight on social networks. By analyzing the most influential pro-Russian Twitter profiles that cover the daily events and impacts of the Russian invasion of Ukraine as well as outside it, the number of followers, people following as well as tweets, time series comparison of publishing, sentiment through polarity and subjectivity, and the classification of individual tweets as malicious by building an SVM (Support Vector Machine) classifier were collected and analyzed. For the purpose of training the SVM classifier, a sample from dataset of 3 million identified malicious tweets deleted by Twitter after being linked to the Russian IRA agency known for operating malicious user accounts, available through the Kaggle data science community, was used. The results of the work show that pro-Russian Twitter profiles have clear and defined influence operations that oscillate through different time periods as a reaction to the dynamics and development of the conflict, and thus certain events become narrative elements that can influence the emotions of the target audience.**

Keywords - twitter; text classification; invasion of Ukraine; data analysis; influence operations;

I. INTRODUCTION

With the Russian invasion of Ukraine on February 24, 2022, a new aspect of the conflict opens up, the fight on social networks. An increased number of pro-Russian Tweets showing significant influence operations aimed at directing readers thinking and behavior toward narratives that suit the Kremlin government. In the first two weeks of the conflict, there was a 381% increase in tweets with the context of Putin rather than Zelensky [1].

Reference [2] states that psychological warfare is part of the information war that the Russian Federation uses with the help of "soft power". Reference [3] defines "soft power" as the power of attraction, where an agenda that is considered legitimate is set. While in contrast "hard power" is the use of force, payment and setting a certain agenda based on them. Hard power is push, while soft power is pull i.e., hard power is like waving a carrot or stick, while soft power is like a magnet. In this paper based on collected pro-Russian tweets we are presenting in which way tweets are disseminated on Twitter platform.

In this paper, pro-Russian Tweeter profiles, which contribute to the pro-Russian agenda with their reactions and announcements, were collected and processed. The aim of this research is to observe and present the trends that characterize the collected tweets, and to process the mentioned tweets at the level of sentiment analysis through subjectivity and objectivity, as well as through positive and negative. In addition to sentiment analysis, tweets were also classified as malicious or not based on a previously defined set of malicious users. All the mentioned features are presented through the passage of time, and connected with the significant events of the Russian invasion of Ukraine extending the understanding of influence operations and its patterns on Twitter.

The remainder of the paper is structured as follows. Process of data collection is presented in Section 2. Time analysis of collected tweets is covered in Section 3. Sentiment analysis as well as time series visualization is outlined in Section 4. SVM malicious user classifier results are obtained in Section 5. Study findings are presented and discussed together with future research in Section 6.

A. Advent of milbloggers

In addition to the aforementioned influence operations, daily reporting on events in Ukraine as well as events related to the conflict using OSINT (Open-source intelligence) sources enables public policy think tanks [4], as well as private users [5] to create their reports. Access to OSINT makes it possible to define the term milblogger, which refers to a creator who addresses the topics and dynamics of the conflict.

While directing the narrative towards Russia's successes in Ukraine, pro-Russian milbloggers also criticize certain decisions of the Russian military command, such as in the case of Alexander Zhychkovskiy and Alexander Khodarkovsky. During which Zhychkovskiy criticized the neglect of reservists on the front in the Zaporozhye region, which lost priority. Zhychkovskiy reported that Russian commanders trapped their own lightly equipped infantry units in areas of intense Ukrainian artillery fire without significant artillery support, and that they did not rotate other units through the areas to replace them. While Khodarkovsky alleges that Russian commanders are not sending reinforcements in time, preventing Russian forces from resting between attacks [6].

One of the most influential pro-Russian millblogger is Rybar, who has over 1.1 million subscribers on the Telegram channel [7], while his English Twitter profile has over 45,000 followers at the time of writing this paper [8]. Reference [9] states that the founder of Rybar, which employs at least 10 people, is 31-year-old military translator Mikhali Zvinchuk. He is a former employee of the press service of the Ministry of Defense of Russia, born in Vladivostok, studied at the Military University in Moscow, specializing in the Arabic language. From 2015-17, he was employed at the Ministry of Defense and helped organize press trips to Syria for Russian journalists. In addition to

transmitting daily news on the field and very detailed war maps, Rybar also serves as a channel for sensitive messages from Moscow. So, Rybar was the first to publish the video of the Russian "nuclear train" that caused panic headlines in Europe [10].

Milbloggers also provide information on which to gather pro-Russian narratives that are part of influence operations. Their reporting, as well as the coverage of events, can therefore create narrative elements that aim to confuse the reader combining high interactivity and the proliferation of the Twitter platform by reaching numerous users.

II. DATA COLLECTION

Using snowball sampling by going to Rybar Twitter followers, 19 active pro-Russian Twitter profiles were selected, which, including the Rybar profile, makes a set of 20 profiles.

TABLE I. OVERVIEW OF SELECTED 20 PRO-RUSSIAN TWITTER PROFILES

Following	Followers	Tweets
82,197	855,124	323,424

For the selected 20 Twitter profiles, tweets were collected from May 6, 2022 until January 22, 2023. The Python programming language and the Tweepy library, which provides access to the Twitter v2 API (application programming interface), were used to collect tweets [11]. While for access to historical data, Academic Research access was used, which was obtained from Twitter for the purposes of writing the paper. From the collected tweets, exclude filters were defined on retweets and replies so that only the author's tweets were collected. This approach aims to prevent a large volume of tweets that are not related to the author.

A. *Analysis of collected tweets*

In addition to collecting the tweet text, the following were also collected: tweet id, publication date, quote count, retweet count, reply count and like count. The number of collected tweets for the specified period is 13,381 tweets.

TABLE II. OVERVIEW OF THE COLLECTED 13,381 TWEETS

Quotes	Retweets	Replies	Likes	Impressions
67,515	718,468	279,683	3,765,282	126,233,781

By analyzing the collected tweets, we can see that 13,381 tweet posts are responsible for 126,233,781 impressions, which indicates a significant influence and reach of each individual posts. While other parameters such as quotes, retweets, replies and likes indicate a high rate of interaction with posts. Each individual tweet is stored in a corresponding column using the Python library openpyxl. For data processing purposes, the pandas data frame was used, while for visualization purposes, Matplotlib was used, both libraries can be installed using the pip package installer for Python.

In addition to using Tweepy library for obtaining relevant tweets relating to milbloggers, two additional datasets were obtained from Kaggle in regard to classifier. The dataset for creating the classifier i.e., training and testing, is based on the Russian Troll Tweets dataset [12] for malicious users while the Sentiment Dataset with 1 Million Tweets is used for non-malicious users [13].

III. TIME ANALYSIS OF TWEETS

The monthly tweet posting interval shown in Fig. 1 demonstrates the increase in the number of posts over the months. A particularly sharp increase was recorded from December to January. The events that took place during the end of 2022 include two successful counter-offensives from the Ukrainian side in the Kharkiv and Kherson regions, in which Ukraine regained a large part of the occupied territory [14]. While on the Russian side during September 21, 2022, the President of the Russian Federation announced partial military mobilization. According to the Defence Minister of the Russian Federation, the approximate number of mobilized persons is at 300,000 [15].

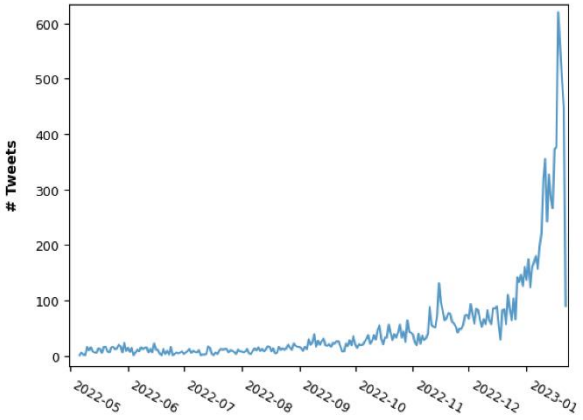


Figure 1. Monthly tweet posting interval

Fig. 2 shows daily interval of posting tweets, with the highest number of posts on Wednesdays and Fridays, and the lowest on Sundays and Mondays. The mentioned distribution is expected since most users who publish tweets are most active during the weekday, and their distribution decreases over the weekend [16].

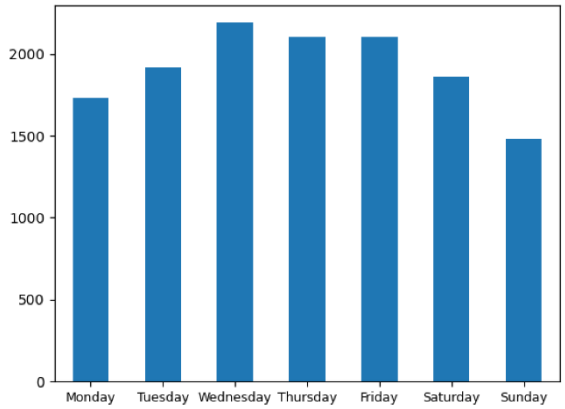


Figure 2. Daily interval of posting tweets

The hourly interval of posting tweets shown in Fig. 3 shows the hourly distribution of tweets where the highest number of tweets was published between 2:00 pm and 6:00 pm, while the lowest number of tweets was published between 3:00 am and 6:00 am.

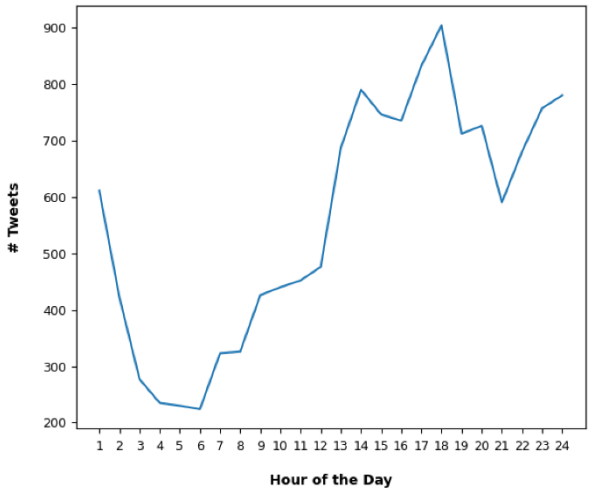


Figure 3. Hourly interval of posting tweets

IV. SENTIMENT ANALYSIS OF TWEETS

Reference [17] states that sentiment classification or sentiment analysis in text classification on social media platform like Twitter is defined as a process of finding out public opinion about an event, product or topic using techniques like machine learning. In it, public opinions are classified into categories like "Positive", "Negative" and "Neutral". Sentiment classification helps organizations to gain insightful knowledge from retrieved data for swift decisions on crucial moments.

Reference [18] points out that sentiment analysis as a field of research, is closely related to (or can be considered a part of) computational linguistics, natural language processing, and text mining. Proceeding from the study of affective state (psychology) and judgment (appraisal theory), this field seeks to answer questions long studied in other areas of discourse using new tools provided by data mining and computational linguistics. Sentiment Analysis has many names. It's often referred to as subjectivity analysis, opinion mining, and appraisal extraction, with some connections to affective computing (computer recognition and expression of emotion).

For the purposes of sentiment analysis, the TextBlob library [19] is used, whose implementation of the SVM classifier used for sentiment analysis shows an accuracy of 88% [20]. TextBlob is a rule-based analysis library that focuses on lexical content and integrates the WordNet corpus for sentiment analysis [21]. In Python the sentiment property of TextBlob returns a named tuple of the form Sentiment(polarity, subjectivity). The polarity score is a float in the range -1.0 and 1.0 while the subjectivity is a float in the range 0.0 and 1.0 where 0.0 is very objective and 1.0 is very subjective. [22]. Tweet polarity histogram shown in Fig. 4 is presenting the polarity of tweets using a visualization line KDE (kernel density estimation), also known as the Parzen's window to estimate the underlying probability density function of a polarity [23].

The visualized polarity in Fig. 4 on the x-axis represents values between -1 and 1, where -1 is the most negative words, and would include words like "disgusting", "distressing" and "miserable", while 1 is the most positive words such as "superb", "greatest" and "magnificent". Apart from the previously mentioned extremes, most of the words do not belong to the mentioned categories, so words like "football", "wood" and "resting" have a neutral value of 0, while words like: "tired", "hard" and "annoyed" have a slightly negative values in contrast to words like: "accomplished", "satisfied" and "gladly" which have slightly positive values.

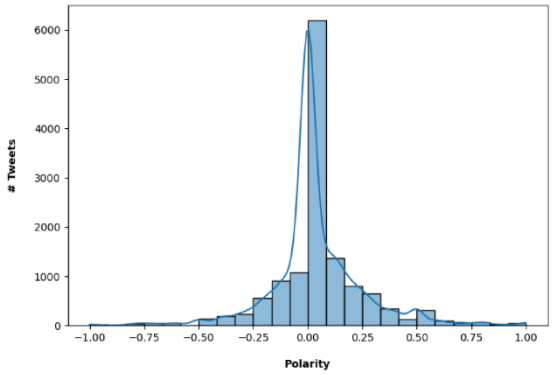


Figure 4. Tweet polarity histogram

That same data is shown in Fig. 5 over time to display correlation between months and polarity associated with each observed month. For more descriptive visualization of time series regarding trends in polarity and subjectivity expanding window and rolling window mean calculations were introduced via calls to pandas.DataFrame.expanding [24] and pandas.DataFrame.rolling [25] methods. For expanding window, number of observations in window is set to value of 1 (default value), while for rolling window, size of the moving window has been set to the time period of 6 hours, both values were calculated by calling:

```
df['sentiment'].expanding(1).mean()
df['sentiment'].rolling('6h').mean()
```

Over time, the change in polarity shows a slight increase, which is evident from the expanding mean which uses more and more observations each time while traversing data. On the other hand, rolling mean uses the same number of observations each time when traversing to define its values. Based on this, the expanding mean value is significantly more flattened than the rolling mean since it takes into account and incorporates new observations that are available. When observing rolling mean we can infer that it is more uniform than the individual representation of each polarity, because it is based on series of averages, while each polarity is raw value without representation of average values.

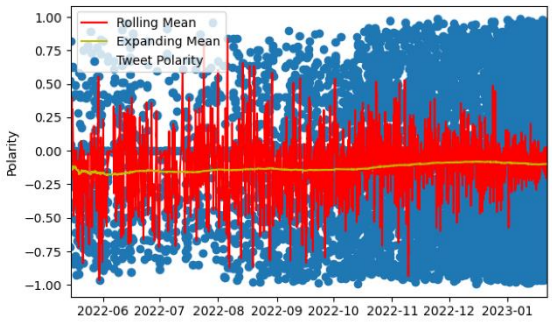


Figure 5. Tweet polarity histogram over time

Fig. 6 shows a histogram of the subjectivity of tweets using the visualization line provided by KDE. Subjectivity is shown on the x-axis with values from 0 to 1, where 0 is very objective, while 1 is very subjective. Thus, words like "great" and "excellent" have a high subjectivity rating, while words like "visit" and "day" have a low subjectivity rating. Low subjectivity indicates that tweets contain more factual information than personal opinions.

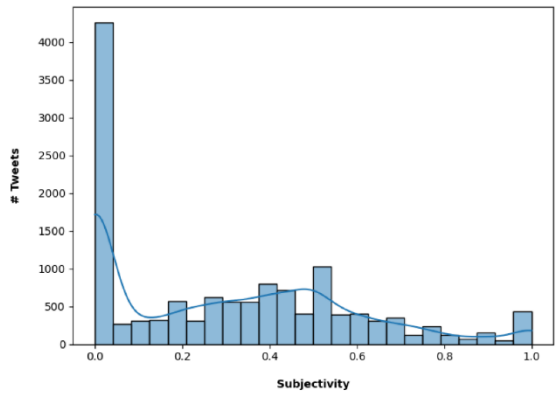


Figure 6. Tweet subjectivity histogram

Fig. 7 is presenting subjectivity over time to display correlation between months and subjectivity associated with each observed month. We can see that over time the subjectivity of tweets increases, which indicates that personal opinions and judgments appear more and more with the passage of time. Increased subjectivity shows a correlation with an increased number of posts, which points to influence operations aimed at triggering readers emotions, therefore each post contains more subjectivity on average.

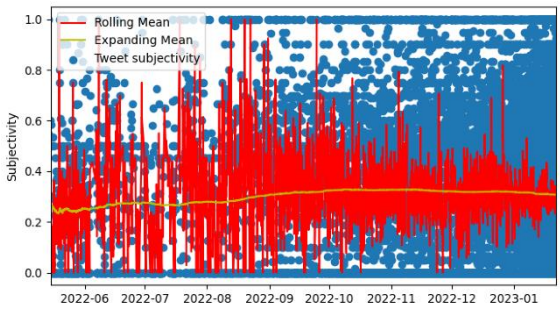


Figure 7. Subjectivity over time

V. SVM MALICIOUS USER CLASSIFICATION

Reference [26] states that SVM (Support Vector Machine) is a supervised machine learning algorithm that performs binary and multiway classification (pattern recognition) of the data into user defined categories. Support Vector Machines maps non-linearly separable training vectors in input space to linearly separable higher dimensional feature space and finds a separating hyper plane with maximal margin in that higher dimensional space. SVM is generally used for text categorization while also achieving good performance in high-dimensional feature space [27].

SVC (Support Vector Classification) implementation based on libsvm was used to classify malicious users. The fit time scales at least quadratically with the number of samples and may be impractical beyond tens of thousands of samples [28] with this limitation in mind, a limit of 20 thousand tweets was imposed for input into the classification model.

From each individual dataset used to create classifier 10,000 tweets were randomly selected, tagging each corresponding tweet with appropriate label, where 1 was used if it belongs to the Russian Troll Tweets dataset, while 0 was used for tweets belonging to Sentiment Dataset with 1 Million Tweets dataset. An additional filter on the Sentiment Dataset with 1 Million Tweets dataset was introduced where the language for selected tweets is specified as English since dataset also contains tweets in other languages such as: French, Spanish, Portuguese and Japanese.

Fig. 8 demonstrates steps that are involved in SVM classifier construction after tweets were appropriately labeled. These steps are separated in three categories: a) data preprocessing, b) model training and c) classification.

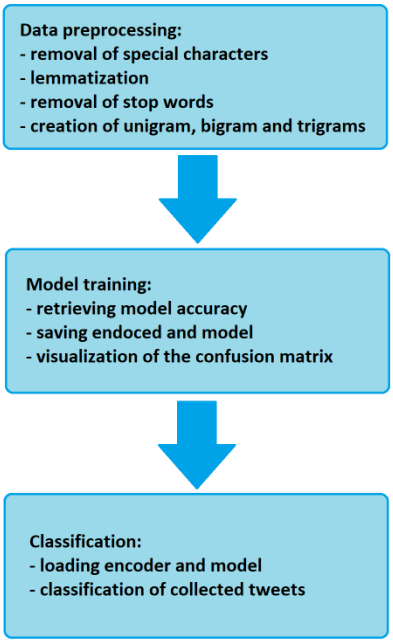


Figure 8. Visualization of SVM classifier construction

Data preprocessing is done in 4 separate steps:

a) Removal of special characters: is done by using regex r'[^a-zA-Z0-9\s]' to remove any special characters and to keep only numbers and alphabet letters inside each tweet.

b) Lemmatization: is achieved by implementation of lemmatization words using NLTK which provides WordNetLemmatizer class which is a slim cover wrapped around the wordnetCorpus. This class makes use of a function called Morphy() to the class to find a root word/lemma [29].

c) Removal of stop words: is done by providing parameter of stop_words to TfidfVectorizer class constructor with value of "english".

d) Creation of unigram, bigram and trigrams: are also provided as a constructor parameter to TfidfVectorizer ngram_range whose value is tuple (1,3).

By using TfidfVectorizer sklearn class collection tweets

TABLE III. OVERVIEW OF CLASSIFIED 13,381 TWEETS

Total	Malicous	Nonmalicous
13,381	3,586	9,795

a. Velika slova

were converted to a matrix of TF-IDF (term frequency–inverse document frequency) features together with removal of stop words while individual words were converted to ngrams. TF-IDF uses the frequency of words to determine how relevant those words are to a given document. The goal of TfidfVectorizer sklearn class is to give higher weightings to terms that appear often in a particular document, but not in many documents. If a word appears often in many of the documents, it is not a good feature for discriminating between classes. Likewise, if a word appears often in some documents and not in others, it is likely a good word for discriminating between classes [30].

TfidfVectorizer when used on 10,000 tweets, that were previously randomly selected and labeled with 1 (tweeets from Russian Troll Tweets dataset), produced vocabulary contains 143,698 words. Words such as: "rt", "trump", "http", "make", "just", "right", "sad", "need", "life", "politics"... showed high TF-IDF scores compared to other words inside this vocabulary.

Train and test split was made with respective 80:20 split, with train dataset containing 16,000 tweets, while test dataset contains 4,000 tweets. TfidfVectorizer produced vocabulary of 238,223 words when it was used on training dataset and 65,112 words when used on test dataset. Confusion matrix shown in Fig. 9 is used to describe the classification summary in a table. It can be interpreted as the summary of the predicted and actual data. As a result, for confusion matrix values tags were provided for: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) with their corresponding values [31]. Based on given values, an accuracy score of 93.95% is achieved, which is obtained by dividing the sum of TN and TP with the size of the test dataset. We can also calculate F1 score by using formula F1 = 2TP / (2TP + FP + FN), that will give us value of 0.9424.

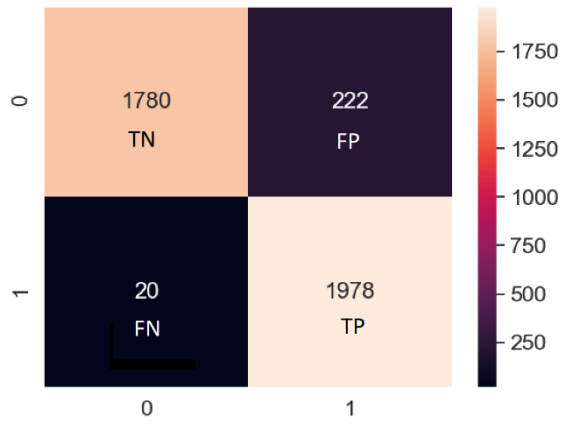


Figure 9. Confusion matrix based on SVC using TfidfVectorizer

In context of this paper malicious tweets are defined as a tweets that are part of Russian Troll Tweets dataset and any new tweets that are classified as such based on training of this dataset. By running the classifier on previously collected tweets, a classification is obtained with 1 for malicious and 0 for nonmalicious tweets, where 26.8% of tweets are classified as malicious tweets.

Creating a histogram Fig. 10 on the obtained data of the SVM classifier, it is possible to visualize the trend of the of malicious tweets, which become less and less malicious with the passage of time.

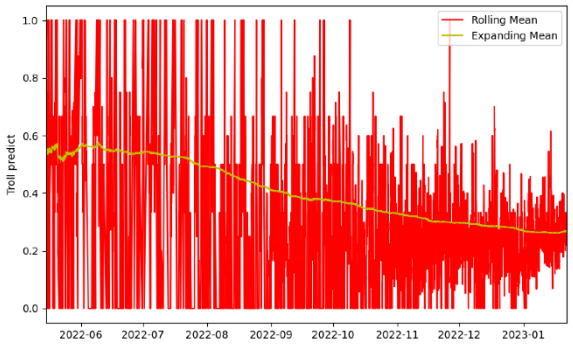


Figure 10. Malicious tweets classification over time

The movement of the trend where the classification of individual tweets as malicious is decreasing shows a clear focus on a controlled campaign that aims to avoid the classification of tweet as malicious with the change of tweet composition from previous campaigns on which the classifier was trained. Such behavior would be expected since the development of the model depends on previously classified malicious users, and by changing their behavior, that is, their posts, it is much more difficult to identify them without the existence of previously published datasets that contains classified malicious users.

VI. CONCLUSION

The Russian invasion of Ukraine on February 24, 2022 triggered a conflict of narratives and influence operations aimed at directing the behavior of readers towards pro-Russian sentiment through the intensity of posts. The growth in the number of tweets covered by pro-Russian milbloggers, especially at the end of the 2022, is related to the Kherson counteroffensive and the Kharkiv counteroffensive, during which a significant area that was previously under Russian occupation was liberated.

Aforementioned events weakened the credibility of the Russian military leadership in the decisions that were made, which affected the public opinion towards the war in Ukraine, as well as the morale of the Russian troops involved in the conflict. Both events caused a more significant start of Russian attacks on critical infrastructure that correspond to the narrative and preparation of Russia for a long war, and thus the preparation of the public through influence operations.

In addition to the increase in the volume of tweets as a stage of preparation for the narrative via Twitter, the approach to the very sentiment of published tweets is also changing, showing a slight increase in subjectivity and increasing polarity with the passage of time. In addition to

the growth of tweets, subjectivity and polarity, the classification of malicious users is in a more significant decline than the previously mentioned growth of other parameters. Which indicates that the new tweets are written much more carefully for the purpose of influence operations, and as such show a lower degree of detection, where the propaganda aspect of the tweets is reduced, but not completely removed which happens through an increased volume that has the possibility of reaching a larger number of users.

Based on this work, numerous possibilities for future work are opened up by monitoring the trend of posting of current users, as well as increasing the pool of Twitter profiles from which tweets are collected. Such research can also encompass sentiment analysis as well as LDA topic modeling to see if detected malicious users have influence on Ukraine, United States of America and European Union political process by changing and shaping their policies in regard to influence operations. This can be done by collecting tweets from members of Ukraine parliament (Verkhovna Rada of Ukraine), United States Congress as well as European Parliament members that in the same time period as influence operations were ongoing.

REFERENCES

[1] Haq E. U, Tyson G, Lee L. H, Braud T, Hui P. Twitter dataset for 2022 russo-ukrainian crisis, 2022, pp. 1-2.

[2] Ślufińska, M. "The Russia-Ukraine War", Information Security Policy, 2022, pp. 69-70.

[3] Nye J. S. “Soft power: the evolution of a concept”, Journal of Political Power, 2021, 14, pp. 202-203.

[4] Clark M, Barros G, Stepanenko K. “Russian Offensive Campaign Assessment”, Institute for the Study of War, 2022

[5] Hoskins A, Shchelin P. “The War Feed: Digital War in Plain Sight”, American Behavioral Scientist, 2022

[6] Kagan, F. W., Barros, G., & Stepanenko, K. “Russian offensive campaign assessment”, Institute for the Study of War, 2022, pp. 2-3.

[7] Rybar, available at: https://t.me/rybar, [accessed: 01.15.2023].

[8] Rybar in English, available at: https://twitter.com/rybar_en [accessed: 01.15.2023].

[9] Unmasking Russia’s influential pro-war ‘Rybar’ Telegram channel, available at: https://thebell.io/en/unmasking-russia-s-influential-pro-war-rybar-telegram-channel/ [accessed: 01.20.2023].

[10] The Bell, available at: https://twitter.com/thebell_io/status/1595106311133597696 [accessed: 01.20.2023].

[11] Tsakiris A. The Portrayal of the 2022 Russian Invasion of Ukraine onSocial Media, 2022, pp. 20.

[12] FIVETHIRTYEIGHT, Russian Troll Tweets, available at: https://www.kaggle.com/datasets/fivethirtyeight/russian-troll-tweets [accessed: 01.27.2023].

[13] TARIQ M., Sentiment Dataset with 1 Million Tweets, available at: https://www.kaggle.com/datasets/tariqsays/sentiment-dataset-with-1-million-tweets [accessed: 01.27.2023].

[14] CHARAP, S. and PRIEBE, M. Avoiding a Long War, 2023, pp. 6.

[15] Zmyvalova E. “The Rights of Indigenous Peoples of Russia after Partial Military Mobilization”, Arctic Review on Law and Politics, 2023, 14, pp. 71.

[16] Honey C. and Herring S. C. “Beyond microblogging: Conversation and collaboration via Twitter”, Hawaii International Conference on System Sciences, 2009, 42, pp. 2.

[17] Shekhawat B. S. Sentiment Classification of Current Public Opinion on BREXIT: Naïve Bayes Classifier Model vs Python’s TextBlob Approach (Doctoral dissertation, Dublin, National College of Ireland), 2019, pp. 2.

[18] Mejova Y. “Sentiment analysis: An overview”, University of Iowa, Computer Science Department, 2009, pp. 5.

[19] TextBlob: Simplified Text Processing, available at: https://textblob.readthedocs.io/en/dev/ [accessed: 01.25.2023].

[20] Falco X, Witten R, Zhou R. SENTIMENT AND OBJECTIVITY CLASSIFICATION, 2009, pp. 23.

[21] Madhu Fadhli I, Hlaoua L, Omri M. N. "Sentiment analysis CSAM model to discover pertinent conversations in twitter microblogs", I. J. Computer Network and Information Security, 2022, 5, pp. 34.

[22] Madhu S. “An approach to analyze suicidal tendency in blogs and tweets using Sentiment Analysis”, I. J. Scientific Research in Computer Science and Engineering, 2018, 6, pp. 36.

[23] Chen, Y.C. “A tutorial on kernel density estimation and recent advances”, Biostatistics & Epidemiology, 2017, 1, pp. 162.

[24] pandas.DataFrame.expanding, available at: https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.expanding.html [accessed: 03.14.2023].

[25] pandas.DataFrame.rolling, available at: https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.rolling.html [accessed: 03.14.2023].

[26] Polavarapu N, Navathe S. B, Ramnarayanan R, Haque A, Sahay S, Liu Y. “Investigation into biomedical literature classification using support vector machines”, Computational Systems Bioinformatics Conference, 2005, pp. 368.

[27] Bholane S. D. and Gore, D. “Sentiment analysis on twitter data using support vector machine”, International Journal of Computer Science Trends and Technology (IJCST), 2016, 4, pp. 368.

[28] C-Support Vector Classification, available at: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html [accessed: 01.27.2023].

[29] Khyani D, Siddhartha B. S, Niveditha N. M, Divya B. M. “An interpretation of lemmatization and stemming in natural language processing”, Journal of University of Shanghai for Science and Technology 2021, 22, pp. 355.

[30] Djuve K. and Burris, J. W. “A case study on the dialect identification of twitter tweets using natural language processing and machine learning”, Journal of Computing Sciences in Colleges, 2019, 34, pp.67.

[31] Adiba F. I, Islam T, Kaiser M. S, Mahmud M. and Rahman M. A. “Effect of corpora on classification of fake news using naive Bayes classifier”. International Journal of Automation, Artificial Intelligence and Machine Learning, 2020, 1, pp. 88.

Syllable and Morpheme Segmentation of Macedonian Language

M. Mitreska and K. Zdravkova
Faculty of Computer Science and Engineering, Skopje, Macedonia
maja.mitreska@students.finki.ukim.mk, katerina.zdravkova@finki.ukim.mk

***Abstract* - Communication is the key to human development. Approximately 5% of the world's population experience some form of hearing disability. Modern assistive devices and technologies can improve the communication skills of hearing impaired people by transcribing the speech into text. The creation of such an application depends on the language specific morphosyntactic properties. It usually starts with the syllabification. The research presented in this paper focuses on the development of an automatic system for rule-based and sonority-based syllable and morpheme segmentation of Macedonian language, which can be easily incorporated into an efficient speech recognition system. The segmentation rules for breaking the words down into syllables and into morphemes were created according to the new orthography of the Macedonian language. For the sonority-based approach, a novel phonological distance measure was introduced capable of efficient syllable clustering. The implementation of the framework is developed in Python using several data structures for optimized performance and CPU usage. Both segmentation strategies were evaluated using the electronic lexicon consisting of more than one million words. A linguistic expert was consulted during the entire process. The consistency of the obtained results promises their sustainability for further speech processing applications.**

Keywords - communication; speech recognition; hearing impairment, word segmentation

I. INTRODUCTION

Communication is a key to human development that plays a vital role in everyday activities. It enables social relations and facilitates the exchange of information and wisdom. Since its development about 100000 years ago, human language has facilitated interaction, supporting considerably the birth and progress of civilization [1].

According to the World Health Organization, over 5% of the world's population have some hearing disorder that impacts their unobstructed cognition, education and employment, causing an annual global cost of around one trillion US$ [2]. The projection is that by 2050, nearly 2.5 billion people will have some kind of hearing loss [2].

Assistive technologies for hearing impaired people are categorized as assistive listening devices (ALD), augmentative and alternative communication (AAC) and alert systems (AS) [3]. ALD either amplifies the sound (for example, hearing loops and personal amplified systems), or wirelessly transmit it (FM, infrared and Bluetooth systems) [4].

AAC is intended for people with speech, language or communication disorder [5]. Powered by machine learning (ML) and deep learning (DL), these NLP based applications embrace word prediction, speech recognition and context processing [6]. AS encompasses notification systems, alert messaging and emergency communication systems intended to alert and protect people from risks and threats [7].

A team of young experts and researchers from the AI innovative company iReason (https://ireason.mk/) intensively work on the creation of assistive software modules for students with communication disorders [8]. One of their main goals are voice bot technologies, which rely on speech processing. To support the transcription of the Macedonian spoken language into text, we focused our research on syllable and morpheme segmentation. Syllable segmentation enables the discovery of the phonological structure of words [9]. The accuracy of speech recognition can additionally be improved by morpheme segmentation, which increases "the capability of understanding and producing new word forms" [10]. According to Yang et al. [11], the following four speech processing tasks: phoneme recognition (PR), automatic speech recognition (ASR), keyword spotting (KS), and query by example spoken term detection (QbE) are crucial to establish a universal performance benchmark. The first two are used to transcribe speech into text, the latter to detect the spoken content [11].

The paper presents the recently developed automatic system for syllable and morpheme segmentation of Macedonian language. In the absence of a larger available corpus of already segmented words, the realization of both tasks was rule-based. The segmentation rules, prefixes and suffixes were extracted from new Macedonian orthography [12]. The sonority-based rules are an original contribution of this paper. The main language resource was the annotated electronic lexicon [13], which is available from the CLASSLA CLARIN knowledge centre for South Slavic languages [14].

The paper continues with the second section, which announces the linguistic background. Third and fourth sections introduce rule-based and sonority-based syllable and morpheme segmentation. Fifth section presents the application development. Sixth section is dedicated to evaluation of the results and identification of the major anomalies of both segmentations. The paper concludes with the further improvements and the impact of word segmentation to speech recognition of Macedonian.

II. LINGUISTIC BACKGROUND

Phonetically, syllables are sequences of sounds "containing one peak of prominence" [15]. Phonologically, they are units of accent placement [15]. Syllables must contain one vowel, which is the voiced, central-oral frictionless sound [18]. For example, the words *apple*, *butterfly*, *frog*, *table*, and *tomato* are divided into the following syllables: *ap-ple*, *but-ter-fly*, *frog*, *tab-le*, and *to-ma-to*, each containing exactly one vowel.

Morphemes are defined as 'the minimal meaning-bearing units of a language' [16]. They express the internal structure of complex words [17]. Morphemes are composed of two separate classes: bases (according to some linguists, roots) and affixes [16]. Those affixes that precede the base are prefixes, those that follow it are suffixes. For example, the word *unhealthy* consists of the prefix *un*, the base *health* and the suffix *y*.

A. Syllable segmentation

Syllable segmentation, also called syllabification or syllabication can be performed implementing a data-driven or a rule-based approach [18]. Data-driven methodology requires a large corpus of already syllabified words [18], [19]. By implementing various machine learning techniques the accuracy can reach 95% [19]. The absence of such a corpus can be bypassed by implementing the rule-based approach. Instead of learning the syllabification from the training set, a concise set of mutually exclusive rules should be established, enabling the correct segmentation [18]. The accuracy is again very high, proving that rule-based segmentation is not inferior to data-driven [20].

B. Morpheme segmentation

Morpheme segmentation methods are usually classified as rule-based, data-driven and hybrid [21]. Rule-based methods detect morpheme boundaries statistically, using for example, letter variety statistics [22] or conditional random fields [23]. Data-based methods can be supervised, semi-supervised and unsupervised [21]. A very successful supervised morpheme segmentation applied convolutional neural networks [24]. Semi-supervised learning that combines several unsupervised segmentation techniques using conditional random fields proved its efficiency for Finish [25]. A good example of a successful unsupervised segmentation was achieved with Bi-LSTM neural network [26]. Hybrid methods can combine rule-based and statistical approaches coupled with unknown morpheme guessing [27] or rule-based morpheme word representation coupled with unsupervised morphological analysis [28].

C. Macedonian language specific phonetic features

The Macedonian language is a phonetic language and the relationship between spoken sounds (phonemes) and written sounds (graphemes) is very strong [29]. Therefore, the linguistic rules of forming syllables, as well as morphemes are very strict [30].

Macedonian language has five vowels: *а* (*a*), *е* (*e*), *и* (*i*), *о* (*o*) and *у* (*u*) and no diphtongs. Most of the words, like: *вода* (Latinized: *voz* / English: *train*), *вода* (*voda* / *water*), *воздушест* (*vozdushest* / *airy*) and *вообразен* (*voobrazen* / *conceinted*) are syllabized with a sequence of one up to five phonemes, one of them compulsory a vowel: *воз*, *во-да* and *воз-ду-шест* and *во-о-бра-зен*.

The sonorant *r* (in Cyrillic, *р*) can also be a syllable-carrier [31] if found in one of the following contexts:

- It appears in the middle of a consonant group. Such words are: *крв* (*krv* / *blood*), *првак* (*prvak* / *champion*), *здрвен* (*zdrven* / *stiff*);
- It is preceded by an apostrophe, which replaces the hard sign existing in some Macedonian dialects and is followed by a consonant at the beginning of the word, such as: *'рбет* (*'rbet* / *spine*), *'ржс* (*'rzh* / *rye*), *'ртење* (*'rtenje* / *germination*);
- It follows a consonant at the end of the word. This situation occurs in a few words of foreign origin, like *масакр* (*masakr* / *masacre*) and *жанр* (*genre*) [12].

The list of prefixes is rather long. They can consist of one syllable only: *без* (*bez*), *ис* (*is*), *по* (*po*), and *се* (*se*), or several syllables: *кусо* (*kuso*), *обез* (*obez*), *прет* (*pret*), and *сино* (*sino*). The list of suffixes is even longer, because the language is highly inflected [32]. They depend on the POS tag. For morpheme segmentation, the suffixes *ски* (*ski*), *ство* (*stvo*) and *ствен* (*stven*), which are nominal or adjectival trigger the corresponding inflections for gender, number and definiteness [33].

Macedonian language is a low-resourced language and it is lacking the manually segmented lexicon. Therefore, the only applicable methods implemented during our research could be rule-based. They comprise two flows: segmentation according to the rules extracted from the orthography and segmentation according to sonority of the phonemes. Both flows were independently done aiming to perform syllable and morpheme segmentations. The following two sections introduce them briefly.

III. RULE-BASED SEGMENTATION

The syllable segmentation was performed by exploiting the phonological distances measures [18] and by implementing a finite set of segmentation rules for the Macedonian [12]. Morpheme segmentation was done implementing the linguistic knowledge originating from the Macedonian orthography [12] in conjunction with the lemmas from the lexicon [13].

A. Macedonian syllabification rules

These are the most simplified rules that were incorporated in the rule-based syllable segmentation of Macedonian language.

1. Each syllable should contain exactly one vowel or the syllable-carrier sonorant *р* (*r*).
2. A syllable can consist of one vowel only independently on its position in the word: *авион* (*avion* / *plane*; *а-ви-он*), *аурора* (*aurora* / *aurora*, *а-у-ро-ра*, *меана* (*meana* / *tavern*; *ме-а-на*) and *тргнаа* (*trgnaa* / *started*; *трг-на-а*).

3. If a consonant group appears at the beginning of the word, then the syllable consists of all the consonants and the vowel or the embedded sonorant *p* (*r*) it is followed by: *здравство* (*здравство* / *health*; *здрав-ство*), *според* (*spored* / *according*; *спо-ред*), *крвав* (*krvav* / *bloody*; *кр-вав*).

4. If a consonant group appears at the end of the word, then the syllable consists of all the consonants and the vowel or the sonorant *p* (*r*) that precede it: *амбиент* (*ambient* / *ambience*; *ам-би-ент*), *радост* (*radost* / *happiness*; *ра-дост*), *накрст* (*nakrst* / *crosswise*; *на-крст*).

5. If a consonant group appears in the middle of the word, then it is divided in half, with the beginning belonging to the first and the end to the next syllable: *коска* (*koska* / *bone*; *кос-ка*), *тетратка* (*tetratka* / *notebook*; *тет-рат-ка*, *поттик* (*pottik* /*motive*; *пот-тик*). Longer consonant groups, like in *исклучи* (*iskluchi* / *turn off*; *ис-клу-чи)* or *оздрави* (*ozdravi* / *recover*) should be treated separately to reduce incorrect segmentation caused by inconsistent division.

6. An exception to the rule 5 are the nouns that end in *ство* (*stvo*) and the adjectives that end in *ски* (*ski*) and *ствен* (*stven*). These three suffixes always remain undivided: *чувство* (*чувство* / *emotion*; *чув-ство*), *градски* (*gradski* / *urban*; *град-ски*), *единствен* (*edinstven* / *unique*; *един-ствен*). The inflections for: gender, plural, definite, distal and proximal definiteness preserve the rule 6.

7. The following suffixes: *штво* (*shtvo*), *шки* (*shki*), and *чки* (*chki*) can be divided according to the rule 5 or can remain undivided obeying the rule 6.

B. *Macedonian morphemes*

Morphemes in the Macedonian language are crucial for word formation and for inflection [12]. For example, the noun *учител* (*uchitel* / *teacher*) consists of two morphemes: the verb *учи* (*uchi* / *to teach*) and the suffix *тел* (*tel*). The noun *надградба* (*dogradba* / *annex*, *extension*) consists of three morphemes: the prefix *над* (*nad*), the verb *гради* (*to build*) and the suffix *ба* (*ba*).

Morpheme segmentation starts with the prefixes. A specific attention was paid to prefixes consisting of one vowel only, like the vowel *и* (*i*) in the adjective *иреален* (*irealen* / *unreal*; *и-ре-а-лен*).

Considering the rich inflectional paradigm of many Macedonian word categories [33] suffixes were searched throughout all of the word. While the suffix *ствен* (*stven*) remains unchanged during inflection, *ски* (*ski*) and *ство* (*stvo*) alter the final vowel for gender and number.

To support the speech recognition, all the prefixes, word bases and suffixes consisting of more than one syllable were additionally syllabized. This is not part of the morpheme segmentation, but it can significantly facilitate the recognition of longer words, particularly those that are part of an accentual whole.

C. *Phonological distance for syllable segmentation*

In 1965, Levenshtein defined phonological distances aiming to correct the mistakes of binary information caused by reversing, losing or adding of some binary value during transmission [34]. Levenshtein's distance inspired the development of two additional measures: the dialect measurement [35] and the measure for language classification [36].

In 1988, Clements introduced the sonority sequencing principle (SSP), according to which the sonority within a syllable rises to the syllable nucleus (in Macedonian, the vowels and the sonorant *p* (*r*)) and fall in sonority thereafter [37]. Although disputed, this principle was applicable to Macedonian language. SSP was combined with the phonological distances in the novel sonority-based segmentation method, which was implemented for both tasks: syllable and morpheme segmentation. The following section introduces the new method in more detail.

IV. SONORITY-BASED SEGMENTATION

The sonority of the Macedonian phonemes is highest for vowels. It is falling down from sonorants towards voiced and voiceless consonants (see Table 1). The list is expanded with the special delimiters S and F.

TABLE I. SONORITY OF MACEDONIAN PHONEMES

Type	**Macedonian phonemes**	
	List of phonemes	***Weight***
Vowels	*а* (*a*), *е* (*e*), *и* (*i*), *о* (*o*), *у* (*u*)	12
Sonorant *p*	*р* (*r*)	5
Sonorants	*ј* (*j*), *л* (*l*), *љ* (*lj*), *м* (*m*), *н* (*n*), *њ* (*nj*)	3
Voiced consonants	*б* (*b*), *в* (*v*), *г* (*g*), *д* (*d*), *ѓ* (*gj*), *ж* (*zh*), *з* (*z*), *ѕ* (*dz*), *џ* (*dzh*)	2
Voiceless consonants	*к* (*k*), *п* (*p*), *с* (*s*), *т* (*t*), *ќ* (*kj*), *ф* (*f*), *х* (*h*), *ц* (*c*), *ч* (*ch*), *ш* (*sh*)	1
Special	‘, spacing (S), fictive consonant (F)	0

The syllabification starts with the preprocessing of the words, which adds a fictive consonant F to phonologically separate the consecutive vowels (Fig. 1).

phonemmes	S	И	Д	Е	F	А	Л	Н	О	S
phonemme sonority	0	12	2	12	0	12	3	3	12	0
triplet difference		10	-22	10	-24	9	-12	-12	9	
		syllable clusters:			4					

phonemmes	S	И	Д	Е	F	А	Л	Н	О	S
phonemme sonority	0	12	2	12	0	12	3	3	12	0
triplet difference		10	-22	10	-24	9	-12	-12	9	
		morpheme clusters:			2					

Figure 1. Segmentation of a word with two consecutive vowel

If the phonemes are denoted with the letter *p*, and their sonority weights with *w*, then the word can be represented with the string:

$$w(p_i),\ i = 1,\ldots,n \qquad (1)$$

n being the number of phonemes and fictive consonants.

The sonority of the spacing S before $w(p_0)$ and after the word $w(p_{n+1})$ is 0. For each phoneme, a triplet difference *TD* is calculated as:

$$TD(p_i)=w(p_i) - w(p_{i-1}) - w(p_{i+1}),\ i = 1,\ldots,n \qquad (2)$$

Triplet difference of the vowels is always positive (Fig. 1 and 2). This property was achieved by denoting the weights 12 to vowels, which guarantee that the difference is bigger than 2, even when the vowel is surrounded by two sonorants, like in the word *anapam* (*aparat* / *device*).

Morpheme segmentation was done by going through all the prefixes, the potential bases matched with the lemmas corresponding to nouns and verbs and finally the suffixes extended with their inflections. To support speech recognition, they were additionally syllabized according to both syllable segmentation methods.

Following SSP, the sonority of the phonemes within a syllable is monotonically increasing towards the syllable nucleus. Syllable border is the phoneme after which the strictly monotonic decrease of the sonority ends (Fig. 1 and 2, highlighted with apricot). According to this rule, the adjective *идеално* (*idealno* / *ideal*) has four syllables: *и* (*i*), *де* (*de*), *ал* (*al*) and *но* (*no*). Its morpheme segmentation consists of the base *идеал* (*ideal*) and the suffix *но* (*no*). Syllabized morpheme segmentation of this word is identical with the syllable segmentation. Syllable and morpheme segmentation of two-syllable words with larger consonant groups like *здравствен* (*zdravstven* / *health*) usually differ. While the sonority-based rule suggests the syllables *здравс* (*zdravs)* and *твен* (*tven*), Macedonian orthography and morpheme segmentation are identical: *здрав* (*zdrav / healthy*) and *ствен* (*stven*).

phonemmes	S	З	Д	Р	А	В	С	Т	В	Е	Н	S
phonemme sonority	0	2	2	5	12	2	1	1	2	12	3	0
phonemme sonority		0	-5	-9	5	-11	-2	-2	-11	7	-9	
		syllable clusters:			2							

phonemmes	S	З	Д	Р	А	В	С	Т	В	Е	Н	S
phonemme sonority	0	2	2	5	12	2	1	1	2	12	3	0
triplet difference		0	-5	-9	5	-11	-2	-2	-11	7	-9	
		morpheme clusters:			2							

Figure 2. Segmentation of a word with longer consonant groups

V. APPLICATION DEVELOPMENT

The proposed segmentations were developed in Python for simple integration and further use in various systems, including its embedding into automation tasks.

The framework consists of two main parts, rule-based and sonority-based syllable and morpheme segmentation. Its inter-process communication is illustrated with the pipeline presented in the Fig. 3.

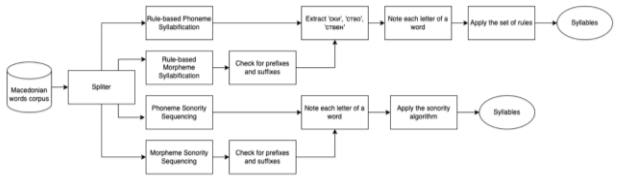


Figure 3. A pipeline of splitting words into their syllables.

Before the segmentation, the dataset was clustered into ten subsets that consist of words with different phoneme structuring. The division of the lexicon was predominantly made to facilitate the evaluation of both segmentations in detail, with the intention of discovering where the new sonority-based approach yields different results than the rule-based approach.

VI. EVALUATION OF BOTH APPROACHES

Since there is no previous work done in this field, and there are no existing records that contain Macedonian words and their respective syllable segmentation for comparison, an annotated lexicon consisting of 1236537 Macedonian words [13], which was run through the four types of segmentation. For evaluation purposes, the two phoneme-based and two morpheme-based approaches were compared with each other aiming to measure how much the segmentations from one approach matches or differs from the other approach. This evaluation method was considered relevant since one of the approaches is based on the rules of syllabification in the Macedonian language and each output that this method obtains is considered a true match. So, the evaluation was made by matching the outputs of the phoneme rule-based segmentation as a reference output and the outputs from the phoneme sonority-based segmentation.

Table II represents the matches between the two implemented approaches for the 10 clusters. The last one is further divided depending on the suffixes that have a special treatment in Macedonian orthography.

TABLE II. EVALUATION OF SEGMENTATION ACCURACY BY MATCHING BOTH APPROACHES

ID	***Cluster type***	***Number of words***	***Frequency***	***Phoneme***	***Morpheme***
1	Alternating vowels and consonants	154476	12.49%	1.000	0.973
2	Two-phoneme consonant group	428537	34.66%	0.992	0.944
3	Three-phoneme consonant group	178426	14.43%	0.628	0.746
4	Four-phoneme consonant group	43860	3.55%	0.774	0.825
5	Sonorant '*р*'	73803	5.97%	0.971	0.927
6	Two consonant groups	380489	30.77%	0.889	0.883
7	Several consonant groups	128129	10.36%	0.857	0.871
8	One vowel group	241637	19.54%	0.914	0.934
9	Two vowel groups	21379	1.73%	0.910	0.956
10.1	Words with '*ски*'	21405	1.73%	0.187	0.957
10.2	Words with '*ство*'	2511	0.20%	0.089	0.937
10.3	Words with '*ствен*'	2273	0.18%	0.000	0.969

The accuracy of the sonority-based approach presents an effective approach that does not rely on a set of fixed rules. More noticeable mistakes can be detected in the third and fourth subsets where the consonant group consists of three or four consonants. Namely, the sonority-based approach usually shifts one of the consonants into the previous syllable, particularly when the sonority of the phonemes is identical. One such word is *авторскана* (*avtorskana* / *that author's*). The phoneme rule-based approach outputs *ав-тор-ска-на*, while the phoneme sonority-based approach outputs *ав-торс-ка-на*.

The same problem can be also noticed in the tenth subset where the segmentation performed with the phoneme approach has low accuracy, however when the segmentation takes into consideration the morphemes, the accuracy significantly increases. This is due to the orthographic rule that the phonemes in the suffixes *ски* (*ski*), *ство* (*stvo*) and *ствен* (*stven*) are never separated, although according to their sonority they must be separated after the phoneme *с* (*s*). A typical example of this case is the word *армиски* (*armiski* / *army*), where the phoneme rule-based approach gives *ар-ми-ски* (*ar-mi-ski*), while the sonority-based approach results in *ар-мис-ки* (*ar-mis-ki*).

The higher accuracy in the morpheme approach is due to the general extraction of the prefixes and suffixes which are most of the time one-syllable morphemes.

An additional interesting case are the words starting in a voiced consonant that are followed by a voiceless consonant. This is quite rare, because in the Macedonian language sonority equalization is usually applied, according to which when two consonants with different sonorities are next to each other, the voiced consonant is transformed into the corresponding voiceless pair [13]. There are some exceptions to this rule that validate the rule-based approach, but may show errors in the sonority-based approach because the segmentation is done by considering and weighing all the letters of the word. The most interesting example of this are the words *вторник* (*vtornik* / *Tuesday*) and *вчера* (*vchera* / *yesterday*), where the rule-based approach produces the segmentations *втор-ник* (*vtor-nik*) and *вче-ра* (*vche-ra*), while the sonority-based segmentations results *в-тор-ник* (*v-tor-nik*) and *в-че-ра* (*v-che-ra*), where the evaluation of syllable number fails. The key reason for this inaccuracy is the fact that the phoneme *в* (*v*) is not a syllable carrier.

Regarding morpheme segmentation, the two approaches give similar results since they pay attention to the detection of prefixes and suffixes, and the difference that is detected is in the additional syllabic definition of the morpheme.

VII. CONCLUSION AND FURTHER WORK

The paper presents two seemingly trivial tasks. They were created using rules and an original sonority-based approach. Both segmentations were exhaustively manually evaluated with hundreds of words automatically segmented and evaluated by the authors and an expert.

The comparison of the results is more than promising. Reaching a segmentation match of 90.12% during syllabification of already morphologically divided words, it will undoubtedly become a valuable contribution to various upcoming language technologies.

A detailed evaluation of the implementation of lexical clustering proves the superiority of syllabized morpheme segmentation compared to the traditional rule-based syllabification.

The authors believe that after fixing the perceived inaccuracies the sonority-based syllabification of morpheme segmentation can be considered the most adequate for word segmentation in the Macedonian language and used for hyphenation within the text processor.

Even without any adjustments, the sonority-based segmentation is more than sufficient to support the Macedonian language speech recognition system that is under construction. Its simplicity will contribute to the efficient recognition of oral language and its translation into text, which is a key component of mobile applications for the hearing impaired. By using such an assistive technology, they will be able to establish uninterrupted two-way communication with people who do not know the Macedonian sign language.

ACKNOWLEDGMENT

Sincere gratitude to Dr. Elka Jačeva Ulčar from the Institute for Macedonian Language, who contributed to the development of the system with her expert advice. This work is partially financed by the Faculty of Computer Science and Engineering (FSCE) at the Ss. Cyril and Methodius University in Skopje.

REFERENCES

[1] WHO, Deafness and hearing loss, retrieved from https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss on 5 January 2023.

[2] S. Miyagawa, S. Ojima, R. C. Berwick, and K. Okanoya. "The integration hypothesis of human language evolution and the nature of contemporary languages." Frontiers in psychology 5, 2014, p. 564.

[3] A. Dhanjal, and W. Singh. "Tools and techniques of assistive technology for hearing impaired people." International conference on machine learning, big data, cloud and parallel computing (COMITCon), pp. 205-210. IEEE, 2019.

[4] C. Nieman, and E. S. Oh. "Hearing loss." Annals of internal medicine 173, no. 11, (2020, ITC81-ITC96.

[5] D. Beukelman, and P. Mirenda. Augmentative and alternative communication. Baltimore: Paul H. Brookes, 1998.

[6] Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor. "Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability." Sensors 19, no. 8, 2019, p. 1911.

[7] K. Zdravkova, V. Krasniqi, F. Dalipi, and M. Ferati. "Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective." Frontiers in Artificial Intelligence, 2022, p. 240.

[8] K. Mishev, A. Karovska Ristovska, O. Rashikj-Canevska, and M. Simjanoska. "Assistive e-Learning Software Modules to Aid Education Process of Students with Visual and Hearing Impairment: A Case Study in North Macedonia.", ICT Innovations 2021, pp. 145-159. Springer, Cham, 2022.

[9] C. Ewen and H. Van Der Hulst: The phonological structure of words: an introduction. Cambridge University Press, 2001.

[10] M/ Creutz: Induction of the morphology of natural language: Unsupervised morpheme segmentation with application to automatic speech recognition. Helsinki University of Technology, 2006.

[11] S. Yang, P. Chi, Y.Chuang, C. Lai, K. Lakhotia, Y. Lin, A. Liu et al. "Superb: Speech processing universal performance benchmark." arXiv preprint arXiv:2105.01051, 2021.

[12] Institute of Macedonian Language "Krste Misirkov", Orthography of the Macedonian language, second edtion, Kultura, 2017, https://pravopis.mk/sites/default/files/Pravopis-2017.PDF

[13] Petrovski, A. Morphological computer dictionary –contribution to Macedonian language resources, PhD thesis, Ss. Cyril and Methodius University in Skopje, 2008, in Macedonian

[14] CLASSLA: Knowledge centre for South Slavic languages, FAQ for Macedonian language resources and technologies, https://www.clarin.si/info/k-centre/faq4macedonian/

[15] J. O'Connor and J. Trim. "Vowel, consonant, and syllable—

[16] J. Krámský: "The word as a linguistic unit." In The word as a linguistic unit. De Gruyter Mouton, 2019.

[17] A. Kovač, and M. Marković: A Rule-Based Syllabifier for Serbian, Proceedings of the Conference on Language Technologies and Digital Humanities, pp. 140-46. 2018.

[18] N. Sanders and S. Chin: Phonological distance measures, Journal of quantitative linguistics 16, no. 1, pp 96-114, 2009.

[19] Y. Marchand, C. Adsett and R. Damper.: Automatic syllabification in English: A comparison of different algorithms." Language and speech 52, no. 1, pp 1-27, 2009.

[20] F. Asahiah: Comparison of rule-based and data-driven approaches for syllabification of simple syllable languages and the effect of orthography." Computer Speech & Language 70: 101233. 2021.

[21] T. Zundi and C. Avaajargal: Word-level morpheme segmentation using transformer neural network, Proceedings of the 19th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, pp. 139-143. 2022.

[22] Ç. Çöltekin: Improving successor variety for morphological segmentation., LOT Occasional Series 16, pp. 13-28. 2010.

[23] H. Liu, M. Li, J. Zhang and L. Chen. :Morpheme segmentation using bilingual features, 2012 International Conference on Asian Language Processing, pp. 209-212. IEEE, 2012.

[24] A. Sorokin and A. Kravtsova: Deep convolutional networks for supervised morpheme segmentation of Russian language, Conference on Artificial Intelligence and Natural Language, pp. 3-10. Springer, Cham, 2018.

[25] T. Ruokolainen, O. Kohonen, S. Virpioja and M. Kurimo: Painless semi-supervised morphological segmentation using conditional random fields., Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, volume 2: Short Papers, pp. 84-89. 2014.

[26] E. Bolshakova and A. Sapin: Bi-LSTM Model for Morpheme Segmentation of Russian Words. Conference on Artificial Intelligence and Natural Language, pp. 151-160. Springer, Cham, 2019.

[27] G. Lee, J. Cha and J. Lee: Syllable-pattern-based unknown-morpheme segmentation and estimation for hybrid part-of-speech tagging of Korean." Computational Linguistics 28, no. 1, pp. 53-70, 2002.

[28] M. Luong, P. Nakov and M. Kan: A hybrid morpheme-word representation for machine translation of morphologically rich languages., arXiv preprint arXiv:1911.08117. 2019.

[29] D. Spasovski, G, Peshanski and G, Madjarov: The influence the training set size has on the performance of a digit speech recognition system in macedonian." In International Conference on ICT Innovations, pp. 205-212. Springer, Cham, 2014.

[30] W. Ghai and N. Singh: Literature review on automatic speech recognition. International Journal of Computer Applications 41, no. 8, 2012.

[31] B. Morén: Consonant–vowel interactions in Serbian: Features, representations and constraint interactions, Lingua 116, no. 8, :pp. 1198-1244, 2006.

[32] K. Zdravkova: Resolving Inflectional Ambiguity of Macedonian Adjectives, LREC 2022 Workshop Language Resources and Evaluation Conference, p. 60. 2022.

[33] M. Bonchanoski and K. Zdravkova: Learning syntactic tagging of Macedonian language, Computer Science and Information Systems 15, no. 3, pp. 799-820, 2018.

[34] V. Levenshtein: Binary codes capable of correcting deletions, insertions, and reversals., Soviet physics doklady, vol. 10, no. 8, pp. 707-710. 1966.

[35] J. Nerbonne and W. Heeringa: Measuring dialect distance phonetically." In Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology. 1997.

[36] T. Dunning: Statistical identification of language. Las Cruces: Computing Research Laboratory, New Mexico State University, 1994.

[37] G. Clements: The sonority cycle and syllable organization, Phonologica, 63-76, 1988.

Multilingual Named Entity Recognition Solution for Optimizing Parcel Delivery in Online Commerce: Identifying Person and Organization Names

M. Pajas*, A. Radovan†, I. Ogrizek Biškupić‡
* BISS, Zagreb, Croatia
† BISS, Zagreb, Croatia
‡ Algebra University College, Zagreb, Croatia
matija.pajas@biss.hr, aleksander.radovan@biss.hr, ivana.ogrizekbiskupic@algebra.hr

***Abstract*—This paper presents a comprehensive solution to enhance parcel delivery in online commerce by implementing multilingual named entity recognition. The solution is designed to accurately identify person and organization names, with a primary emphasis on correctly identifying recipients. The ultimate goal is to use this information to automatically validate recipients and select the most accurate one to improve data accuracy and reliability for parcel delivery. The process begins by collecting a large dataset of online commerce data, including customer search queries, and annotating it with person and organization names. The data is then preprocessed, cleaned to eliminate irrelevant information, and prepared for training a named entity recognition model. Next, the model is trained and evaluated using this data to ensure its ability to identify named entities and extract recipients from queries accurately. The process employs an iterative training process and data generation techniques, while also addressing the issue of noisy data and iterative training introducing unwanted patterns by retraining the model on the subset of the original annotated dataset. Our experiments conclude a consistent increase of $F1$ score over the baseline and best iteration using this method of training and fine-tuning.**

***Keywords*—named entity recognition, parcel delivery, data reliability, person and organization names, multilingual, natural language processing**

I. INTRODUCTION

The increasing demand for online commerce has led to a need for dependable and efficient parcel delivery services. One way of reducing delivery expenses is to keep customer data as up-to-date and accurate as possible. However, there is a lot of noise in our database, especially in fields with recipient names that could be either personal or corporate, which comes from a vast number of accessible user inputs from various online shops, so the records and fields are not standardized. The noise appears in the form of misspelled words, irrelevant recipient information, and even multiple recipient names in a single record, making it challenging to match recipients with their correct addresses, leading to many duplicate records. To address it, this paper proposes a solution that leverages the latest advancements in natural language processing and machine learning to enhance parcel delivery in online commerce. The solution involves implementing a multilingual named entity recognition (NER) system that can accurately identify person and organization names, focusing on correctly identifying the intended recipient. The experiments and model selection for our named entity recognition task was performed using the FLAIR framework [1]. The FLAIR framework utilizes its proposed contextual string embeddings, known as *Flair Embeddings* [2] in the framework, which have proven to be highly effective for our downstream NER task. Additionally, self-supervised learning techniques [3], [4] and weak supervision were employed using Snorkel [5] to generate the training data as well as handwritten rules, and heuristics to generate additional synthetic data.

A. Recent research and challenges in real-world data

Named entity recognition involves identifying named entities, such as organizations, places, or individuals, from a large text corpus to extract knowledge and information automatically. Since early NER research papers were published using carefully handcrafted rules, such as Grishman and Sundheim [6] and Rau [7], the field has gained significant popularity. Collobert and Weston [8] introduced the first neural network architecture for NER, and since then, various neural network-based approaches have been proposed and published. The most significant development is the introduction of word and character embeddings used for word representation, improving sequence tagging tasks. This technique is now widely used in almost all NER studies [9]. Academic research in natural language processing (NLP) often concentrates on developing and improving recent neural network architectures, intending to slightly increase F1 scores on standardized datasets such as [10]. Although these efforts are crucial for advancing the scientific understanding of NER, they need to pay more attention to the practical application of these solutions on real-world data. Standardized datasets used for evaluation are often a good sign of a well-performing NER model and provide a common benchmark. Still, they need to reflect the complexity and diversity of real-world data that is often noisier and may not consider domain-specific knowledge

that is critical for effective NER. This can result in NER models that perform well on standardized datasets but are unsuited for use in real-world data without adaptations.

The patterns that emerge in the data can vary greatly depending on the dataset being modeled. Models that performed great for some datasets, don't necessarily perform well on arbitrary datasets without necessary data processing or adaptation of model parameters. For instance, in the case of parcel delivery, if the data contains a large number of misspelled names or irrelevant information, it becomes difficult to extract the correct recipient name in automated processes, resulting in a higher number of failed deliveries and increased costs. On the other hand, if the data is well-structured and standardized, it becomes much easier to accurately identify and match recipients with their addresses, reducing the number of failed deliveries and lowering costs.

More often than not, NER tasks deal with unstructured and non-standardized text to extract useful information. Hence, it is essential to understand the structure and patterns of the data to build and train models effectively. Cheng et al. [11] proposed an end-to-end solution called *TripleLearn* to address a similar issue in their NER task of extracting brand and product types, which takes advantage of three different datasets: Golden training data (manually annotated data considered the best and unbiased source of labels), Noisy training data (generated from noisy customer data and product catalog), and Synthetic training data (generated using heuristics and rule-based generation). The authors also identify several related papers that use neural networks and a large amount of unlabeled data for commercial search engines [12]–[15]. They also demonstrate that the F1 score improves through iterative training. The exact details of how TripleLearn works and how it achieves improved performance are described in the paper. We extend the idea by incorporating the FLAIR, and Snorkel frameworks [1], [5] and tackle the problem of the iterative training process introducing new patterns that may not be relevant to the original dataset by retraining the best iteration on a subset of the original Golden Dataset, aiming to restore model's predictive ability on the original dataset. Using this method helps to restore knowledge about the initial problem while also retaining the insights gained from subsequent training iterations, thus leading to improved F1 scores and model accuracy. On GitHub [16], you can discover the demonstration version of the model as well as the outcome of the evaluation on a held-out test dataset.

II. Iterative training

A. Training dataset

Our final training dataset consists of various data sources, with the Recipient Delivery Database containing the primary dataset as depicted in Fig. 1. We focus on recipient names, including names in various languages, the most prominent being German, French, and Italian. The initial and most accurate Golden Dataset was created by sampling approximately 3000 data points from the database and having them annotated by domain experts. In addition to personal names, recipient names also include company names that are named in their respective language depending on the region. It is common to find person and company names in single data records, often combined with unnecessary information like phone numbers or addresses that were meant to be saved in other database fields. Moreover, the records often contain random junks or encoding errors that can adversely affect performance. Our goal is to extract entities of interest, person (PER) labels and organization (ORG) labels, in order to keep the fields containing recipient names as clean as possible. We aim to reduce this effect further through data preprocessing which leaves us with lowercased sentences containing only standard punctuation, letters, and numbers. That removes most of the noise and provides a standard data form for modeling.

After training the initial model using Golden Dataset, we use the model to data mine the underlying structure of a bigger database sample. We mainly consider the sequences in which entities are present in records to sample later datasets accordingly. Using this sequence distribution, we leverage Knowledge Bases containing the most common person and company names to generate Synthetic Dataset to mimic the rules and patterns of the main database to our best ability while also being utilized in the later process of generating new training datasets using the framework described below. By incorporating synthetic data we aim to cover a variety of different entities that are not necessarily contained in the dataset so the model can better learn the underlying entity patterns in the dataset.

B. Generating training dataset

The Snorkel framework [5] is a powerful tool designed to facilitate machine learning using weak supervision. Weak supervision refers to using multiple sources of noisy, imperfect data to train a model [17]. The Snorkel framework enables users to train machine learning models using weak supervision by generating training data from various sources of imperfect data. These sources can include different heuristic rules or comparisons with sources of truth, each of which will vary in their accuracy and coverage over the dataset being labeled. To accomplish this, the framework includes creating labeling functions (LFs), which label or abstain from labeling data points. These labels are then used to train a discriminative model using agreements and disagreements of LFs to probabilistically generalize training labels beyond what the LFs explicitly cover. In other words, the framework uses a combination of LFs to train a model that can generalize beyond the training set, despite the noise and imperfections in the data.

As part of this process, we also incorporate the model trained on the Golden Dataset as an (LF). By incorporating a NER model trained on Golden dataset into the Snorkel framework as an LF, training of the discriminative model can take advantage of the pre-existing knowledge and

TABLE I: Coverage and empirical accuracy of labeling functions (LFs) for sample size N=10000. Each LF has a certain polarity, labeling tokens as person (PER), organization (ORG) or both.

LF Function	Polarity	Coverage	Emp. Acc.
is company	ORG	0.1733	0.95
is company suffix	ORG	0.2376	0.96
is name prefix	PER	0.0364	0.7
is first name	PER	0.6976	0.78
is last name	PER	0.6148	0.73
pre-trained model	ORG and PER	0.9913	0.88

TABLE II: Training results showing training dataset size, accuracy, validation, and test $F1$ scores for every training iteration.

Iteration	Dataset size	Train Dev $F1$	Held-out Dev $F1$	Held-out Test $F1$
1	2944	0.993	0.928	0.916
2	27944	0.996	0.935	0.925
3	102944	0.975	**0.943**	**0.947**
4	202944	0.958	0.922	0.914
fine-tuned	2502	0.998	**0.964**	**0.953**

accuracy of the model. The idea is to secure correct labels with agreements between the current NER model and other LFs, but also challenge the outputs of the NER model in cases where its labels are inaccurate, but LFs disagree with it. This approach leverages the model's strengths while minimizing weaknesses, leading to a more robust and accurate discriminative label model. Finally, it's worth noting that to sample weakly labeled data, only high-confidence labels are considered, specifically those with a confidence score over 0.8. This ensures that the training data is of high quality and can be used to train a more accurate model. In the subsequent training iterations, a freshly trained model is employed as an LF, and the process is repeated.

We utilize a variety of other labeling functions (LFs) as sources of weak labels to generate training data in addition to the model itself. Table I lists several best LFs with their respective coverage and empirical accuracy, including LFs that identify common first names, last names, and companies while other LFs leverage expert knowledge of name and company prefixes and suffixes. For the specific scenario shown in the table, the model was trained solely on the Golden Dataset, meaning that the measurements provided in Table I pertain to the initial training iteration only.

C. Training

Our proposed training process is shown in Fig. 1. Inside the red rectangle is a repeatable training process where training data is generated through Snorkel framework before each training iteration. The process feeds each model with new weakly labeled data leveraging previous models knowledge and the statistical advantage of expert knowledge LFs ensemble combined to train the discriminative label model that has a final say in what the labels for each sentence should be.

The training starts by annotating a sampled Golden Dataset (1.) and starting the initial model training (2.) using it. After the model is trained, using Snorkel and synthetic data generation rules in parallel (3.) a combined training dataset is generated (4.) which is then combined with a Golden Dataset, starting a new training iteration (5.), closing the loop. Each iteration begins with a preprocessing and stratified sampling of all available datasets by class distribution and the most common patterns of entities. Particularly, we consider the order in which entities are placed inside each data record. We found the model is sensitive to the order of entities, and by sampling in this way, we avoid situations in which the model is classifying, for example the first entity as a PER instead of ORG based on the order bias. Furthermore, this process ensures that the sample drawn from each dataset is representative of the dataset's overall distribution as the order of words in a sentence can significantly alter its meaning. The resulting samples are then merged to form a single data frame labeled using Snorkel LFs and stratified sampled in the ways described.

The importance of data quality in training deep learning models must be considered. It has significant implications for the performance of the model. When it comes to weakly supervised learning, the initial dataset is particularly important, as it is used to develop the model and establish its rules. The concern regarding the introduction of noisy data patterns through the model's self-feeding of data persists. This could cause a deviation from the model's primary task of accurately predicting outcomes. Although we leverage data preprocessing and stratified sampling to minimize that effect, we observe that after a few training iterations, the models performance inevitably starts decaying due to poor quality of automatically generated labels. Another step we introduce to address this is using fixed contextual word embeddings as opposed to letting them fine-tune during the training as it is found to be preferable according to [18], because our model uses Flair Embeddings that are based on a language model similarly to ELMo that was used for embedding inputs for LSTM network in Arora et al. [18].

A similar observation in Cheng et al. [11] was made during their training experiments. The $F1$ score stopped improving after a certain number of iterations and started dropping significantly compared to previous iterations. In Table II, we can see that the $F1$ score drops below the initial score of a model trained on the Golden Dataset. The model started to learn self-created bias towards unknown patterns created during automatic data generation, which led to a decline in the models performance. This can be attributed to the fact that the model was trained on a combination of weakly supervised data, which unavoidably introduced noise and inconsistencies in the data that overwhelmed and completely declined the gain in accuracy after a certain point. To address this, we introduce one final step after iterative training before concluding the final score and deciding which model is going to be deployed.

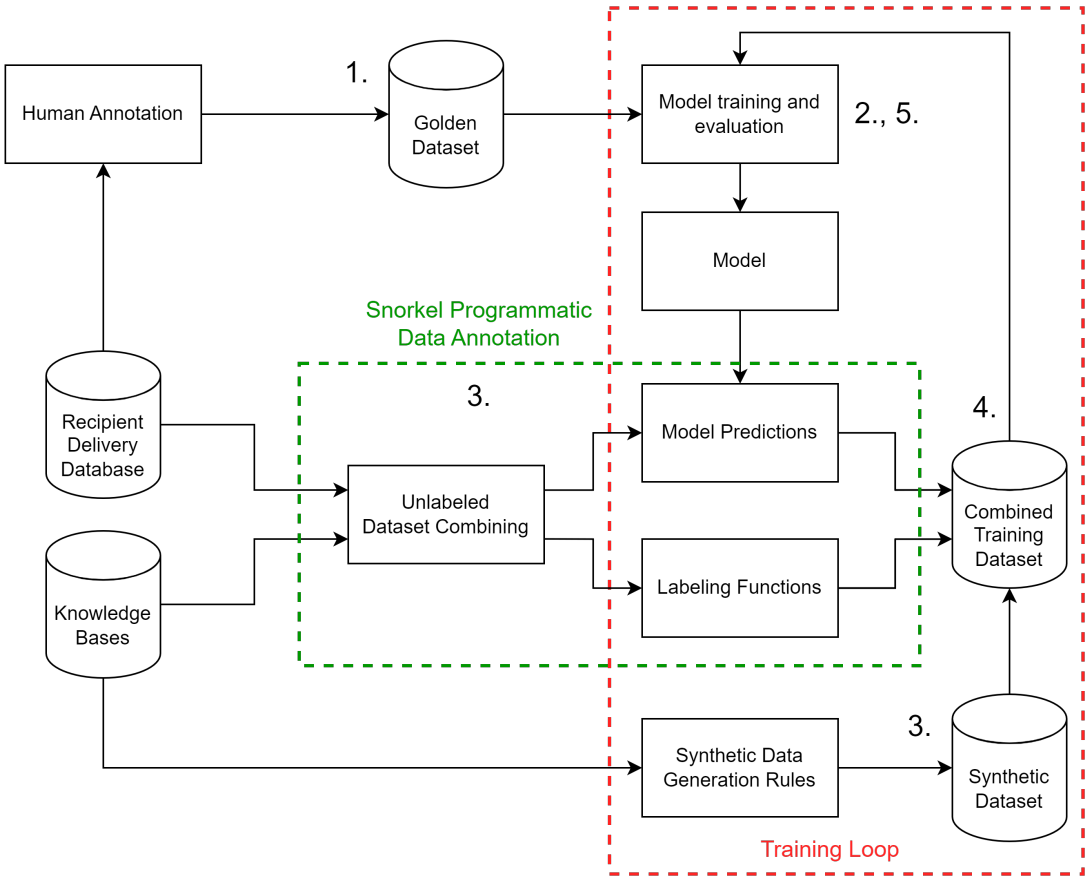


Fig. 1: The process of creating a training set utilizing the Snorkel framework involves the stratified sampling of all datasets and merging them with labeled Golden Data to form a single data frame. This data frame is used with Snorkel to apply expert knowledge rules and compare tokens with knowledge bases while leveraging a pre-trained model developed using Golden Dataset. Once the model is trained on the Combined Training Dataset, it is utilized for Programmatic Annotation in the next training iteration.

III. Model fine-tuning

To fine-tune a model, first, we choose a model from training that is just one iteration before $F1$ score on our held-out validation set starts dropping. In our case, this is the third iteration as seen in Fig. 2. Finally, we proceed with fine-tuning the best iteration using 85% of the original Golden Dataset with a very small learning rate and up to 10 epochs in order to avoid overfitting and ultimately negatively affecting the model.

To select an appropriate learning rate, we choose a value from the optimal learning rate range where loss decreases the most and before it explodes as the learning rate becomes too large. This can be achieved by plotting loss as a function of learning rate and selecting the highest value of learning rate from the optimal range. This technique is described in Smith [19]. The technique can improve the performance of the model by finding an appropriate learning rate and avoiding problems such as slow convergence or overshooting the optima of the loss function. In Fig. 3, we can observe the plot of loss as a function of learning rate for our selected model. The steep optimal curve in the plot indicates that the model can still learn from the initial Golden Dataset. This finding further strengthens our suspicion that models trained in this way can in fact benefit from retraining on the Golden Dataset to address introduced bias through weakly labeled data generation. For our particular case shown in Fig. 3 we choose value 0.001 as our learning rate and fine-tune the model for 10 epochs. This results in a $F1 = 0.953$ on our held-out test dataset. The results are shown in Table II.

A. Model Architecture and Word Embeddings

Our approach employs a BiLSTM-CRF neural network which has demonstrated exceptional accuracy on NER datasets [20], [21]. We utilize contextual word embeddings, known as Flair Embeddings, accessible within the Flair framework [1], to generate character-based word embeddings that serve as inputs to the BiLSTM-CRF classifier, as shown in Fig. 4. The BiLSTM layer is a neural network that can process sequential data such as text, which works great for identifying sequences like named entities in sentences. This layer is designed to learn patterns and relationships between the input features, allowing it to capture complex patterns in the data. The CRF layer is a probabilistic graphical model used to

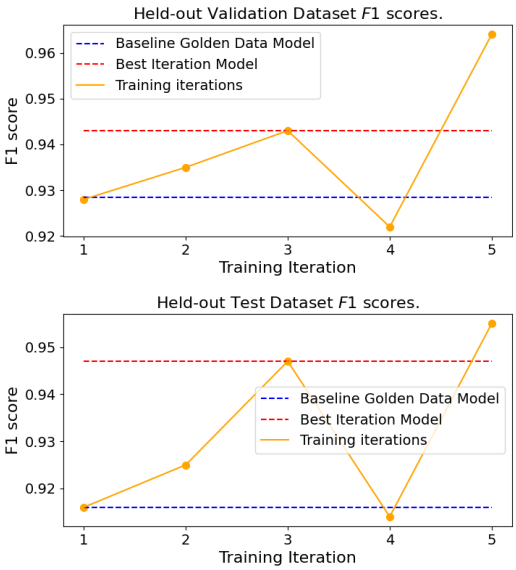


Fig. 2: Results from Table II include $F1$ scores for validation and test datasets (orange) per iteration, a baseline using only Golden Dataset (blue), and the best iteration's $F1$ score (red). The final measurement is from a fine-tuned best iteration model.

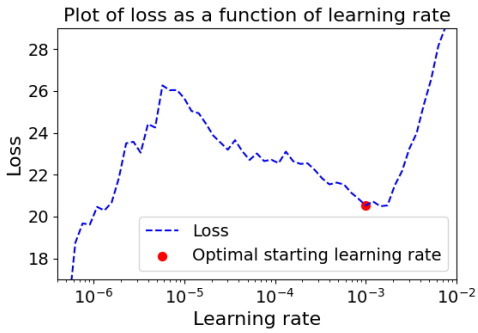


Fig. 3: A plot of loss against the learning rate using Adam optimizer. The red point shows the optimal starting learning rate for a fine-tuning process.

predict a sequence of labels based on the input features allowing the classifier to take into account the context of the words. The whole architecture is implemented through the Flair framework, which is built on top of PyTorch.

IV. Results

In Fig. 2, we observe that our final fine-tuned model, which is trained solely on the Golden Dataset, outperforms the baseline model and all other models in the iterative training process. To achieve this performance, we use our proposed solution, which involves generating additional data using our proposed techniques and fine-tuning the model from the best training iteration using a small learning rate and a subset of the initial Golden Dataset. This allows the model to relearn concepts from the initial training, while also improving the overall score and achieving higher $F1$ score than all the previous

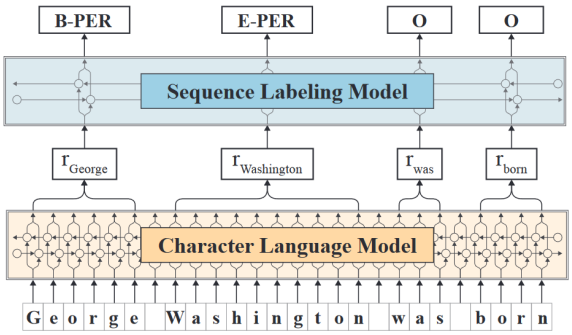


Fig. 4: The proposed approach involves using a pre-trained bidirectional character language model to obtain contextual embeddings of words in a sentence. These embeddings are then passed into BiLSTM-CRF sequence labeler [2].

training iterations. By fine-tuning the model on the original Golden Dataset, we can potentially restore its knowledge about the initial data while also retaining insights gained from subsequent training iterations and patterns imposed by noisy data generation.

The training results are summarized in Table II, where we observe that the iterative training process improved the $F1$ score by approximately 3.1% compared to the baseline. However, fine-tuning the best iteration on the original Golden Dataset resulted in an additional increase of approximately 0.6% in the $F1$ score. Consequently, the overall increase in $F1$ score from the baseline is 3.7%.

V. Future Work

There are several avenues for future work related to our proposed approach. One promising direction is to explore knowledge transfer techniques to improve the initial performance of the model. This could involve leveraging knowledge from related tasks or domains to improve the accuracy of the model in the target task.

Another potential area of investigation is to explore the use of other advanced embedding models based on transformers and language models, or fine-tuning the existing ones. Fine-tuning these models could lead to even higher accuracy and better performance as the words for specific domain case would have better vector representations.

Finally, it would be interesting to investigate how our approach would affect other popular architectures such as BiGRU-CRF or CNN architectures. By experimenting with different model architectures, we can gain a better understanding of the generalizability and robustness of our proposed approach.

It would also be worth exploring the applicability of our approach to other domains. While our approach was developed and evaluated on a specific domain of identifying person and organization names of one particular multilingual country's delivery recipients, it has the potential to be adapted to other domains where the dataset is also multilingual, and similar challenges exist.

VI. Conclusion

Refitting a model on Golden Dataset in our proposed process of fine-tuning its parameters on the original dataset after iterative training, has a positive impact on its performance. It is important to keep in mind that the model would still retain some knowledge from training iterations, and this could potentially impact its performance on the original task. In our case, the retaining of knowledge from data generation techniques and training iterations shows an overall improvement of $F1$ score by 3.7% from the baseline, or by 0.6% from the best iteration, reaching $F1 = 0.953$. During our experiments this proved to be a consistent phenomenon.

One disadvantage of this method is its high computational cost due to training multiple iterations. However, our approach demonstrates that it is feasible to achieve state-of-the-art results in scenarios where a Golden Dataset annotated by human domain experts is a scarce resource. By leveraging a smaller dataset to establish initial knowledge and using an iterative training process that generates new data to train the model, we can overcome the challenge of the absence of human annotations. Furthermore, we can improve model accuracy by fine-tuning the model on the initial small Golden Dataset, which helps to restore the model's knowledge of the initial dataset. In addition to this, the knowledge retained from training iterations has proven to be beneficial as the overall accuracy of the model further improves after refitting. Although our approach comes with a high computation cost, we demonstrate that it is a viable alternative in situations where obtaining a large enough Golden Dataset for traditional supervised learning is not feasible.

Acknowledgment

We thank BISS d.o.o. for allowing us to use the hardware and equipment for training and testing purposes. Our gratitude goes to Dr. Matej Mihelčić, Matija's mentor, for providing guidance towards the final version of this paper.

References

[1] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, and R. Vollgraf, "FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 54–59. [Online]. Available: https://aclanthology.org/N19-4010

[2] A. Akbik, D. Blythe, and R. Vollgraf, "Contextual String Embeddings for Sequence Labeling," in *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, Aug. 2018, pp. 1638–1649. [Online]. Available: https://aclanthology.org/C18-1139

[3] A. Jaiswal, A. R. Babu, M. Z. Zadeh, D. Banerjee, and F. Makedon, "A Survey on Contrastive Self-Supervised Learning," *Technologies*, vol. 9, no. 1, p. 2, Mar. 2021, number: 1 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2227-7080/9/1/2

[4] D. Hendrycks, M. Mazeika, S. Kadavath, and D. Song, "Using Self-Supervised Learning Can Improve Model Robustness and Uncertainty," Oct. 2019, arXiv:1906.12340 [cs, stat]. [Online]. Available: http://arxiv.org/abs/1906.12340

[5] A. Ratner, S. H. Bach, H. Ehrenberg, J. Fries, S. Wu, and C. Ré, "Snorkel: Rapid Training Data Creation with Weak Supervision," *Proceedings of the VLDB Endowment*, vol. 11, no. 3, pp. 269–282, Nov. 2017, arXiv:1711.10160 [cs, stat]. [Online]. Available: http://arxiv.org/abs/1711.10160

[6] R. Grishman and B. Sundheim, "Message Understanding Conference- 6: A Brief History," in *COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics*, 1996. [Online]. Available: https://aclanthology.org/C96-1079

[7] L. Rau, "Extracting company names from text," in *The Seventh IEEE Conference on Artificial Intelligence Application [1991] Proceedings*, vol. i, Feb. 1991, pp. 29–32.

[8] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," Jan. 2008, pp. 160–167.

[9] V. Yadav and S. Bethard, "A Survey on Recent Advances in Named Entity Recognition from Deep Learning models," Oct. 2019, arXiv:1910.11470 [cs]. [Online]. Available: http://arxiv.org/abs/1910.11470

[10] E. F. T. K. Sang and F. De Meulder, "Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition," Jun. 2003, arXiv:cs/0306050 version: 1. [Online]. Available: http://arxiv.org/abs/cs/0306050

[11] X. Cheng, M. Bowden, B. R. Bhange, P. Goyal, T. Packer, and F. Javed, "An End-to-End Solution for Named Entity Recognition in eCommerce Search," Dec. 2020, arXiv:2012.07553 [cs]. [Online]. Available: http://arxiv.org/abs/2012.07553

[12] D. Putthividhya and J. Hu, "Bootstrapped Named Entity Recognition for Product Attribute Extraction," in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Edinburgh, Scotland, UK.: Association for Computational Linguistics, Jul. 2011, pp. 1557–1567. [Online]. Available: https://aclanthology.org/D11-1144

[13] A. More, "Attribute Extraction from Product Titles in eCommerce," Aug. 2016, arXiv:1608.04670 [cs]. [Online]. Available: http://arxiv.org/abs/1608.04670

[14] M. Wen, D. K. Vasthimal, A. Lu, T. Wang, and A. Guo, "Building Large-Scale Deep Learning System for Entity Recognition in E-Commerce Search," in *Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, ser. BDCAT '19. New York, NY, USA: Association for Computing Machinery, Dec. 2019, pp. 149–154. [Online]. Available: https://doi.org/10.1145/3365109.3368765

[15] A. Ahmed, C.-y. Wu, G. R. Kumar, and R. Datta, "Predicting Latent Structured Intents from Shopping Queries," 2017.

[16] "Github:mpajas/MIPRO-2023." [Online]. Available: https://github.com/mpajas/MIPRO-2023

[17] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, Jan. 2018. [Online]. Available: https://doi.org/10.1093/nsr/nwx106

[18] G. Arora, A. Rahimi, and T. Baldwin, "Does an LSTM forget more than a CNN? An empirical study of catastrophic forgetting in NLP," in *Proceedings of the The 17th Annual Workshop of the Australasian Language Technology Association*. Sydney, Australia: Australasian Language Technology Association, 2019, pp. 77–86. [Online]. Available: https://aclanthology.org/U19-1011

[19] L. N. Smith, "Cyclical Learning Rates for Training Neural Networks," Apr. 2017, arXiv:1506.01186 [cs]. [Online]. Available: http://arxiv.org/abs/1506.01186

[20] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural Architectures for Named Entity Recognition," Apr. 2016, arXiv:1603.01360 [cs]. [Online]. Available: http://arxiv.org/abs/1603.01360

[21] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF Models for Sequence Tagging," Aug. 2015, arXiv:1508.01991 [cs]. [Online]. Available: http://arxiv.org/abs/1508.01991

Augmentacija podataka za klasifikaciju kratkih tekstova

I. Hrga
Sveučilište u Rijeci/Fakultet informatike i digitalnih tehnologija, Rijeka, Hrvatska
ingrid.hrga@gmail.com

***Sažetak* - Augmentacija podataka postala je neizostavni korak u procesu učenja većine sustava temeljenih na dubokim neuronskim mrežama. Velike količine podataka potrebne da bi sustav uspješno naučio rješavati zadatke često je teško pribaviti u dovoljnoj količini i kvaliteti. Zbog toga se pribjegava različitim postupcima kojima se primjenom raznovrsnih transformacija može višestruko povećati osnovni skup podataka. U području računalnog vida već su uvriježene brojne tehnike koje relativno lako transformiraju sliku bez da se promijeni što ta slika predstavlja. Međutim, u području obrade prirodnog jezika takve su transformacije znatno manje zastupljene. Već i promjena samo jednog slova može znatno promijeniti smisao rečenice, stoga je potrebno puno više opreza prilikom odabira pogodnih transformacija. U ovom radu uspoređuju se postupci za augmentaciju tekstualnih podataka. Uz prikaz njihovih prednosti i nedostataka, analizira se kako se promjena pojedine tehnike odražava na rezultate klasifikacije kratkih tekstova.**

Ključne riječi - augmentacija podataka; klasifikacija teksta; konvolucijske neuronske mreže; WordNet sinonimi; kontekstualni vektori riječi

I. UVOD

Otkako su duboke neuronske mreže pokazale sposobnost rješavanja najraznovrsnijih zadataka iz područja računalnog vida ili obrade prirodnog jezika [1, 2], sve se više autora bavi problemom učenja iz ograničenih skupova podataka [3]. Nedovoljna količina podataka za učenje u odgovarajućoj kvaliteti dovodi do problema prenaučenosti (eng. *overfitting*), što se negativno odražava na sposobnost generalizacije modela.

Postoje različiti načini za ublažavanje prenaučenosti [4], stoga se neki autori usredotočuju na razvoj naprednijih tehnika za prijenos znanja [5], drugi se usmjeravaju na unaprjeđenje metoda za regularizaciju modela [6]. Međutim, veće količine podataka za učenje u pravilu predstavljaju najpoželjnije rješenje problema prenaučenosti, u čemu važnu ulogu ima augmentacija podataka [7].

Augmentacija podataka obuhvaća tehnike za povećanje veličine i raznovrsnosti osnovnog skupa bez potrebe za prikupljanjem novih podataka [7]. Takve tehnike primjenom različitih transformacija stvaraju varijacije postojećih podataka. Osim toga, u svrhu proširenja osnovnog skupa mogu se koristiti i sintetički primjeri dobiveni generativnim modelima ili primjenom ostalih naprednih tehnika koje stvaraju sasvim nove podatke [8]. Međutim, bez obzira na to koji se način za proširenje osnovnog skupa koristi, važno je sačuvati vezu između podataka i dodijeljenih im oznaka, poput klasa. U protivnome, u podatke se uvodi nepotreban šum, što u konačnici, unatoč većem skupu za učenje, može rezultirati lošijim performansama modela.

Pored navedenoga, augmentacija podataka može pomoći i u situacijama kada postoji znatna neravnoteža među pojedinim klasama. Manjinska klasa tad se augmentira transformiranim ili sintetičkim primjerima, čime se pridonosi uravnoteženju skupa podataka [9]. Augmentacija smanjuje potrebu za prikupljanjem i označavanjem podataka, što je često dugotrajan i zamoran proces, a ponekad nije niti moguć, npr. ako zahtijeva specifično ekspertno znanje ili se provodi nad osjetljivim podacima. Generiranjem sintetičkih primjera iz distribucije jednake ili slične onoj osnovnog skupa, moguće je nadopuniti ili zamijeniti originalne podatke.

Neupitna je korisnost augmentacije te je ona postala neizostavni korak procesa učenja u području računalnog vida. Pomoću jednostavnih transformacija slika poput rotacije, translacije, promjena boje [10] te kompleksnijih poput prekrivanja dijelova slike [11] ili interpolacijom više različitih slika [12], postižu se bolje performanse modela uz veću robusnost na varijacije u podacima [7].

Međutim, u radu s tekstualnim podacima takvi su postupci znatno manje zastupljeni. Za razliku od slika, gdje većina promjena vrijednosti pojedinih piksela ili objekata neće utjecati na to što ta slika predstavlja (npr. mačka sa zelenim krznom i dalje će biti prepoznata kao mačka), u radu s tekstom već i promjena jednog znaka može promijeniti smisao čitave rečenice. Diskretna priroda jezika onemogućava stvaranje postepenih varijacija. Osim toga, poredak riječi u rečenici utječe na smisao rečenice, a smisao također ovisi i o kontekstu. Promjenom konteksta može se znatno promijeniti značenje riječi. Kompleksna struktura teksta i veza među riječima čine augmentaciju znatno izazovnijom te je potrebno puno više opreza prilikom odabira prikladnih transformacija.

U ovom radu uspoređuju se odabrane tehnike za augmentaciju tekstualnih podataka. Dan je pregled postupaka temeljenih na promjenama pojedinih znakova, zamjenama riječi baziranih na WordNet [13] sinonimima te tehnikama koje manipuliraju reprezentacijom teksta u vektorskom prostoru. Uz prikaz njihovih prednosti i nedostataka, empirijski je provjereno kako se primjena

pojedine tehnike odražava na rezultate klasifikacije kratkih tekstova.

Rad je strukturiran na sljedeći način: nakon uvodnog dijela, dan je kratki pregled najčešće korištenih tehnika za augmentaciju tekstualnih podataka prikladnih za problem klasifikacije, uz dodatni prikaz mreža koje se u tu svrhu koriste. U trećem poglavlju opisani su podaci, odabrane transformacije i ostale postavke eksperimenta. U četvrtom poglavlju prikazani su rezultati klasifikacije modela treniranih na originalnim podacima, onim augmentiranim te njihovim kombinacijama. Na kraju je dan zaključak, uz nekoliko smjernica za buduća istraživanja.

II. AUGMENTACIJA PODATAKA I KLASIFIKACIJA

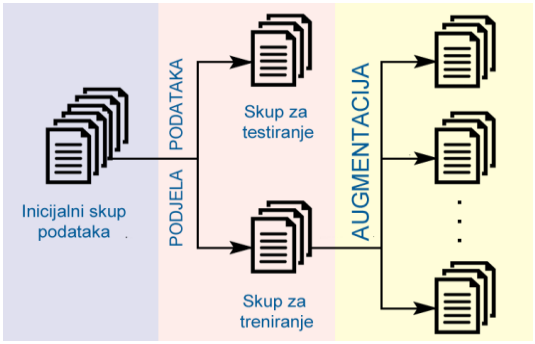
A. *Augmentacija podataka*

Većina transformacija koje se u svrhu augmentacije koriste u području obrade prirodnog jezika inspirirana je augmentacijama uvriježenima u području računalnog vida. Na primjer, tehnikama koje mijenjaju pojedine znakove odgovaraju promjene vrijednosti pojedinih piksela, dok se promjene na razini riječi mogu poistovjetiti s promjenama pojedinih objekata na slici. Međutim, za razliku od računalnog vida, gdje je augmentacija podataka tijesno povezana s procesom učenja, tekstualne se transformacije uglavnom pripremaju unaprijed, prije samoga učenja (Slika 1).

Tehnike augmentacije namijenjene radu s tekstualnim podacima mogu se kategorizirati na više načina [3, 9, 14]: 1) s obzirom na granularnost jedinica transformacije mogu se podijeliti na one koje se primjenjuju na pojedinim znakovima, riječima, frazama ili rečenicama, 2) na temelju reprezentacije teksta dijele se na one koje izravno mijenjaju sirove podatke te one koje manipuliraju njihovom vektorskom reprezentacijom, 3) augmentacije mogu biti bazirane na pravilima, na unaprijed treniranom modelu ili dobivene učenjem novog modela, 4) mogu biti ovisne o jeziku ili neovisne o jeziku, 5) prilagođene određenom zadatku ili univerzalne, 6) determinističke ili stohastičke.

Najjednostavnije su tehnike koje mijenjaju pojedine znakove. Takve promjene mogu biti neovisne o jeziku, poput brisanja, zamjene ili ubacivanja slučajno odabranog znaka, ili mogu ovisiti o jeziku teksta, poput simulacije pogrešaka prilikom tipkanja na temelju rječnika pogrešaka. Ovakve promjene, zbog uvođenja šuma stvaranjem nepostojećih riječi, rezultiraju velikim vokabularom, a iako su jednostavne za implementaciju, u radovima nisu značajnije zastupljene.

S druge strane, autori glavnine radova predlažu transformacije na razini pojedinih riječi ili pojavnica (eng. *token*) među kojima su najzastupljenije zamjene slučajno odabranih riječi njihovim sinonimima tj. riječima jednakog ili bliskog značenja. Moguće su i zamjene antonimima tj. riječima suprotnog značenja, međutim, takve promjene mijenjaju smisao rečenice, stoga nisu prikladne za problem klasifikacije.



Slika 1. Shema općenitog postupka augmentacije podataka. Nakon podjele podataka, transformiraju se samo podaci za treniranje dok se podaci za testiranje koriste u izvornom obliku.

U literaturi su zastupljeni različiti pristupi pri odabiru sinonima. U [15] autori koriste WordNet, leksičku bazu međusobno povezanih sinonima za engleski jezik. Takve su zamjene brze i jednostavne, međutim, baze sinonima ne postoje za sve jezike ili ne pokrivaju dovoljan broj riječi, o čemu ovisi i kvaliteta augmentacije. Npr. WordNet obuhvaća 117000 skupova sinonima (eng. *synsets*), dok hrvatska inačica, Crown [16] u verziji 2.0 obuhvaća samo 23122 sinonimska skupa. Zato drugi autori [17] predlažu odabir sinonima na temelju kosinusne sličnosti vektorskih reprezentacija riječi. U tu se svrhu najčešće koriste word2vec [18], GloVe [19] ili fasttext [20] vektori. Nedostatak ovih pristupa je to što ne uzimaju u obzir kontekst te će riječ u slučaju polisemije, tj. u slučaju kada ona ima više značenja, biti jednako reprezentirana (vektorom istih vrijednosti) za svako od značenja. Npr. riječi glava i lubanja dovoljno su bliske, ali zamjena "glavica luka" → "lubanja luka" ne bi bila prihvatljiva. Zbog toga se prednost često daje kontekstualnim vektorima koje stvaraju jezični modeli unaprijed trenirani na iznimno velikim korpusima, s ciljem predviđanja prikrivene riječi u rečenici. Njihova je prednost što pri tome kao kontekst uzimaju čitavu rečenicu. To znači da će pojedina riječ imati poseban vektor za svako značenje, a za zamjenu bit će odabrana ona riječ koja odgovara smislu čitave rečenice. Primjeri takvih jezičnih modela su BERT [2] ili DistillBERT [21] kao njegova umanjena verzija.

Jedna od ranijih primjena augmentacije može se pronaći u [22]. Autori u [23] koriste augmentaciju sinonimima za treniranje konvolucijskih neuronskih mreža, dok su autori u [15], pod nazivom *Easy Data Augmentation (EDA)*, obuhvatili kombinaciju slučajnih zamjena sinonimima s ubacivanjem sinonima, te brisanjem slučajno odabranih riječi ili zamjenom njihovog poretka. Ovakve jednostavne transformacije pokazale su se učinkovitima u treniranju jednostavnijih modela na malim skupovima podataka.

Druga skupina popularnih tehnika stvara varijacije rečenica parafraziranjem, pri čemu se u tu svrhu najčešće koristi prevođenje [24]. Augmentirana rečenica dobije se prevođenjem izvorne rečenice na pivotalni jezik, a zatim se takav prijevod ponovno prevodi na izvorni jezik. Moguće je koristiti jedan pivotalni jezik ili više različitih jezika, čime se augmentirana rečenica dodatno udaljava od izvorne.

Potpuniji pregled dodatnih tehnika može se pronaći u [3, 9, 14].

B. *Klasifikacija teksta*

Klasifikacija teksta je postupak kojim se nekom tekstu dodjeljuje oznaka iz konačnog skupa unaprijed poznatih oznaka. Koristi se u zadacima poput analize sentimenta, gdje se na temelju izraženog stava tekst označava polaritetom (npr. pozitivan ili negativan), zatim u zadacima filtriranja informacija na relevantne i irelevantne, označavanju imenovanih entiteta (eng. *Named Entity Recognition - NER*) poput osoba, lokacija ili organizacija, te označavanju vrsta riječi u rečenici (eng. *Part-of-Speech (POS) tagging*), uz brojne druge zadatke.

Veliki jezični modeli bazirani na arhitekturi transformera [25] pokazuju zadivljujuće rezultate u raznim zadacima obrade jezika [2, 26], međutim, njihova primjena nije uvijek preporučljiva niti moguća zbog velikih zahtjeva za resursima. Stoga se kao manje zahtjevna alternativa i dalje koriste inačice konvolucijskih neuronskih mreža (CNN), poput onih usmjerenih na riječi [27] ili usmjerenih na pojedine znakove [23]. Drugu grupu popularnih modela čine oni temeljeni na povratnim mrežama, poput LSTM [28] ili BiLSTM [29] koja procesiranje sekvencijalnih podataka vrši u oba smjera, unaprijed i unazad. Osim toga, korpusi na kojima su trenirani veliki jezični modeli već sadrže brojne varijacije riječi i njihovih kombinacija, zbog čega se pokazalo da augmentacija podataka u tom slučaju nije potrebna, a može dovesti i do lošijih rezultata [30]. Raniji radovi [29] već su ukazali na prednosti jednostavnijih modela u situaciji kad se raspolaže s malom količinom podataka.

Ovaj je rad najsličniji [31] u kojem su autori usporedili različite tehnike augmentacije, ali na problemu klasifikacije slika. Osim toga, ovdje nije korišten prijenos znanja već su mreže trenirane od početka. Ovaj rad ima određenih sličnosti i s [3], u kojem su autori usporedili tehnike za augmentaciju teksta, ali koristeći BERT-base model, te [15] u kojem su autori koristili samo WordNet za odabir sinonima. Za razliku od toga, ovdje je uspoređeno više različitih pristupa u odabiru sinonima kako bi se utvrdilo odražava li se dodatna kompleksnost pozitivno i na rezultate klasifikacije.

III. METODE I EKSPERIMENT

A. *Podaci*

Kao podloga za transformacije upotrijebljen je „Subjectivity" [32], standardni skup za provjeru performansi klasifikatora. Skup se sastoji od 10000 recenzija i sažetaka filmova na engleskom jeziku. Podaci su prikupljeni iz dva izvora: Rotten Tomatoes[1] i Internet Movie Database[2]. Prvi izvor smatra se subjektivnim dok se drugi smatra objektivnim, a zadatak je klasificirati rečenice na objektivne i subjektivne. Skup je uravnotežen te ima po 5000 primjera svake od navedenih dviju klasa. S obzirom da nije unaprijed podijeljen na podskupove za treniranje, validaciju i testiranje modela, u podjeli podataka slijedio se primjer iz [15] te je slučajnim odabirom izdvojeno 1000 primjera za testiranje, a ostatak je korišten za treniranje.

[1] http://www.rottentomatoes.com

[2] http://www.imdb.com

Iz podskupa za treniranje slučajnim su odabirom dodatno izdvojena dva manja, veličine 100 i 1000 primjera, kako bi se simulirala situacija s ograničenim podacima. Pri tome je svaki podskup pravi podskup tj. svi primjeri iz manjega nalaze se u većem, a uz to je sačuvan i postojeći omjer među klasama. Tijekom procesa učenja izdvajalo se 20% primjera za validaciju. Statistike skupova za treniranje dane su u Tablici I.

B. *Tehnike augmentacije*

U eksperimentu je korišteno sedam transformacija generiranih pomoću knjižnice nlpaug [33]. To su:

Na razini pojedinog znaka:

1. Keyboard" (u daljnjem tekstu KB) - simulacija pogrešaka prilikom tipkanja zamjenom slučajno odabranih znakova (1-3) drugim bliskim znakovima.

Na razini riječi:

2. „Synonym" (WN) - zamjena slučajno odabranih riječi (1-3) njihovim sinonimima. Za odabir sinonima korištena je leksička baza WordNet.
3. „Word embedding" (WE) - zamjena slučajno odabranih riječi (1-3) sinonimima. Sinonimi su birani na temelju kosinusne sličnosti riječi reprezentiranih pomoću GloVe vektora riječi.
4. „Contextual word embedding" (CWE) - zamjena slučajno odabranih riječi (1-3) sinonimima koji odgovaraju kontekstu rečenice, za što je korišten model DistillBERT.
5. „Delete word" (DW) - brisanje slučajno odabranih riječi (1-2).
6. „Swap word" (SW) - zamjena poretka slučajno odabranih riječi (1-3).

Na razini rečenice:

7. „Backtranslation" (BT) – parafraziranje rečenice prevođenjem s engleskog jezika na njemački te ponovno na engleski. Za ovu kombinaciju jezika pokazano je da daje dobre rezultate [34], a za prijevod su korišteni modeli Facebook FAIR's WMT19 (wmt19-en-de i wmt19-en-de) [35].

Primjeri pojedinih augmentacija dani su u Tablici II.

C. *Model i eksperimentalne postavke*

S obzirom na to da cilj rada nije postizanje najboljih performansi na zadatku klasifikacije, već izmjeriti relativni dobitak u točnosti primjenom različitih tehnika augmentacije, za klasifikaciju je upotrijebljena jednostavna konvolucijska mreža prilagođenu radu s

TABLICA I. STATISTIKE SKUPOVA ZA TRENIRANJE

Veličina skupa za treniranje	Prosječna duljina rečenice	Veličina vokabulara
100	23,98	1833
1000	24,11	5956
9000	24,62	21853

TABLICA II. PRIMJERI ODABRANIH TRANSFORMACIJA

Vrsta transformacije	Rezultat
Originalna rečenica	the movie begins in the past where a young boy named sam attempts to save celebi from a hunter .
KB	the mo**D**ie be**r**ins in the past where a y**oJ**ng boy named sam at**f**empts to save **v**elebi from a hunter.
WN	the movie begins in the past where a young **male child** named sam attempts to **relieve** celebi from a **huntsman**.
WE	the **films continues** in the past **outside** a young boy named sam attempts to save celebi from a hunter.
CWE	**same** movie **ends** in **centuries** past where **handsome** young boy named sam attempts to **differentiate himself from monster hunter.**
DW	the _ begins in the past where a young boy named sam attempts to save celebi from _ hunter.
SW	the movie begins **the in where past** a young boy named sam attempts **save to** celebi from a hunter.
BT	**T**he film begins in the past **when** a **little** boy named **S**am **tries** to save **celebs** from a hunter.

KB – "Keyboard", WN - "Synonym", WE - "Word embedding", CWE - "Contextual word embedding", DW - "Delete word", SW - "Swap word", BT - "Backtranslation"

tekstualnim podacima. Iako određene varijante povratnih mreža, kao što je to BiLSTM, daju ponešto bolje rezultate u odnosu na CNN [29], učenje takvih mreža traje duže te je iz tog razloga odabrana CNN mreža.

Mreža se sastoji od dva konvolucijska sloja koji su praćeni slojevima za sažimanje po maksimumu (eng. *max pooling*), odnosno globalnim sažimanjem po maksimumu (eng. *global max pooling*), u skladu s preporukom u [36]. Veličina filtera je 5 dok je broj mapa značajki 128. Za reprezentaciju teksta korišteni su GloVe vektori riječi (glove6B) dimenzionalnosti d=300. Prednost GloVe vektora u odnosu na word2vec je to što uzimaju u obzir supojavljivanje riječi (eng. *co-occurrence*), a u odnosu na fasttext vektore manje su memorijski zahtjevni. Na kraju mreže nalazi se linearni sloj uz sigmoidnu aktivacijsku funkciju za binarnu klasifikaciju. U konvolucijskim slojevima aktivacijska funkcija je ReLU, a za regularizaciju je upotrijebljen dropout [6] uz p=0.5. Za optimizacijski algoritam odabran je Adam [37]. Navedene vrijednosti hiperparametara predstavljaju uobičajene postavke korištene u literaturi [15].

Mreža je trenirana 20 epoha uz ranije zaustavljanje ako kroz tri epohe nije došlo do smanjenja validacijske pogreške. Svaki je eksperiment ponovljen tri puta te je izračunata prosječna točnost na skupu za testiranje.

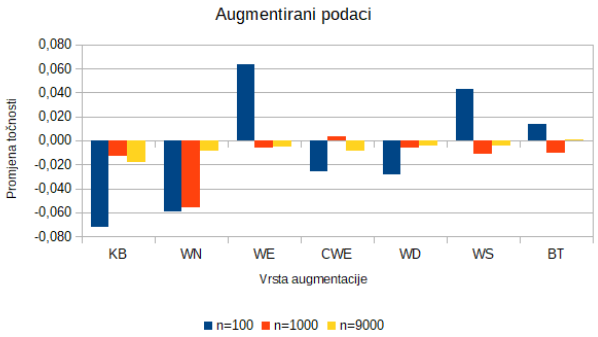
Mreža je trenirana na podskupovima veličine 100, 1000 i 9000 primjera, na originalnim podacima, podacima dobivenim pojedinim transformacijama te kombinacijom originalnih i augmentiranih podataka u dva omjera: u prvom je slučaju za svaki originalni primjer kreirana n=1 transformacija, dok je u drugom slučaju n=3.

IV. REZULTATI

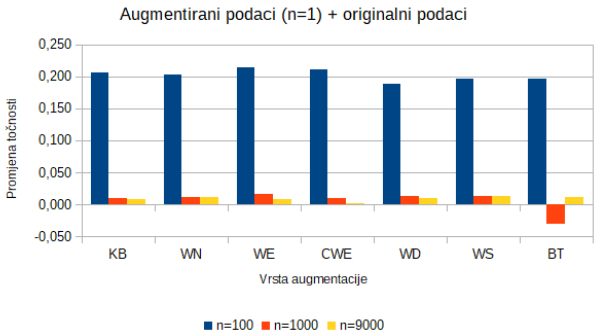
Točnost modela treniranog samo na originalnim podacima veličine 100, 1000 i 9000 primjera iznosi redom: 58,0% (±4,5), 86,5% (±0,7), te 91,4% (±0,3). Najbolji rezultat iznosi 93,6% (±2,3) te je ostvaren modelom treniranim na najvećem skupu koji se sastojao od kombinacije originalnih i KB augmentiranih podataka u omjeru 1:3. Rezultatom slijede kombinacije originalnih podataka s WS (92,6%, ±0,8), i WN augmentacijama u omjeru 1:1 (92,5, ±0,8).

Rezultati klasifikacije izraženi kao razlika točnosti u odnosu na modele trenirane samo na osnovnom skupu prikazani su na slikama 2-4. Može se uočiti da su razlike izraženije na malom skupu (n=100) dok se s povećanjem veličine skupa za treniranje one gube.

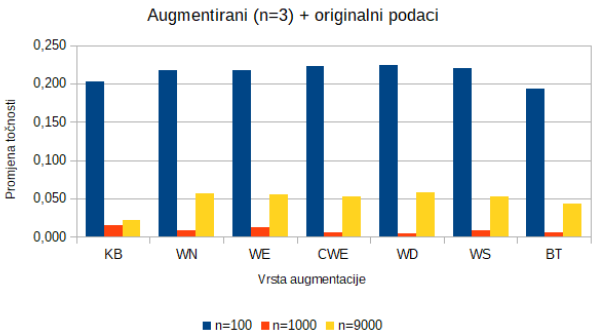
Ako se promatraju rezultati treniranja samo na augmentiranim podacima (Slika 2), može se primijetiti da su oni gotovo u pravilu lošiji od rezultat modela treniranog samo na osnovnom skupu. Iako se ne radi o značajnijim odstupanjima, ipak su te razlike ponešto veće na najmanjem skupu. Međutim, ako se augmentiranim podacima dodaju i originalni (Slika 3), tada dolazi do jasnog poboljšanja točnosti, posebno na najmanjem skupu.



Slika 2. Razlika točnosti modela treniranih na augmentiranim podacima u odnosu na točnost klasifikacije modela treniranih samo na osnovnom skupu veličine 100, 1000 i 9000 primjera.



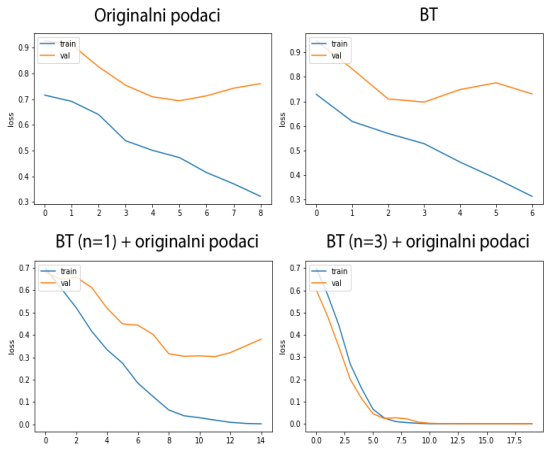
Slika 3. Razlika točnosti modela treniranih na skupovima originalnih podataka kojima su dodani augmentiranim podaci u jednakom omjeru, u odnosu na točnost klasifikacije modela treniranih samo na osnovnom skupu veličine 100, 1000 i 9000 primjera.



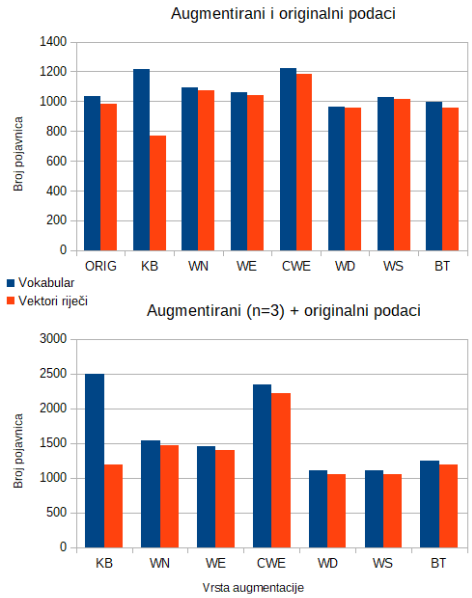
Slika 4. Razlika točnosti modela treniranih na skupovima originalnih podataka kojima su dodani augmentiranim podaci u omjeru 1:3, u odnosu na točnost klasifikacije modela treniranih samo na osnovnom skupu veličine 100, 1000 i 9000 primjera.

Ipak, daljnjim dodavanjem augmentiranih podataka, premda generiranih na način da se postigne što veća raznolikost, nije značajnije pridonijelo većoj točnosti klasifikacije (Slika 4). Iako je s porastom omjera augmentacije s 1:1 na 3:1 došlo do kretanja u pozitivnom smjeru, značajniji je učinak postignut samo na najvećem skupu. Može se primijetiti da su dodatni podaci temeljeni na kontekstualnim vektorima pokazali ponešto povoljniji učinak od ostalih tehnika augmentacije, ali samo u kombinaciji s originalnim podacima. Međutim, generiranje takvih podataka trajalo je nekoliko desetaka puta duže u odnosu na pronalaženje WordNet sinonima. Zanimljivo je da su rezultati na skupu srednje veličine najmanje reagirali na tehnike augmentacije, neovisno o veličini skupa za učenje.

S druge strane, ako se uzmu u obzir krivulje validacijskog gubitka, može se uočiti jasno poboljšanje, odnosno smanjenje prenaučenosti s proširenjem skupa augmentiranim podacima. Slika 5 prikazuje krivulje za primjer augmentacije prevođenjem. Može se uočiti da je u slučaju treniranja samo na originalnim podacima do prenaučenosti dolazilo već tijekom prvih nekoliko epoha.



Slika 5. Krivulje validacijskog gubitka za veličinu osnovnog skupa od 100 primjera i modele trenirane a) samo na originalnim podacima (gore lijevo), b) samo na augmentiranim podacima gore desno), c) na kombinaciji originalnih i augmentiranih podataka (n=1) (dolje lijevo), d) na kombinaciji originalnih i augmentiranih podataka (n=3) (dolje desno). BT - „Backtranslation“.



Slika 6. Veličina vokabulara i pokrivenost vokabulara vektorima riječi skupa za učenje od 100 primjera. U gornjem prikazu uspoređuju se vokabulari originalnih podataka i njihovim transformacija, a u donjem dijelu kombinacije originalnih i augmentiranih podataka (omjer 1:3).

Ako se promatraju pojedine tehnike augmentacije, može se vidjeti da su one prilično ujednačene. Iznenađujuće, parafraziranje prevođenjem s engleskog na njemački i obratno, pokazalo je lošije rezultate od očekivanog. A kad se uzme u obzir da je generiranje takvih podataka trajalo 100-200 puta duže nego npr. KB augmentiranih, postavlja se upitnost primjene takve tehnike, barem u slučaju općenitih tekstova. S druge strane, s obzirom na to da kvaliteta augmentacije ovisi o odabranom modelu, ali i o odabranoj kombinaciji jezika, trebalo bi ispitati dodatne modele te kombinacije leksički udaljenijih jezika, npr. engleskog i japanskog, što bi se moglo pozitivno odraziti na raznovrsnost rečenica.

Slika 6 prikazuje odnos veličine vokabulara i pokrivenosti riječi iz vokabulara GloVe vektorima. S obzirom na to da su korišteni vektori koji prepoznaju samo 400000 riječi, to se odrazilo i na uspješnost tehnika augmentacije. Može se vidjeti da je najveći vokabular dobiven s augmentacijom KB, što je očekivano jer se kod takve augmentacije zamjene vrše znakovima a ne čitavim riječima. Tako se stvara veliki broj nepostojećih riječi koje neće biti pronađene među GloVe vektorima. S druge strane, CWE augmentacija je također stvarala vokabular veći od prosjeka, ali to se proširenje dobilo stvarnim riječima. Moguće da je ta dodatna raznolikost doprinijela nešto većoj točnosti klasifikacije. Iako, promatrajući WD augmentaciju, može se uočiti ispodprosječni vokabular, ali rezultat koji parira najboljem u slučaju kombinacije s originalnim podacima u omjeru 3:1.

V. ZAKLJUČAK

U radu su uspoređene odabrane tehnike augmentacije na primjeru klasifikacije teksta pri čemu su korišteni skupovi za učenje različitih veličina. Rezultati su pokazali da je augmentacija neupitno korisna na skupovima jako ograničene veličine, ali da se ta korist gubi s porastom broja dostupnih primjera za učenje. Također, postoji

granica iznad koje više nije preporučljivo dodatno povećavanje osnovnog skupa jer dobitak u performansama uglavnom ne kompenzira dodatno vrijeme potrebno za pripremu i učenje. Tehnike augmentacije, iako međusobno različite, pokazale su prilično ujednačene rezultate. Stoga se za male skupove općenitog sadržaja, osim jednostavnijih modela, može preporučiti i jednostavnije tehnike augmentacije. U ovom slučaju kompleksnije tehnike nisu opravdale dodatne zahtjeve za resursima. Međutim, određeno ograničenje istraživanja je oslanjanje na vektore riječi relativno malog kapaciteta. Stoga bi se za potpuniju sliku istraživanje trebalo proširiti i na ostale varijante reprezentacije teksta. Uz to, iako su u radu korišteni podaci na engleskom jeziku, odabrane tehnike augmentacije mogle bi se primijeniti i u slučaju drugih jezika, poput hrvatskog, uz pretpostavku da postoje odgovarajući resursi, kao što su to baze sinonima ili modeli za strojno prevođenje, a o kvaliteti navedenih resursa ovisit će i kvaliteta augmentacije. Osim toga, buduća istraživanja mogla bi se usmjeriti i na multimodalnu augmentaciju, što predstavlja još nedovoljno istraženo područje.

REFERENCE

[1] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 10684–10695.

[2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[3] J. Chen, D. Tam, C. Raffel, M. Bansal, and D. Yang, "An empirical survey of data augmentation for limited data learning in NLP," *arXiv preprint arXiv:2106.07499*, 2021.

[4] M. M. Bejani and M. Ghatee, "A systematic review on overfitting control in shallow and deep neural networks," Artificial Intelligence Review, pp. 1–48, 2021.

[5] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[6] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[7] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.

[8] S. I. Nikolenko, "Synthetic data for deep learning," arXiv preprint arXiv:1909.11512, 2019.

[9] S. Y. Feng *et al.*, "A survey of data augmentation approaches for NLP," *arXiv preprint arXiv:2105.03075*, 2021.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[11] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," *arXiv preprint arXiv:1708.04552*, 2017.

[12] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization," *arXiv preprint arXiv:1710.09412*, 2017.

[13] G. A. Miller, "WordNet: a lexical database for English," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[14] B. Li, Y. Hou, and W. Che, "Data augmentation approaches in natural language processing: A survey," *AI Open*, vol. 3, pp. 71–90, 2022.

[15] J. Wei and K. Zou, "Eda: Easy data augmentation techniques for boosting performance on text classification tasks," *arXiv preprint arXiv:1901.11196*, 2019.

[16] I. Raffaelli, M. Tadic, B. Bekavac, and Ž. Agic, "Building croatian wordnet," in *Proceedings of GWC*, 2008, pp. 349–360.

[17] W. Y. Wang and D. Yang, "That's so annoying!!!: A lexical and frame-semantic embedding based data augmentation approach to automatic categorization of annoying behaviours using# petpeeve tweets," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, 2015, pp. 2257–2563.

[18] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[19] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[20] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017.

[21] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.

[22] S. Kolokiyets Oleksandr andBethard and M.-F. Moens, "Model-portabilty experiments for textual temporal analysis," in *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*, 2011, vol. 2, pp. 271–276.

[23] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," *Advances in neural information processing systems*, vol. 28, 2015.

[24] R. Sennrich, B. Haddow, and A. Birch, "Improving neural machine translation models with monolingual data," *arXiv preprint arXiv:1511.06709*, 2015.

[25] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[26] P. He, X. Liu, J. Gao, and W. Chen, "Deberta: Decoding-enhanced bert with disentangled attention," *arXiv preprint arXiv:2006.03654*, 2020.

[27] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1746–1751.

[28] J. Schmidhuber, S. Hochreiter, and others, "Long short-term memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] A. Adhikari, A. Ram, R. Tang, and J. Lin, "Rethinking complex neural network architectures for document classification," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4046–4051.

[30] A. Ezen-Can, "A Comparison of LSTM and BERT for Small Corpus," arXiv preprint arXiv:2009.05451, 2020.

[31] I. Hrga and M. Ivasic-Kos, "Effect of Data Augmentation Methods on Face Image Classification Results," in *Proceedings of the 11th International Conference on Pattern Recognition Applications and Methods - Volume 1*, 2022, pp. 660–667.

[32] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," in *Proceedings of the ACL*, 2004.

[33] E. Ma, "NLP Augmentation." 2019. [Online]. Available: https://github.com/makcedward/nlpaug

[34] M. Fadaee, A. Bisazza, and C. Monz, "Data augmentation for low-resource neural machine translation," arXiv preprint arXiv:1705.00440, 2017.

[35] N. Ng, K. Yee, A. Baevski, M. Ott, M. Auli, and S. Edunov, "Facebook FAIR's WMT19 news translation task submission," arXiv preprint arXiv:1907.06616, 2019.

[36] Y. Zhang and B. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification," *arXiv preprint arXiv:1510.03820*, 2015.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

MACHINE LEARNING APPLICATIONS

Using Generated LIBS Data as a Base for Neural Network Architecture Development

J. Lehtonen, T. Aaltonen
Satakunta University of Applied Sciences, Pori, Finland
juuso.2.lehtonen@samk.fi, toni.aaltonen@samk.fi

***Abstract* - Laser-Induced Breakdown Spectroscopy (LIBS) is a rapid atomic spectroscopy technique used to measure the concentration of elements in samples. This research paper uses supervised deep learning with a LIBS database provided by National Institute of Standards and Technology (NIST) to find out what kind of neural network architecture can best predict element concentrations from metal alloy samples. Accuracy of predictions was used to evaluate the network architectures. Because collecting and storing real samples of metal alloys with a wide range of element concentrations requires a lot of resources, the network's ability to predict outside of training set's range of concentrations was also tested. The goal of the research was to find out if NIST database together with machine learning can be used to reduce the amount of real alloy samples needed for LIBS calibration.**

***Keywords* - LIBS; NIST; neural network; supervised learning; deep learning; regression**

I. Introduction

In this paper we explore what kind of neural network architecture is best suited to predict element concentrations from laser-induced breakdown spectroscopy (LIBS) data generated from the National Institute of Standards and Technology (NIST) Atomic Spectra Database (ASD) [1]. We also test if the network can predict concentrations outside of the training range. As the generated data is similar to measured data, we try to find the best kind of network without the need for large number of samples and measurements [2].

Several studies have been made with machine learning , LIBS and NIST data [3][4]. Many of the studies done earlier use partial least squares algorithms or support vector machines. NIST data hasn't been widely utilized in training the algorithms and is often used for selecting the spectral lines that are most interesting for the study.

II. Generated LIBS Spectra

The dataset used for the tests was acquired through NIST website using SimulatedLibs [5] with modifications to acquire the spectra faster. The modifications were using a session for the requests and using multiprocessing. Generated LIBS spectrum line intensities are determined by electron temperature, electron density and the elemental composition of the sample. When electron temperature and density stay static the spectral line intensities change linearly (Fig 1.) as a function of elemental concentration when the data is generated on the website [6]. Changing the composition affects the intensities but has little to no effect on the relative

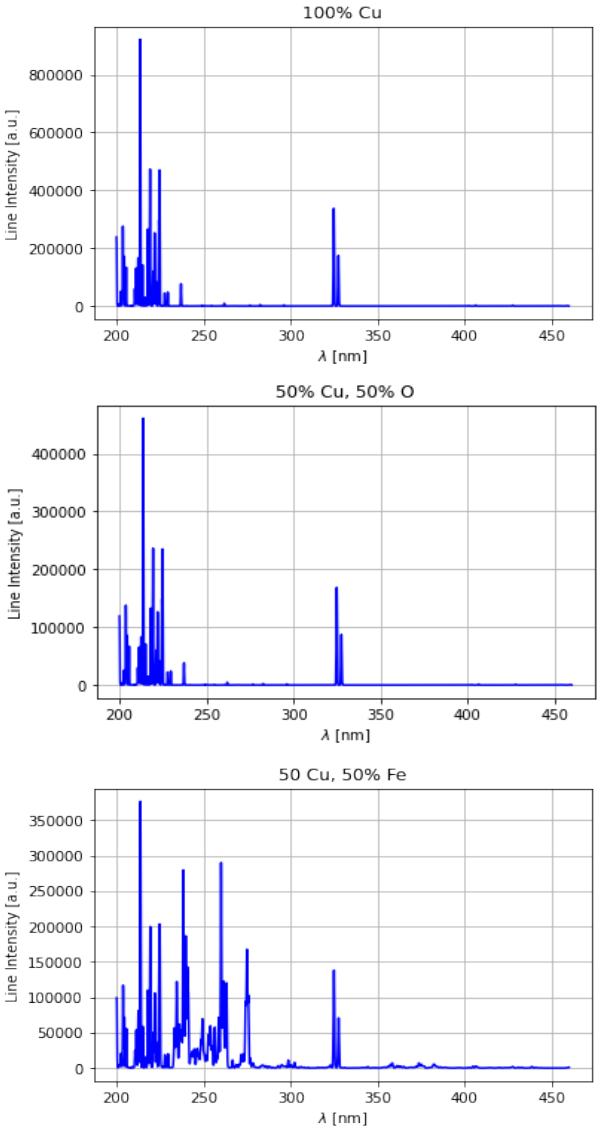


Figure 1: The relative line intensities stay the same when elemental composition changes unless the elements have overlapping lines. The line intensities from 100% copper (Cu) sample to 50% sample are halved and lines from oxygen (O) are not visible in the plot since they are less intense compared to Cu with the chosen electron temperature and density.

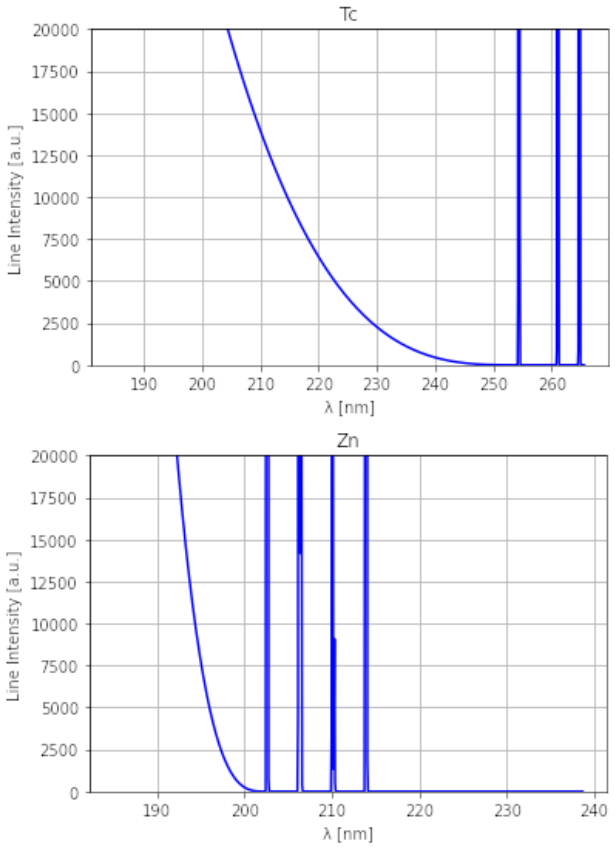


Figure 2: Some of the samples have erroneous data, possibly a result of cubic spline interpolation. The erroneous data is usually found at the start or end of the spectrum.

intensities between lines where elements have no overlapping spectral lines.

The generated dataset consists of 24 elements, C, Mg, Al, Si, Ti, V, Fe, Co, Ni, Cu, Mn, Li, Zn, O, P, Pt, Au, Ag, Ir, W, Pd, Sn, Tc, and Y, with combinations of concentrations from 100% to 1% with two elements, and combinations of three element compositions of [33, 34, 33], [40, 20, 40], [20, 40, 40], [40, 40, 20], [60, 20, 20], [20, 60, 20], [20, 20, 60], [80, 10, 10], [10, 80, 10], [10, 10, 80], [15, 30, 55], [30, 55, 15], [55, 15, 30]. The combinations add up to a total of 53,660 samples. The samples are of electron temperature 1 eV and electron density of 1e17 cm-3. The data was generated between 185 nm and 460 nm wavelengths, with a resolution of 0.067 nm, because a spectrometer used in our other experiments has an average resolution of 0.067 nm. The labels are generated are 24 numbers, one for each element, between 0 and 1.

The data has some abnormalities possibly due to cubic spline interpolation used in the generation process. Some of the samples have “tails” that are not spectral lines (Fig 2.). The data was not cleaned of these abnormalities prior to the tests, and they might affect the results.

III. Methodology

TensorFlow [7] was used to create and train the networks. The networks consisted of two fully-connected hidden layers of varying widths, input layer of 4094 and output layer of 24, one for each of the elements. The hidden layers use ReLU [8] activation function, and the output layer uses linear activation. The linear activation was chosen instead of softmax, because in real measurements there can be elements present which are not included in the datasets. A combination of 16, 32, 64, 128, 256, 512 widths for both hidden layers were tested, the 16 width layers were dropped from the tests after some time because the other sized networks consistently showed significantly better results. Different numbers of epochs and batch sizes were also tested. Epoch sizes from 5 to 40 with a step size of 5 were tested. Tested batch sizes were 2, 4, 8, 16, 32, 64, 128 and 256 [9]. Tests were also run with kernel and/or bias regularization applied to both hidden layers using default settings of TensorFlow. Altogether, more than 12800 networks were trained and tested. The initialization of weights used was the default in TensorFlow, glorot uniform [10]. Adam [11] with the default hyperparameters from TensorFlow was used as an optimizer. The data was randomly divided into 60:40 training:test sets. Mean squared error (MSE) was used to evaluate the networks. The networks were only trained once with each configuration of epochs, batch size and different width layers because of the time it takes to train each network and the number of combinations was already sizable, this results in randomness in measured MSE because of the randomness of weight initialization. Dataset was scaled to 0-1 range, which preserves the relative intensities of the peaks.

For testing how well a network can learn outside of the training range, the test set was changed to copper samples which had less than 70% of copper or more than 90% of copper and the rest of the data was used as a training set. This resulted in a training set with 48829 samples and test set of 4831 samples. This was done to see if a network could be trained to extrapolate outside of the range of real samples, which are difficult to get and store in large numbers.

Inference time of the networks were tested with 1000 samples picked randomly from 2000 samples. A prediction was made for each of the 1000 samples and the time it took to run inference was recorded. These tests were only done for the networks without bias and kernel regularization to reduce the time to run the tests since doing 1000 inferences for a network took approximately 45 to 60 seconds. The tests were run on a GPU-server using a NVIDIA RTX A5000, 10 cores of AMD Epyc 7413 processor and 50 GB random access memory.

IV. Results

The average inference time of 1000 independent predictions for the networks was 52 seconds with a standard deviation of 2.9 seconds, the fastest time being 41.7 seconds and the slowest 82.8 seconds as seen on table 1. The tests might have been affected by some

Table 1: The data sorted by inference time, with 5 of the fastest and slowest results showing. The three slowest times are clearly slower than the rest of the tests.

Epochs	Batch size	First layer	Second layer	Inference time (1000 samples)
5	2	16	16	41.681572 s
5	2	16	32	42.710497 s
5	2	16	64	44.343524 s
5	2	32	16	44.729965 s
5	2	16	256	44.830615 s
…	...	...	...	...
30	4	32	16	64.053025 s
30	4	64	128	64.097452 s
35	4	128	256	67.813859 s
20	128	32	32	72.505119 s
30	4	64	256	82.783849 s

unknown software or hardware issues. The inference time was tested to see if the size of the network matters for practical use. The inference tests were only run with networks not using regularization to save time. The larger networks achieve lower MSE than the smaller networks and benefit more from longer training as seen in table 2. The networks' ability to generalize to the test set also seems to benefit from having a larger batch size as the 100 lowest test MSE only have one network with less than 16 batch size.

Almost all the networks with kernel regularization perform worse than without it, but for most of the networks the test set MSE is lower than minimum and final loss from training. The minimum and final loss are often the same which could mean that the networks with kernel regularization could benefit from more training. The worst MSE achieved was 0.024501 as shown in table 3.

Testing the networks' ability to generalize outside of the training concentration showed that the network can predict outside of training range up to a point. The test was conducted with the best performing network from the first test. The training loss after 40 epochs was 0.00077266 and the test set MSE was 0.004. A considerably larger training set most likely helped to bring the training loss down. The closer the tested values are to the training range of 70-90% the better the predictions are. The largest error in predicting the copper concentration was 85 percentage points, which came from a sample with 10% copper, 80% carbon and 10% phosphorous. In the worst 20 predictions, 13 of the samples have copper and oxygen in them and the other 7 have copper with 2 other elements in them as shown in table 4. All the samples with only copper and oxygen in the bottom 20 predictions have 15% or less copper. 133 predictions out of 4831 are less than 1 percentage point off. When looking at the most accurate predictions for copper we can see that prediction of other elements in those samples was off tens of percentage points in some cases, as shown in table 5. Five of the best performing networks were trained again to see if the performance was consistent. Each of the five

Table 2: 20 networks with the lowest MSE. Larger networks and longer training achieve lower MSE with the test set.

Epochs	Batch size	First layer	Second layer	Kernel regularization	Bias regularization	Minimum loss during training	Final loss	Test set MSE
40	32	512	128	No	No	8.13e-4	8.18e-4	8.77e-4
40	256	512	256	No	No	8.12e-4	8.20e-4	9.15e-4
40	256	512	512	No	Yes	8.02e-4	8.02e-4	9.25e-4
35	64	512	128	No	No	8.33e-4	8.33e-4	9.31e-4
35	128	512	128	No	No	8.43e-4	8.43e-4	9.51e-4
35	128	512	256	No	No	8.03e-4	8.25e-4	9.58e-4
35	32	512	128	No	No	8.35e-4	8.35e-4	9.60e-4
40	64	128	512	No	Yes	8.29e-4	8.29e-4	9.63e-4
25	128	512	512	No	No	8.80e-4	8.80e-4	9.76e-4
40	16	128	512	No	No	8.93e-4	8.93e-4	9.78e-4
40	64	64	512	No	No	8.79e-4	8.83e-4	9.79e-4
35	64	128	512	No	Yes	8.68e-4	8.68e-4	9.80e-4
40	256	512	128	No	Yes	8.89e-4	8.89e-4	9.82e-4
25	128	512	512	No	Yes	8.77e-4	8.77e-4	9.84e-4
30	64	512	128	No	No	8.84e-4	8.96e-4	9.87e-4
30	32	512	512	No	No	8.51e-4	8.56e-4	9.88e-4
30	128	512	256	No	No	8.51e-4	8.51e-4	9.91e-4
40	32	512	256	No	No	7.94e-4	7.94e-4	9.92e-4
40	64	512	256	No	No	7.90e-4	7.90e-4	9.96e-4
25	32	512	512	No	No	9.10e-4	9.10e-4	9.97e-4

Table 3: Networks trained with batch size 2 or 4 perform consistently worse than networks trained with larger batch size.

Epochs	Batch size	First layer	Second layer	Kernel regularization	Bias regularization	Minimum loss during training	Final loss	Test set MSE
15	2	32	64	Yes	No	0.021380	0.021814	0.022002
5	2	512	256	Yes	Yes	0.021901	0.021957	0.022210
30	2	64	512	Yes	Yes	0.021889	0.027678	0.022454
25	2	128	512	Yes	Yes	0.021906	0.027999	0.022508
20	2	128	512	Yes	Yes	0.021904	0.025114	0.023366
25	2	128	256	Yes	Yes	0.021844	0.023886	0.024065
5	2	256	512	Yes	Yes	0.021934	0.021978	0.024501

networks were trained 20 times and the test set MSE was recorded for each run. In figure 3 we can see that the previously third best network is performing better than in the single training run while the order of the other networks has stayed the same. The overall MSE seems to be lower in the new tests, but the code was not changed.

V. CONCLUSION AND FURTHER RESEARCH

Predicting concentration from generated data with a neural network is possible, but gets harder for data that is further away from the provided training range. With the data tested the network size can be quite small, and the inference times are not affected much by the network size at such a small scale. Further research can be done with larger datasets and with varying electron temperature and density. In the future the data should also be cleaned from the anomalies caused by interpolation. The networks could benefit from more epochs. The large prediction error with some of the elements suggests that more careful selection of elements for the dataset could be beneficial when using the generated data as a base for the models used with real samples.

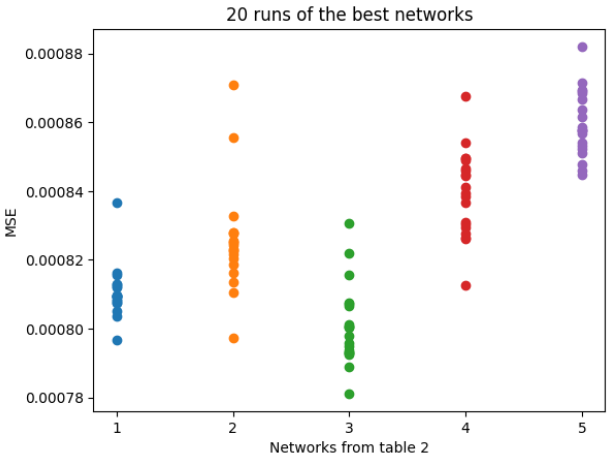


Figure 3: 5 of the best performing networks were trained 20 times each to see how much the results vary. X-axis shows the ranking of the networks from the table 2 tests. Blue dots were the best network in single training run, but the green dots show better results in the tests with multiple training runs.

Table 4: 20 of the predictions with the largest error for copper. Error is calculated by taking an absolute of model prediction of copper – copper label, first value in the element columns are label values and second values are the predictions, both multiplied by 100. Only elements included in the labels are shown.

Error	Cu	O	P	C	Si
0.852328	10 / 95	0 / -8	10 / 4	80 / 49	0 / -2
0.806924	10 / 91	0 / -5	80 / 16	10 / 12	0 / -1
0.799032	15 / 95	0 / -8	30 / 6	55 / 42	0 / -2
0.758974	1 / 77	99 / 21	0 / -3	0 / 0	0 / 1
0.749208	2 / 77	98 / 21	0 / -3	0 / 0	0 / 1
0.745582	20 / 95	0 / -6	40 / 7	40 / 25	0 / -1
0.745178	20 / 95	0 / -8	20 / 5	60 / 42	0 / -2
0.743142	20 / 94	0 / -6	60 / 9	20 / 13	0 /-1
0.739396	3 / 77	97 / 21	0 / -3	0 / 0	0 / 1
0.729482	4 / 77	96 / 21	0 / -3	0 / 0	0 / 1
0.719343	5 / 77	95 / 21	0 / -3	0 / 0	0 / 1
0.709479	6 / 77	94 / 21	0 / -3	0 / 0	0 / 1
0.699579	7 / 77	93 / 21	0 / -3	0 / 0	0 / 1
0.689445	8 / 77	92 / 21	0 / -3	0 / 0	0 / 1
0.679562	9 / 77	91 / 21	0 / -3	0 / 0	0 / 1
0.669623	10 / 77	90 / 21	0 / -3	0 / 0	0 / 1
0.659569	11 / 77	89 / 21	0 / -3	0 / 0	0 / 1
0.649558	12 / 77	88 / 21	0 / -3	0 / 0	0 / 1
0.642394	10 / 74	80 / 23	0 / -7	0 / -2	10 / 45
0.639570	13 / 77	87 / 21	0 / -3	0 / 0	0 / 1

Table 5: 5 of lowest absolute errors for copper prediction with the element label and prediction multiplied by 100. Only elements included in the labels are shown.

Error	Cu	Mg	Mn	Tc	Zn
1.00e-4	63 / 63.00	37 / 26.92	0 / -0.34	0 / -0.98	0 / -0.60
1.47e-4	64 / 63.99	0 / -0.61	36 / 28.13	0 / -2.04	0 / -1.58
2.92e-4	20 / 19.97	20 / 14.25	0 / 0.00	60 / 37.63	0 / -0.69
5.48e-4	61 / 61.05	0 / 0.00	0 / -0.11	0 / -0.99	39 / 29.56
7.54e-4	64 / 63.92	36 / 26.69	0 / -0.30	0 / -0.94	0 / -0.56

REFERENCES

[1] Kramida, A., Ralchenko, Yu., Reader, J. and NIST ASD Team, 2022, NIST Atomic Spectra Database (version 5.10), [Online]. Available: https://physics.nist.gov/asd [Fri Feb 03 2023]. National Institute of Standards and Technology, Gaithersburg, MD. DOI: https://doi.org/10.18434/T4W30F

[2] Poggialini, Francesco & Campanella, Beatrice & Legnaioli, Stefano & Pagnotta, Stefano & Raneri, Simona & Palleschi, Vincenzo. (2020). "Improvement of the performances of a commercial hand-held laser-induced breakdown spectroscopy instrument for steel analysis using multiple artificial neural networks", Review of Scientific Instruments. 91. 073111. 10.1063/5.0012669.

[3] Thomas F. Boucher, Marie V. Ozanne, Marco L. Carmosino, M. Darby Dyar, Sridhar Mahadevan, Elly A. Breves, Kate H. Lepore, Samuel M. Clegg, 2015, "A study of machine learning regression methods for major elemental analysis of rocks using laser-induced breakdown spectroscopy", Spectrochimica Acta Part B: Atomic Spectroscopy, DOI: https://doi.org/10.1016/j.sab.2015.02.003.

[4] Tianlong Zhang, Shan Wu, Juan Dong, Jiao Wei, Kang Wang, Hongsheng Tang, Xiaofeng Yang, Hua Li, 2015, "Quantitative and classification analysis of slag samples by laser induced breakdown spectroscopy (LIBS) coupled with support vector machine (SVM) and partial least square (PLS) methods", DOI: https://doi.org/10.1039/C4JA00421C

[5] MKastek, 2022, SimulatedLIBS v1.0.0 https://github.com/MKastek/SimulatedLIBS

[6] NIST, 03.02.2023, NIST LIBS Database, website source code lines:964,966 , https://physics.nist.gov/PhysRefData/ASD/LIBS/js/saha_lte.js

[7] Abadi et al., 2016, "TensorFlow: A system for large-scale machine learning", DOI: https://doi.org/10.48550/arXiv.1605.08695

[8] Raman Arora, Amitabh Basu, Poorya Mianjy, Anirbit Mukherjee, 2016, "Understanding deep neural networks with rectified linear units", DOI: https://doi.org/10.48550/arXiv.1611.01491

[9] Dominic Masters, Carlo Luschi, 2018, "Revisiting small batch training for deep neural networks", DOI: https://doi.org/10.48550/arXiv.1804.07612

[10] Boris Hanin, David Rolnick, 2018, How to Start Training: The Effect of Initialization and Architecture, DOI: https://doi.org/10.48550/arXiv.1803.01719

[11] Diederik P. Kingma, Jimmy Ba, 2014, "Adam: a method for stochastic optimization", DOI: https://doi.org/10.48550/arXiv.1412.6980

Imitation Learning for Financial Applications

Sven Goluža*, Tessa Bauman†, Tomislav Kovačević‡ and Zvonko Kostanjčar§,
University of Zagreb, Faculty of Electrical Engineering and Computing
Laboratory for Financial and Risk Analytics
Unska 3, 10000 Zagreb, Croatia
*sven.goluza@fer.hr, †tessa.bauman@fer.hr, ‡tomislav.kovacevic@fer.hr and §zvonko.kostanjcar@fer.hr

***Abstract*—Algorithms for solving decision-making problems under uncertainty are often employed in complex environments such as financial markets. Reinforcement learning (RL), a mathematical framework for sequential decision-making, is widely used in many financial decision-making problems, such as portfolio optimization, optimal execution, and market making. However, RL methods require a manual design of the reward function, which can be challenging in noisy market environments. Moreover, in such environments, it is often easier to demonstrate the desired behavior rather than manually engineer it. These problems can be solved using imitation learning (IL) algorithms, which extract knowledge from expert demonstrations by directly imitating the desired behavior or learning the expert's reward function. This paper introduces the foundational methods in IL, outlines the differences with familiar frameworks, and provides a survey of the current IL applications in finance.**

***Keywords*—imitation learning, reinforcement learning, quantitative finance, machine learning**

I. Introduction

With the advances in information technology and the ever-increasing amount of data, the number of potential applications using artificially-intelligent agents capable of mimicking human behavior and making decisions has increased rapidly in the financial industry. A popular approach to learning autonomous behavior is reinforcement learning (RL) [1]. RL agents learn by interacting with the environment through trial and error learning and receiving feedback for their actions, thus improving their behavior using their experience. Despite the recent success in many fields, especially with advances in deep reinforcement learning [2], [3], RL algorithms have shown shortcomings in complex and dynamic environments such as financial markets. Learning by interaction in the context of markets usually involves historical data, which is highly stochastic and noisy. Under these circumstances, data-driven methods frequently fail to generalize and are often less effective than experts. Moreover, manually designing a reward function proves to be a sensitive issue as it may require a significant amount of experiments and feature engineering [4]. All of the mentioned problems have elevated the study of imitation learning (IL) [5].

Imitation learning (also known as apprenticeship learning or learning from demonstrations) deals with learning behaviors from expert demonstrations. Historically, it is an interdisciplinary research area with multiple existing taxonomies stemming from robotics, dynamic systems, and other engineering-based methods [6]. It has been applied to control domain problems such as navigation [7] and robotic manipulation [8]. Progress in high-dimensional computation and machine learning, particularly deep learning, has given rise to potential applications of AI-based IL methods in the financial industry.

Traditionally, the two main branches of study within IL are behavioral cloning [9] and inverse reinforcement learning [10]. Behavioral cloning (BC) methods learn to perform a task by directly replicating observed expert behavior. They are closely related to a traditional supervised learning setting as they learn a mapping from features (states) to labels (actions). On the other hand, inverse reinforcement learning (IRL) methods learn a reward function from the observed expert behavior, leveraging the learned reward function later on through RL. However, due to a rise in research on generative adversarial networks [11], and their connection to IRL [12], a new IL branch based on the generative adversarial imitation learning algorithm [13] started to develop. Given that this algorithm does not belong to either BC or IRL (as it follows the IRL structure but does not recover the reward function), researchers nowadays tend to form an additional IL branch often referred to as adversarial imitation learning (AIL) [14]. Even though IRL and AIL are closely related, this paper will follow the newly formed taxonomy of three separate branches.

II. Reinforcement learning

Reinforcement learning is a machine learning approach for learning optimal decisions by interacting with an environment in a trial and error fashion. The environment dynamics are described using Markov Decision Processes (MDPs). Formally, an MDP is a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \gamma, \mathcal{R}, p_0 \rangle$:

- $\mathcal{S}$ is a set of states,
- $\mathcal{A}$ is a set of actions,
- $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \mapsto [0, 1]$ is a transition probability matrix,
- $\gamma \in [0, 1]$ is a discount factor,
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$ is a reward function,
- p_0 is an initial state distribution.

At each step t, the agent makes an action A_t based on the environment's current state observation S_t. Performing an action A_t in state S_t determines the next state S_{t+1} and the reward R_{t+1}. This iterative process is repeated, and sequences of states, actions, and rewards are acquired $(S_0, A_0, R_1, S_1, A_1, R_2, ..., S_t, A_t, R_{t+1})$. The

state-action sequence is referred to as a trajectory. MDP states have Markov property, meaning that the next state only depends on the current state and the current action, not on the history of all the states and actions that were taken before.

An agent's behavior is determined by its policy π, which can either be a deterministic or stochastic function. RL is based on the reward hypothesis, which states that all goals can be described by the maximization of cumulative reward, which is mathematically expressed as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}.$$

The agent's primary goal is to find a policy π^* that maximizes the expected return $\mathbb{E}[G_t]$, i.e., the expected cumulative reward over all future time steps, for each given state:

$$\pi^* = \arg\max_{\pi \in \Pi} \mathbb{E}[G_t | S_t = s], \forall s \in \mathcal{S}.$$

Any data-driven method well suited to solving problems formulated in terms of an MDP and yielding an optimal policy can be considered an RL method. Numerous RL methods are broadly categorized into value-based, policy-based, and actor-critic. Another categorization is based on the use of model dynamics and transition probabilities during training: model-free and model-based methods. Further information on RL algorithms can be found in [1].

RL and IL algorithms handle decision-making tasks and differ from supervised learning in terms of sequential and non-i.i.d data. As discussed in Section I, IL is a useful tool for dealing with the limitations of RL in complex environments. It is a common assumption that having prior knowledge about a problem is more effective than learning solely through trial and error. IL methods incorporate prior knowledge about a specific problem by leveraging experts and collecting their demonstrations as a set of trajectories. IL approach assumes this is a more effective way of addressing behavioral learning, as it's often easier to demonstrate the optimal behavior for a task than mathematically formulate it.

III. Imitation Learning

Imitation learning is an approach for learning the desired behavior from expert demonstrations. This approach is helpful in scenarios where it is more practical for an expert to demonstrate the ideal behavior rather than compute the optimal policy or define the reward function manually. An expert can either be a human or any agent capable of optimally performing a task.

A. Behavioral Cloning

Behavioral cloning methods are direct IL methods that learn a mapping from states to actions without recovering the reward function [9]. These methods are formulated as a supervised learning problem where the dataset is a set of state-action pairs $\mathcal{D} = \{(s_i, a_i)\}_{i=1}^N$ obtained from the expert, and the goal is to find the optimal policy π^* [15]. The policy refers to a regressor or classifier here. BC methods are efficient because of their ability to learn without interacting with the environment, as they are not formulated as an MDP.

Algorithm 1: Dataset Aggregation (DAgger)

1: Initialize $\mathcal{D} \leftarrow \emptyset$, initialize $\hat{\pi}_1$ to any policy in Π
2: **for** iteration $i = 1$ to N **do**:
3: Let $\pi_i = \beta_i \pi^* + (1 - \beta_i)\hat{\pi}_i$
4: Sample T-step trajectories using π_i
5: Get dataset $\mathcal{D}_i = \{(s, \pi^*(s))\}$ of visited states by π_i and actions given by expert
6: Aggregate datasets: $\mathcal{D} \leftarrow \mathcal{D} \cup \mathcal{D}_i$
7: Train classifier $\hat{\pi}_{i+1}$ on $\mathcal{D}$
8: **endfor**
9: **return** best $\hat{\pi}_i$ on validation set

There are some tradeoffs, however, as supervised learning methods assume actions in the expert's trajectories are i.i.d. BC methods violate that property because actions (labels) influence future states (features) and thus influence future data distribution. This distribution mismatch problem leads to compounding errors when the agent finds himself in a state he has never been to before. Ideally, at each step t, the agent finds himself in a state drawn from the state distribution of expert trajectories ($s_t \sim p_{\pi^*}(s_t)$). In the setup described above, the agent will find himself in the state distribution resulting from taking action induced by the learned policy in the previous step ($s_t \sim p_{\pi_\theta}(s_t)$). A supervised approach to imitation learning disables the agent from learning to recover from failures and fails to generalize to unseen scenarios.

BC methods tend to solve this problem by changing $p_{\pi_\theta}(s_t)$ using data augmentation techniques. While using standard data augmentation techniques such as synthetic data and expanded state spaces helps, most BC methods build on specific data augmentation technique which interactively queries the experts in additional data points resulting from applying the current policy [16]. The idea is to start from an expert policy and iteratively and slowly replace it with the learned policy. Using this approach, the agent is retraining under the new state distribution after each policy change, and thus BC algorithms reduce imitation problem to supervised learning.

Dataset Aggregation (DAgger) [16] is an iterative BC algorithm that learns a deterministic policy using the dataset aggregation technique. It is shown in detail in Algorithm 1. This data-efficient approach combines the presence of the expert with expanding the space of possible trajectories to improve performance in unseen scenarios and recover from failure. The main drawbacks of the algorithm are frequent expert querying and the computational burden of iteratively updating the agent's policy. Later research papers such as [17]–[19] were proposed to improve DAgger.

B. Inverse Reinforcement Learning

Inverse reinforcement learning methods are indirect IL methods that learn a reward function from expert demonstrations [20]. As the reward hypothesis states that all goals can be described by the reward function formulation, RL assumes that the reward function, not policy, is the most robust and transferable definition of the learning task [21]. IRL methods describe the environment using an MDP, and the input to an IRL method consists of a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \gamma, p_0, \mathcal{D} \rangle$, where $\mathcal{D} = \{\tau_1, \tau_2, ..., \tau_n\}$ is a set of expert demonstrations, and τ_i represents a trajectory. The goal is to recover the reward function $\mathcal{R}$. All IRL methods follow the same iterative template, which allows for the simultaneous improvement of the learned policy and the learned reward function [20]:

1) Collect expert demonstrations $\mathcal{D} \sim \pi^*$.
2) Initialize the reward function r_w.
3) Learn the policy π^L by using RL with respect to the current r_w.
4) Update parameters w to minimize the divergence between the expert policy π^* and the learned policy π^L.
5) Repeat the previous two steps until divergence is reduced to the desired level.

The problem of determining the reward function is ill-posed because a policy can be optimal for multiple reward functions [4]. One of the first IRL algorithms which addressed the solution ambiguity is presented in [21]. The authors assume the reward function is linear in parameters and use the concept of feature expectations to resolve the ambiguity. Feature expectations are the expected value of state (and/or action) features, defined in the reward function, accumulated over the trajectory. Furthermore, they are estimated by sampling trajectories and used to express the value of a policy. The authors find a policy π^L and its corresponding r_w such that the difference between the value of the expert policy π^* and π^L is smaller than a predefined margin ϵ. IRL methods that optimize different margins are referred to as the maximum margin methods [4], [21], [22]. Minimizing the margin based on feature expectations is a heuristic approach and therefore has significant drawbacks when expert demonstrations are suboptimal. This leads to the case where no single reward function makes the expert policy both optimal and better than all the other policies.

To tackle these issues, [23] introduced the method based on the probabilistic model of behavior. This method chooses a distribution $p(\tau|w)$ that maximizes the entropy among the distributions that match the feature expectations of the demonstrator. Maximum entropy method builds on maximum margin methods and alleviates some of their fallbacks. However, this method still requires solving the full RL problem for optimal policy in the inner loop, carefully designing features to impose the structure on the reward function, and a model of the environment. To apply these methods in real-world settings, especially in finance, the problem of learning the expressive reward function under unknown dynamics needs to be solved.

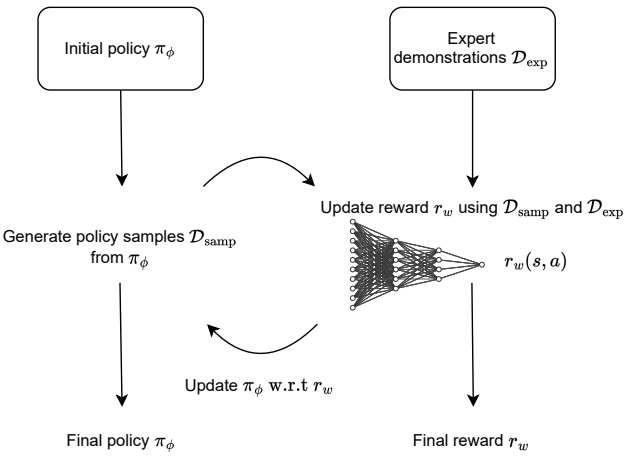


Fig. 1: Guided cost learning (GCL) algorithm

Guided cost learning (GCL) algorithm was proposed in [24], combining sample-efficient handling of unknown dynamics with neural network representation of reward function to handle existing problems in the literature. The visual guide to the algorithm is given in Fig. 1. Later research like [25]–[27] explored ideas such as ranking the expert demonstrations and integrating IRL methods with other methods such as self-supervised learning.

Most of the problems found in BC and IRL methods (robustness, unknown dynamics, suboptimal demonstrations) have been tackled with generative adversarial imitation learning (GAIL) [13], which combines the robustness of IRL and the efficiency of BC to improve IL methods further.

C. Adversarial Imitation Learning

The connection between IRL and GANs was presented in [12], where authors conclude that sample x represents a trajectory τ, generator G represents an agent's policy $\pi^L \sim p(\tau)$, and discriminator D is used as a proxy for reward [12].

GAIL algorithm [13] leverages the mentioned connection between IRL and GANs by training a policy that reproduces the expert behavior and a discriminator that distinguishes trajectories induced by the agent's policy from trajectories demonstrated by the expert. GAIL fits a parameterized policy π^L_θ, and a discriminator network $D_w : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$. The discriminator network represents the cost function, the negative of the reward function. The iterative algorithm is shown in detail in Algorithm 2. At the end of an iterative process, GAIL does not yield the reward function and, as such, can not be considered an IRL method. Ever since the introduction of GAIL, the research community has made further proposals on building upon it and applying it to real-world scenarios [28]–[33].

Algorithm 2: Generative Adversarial Imitation Learning (GAIL)

1: **Input:** expert trajectories $\tau_{\text{exp}} \sim \pi^*$, initial policy and discriminator parameters θ_0, w_0;
2: **for** iteration $i = 1$ to N **do**:
3: Sample trajectories $\tau_i \sim \pi^L_{\theta_i}$
4: Update discriminator parameters from w_i to w_{i+1} with gradient

$$\mathbb{E}_{\tau_i}[\nabla_w \log(D_w(s,a))] + \mathbb{E}_{\tau_{\text{exp}}}[\nabla_w \log(1 - D_w(s,a))]$$

5: Update policy parameters from θ_i to θ_{i+1}, using the TRPO rule with cost function $\log(D_{w_{i+1}}(s,a))$
6: **endfor**
7: **return** optimized policy parameters θ

IV. Applications in Finance

Imitation learning applications in finance can be broadly classified into identifying investment behavior and learning investment strategies. Important mechanisms that drive financial markets, such as price discovery, market liquidity, and order flow, can be explained by studying investment behavior and evaluating its economic impact. Those studies can also be of interest to market regulators and operators to counter possible fraudulent practices and maintain the transparency and functioning of financial markets. The other group of IL applications in finance considers the problem of investment strategy design across different assets and periods. They do so by observing trading history data from experts, which may be human (such as fund managers or prop traders) or an oracle with complete information (access to past and future data). Although the ML literature has recently become more prominent in finance, the applications of IL in finance have yet to be studied in depth [34].

One of the first applications was the identification of high-frequency trading strategies based on order book data [35]. The authors collected up to a month of order book audit trail data from the CME exchange for the E-Mini S&P 500 index futures. They modeled trader behavior as an MDP and used order book trails as demonstration data to infer traders' policies and reward functions. The state space was defined by inventory positions and order volume imbalance at the first five levels of the limit order book (LOB). Action space was defined by three actions (placing a limit order (LMT), canceling an existing limit order (CXL), and placing a market order (MKT)) directed at any of the ten buckets that LOB was divided into. The authors used model-based IRL methods which estimated the state transition probability matrix. Reward functions offered useful features for supervised learning, such as the classification of traders, and unsupervised learning, such as the categorization of traders. They introduced the gaussian process IRL (GPIRL) method to learn traders' reward functions and compared it to linear IRL (LNIRL) from [4] in terms of trading strategy identification. They used an SVM classifier to identify traders based on reward functions recovered from mentioned IRL methods and showed that GPIRL outperforms LNIRL. The authors showed that IRL methods are superior to statistical approaches, as they aim to learn the objectives of traders in various market conditions. That makes IRL methods more informative and robust and provides a solid foundation for identifying the behavioral patterns of traders.

In contrast to previous research that relied on labeled market participant data, [36] aimed to model traders' behavior using only unlabeled market data. The authors proposed a neural network-based multi-modal IL model which performs latent segmentation of stock trading strategies by their reward function. Each segment represents an individual strategy and has its reward function. An IL model is defined for each segment and trained to predict which trading segment was most likely to have submitted a particular order at a particular time. Finally, the model predicts order submission probability by combining outputs from each segment. State space features included price series and order book features at ten levels above and below the mid-market price. Action space consisted of order type (LMT, MKT, CXL), discretized order price, and volume. Their proposed model outperformed other IL benchmarks, such as GAIL, in terms of accuracy on historical order book data obtained from the Tokyo Stock Exchange. Thus, it provides a meaningful interpretation of traders' behavior. The limitation of their study is that the segment classification ability depends significantly on the formulation of the reward function.

In [37], the authors designed an investor sentiment trading system using IRL. They modeled the financial market dynamics as MDP and treated investor sentiment as actions in different market states. They extracted the reward function using GPIRL and leveraged the learned reward function in supervised learning to construct a trading strategy. The hypothesis was that the market sentiment reward function could be an effective feature space for predicting future financial market trends as it filters out irrelevant news sentiment signals. News sentiment from the Thomson Reuters news analytics database was used to proxy investor sentiment toward the U.S. market. State variables included discretized log return and volatility to describe the financial market conditions. At the same time, action space consisted of investor sentiment shocks to the market, categorized into three categories: positive, neutral, and negative. From an academic standpoint, this research offers a novel method for uncovering the underlying connection between the financial market and investor sentiment. From a practical standpoint, this research shows that an IRL-based trading strategy is superior to benchmark strategies based on S&P 500 index and market-based ETFs, as well as a few other news sentiment-based strategies.

In [38], a trading strategy combining BC and RL was proposed. The authors modeled the market, consisting of minute-level frequency data, as a partially observable MDP (POMDP). They employed a recurrent deterministic

policy gradient (RDPG) model for solving the POMDP. Observation space features consisted of cumulative account profits, price series, and technical indicators. The action space was defined as a continuous probability vector corresponding to taking the trading positions, but actions (short/long) are discretized to $\{-1, 1\}$. Differential Sharpe ratio was used as the reward function. The model was improved with BC by using the dual thrust strategy as an oracle, having access to all historical data. This IL technique enhances the trading agent's financial domain knowledge. BC was added to the model by incorporating two separate modules. Firstly, the agent was pre-trained using demonstrations from the dual thrust strategy, and secondly, the intra-day oracle was introduced. The oracle takes a long position at the lowest price and a short position at the highest price in a given day. Its actions were incorporated into the policy gradient to reduce the inefficient exploration phase. The model was backtested on IF and IC futures data, representative of stock index futures in China, and outperformed other benchmarks (long-only, short-only, and DDPG) in terms of Sharpe ratio and total return. The authors demonstrated that the BC modules significantly enhance the model's profitability metrics through an ablation study, revealing the impact of each component on performance. The final study demonstrates the model's ability to generalize by evaluating its performance on markets different from those it was trained on.

Goal-based wealth management problem (optimization of retirement plans or target-dated funds) was explored in [39] with IRL and RL. G-learner, the algorithm based on G-learning (probabilistic extension of Q-learning), was used in the RL wealth management problem. In contrast, the GIRL (G-learning IRL) model was used to infer the rewards of financial agents. The state space was defined as the vector of dollar values of asset positions and the action space as the vector of dollar-valued changes in asset positions. The authors defined a parameterized reward function that imposes a penalty for the underperformance of a portfolio relative to a target, along with regularization terms. The GIRL algorithm has been used to infer the reward function parameters, given a history of positions in an investment portfolio and an agent's allocation decisions. GIRL imitates the G-learner by minimizing a loss function over the state-action trajectories generated by the G-learner. Simulated models of equity returns were utilized to showcase the application of the G-learner and GIRL algorithms in goal-based wealth management. The results show that the G-Learner outperforms the benchmark equally-weighted portfolio strategy and that GIRL's learned parameters are close to the original G-learner parameters.

A combination of IRL and RL for optimal asset allocation was proposed in [40]. The IRL model learned the investment strategies of portfolio managers using their trading histories, and the RL model was used to improve those strategies. The T-REX algorithm [27] is utilized in the IRL step, incorporating the portfolio performance ranking of managers into its optimization process and eliminating the assumption of an optimal or near-optimal policy in the demonstrated trading history. The state and action vectors are the same as in [39]. The reward function is a quadratic function of the state and action, consisting of only four parameters due to the limited data available in portfolio management applications. It incorporates penalty for under-performance of the traded portfolio relative to its moving target, transaction costs, and trades constraint. The recovered reward function was used in the G-learner algorithm to further optimize the investment policy. The dataset included 18 mutual funds' data with monthly holdings, trades, and cashflows, but due to data constraints, all stocks within each portfolio have been grouped by industry sector. Experiments showed that the combination of IRL and RL outperforms the majority of fund managers. Additionally, inferred reward parameters can be used to group different trading strategies and gain a deeper understanding of fund managers' actions.

In [41], the authors introduced an RL framework for optimal order execution. They improved noisy and imperfect data in their RL framework by using the learning-based oracle teacher with access to all historical data. Model over-fitting was prevented by training universal strategy on various instruments. State space consisted of the elapsed time, remaining inventory to execute, price, and volume of each time step. Action corresponded to the discrete proportion of the target order, and the reward function consisted of trading profitability and market impact penalty. Two agents were being trained: a student and a teacher. The teacher agent, who has access to perfect information, learns a policy that acts as an oracle. This policy then guides the learning of the student agent, which is the agent used in real-time order execution. Finally, the student agent is tested and compared to benchmarks in the China A-shares market using a dataset that includes minute-level market information. Experiments showed that the universal IL model outperforms all the other models and that teacher guidance significantly boosts the model's performance.

V. Conclusion

Machine learning methods, combined with growth in computational power and the amount of data, have enabled significant development of imitation learning in the last decade. IL algorithms aim to learn an optimal policy given expert demonstrations, and they were shown to be effective in financial applications. IL methods were found to outperform other statistical approaches in identifying investment behavior. IL also learned investment strategies across different assets and periods, outperforming baselines such as market indices and machine learning based models. These presented methods can serve as a future benchmark in developing IL models. Selection of a labeling algorithm as an oracle, investors' suboptimal behavior, insufficient training data, and feature representation are some of the open IL research points in quantitative finance.

ACKNOWLEDGMENT

This work was supported in part by the Croatian Science Foundation under Project 5241, and European Regional Development Fund under Grant KK.01.1.1.01.0009 (DATACROSS).

REFERENCES

[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[3] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, "Mastering the game of go without human knowledge," *nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[4] A. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proceedings of the Seventeenth International Conference on Machine Learning*, vol. 0, 2000.

[5] A. Hussein, M. M. Gaber, E. Elyan, and C. Jayne, "Imitation learning: A survey of learning methods," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.

[6] T. Osa, J. Pajarinen, G. Neumann, J. A. Bagnell, P. Abbeel, J. Peters *et al.*, "An algorithmic perspective on imitation learning," *Foundations and Trends® in Robotics*, vol. 7, no. 1-2, pp. 1–179, 2018.

[7] F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 4693–4700.

[8] B. Fang, S. Jia, D. Guo, M. Xu, S. Wen, and F. Sun, "Survey of imitation learning for robotic manipulation," *International Journal of Intelligent Robotics and Applications*, vol. 3, pp. 362–369, 2019.

[9] M. Bain and C. Sammut, "A framework for behavioural cloning," in *Machine Intelligence 15*, 1995.

[10] S. Russell, "Learning agents for uncertain environments," in *Proceedings of the Annual ACM Conference on Computational Learning Theory*, 1998, pp. 101–103.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[12] C. Finn, P. Christiano, P. Abbeel, and S. Levine, "A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models," *arXiv preprint arXiv:1611.03852*, 2016.

[13] J. Ho and S. Ermon, "Generative adversarial imitation learning," in *Advances in Neural Information Processing Systems*, 2016.

[14] B. Zheng, S. Verma, J. Zhou, I. W. Tsang, and F. Chen, "Imitation learning: Progress, taxonomies and challenges," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–16, 2022.

[15] S. Ross and D. Bagnell, "Efficient reductions for imitation learning," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2010, pp. 661–668.

[16] S. Ross, G. J. Gordon, and J. A. Bagnell, "No-regret reductions for imitation learning and structured prediction," *Aistats*, vol. 15, pp. 627–635, 2011. [Online]. Available: http://proceedings.mlr.press/v15/ross11a/ross11a.pdf

[17] S. Ross and J. A. Bagnell, "Reinforcement and imitation learning via interactive no-regret learning," *arXiv preprint arXiv:1406.5979*, 2014.

[18] H. Le, A. Kang, Y. Yue, and P. Carr, "Smooth imitation learning for online sequence prediction," in *International Conference on Machine Learning*. PMLR, 2016, pp. 680–688.

[19] M. Laskey, J. Lee, R. Fox, A. Dragan, and K. Goldberg, "Dart: Noise injection for robust imitation learning," in *Conference on robot learning*. PMLR, 2017, pp. 143–156.

[20] S. Arora and P. Doshi, "A survey of inverse reinforcement learning: Challenges, methods and progress," *Artificial Intelligence*, vol. 297, p. 103500, 2021.

[21] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*, pp. 1–8, 2004.

[22] N. D. Ratliff, J. Andrew Bagnell, and M. A. Zinkevic, "Maximum margin planning," in *ACM International Conference Proceeding Series*, vol. 148, 2006, pp. 729–736.

[23] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 3, 2008.

[24] C. Finn, S. Levine, and P. Abbeel, "Guided cost learning: Deep inverse optimal control via policy optimization," in *33rd International Conference on Machine Learning, ICML 2016*, vol. 1, 2016, pp. 95–107.

[25] A. Nair, D. Chen, P. Agrawal, P. Isola, P. Abbeel, J. Malik, and S. Levine, "Combining self-supervised learning and imitation for vision-based rope manipulation," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 2146–2153.

[26] S. Reddy, A. D. Dragan, and S. Levine, "SQIL: Imitation Learning via Reinforcement Learning with Sparse Rewards," *arXiv preprint arXiv:1905.11108*, may 2019. [Online]. Available: https://arxiv.org/abs/1905.11108v3

[27] D. S. Brown, W. Goo, P. Nagarajan, and S. Niekum, "Extrapolating beyond suboptimal demonstrations via inverse reinforcement learning from observations," in *36th International Conference on Machine Learning, ICML 2019*, vol. 2019-June, 2019.

[28] J. Fu, K. Luo, and S. Levine, "Learning Robust Rewards with Adversarial Inverse Reinforcement Learning," *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, oct 2017. [Online]. Available: https://arxiv.org/abs/1710.11248v2

[29] Y. Li, J. Song, and S. Ermon, "Infogail: Interpretable imitation learning from visual demonstrations," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[30] Z. Wang, J. S. Merel, S. E. Reed, N. de Freitas, G. Wayne, and N. Heess, "Robust imitation of diverse behaviors," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[31] Y. Ding, C. Florensa, M. Phielipp, and P. Abbeel, "Goal-conditioned imitation learning," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[32] W. Sun, A. Vemula, B. Boots, and D. Bagnell, "Provably efficient imitation learning from observation alone," in *International conference on machine learning*. PMLR, 2019, pp. 6036–6045.

[33] G. Zuo, K. Chen, J. Lu, and X. Huang, "Deterministic generative adversarial imitation learning," *Neurocomputing*, vol. 388, pp. 60–69, 2020.

[34] M. F. Dixon, I. Halperin, and P. Bilokon, *Machine learning in Finance*. Springer, 2020, vol. 1170.

[35] S. Y. Yang, Q. Qiao, P. A. Beling, W. T. Scherer, and A. A. Kirilenko, "Gaussian process-based algorithmic trading strategy identification," *Quantitative Finance*, vol. 15, no. 10, pp. 1683–1703, 2015.

[36] I. Maeda, D. DeGraw, M. Kitano, H. Matsushima, K. Izumi, H. Sakaji, and A. Kato, "Latent segmentation of stock trading strategies using multi-modal imitation learning," *Journal of Risk and Financial Management*, vol. 13, no. 11, p. 250, 2020.

[37] S. Y. Yang, Y. Yu, and S. Almahdi, "An investor sentiment reward-based trading system using gaussian inverse reinforcement learning algorithm," *Expert Systems with Applications*, vol. 114, pp. 388–401, 2018.

[38] Y. Liu, Q. Liu, H. Zhao, Z. Pan, and C. Liu, "Adaptive quantitative trading: An imitative deep reinforcement learning approach," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 02, 2020, pp. 2128–2135.

[39] M. Dixon and I. Halperin, "G-learner and girl: Goal based wealth management with reinforcement learning," *arXiv preprint arXiv:2002.10990*, 2020.

[40] I. Halperin, J. Liu, and X. Zhang, "Combining reinforcement learning and inverse reinforcement learning for asset allocation recommendations," *arXiv preprint arXiv:2201.01874*, 2022.

[41] Y. Fang, K. Ren, W. Liu, D. Zhou, W. Zhang, and Y. Yu, "Imitation from learning-based oracle for universal order execution in quantitative finance," 2021, unpublished.

The Comparison of Different Feature Extraction Methods in Musical Instrument Classification

N. Rodin*, D. Pinčić*, K. Lenac*, D. Sušanj*
* University of Rijeka, Faculty of Engineering, Rijeka, Croatia
dsusanj@riteh.hr

***Abstract*—In this paper, we analyze four different methods for audio feature extraction and compare their efficiency in the context of musical instrument classification. We study spectrograms, Mel spectrograms, Linear-Frequency Cepstral Coefficients (LFCCs) and Mel-Frequency Cepstral Coefficients (MFCCs) in combination with three different Deep Learning architectures: VGG-16, ResNet-34 and a custom CNN. We investigate the behavior of our models in two different classification scenarios to determine a possible correlation between the number of classes and the efficiency of each method. For this purpose, we took samples from the London Philharmonic Orchestra dataset and ran the experiment for three and fifteen classes of musical instruments belonging to three different instrument families: Woodwinds, Strings and Brass.**

***Keywords*—Musical Instrument Classification, CNN**

I. INTRODUCTION

Music has been a fundamental part of human identity and culture since prehistoric times. Over thousands of years, musical instruments have evolved to a high level of diversity and intricacy, aiming to create unique sounds and trigger specific emotional responses in listeners. From an analytic point of view, the skill of distinguishing between different instruments playing together in a complex musical piece carries an important role. Inspired by this fascinating task, our paper explores different ways of capturing audio features of classical instruments and categorizing them into correct instrument classes. A carefully selected preprocessing method combined with a suitable Deep Learning architecture establishes a foundation for more complex tasks such as music recommendation systems, melody extraction, and music information retrieval. Having that in mind, the goal of this study is to provide an overview of several preprocessing methods and their performance in different classification environments.

In the past years, audio feature extraction and signal processing have become more popular and more researched topics. There have been several studies that tried to compare the efficiency of the existing audio preprocessing methods. J. D. Deng *et al.* [1] analyzed three types of audio features used in musical instrument classification, including perception-based features, MPEG-7 timbral features, and MFCC features. They were compared using different machine learning techniques. In the end, the results showed that MFCC features gave the best classification performance. In addition, A. Eronen [2] compared Linear Prediction Cepstral Coefficients (LPCCs) with MFCCs and concluded that MFCCs gave the best accuracy results when dealing with instrument family classification.

Recently, many audio classification researchers have been using MFCCs as their feature extraction method of choice. This is not only common in studies dealing with musical instrument classification [3], [4], but also fields like speech recognition [5], heart sound classification [6], and animal sound classification [7]. S. K. Mahanta *et al.* [8] combined MFCCs with an artificial neural network and got 97% accuracy on the full London Philharmonic Orchestra dataset. Similar experiment was conducted by S. Prabavathy *et al.* [9] when they used MFCCs in combination with a convolutional neural network to reach the accuracy of 97.5%. Even though MFCCs seem to be more popular, some works such as [10] still opt for Mel spectrograms. Interestingly, R. Profeta and G. Schuller [11] managed to design a new CNN autoencoder-based filter bank that outperformed spectrograms and Mel spectrograms.

Unlike other works, our study encompasses the analysis of four common feature extraction methods that are based on the short-time Fourier transform. Another novelty is the use of multiple CNN architectures and different-sized datasets for comparison.

II. DATASET

For this study, we decided to use a slightly modified version of the London Philharmonic Orchestra dataset [12]. The original dataset consists of 20 instrument classes, including standard orchestral instruments, guitar, mandolin, banjo, saxophone, and various percussion instruments. Each class contains a number of high-quality audio samples, ranging from 74 (banjo) up to 1502 samples (violin).

Considering the significant imbalance of data samples per class, we decided to focus on standard orchestral instruments, thus eliminating the classes with insufficient number of samples. This resulted in a dataset of 15 classes, containing only traditional instruments found in classical orchestral music, with a total of 12 541 audio samples. The distribution of samples in the modified dataset can be observed in Fig. 1.

Seeing that all of the examined instruments belong to three instrument families (Woodwinds, Strings and Brass),

This research was supported by University of Rijeka, Rijeka, Croatia under the grant uniri-tehnic-18-295.

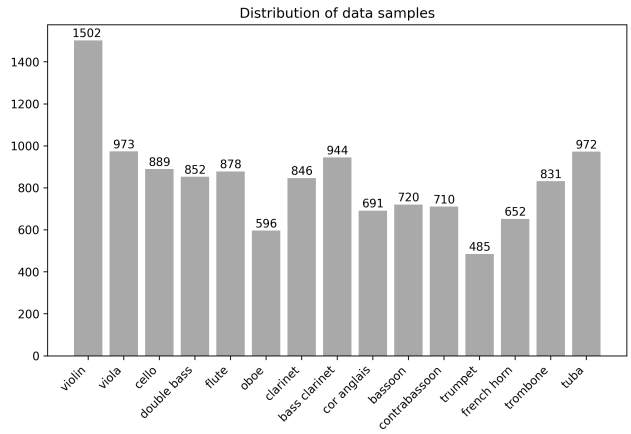
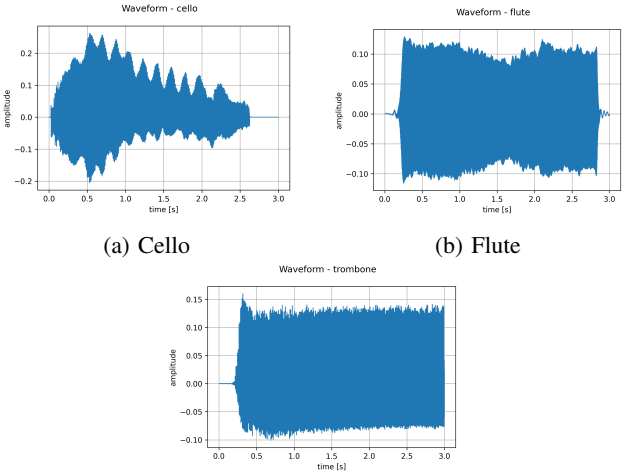


Fig. 1: Distribution of samples in the modified dataset



(a) Cello (b) Flute

(c) Trombone

Fig. 2: Waveforms of single notes being played on different instruments

we decided to create an additional experiment with a reduced dataset containing only three classes of choice, each representing one instrument family. For this purpose, we opted for cello, flute, and trombone, as they appeared to have a reasonably balanced number of samples. The total number of samples in the smaller dataset was 2 598.

The audio samples consisted of single tones and simple phrases played on different musical notes ranging from A1 to G7. In addition, samples were played with a wide range of techniques and dynamics.

The duration of audio files varied from less than a second to over a minute, which meant we had to decide upon a fixed length and perform trimming or padding, depending on the sample. The chosen duration was 3 seconds, as proposed by [8], stating that all of the 'onsets and attacks of the sound take place within the first 3 seconds'. To facilitate data handling, we converted all data samples from the original MP3 audio file format to WAV format. Additionally, we generated a CSV file containing file names, folder names, and corresponding class labels for each audio sample to make annotation easier during preprocessing and training.

III. METHODS

A. *Preprocessing*

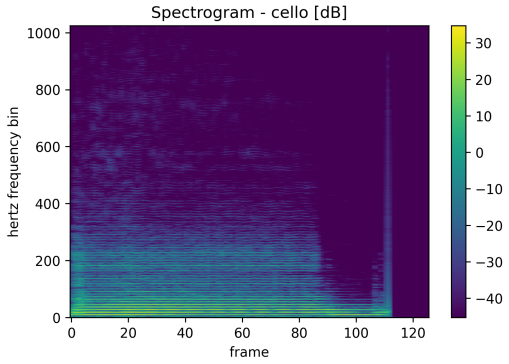
When it comes to processing and feeding audio data into neural networks, several concepts need to be considered. The sound produced by a musical instrument is an oscillation of air molecules caused by the vibration of the sound source (e.g., strings, reeds, or lips). This periodic alternation of high and low air pressure determines a sound wave, which can be represented by a waveform (Fig. 2). A waveform is a visual representation of the raw audio signal in the time domain, having amplitude values on the y-axis, and time on the x-axis.

In the real world, audio signals are rarely just pure sine or cosine waves. Nevertheless, it is possible to approximate complex periodic signals using a sum of elementary trigonometric functions of different wavelengths, frequencies, and amplitudes. This is known as the Fourier series. We heavily rely on this concept during sound analysis when we try to extract information about the signal's constituent parts to gain a better understanding of its nature and origin.

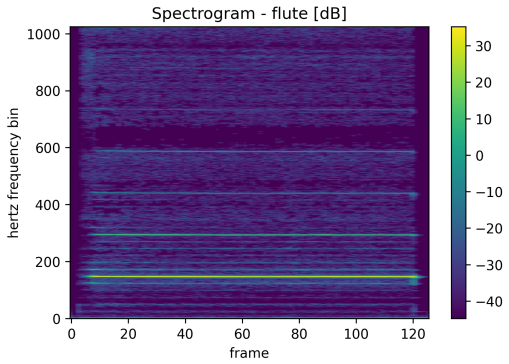
In the following text, we describe four sound preprocessing methods, each focusing on different components of audio signals and human perception of sound.

Spectrograms. One of the most observed wave features of audio signals is their frequency. Knowing that complex sound is composed of many sinusoids with different frequencies, a lot can be learned by extracting a frequency-power spectrum of the signal. A spectrum represents a distribution of frequencies over the entire time domain of the signal. It is performed by the fast Fourier transform (FFT), and it gives information about how much each frequency component (aka harmonic) contributes to the overall waveform of the signal. It is generally presented as a two-dimensional graph in the frequency domain, having magnitude on the y-axis and frequencies measured in hertz on the x-axis. However, considering that this type of representation captures the frequency distribution of the entire waveform, the information about frequencies at each given moment in time is lost.

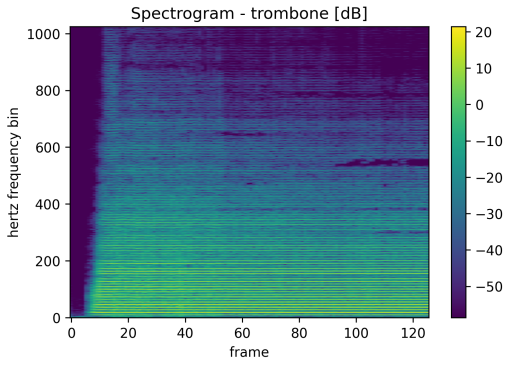
To preserve time domain features, a more profound transform is applied to the raw signal. This is called a short-time Fourier transform (STFT). STFT consists of many FFTs which are computed throughout several time intervals. These time intervals are called frames, and their lengths are defined by a frame size measured in samples. Another important parameter is the hop length, which determines the distance in samples between each frame. Setting a hop length that is smaller than the frame size causes frames to overlap, thus minimizing spectral leakage at the edges of the transformed signal. Finally, as a result of STFT, we get a spectrogram. Spectrograms are three-dimensional representations of sound in both frequency and time domains, with frames on the x-axis, frequency on



(a) Cello



(b) Flute



(c) Trombone

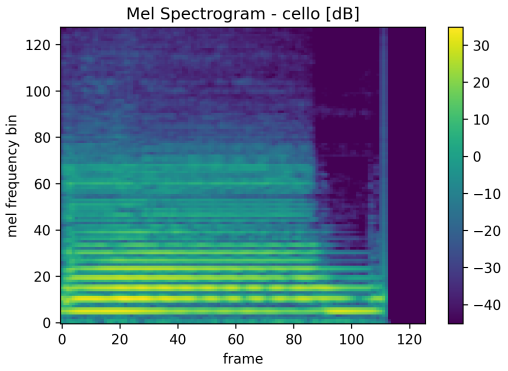
Fig. 3: Spectrograms

the y-axis, and magnitude on the z-axis, often visualized by a color bar.

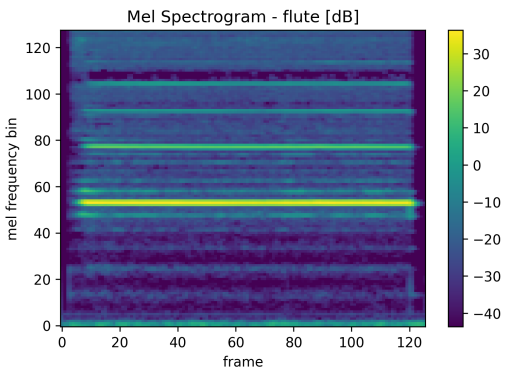
Fig. 3 shows spectrograms of a single note being played by cello, flute, and trombone. In our case, a sample rate of 22.05 kHz was used for all audio samples. To perform STFT, we used a frame size of 2048 samples and a hop length of 512. This resulted in a 1025×126 sized matrix which was going to be treated as input data for our classification models. The y-axis shows frequency bins instead of continuous frequency. Frequency bins are simply intervals of continuous frequency whose width can be calculated as the sample rate divided by the frame size of the FFT. In our case, having a sample rate of 22.05 kHz, and a frame size of 2048, each bin represents around 10.77 Hz in the frequency domain.

Another point worth emphasizing is the measurement unit on the magnitude axis, which has to do with the fact that the human auditory system is logarithmic in nature. This means that we perceive certain features of sound, such as pitch and loudness, in a nonlinear fashion. For this reason, we decided to represent magnitude in decibels (dB), which is a measure of intensity level based on the logarithmic scale. This was used throughout all preprocessing methods, in an attempt to mimic the way humans perceive sound.

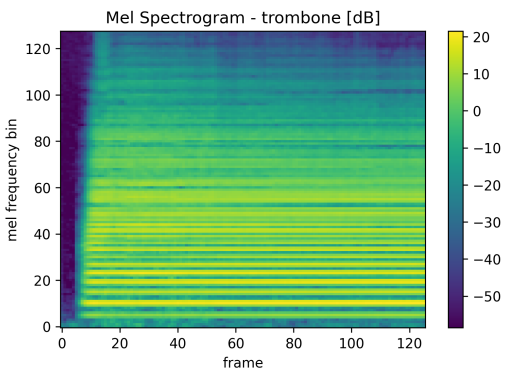
Mel spectrograms. Another important characteristic of sound perception is pitch. Despite being closely correlated to frequency, pitch is a subjective measure that portrays nonlinear tendencies. Because of this, the same differences in the lower frequency range are perceived as further apart



(a) Cello



(b) Flute



(c) Trombone

Fig. 4: Mel spectrograms

than those in the higher frequency range. To create a perceptually relevant frequency representation of sound, we apply the logarithmic scale on the frequency values of the spectrogram, which results in a Mel spectrogram. Equal distances between values on the mel scale have the same perceptual distance in pitch, with 1000 mels being equal to 1000 Hz, by agreement. When creating a Mel spectrogram, it is important to define the number of mel bands to be used during conversion. Mel bands define the range of perceptually relevant frequencies. More specifically, they determine which points of the signal's frequency range will be converted into mel scale. In our case, 128 mel bands were used, which ultimately resulted in data samples of size 128×126.

Fig. 4 shows Mel spectrograms of cello, flute, and trombone. Equation (1) is used to convert frequency expressed in hertz to frequency expressed in mels.

$$m = 2595 \cdot \log(1 + \frac{f}{700}) \tag{1}$$

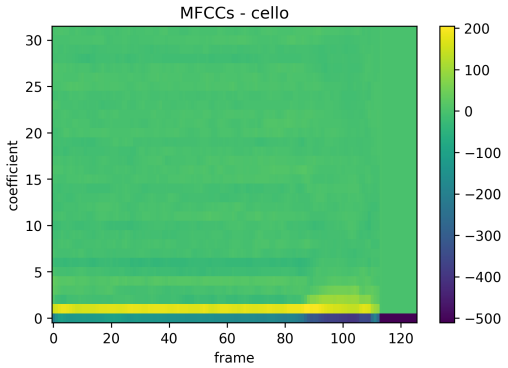
MFCCs. Another highly subjective aspect of sound is timbre. It is often described as the *color* of sound that is unique to each and every instrument. Timbre is what acoustically sets apart two instruments playing the same note at the same intensity, for the same duration of time.

Features that are commonly used to extract timbral and textural aspects of sound are Mel-Frequency Cepstral Coefficients or MFCCs. They convey information about the spectral envelope, spectral detail, harmonic content, and amplitude and frequency modulation of the signal (aka *tremolo* and *vibrato*). Some of these elements can easily be mapped onto characteristics of speech generation, specifically glottal pulse, formants, and frequency response of the vocal tract. For this reason, MFCCs are often used in speech recognition tasks [5].

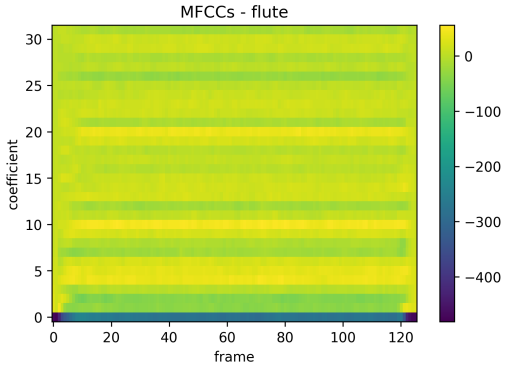
The signal's cepstrum is computed as an inverse Fourier transform applied on the log amplitude spectrum of the signal. This is given by the expression in (2). Cepstral coefficients with positive values signalize that spectral energy is mostly present in the lower frequency regions, whereas negative values indicate higher frequencies.

$$C[x(t)] = F^{-1}(\log(F[x(t)]) \tag{2}$$

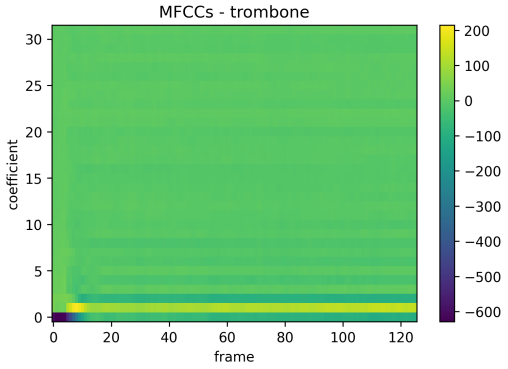
When extracting audio features using the MFCC method, several parameters have to be defined. To stay consistent with the rest of examined methods, we used 22.05 kHz as the sampling rate, 2048 samples for the frame size, the hop length of 512 samples, and 128 mel bands. Due to the input size restriction of our VGG classification model, 32 MFCC coefficients were chosen instead of 13, which is a number traditionally used in machine learning. The reason behind using the lowest possible value of MFCCs is the fact that they contain the most relevant timbral information, specifically the characteristics of the spectral envelope. As a result of preprocessing using the MFCC method, we get data points of size 32×126. A visual representation of MFCCs for cello, flute and trombone can be seen in Fig. 5.



(a) Cello



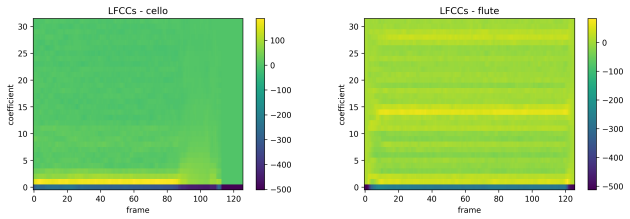
(b) Flute



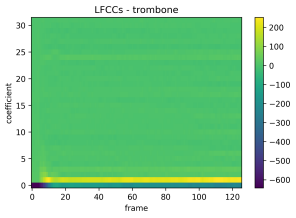
(c) Trombone

Fig. 5: MFCCs

LFCCs. Linear-Frequency Cepstral Coefficients, or LFCCs, are a feature extraction method similar to MFCCs. The only major difference is that they use a linear frequency scale instead of the logarithmic one. This way, audio signals are being treated in disregard to human perception of pitch, therefore preserving a better resolution in the higher frequency regions of the sound [13]. To perform preprocessing, we used 32 LFCC coefficients, due to the previously mentioned reasons. Since LFCCs don't use mel scaling, there was no need to define the number of mel bands. All of the other parameters remained the same as in previous methods. The resulting matrix had the same dimensions as the one derived by the MFCC method, i.e.,



(a) Cello (b) Flute



(c) Trombone

Fig. 6: LFCCs

32×126. LFCCs of a single note played by cello, flute and trombone are visualized in Fig. 6.

B. Classification

To get a better understanding of how each preprocessing method contributes to the overall efficiency of instrument classification, we decided to compare them using three different Deep Learning architectures.

The first architecture was a simple, custom-made convolutional neural network with four convolutional blocks, one flatten layer, one linear layer and a softmax layer at the end. Each convolutional block contained one 2D convolutional layer, a ReLu activation function, and a 2D max pooling layer. The second architecture was the standard VGG-16 convolutional neural network which consists of 16 convolutional layers. The first Visual Geometry Group (VGG) architecture was proposed by [14]. Lastly, we used ResNet-34 as our third architecture. ResNet-34 is a state-of-the-art image classification model that consists of 34 convolutional layers, and was first presented in [15].

The reason behind using several Deep Learning architectures in our experiments was to compare the preprocessing methods with a more general approach, as well as to eliminate any potential biases.

Furthermore, the dataset was split randomly into 80% of training samples and 20% of test samples. All of the methods were trained and tested on the same data samples. The models were trained for 30 epochs, using a batch size of 128. We used the cross-entropy loss function and the Adam optimizer with parameters Beta1 = 0.9 and Beta2 = 0.999. A learning rate of 0.00001 was used for all methods and models, and weight decay was set to None. The models were evaluated after the last epoch using test samples. Additionally, the experiment was repeated several times for each method and the best accuracy results were saved for analysis.

IV. RESULTS

The accuracies of the trained models for the smaller and larger dataset are shown in Table I and Table II, respectively. The tables display the best-achieved accuracy results for each method.

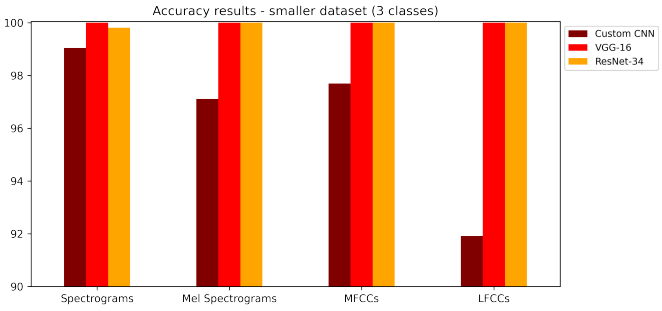
TABLE I: Results - smaller dataset (3 classes)

	Custom CNN	VGG-16	ResNet-34
Spectrograms	99.04%	100.0%	99.81%
Mel spectrograms	97.11%	100.0%	100.0%
MFCCs	97.69%	100.0%	100.0%
LFCCs	91.91%	100.0%	100.0%

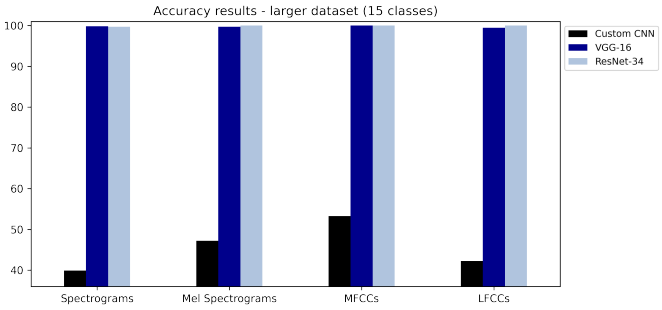
TABLE II: Results - larger dataset (15 classes)

	Custom CNN	VGG-16	ResNet-34
Spectrograms	39.91%	99.84%	99.72%
Mel spectrograms	47.17%	99.72%	100.0%
MFCCs	53.27%	100.0%	100.0%
LFCCs	42.22%	99.44%	100.0%

When observing the results for the smaller dataset, we notice an interesting occurrence of 100% accuracy throughout majority of the cases. This becomes particularly apparent with more complex Deep Learning architectures (VGG and ResNet). In this case, all 3 classes were noticeably distinct, which resulted in deep learning architectures being able to successfully classify the instruments in all cases, thanks to their high generalization ability, especially on small dataset with distinct samples. The dataset contains simple musical notes and phrases played similarly within each class, resulting in low intra-class variance. This allowed models to focus on inter-class



(a) Smaller dataset (3 classes)



(b) Larger dataset (15 classes)

Fig. 7: Visual representation of accuracy results

differentiation, leading to effective performance across all methods.

Judging by the accuracy of the custom CNN, spectrograms seem to have given the best results when it came to the smaller dataset. They were followed by MFCCs and Mel spectrograms, while LFCCs produced noticeably lower results. However, it is important to note that during training, spectrograms required a lot more memory and computational power than other methods. This may be due to a significant difference in input size that we get as a result of preprocessing with spectrograms.

If we look at the larger dataset, the efficiency of each method becomes more apparent. Although comparing the methods in combination with VGG and ResNet doesn't provide us with much information, the results of the custom CNN display clear differences - MFCCs seem to outperform other methods, with Mel spectrograms taking the second place. LFCCs along with spectrograms produce poor results when tested on a dataset with less distinguishable classes.

These results can be explained through the nature of each feature extraction method. Both MFCCs and Mel spectrograms use the mel scale instead of linear scale to preserve the perceptual relevance of the frequency spectrum. MFCCs, being built on top of Mel spectrograms, extract additional information about timbral and textural aspects of sound. Therefore, seeing that our dataset consists of instruments playing the same notes with similar intensities, timbre becomes the most important mean of distinguishing between closely related classes. This is why MFCCs tend to outperform other feature extraction methods in the task of musical instrument classification. MFCCs combined with strong ResNet or VGG architectures serve as an excellent cornerstone for building successful audio classification models.

V. CONCLUSION

The goal of this paper was to provide an overview of four preprocessing methods and their efficiency in the task of musical instrument classification. We ran several experiments and compared spectrograms, Mel spectrograms, MFCCs and LFCCs as tools for audio feature extraction. The preprocessed data was fed into three types of Deep Learning architectures, and the experiments were repeated for three and fifteen classes of musical instruments. Spectrograms showed an unusually high performance when it came to a simpler dataset. However, as the number of classes grew higher, their performance significantly dropped. On the other hand, the rest of the preprocessing methods portrayed consistency in both scenarios. The results showed that MFCCs performed the best job of capturing crucial audio features, especially when it came to less distinguishable classes.

In conclusion, a thoughtful selection of the preprocessing method can significantly impact the results of classification. For this reason, it is important to understand which features of sound are relevant to the given problem. Although a strong Deep Learning architecture can partially alleviate the effects of preprocessing, it is still important to pay attention to the nature of each feature extraction method.

REFERENCES

[1] J. D. Deng, C. Simmermacher, and S. Cranefield, "A study on feature analysis for musical instrument classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 2, pp. 429–438, 2008.

[2] A. Eronen, "Comparison of features for musical instrument recognition," in *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*, 2001, pp. 19–22.

[3] S. Chakraborty and R. Parekh, *Improved Musical Instrument Classification Using Cepstral Coefficients and Neural Networks*, 09 2018, pp. 123–138.

[4] K. Racharla, V. Kumar, C. B. Jayant, A. Khairkar, and P. Harish, "Predominant musical instrument classification based on spectral features," in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2020, pp. 617–622.

[5] W. Han, C.-F. Chan, C.-S. Choy, and K.-P. Pun, "An efficient mfcc extraction method in speech recognition," in *2006 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2006, p. 4.

[6] M. Deng, T. Meng, J. Cao, S. Wang, J. Zhang, and H. Fan, "Heart sound classification based on improved mfcc features and convolutional recurrent neural networks," *Neural Networks*, vol. 130, pp. 22–32, 2020.

[7] E. Şaşmaz and F. B. Tek, "Animal sound classification using a convolutional neural network," in *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, 2018, pp. 625–629.

[8] S. K. Mahanta, A. F. U. R. Khilji, and P. Pakray, "Deep neural network for musical instrument recognition using mfccs," *Computación y Sistemas*, vol. 25, no. 2, pp. 351–360, 2021.

[9] S. Prabavathy, V. Rathikarani, and P. Dhanalakshmi, "An enhanced musical instrument classification using deep convolutional neural network," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, 11 2019.

[10] D. Szeliga, P. Tarasiuk, B. Stasiak, and P. S. Szczepaniak, "Musical instrument recognition with a convolutional neural network and staged training," *Procedia Computer Science*, vol. 207, pp. 2493–2502, 2022, knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 26th International Conference KES2022.

[11] R. Profeta and G. Schuller, "End-to-end learning for musical instruments classification," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 1607–1611.

[12] London philharmonic orchestra dataset. [Online]. Available: https://philharmonia.co.uk/resources/sound-samples/

[13] X. Zhou, D. Garcia-Romero, R. Duraiswami, C. Espy-Wilson, and S. Shamma, "Linear versus mel frequency cepstral coefficients for speaker recognition," in *2011 IEEE Workshop on Automatic Speech Recognition Understanding*, 2011, pp. 559–564.

[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

OTHER TOPICS

Driven by Artificial Intelligence (AI) – Improving Operational Efficiency and Competitiveness in Business

Katarina Šiber Makar
Juniper Group, Palma de Mallorca, Spain
ALGEBRA University College, Zagreb, Croatia
katarina.sibermakar@in2.hr

Abstract **- Artificial intelligence has already become a necessary factor of advanced digitalization of companies and is one of the main prerequisites for global competitiveness and highly efficient operations. With cloud computing wider usage, decreased costs of AI technology and openAI, advanced AI technology and algorithms have become easily accessible to everyday users and even small to mid-size companies, which has widely extended its application. Based on the relevant research studies, this paper analyzes areas where AI can provide added value with its application in different industries like healthcare, retail, telecommunications, manufacturing, etc. Hence, an overview on different areas of AI implementation potential, like marketing and sales, product and service development, finance, supply chain management, customer experience, maintenance, quality control, services' and products' personalization, etc. and its benefits and challenges will be analyzed in detail. Main focus of this paper is on the organizations that have already started digitalizing their business from end to end and have exact benefits from applying AI by increasing revenues and decreasing costs.**

Keywords - artificial intelligence, machine learning, process automation, use case

I. INTRODUCTION

Nowadays, many companies go through intensive digital transformation to transform their businesses, improve their efficiency, automate its operations, improve customer experience, create new digital or digitally enhanced products and services and new digital business models. Artificial intelligence (AI) is one of digital technologies often seen as part of extensive digital transformations.

Artificial Intelligence refers to the simulation of human intelligence in machines and algorithms that are designed to think and act like humans. These intelligent machines and algorithms can be trained to perform tasks that would normally require human intelligence such as visual perception, speech recognition, decision-making, and language translation. AI systems use algorithms and statistical models to perform these tasks and continuously improve them through machine learning. The goal of AI is to create systems that can operate autonomously and work, reason, learn, and advance like humans.

After the initial appliance of AI that was reserved for corporations with significant investments in supercomputers and digital technologies, we are increasingly seeing applications of AI in creative domains, such as science, art and computer programming. With decreased cost of AI associated technology, like cloud computing and OpenAI, availability of massive amounts of data and ability to process the data, AI is experiencing much wider usage than before.

II. AI BENEFITS

AI can bring many benefits to businesses. Some of the most frequently reported benefits are: increased efficiency and productivity (AI can automate repetitive and time-consuming tasks), improved decision making (AI can analyze vast amounts of data in real-time, providing insights and recommendations that can support better business decisions), enhanced customer experience (AI-powered chatbots and virtual assistants can provide 24/7 support to customers, improving their experience, understanding their needs better, and increasing brand loyalty), better target marketing (AI can analyze customer data to identify patterns and preferences, allowing businesses to deliver more personalized and effective marketing campaigns, products and services), cost savings and increased productivity (automation and increased efficiency through AI can reduce costs and increase margins), risk reduction (AI can help businesses identify potential risks, fraud and ensure compliance with regulations, helping them avoid costly fines and reputational damage), competitive advantage (AI can provide businesses with a competitive advantage by giving them access to new and innovative technologies that can help them differentiate themselves from their competitors).

According to results of McKinsey and Company Survey [1] AI initiatives influence both cost decrease and revenue increase. Cost decreases are most often reported in supply chain management (logistic-network optimization and inventory and parts optimization), service operations optimization and predictive service and interventions, strategy and corporate finance, risk management (risk modeling and analytics as well as fraud and debt analytics), and manufacturing, while the greatest revenue increase is visible in marketing and sales

(customer segmentation and customer-service analytics), product and service development (AI enhancements of products and product-feature optimization), and corporate finance. 25% of respondents claim that at least 5% of their EBIT is generated by the application of AI, 48% claim it is less than 5%.

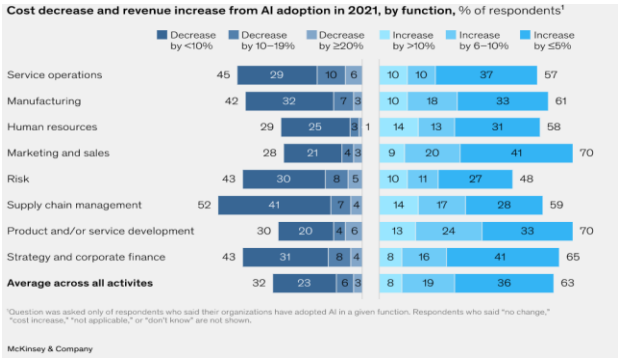


Figure 1. Cost reduction and revenue growth through AI implementation in 2021 by function [1]

CompTIA report [2] shows that adopting AI can have positive benefits for business, including (but not limited to): reduced operational time, greater business insights, reduced human errors, tasks automation, enhanced productivity, better customer service.

According to Statista [3], as of 2020, more than one third of financial services companies worldwide reported an increase in revenue of more than 20 percent because of the adoption of AI technologies. Moreover, 20 percent of the respondents reported a revenue increase between 5 and 10 percent. Many industries are reporting financial effects of AI implementation, e.g., the estimated impact attributed to AI in clinical and operational savings worldwide by the year 2022 is more than $150 billion, according to [4].

III. AI APPLIANCE AND USE CASES

According to McKinsey and Company [1], the most widely implemented capabilities of AI are robotic process automation, computer vision, natural-text language understanding, deep learning, recommendation systems, knowledge graphs and digital twins.

OpenAI has helped in wider popularization of AI. Some of OpenAI releases that have already experienced adoption for different purposes are Generative Pre-trained Transformer (GPT) - a neural network, or a machine learning model, created to function like a human brain, trained on input (large data sets) to produce outputs (e.g., answers to users' questions), Dall-E - a generative AI model that analyzes natural language text from human users and then generates images based on what is described in the text, PT-3 - language model that analyzes human-generated text to learn to generate similar text on its own, DALL-E - analyzes text-based descriptions of images that users want to produce and then generating those images exactly as described, ChatGPT - AI chatbot designed for generating humanlike text and producing answers to users' questions, Codex - trained on billions of lines of code in various programming languages to help software developers simplify coding processes, Whisper - automatic speech recognition trained on a multitude of audio data in order to recognize, transcribe and translate speech in about 100 different languages, including technical language and different accents [5].

AI is not equally accepted across all industries. According to [4], industries that have embraced the most usage of AI capabilities are retail (40% considers industry to have mature AI practices), financial services (38% considers industry to have mature AI practices), telecommunications (37% considers industry to have mature AI practices), computers, electronics, and technology (35% considers industry to have mature AI practices). According to the same research, education (10%) and government (16%) were the laggards. Media (29%), healthcare and life sciences (28%,) were in the middle, as were defense (26%) and manufacturing (25%).

According to [1], most widely adopted AI use cases are in the following areas: service operations optimization, creation of new AI-based products and AI-based enhancements of existing products, customer service analytics, customer segmentation, customer acquisition and lead generation, customer service (contact center) automation, product feature optimization, risk modeling and analytics and predictive service (maintenance) and intervention, as shown in Figure 2.

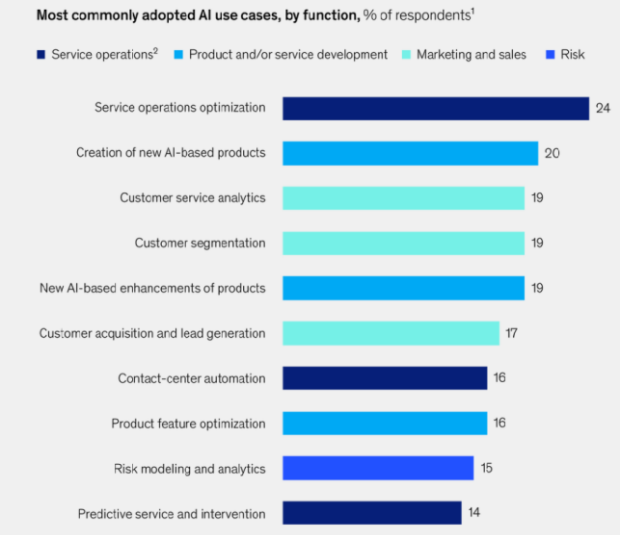


Figure 2. Most commonly adopted AI Use Cases by function [1]

A. *Telecommunication use cases*

According to [6], innovative telecom providers use AI and machine learning techniques for several use cases: to optimize network performance (network planning and optimization, predictive maintenance), improve efficiency of customer support operations, improve customer satisfaction and retention, improve efficiency of sales (personalized offers), provide personalized customer experience, streamline their business processes for higher profit, increase operational efficiency and reduce time to market.

AI-powered customer service solutions are often represented by virtual assistants or a chatbot interface, but these algorithms can also work in the background, making customer service more cost-efficient. For example, AI can support analysis of extensive background data to help customer agents in identifying the root cause of a customer's problem and find the appropriate solution more quickly.

Customer service is a critical component of the telecom industry, and the use of chatbots to increase or replace human call centers is one of the key AI applications in telecoms to date [7]. Many customer actions require specific support from customers, e.g., changing the bill plans or refunds, making payments, raising complaints, etc. Chatbots provide the customers with a single point of contact and are available 24/7, without increasing the number of agents, and can have a significant positive impact on customer experience, on the other hand. Chatbots implementation reduces costs by enhancing efficiency.

AI algorithms in customer service can act as virtual assistants or gateways between customer requests and contact center/live chat (help customers with troubleshooting, account information, or general questions about services), can route customer requests to the proper agent or route a prospect with buying intent directly to the sales agent, analyze customer support requests together with network data to find the solution to customer's problems more efficiently, etc. The assistant can either provide users with helpful tips or in the case of more complex requests, refer them to live chat or the most appropriate agent (based on agents' experience), thus offloading work from customer service and enabling agents to focus on more demanding customer complaints and requests.

One of the major sources of customer calls in telecoms, according to McKinsey and Company [8], is billing inquiries. An AI supported algorithm would consider the primary driver of the billing issue, customer's billing history, lifetime value, and propensity to call based on a bill change, and then take any number of different actions. Some customers might be satisfied with an explanation included with their bill; another might require a retroactive data package to be applied. Some customers might have a high likelihood to choose an upgrade or take some other revenue-enhancing action, in which case the preference would be for them to call customer service. Such an AI based solution would, among other activities, include clustering different customer profiles to identify their propensity to call and likely revenue and customer lifetime value impact of their call. It would need to predict the impact of each possible action on lifetime value and choose the best option. Once in place, AI solutions can be augmented with a machine-learning feedback loop to reflect the effectiveness of the actions taken, thus incrementally making solutions more precise in its decisions.

Improving efficiency of customer service operations use case is often closely connected to predictive maintenance use case. AI supported customer service can reduce number of calls by proactively troubleshooting issues (for example, a router that is slowing down could be identified and repaired before the customer even notices; it can be reset or re-configured remotely at a time when the customer isn't actively using the device and without customer even knowing a problem might occur). If the problem would be such that requires customer intervention, the AI supported algorithm could predict the customer's propensity to call regarding potential issue before alerting the customer or prepping the necessary information to reduce the length of the eventual call. For an issue that would require on-site resolution, a technician could be dispatched before customers notice the slower network speed and call to complain.

Measures like these can help telcos drastically reduce call volumes and enable agents to spend more time on calls that require direct customer interaction or can create significant added value.

AI can help telecoms improve customer retention and receive higher profit per user by providing personalized user experience. Some of the examples of AI use cases in this area are personalizing recommendations based on a user's behavioral patterns and preferences, predicting which call & data package best suits different types of users, making relevant upsell and cross-sell offers to the right users at the right time, thus increasing the sales success rate. Analyzing social media, brand coverage, and customer sentiment to learn what drives customers to the service provider and what drives them to leave is important, according to [6]. The main goal is to prevent calls that do not add value and take those that do create it.

One of the most frequently applied AI use cases in the telecom industry is matching customers with best-suited products (data packages), which truly suits their need. Self-learning algorithms accumulate insight into which packages match different customer types (based on pattern usage, billing/payment history, price sensitivity evaluation, quality sensitivity evaluation, etc.), easing the burden on call operators and making the sales process far more efficient [6]. Using AI solutions to improve efficiency of customer support operations can significantly decrease call abandonment rate and improve customer experience.

AI-enabled networks (network optimization use case) are capable of self-analysis and self-optimization, thus having greater agility and precision. Some of the examples of AI-enabled network capabilities are: restarting cell sites or towers based on their behavior, e.g. if they're not connecting to the network, algorithms pointing out parts of networks with a large number of users who would benefit from network improvements, optimizing the behavior of the network based on weather data, daily movements and real-time usage data, enhancing network utilization and customer satisfaction through dynamic resource allocation.

AI analyses network data to optimize network performance, reduce downtime, and minimize network disruptions. Optimization network teams use neuron network to analyze drive tests and calls quality based on psychoacoustic characteristics of human ear. AI algorithms can change parameters of network 24/7 to optimize characteristics for expected capacity during big

events or coverage while load in the network is low. AI algorithms can support telecoms to plan and design their networks more efficiently, reducing costs and improving network performance by suggesting optimal infrastructure positioning, based on user behaviors or seasonal traffic specifics.

AI algorithms can be used for fraud detection use cases (in telecommunications and other industries), they can detect unusual patterns and activities that might indicate fraud, reducing the time and effort required to identify and investigate fraudulent transactions.

B. Healthcare use cases

According to [9], some of the most common AI use cases in healthcare are improving medical diagnosis, speeding up drug discovery, transforming patient experience, managing healthcare data, performing robotic surgery, and detecting fraud.

One of the often-used AI capabilities is diagnostic imaging analysis. AI algorithms can assist radiologists in analyzing medical images such as X-rays, MRIs, and CT scans to detect abnormalities and make more accurate diagnoses. After receiving a medical image, the diagnostician analyzes it, reveals abnormalities or signs of the disease. To establish a diagnosis, the doctor must analyze several studies in a graphical form. AI-based systems do not only recognize medical graphics, but they also analyze medical images and report the features they find (sometimes specifics not visible to human eye). The diagnosis will be more accurate if the patient has several pictures taken at different times (then the neural network will analyze the dynamics of the disease). Deep learning-based computer-aided diagnosis system for melanoma is a god example of usage of deep learning algorithms for melanoma detection in dermoscopy images [10].

Humans vary widely at the genetic, biochemical, physiological, exposure and behavioral levels, especially with respect to disease processes and treatment responsiveness. This leads to the conclusion that there is often a need to tailor and personalize medicines to individual patients. Given how data-intensive is to reveal appropriate intervention targets and strategies for personalizing medicines, AI can play an important role in the development of personalized medicines. Personalizing a medicine to a patient requires a very deep understanding of that patient's condition, circumstances, genetics, and medical history, which requires generating, understanding, and processing massive amounts of data, such as DNA sequencing, imaging protocol, proteomics, wireless health monitoring devices inputs, etc. Essentially, the data produced needs to be organized so that analyses can be pursued to identify features that the patient possesses that may indicate the optimal intervention [11].

AI algorithms can help create personalized treatment plans for patients based on their medical history, genetic data, and other relevant factors. There are already many studies and appliances of personalized medicine. As an example, a study conducted by the Memorial Sloan Kettering Cancer Center found that AI algorithms could help identify the most effective cancer treatment options for individual patients and create Personalized treatment plans [11].

One of the most relevant and most promising AI use cases in healthcare is early disease detection. By using the patient's family and personal history, medical background and records, clinical and other data, the AI solution can help diagnose the disease at a very early stage.

AI-controlled robotic surgeries have brought revolution in the surgical sector in the form of collaborating robots and are already widely used. The main difference may be noticed in terms of depth and cutting rate. In general, the expertise of the surgeon significantly affects the outcome of surgery, especially one involving a novel or complicated procedure. Even the most talented and most experienced surgeons operate more efficiently when AI is used to support surgeries. AI robots, due to their precision, reduce the possibility of tremors or other unintentional movements during surgery. AI-controlled robots can conduct fundamental tasks like precise cutting and sewing while functioning with more accuracy and miniaturization [12].

AI algorithms are used to predict when medical equipment will fail and schedule maintenance to prevent equipment downtime and ensure patient safety. For instance, GE Healthcare has developed a predictive maintenance solution that uses AI to reduce downtime and improve equipment performance. AI is used in drug discovery and development. analyses large amounts of genetic, biochemical, and clinical data to help identify new drug targets and optimize drug development. A study conducted by the University of Cambridge [13] found that AI algorithms could be used to predict the success of drug candidates and help prioritize drug development.

Clinical decision support AI algorithms are used to analyze patient data in real-time and provide doctors with relevant information to support clinical decision-making. IBM Watson Health has already developed such an AI supported clinical decision support system.

AI-powered chatbots can improve patient care and assist patients in scheduling appointments, answer common questions, and even triage symptoms. For example, Babylon Health has developed a chatbot that uses AI to triage symptoms and provide patients with relevant information and advice.

Healthcare claims AI algorithms are used to detect fraud, waste, and abuse to ensure the efficient use of healthcare resources. A study conducted by the National Health Care Anti-Fraud Association found that AI algorithms could be used to detect fraud in healthcare claims with an accuracy rate of over 90% [14].

C. Retail use cases

According to [15], top retail use cases focus on inventory demand forecasting and management, personalization and marketing, pricing strategy, social media monitoring, sentiment analysis and call centers improvements. The retail market is embracing the use of AI due to its ability to hyper personalize and make recommendations based on product selection. AI algorithms are often used in Retail industry to increase

Net promotor score and for customers profiling. AI algorithms can analyze customer data to provide personalized recommendations, advertisements, and experiences. For example, Amazon uses AI to personalize the shopping experience for its customers through personalized product recommendations.

AI can help in predicting churn, determining floor and shelf space for new products, store layout optimization (using heatmaps), optimizing pricing, real-time inventory management (predict inventory needs in real-time and use automated system ordering before items run out to ensure that store shelves are always stocked with the products customers want), etc. AI algorithms help retailers optimize their inventory levels by predicting demand, reducing waste, and lowering costs.

Store operations and logistics are some of the largest cost drivers for retailers. For that reason, many use cases are focused on this segment, e.g., frictionless check out, picker routing, automated task dispatch and shelf checking. Retailers are leveraging AI and ML capabilities to gain efficiency and productivity for their employees by automating tasks across stores and distribution centers [16].

Many retailers are using AI for recommendations to customers based on their buying behaviors. The physically established retail stores will get users to create their accounts, just like users have their online accounts (retailers will motivate customers to participate by offering pickups, deliveries, loyalty points, discounts, etc.). This enables the storekeepers to personalize the preferences and buying patterns of each customer who walks into the store: categories of products bought, combinations bought, preferred brands, amount spent and frequency of buying. This information not only brings products recommendation but helps predicting customer's needs and anticipating their arrival to the physical or online store and particular purchase too [17].

In a similar manner retailers are using AI for customer profiling, AI can be used for product profiling too. Profiling depends on the product type, brand, cost, combinations (if exists) in which customers prefer to buy this product along with other products, the location of the product in physical store, the minimum quantity that is always available with the store, product popularity, etc. The target is that the product is sold along with other products that customers tend to buy and that the optimal quantity of product is always maintained in the stock [17].

Many operational activities in retail revolve around the pricing of the product. Different customers segments tend to behave in different manner: some will wait for the price to drop for non-urgent items, some for products to arrive in combination with another product at an optimal price, some customers are more price sensitive than others, etc. AI can be used to identify an optimal price/product combination for a set of customers which will drive better sales and profitability.

AI systems and history analyses of vending machines troubleshooting can help retailers in predictive maintenance by optimizing maintenance schedules for their equipment, reducing downtime, and saving costs. Coca-Cola is one of the companies that uses AI for predictive maintenance on their vending machines.

Retailers can use AI to better target their marketing efforts, develop a more efficient supply chain and better calculate pricing for optimal returns.

D. Other industries use cases

AI has wide application across many industries including: predictive maintenance, quality control (AI algorithms that inspect products for defects, reduce waste and increase productivity), supply chain optimization (AI algorithms analyze data on suppliers, demand, and inventory to optimize the supply chain and reduce costs), process optimization (AI algorithms analyze data on production processes to identify bottlenecks and inefficiencies, leading to improved production times and reduced costs), energy optimization (AI algorithms analyze energy usage data to identify ways to reduce energy consumption and improve energy efficiency), product recommendation (AI algorithms analyze user behavior and preferences to recommend relevant and personalized content), image and video recognition (AI algorithms automatically recognize and categorize images and videos, such as identifying objects, scenes, locations or faces), sentiment analysis (AI algorithms analyze the sentiment expressed in text data, e.g. social media posts or news articles, to understand public opinion and mood, as well as content popularity), speech-to-text conversion (AI algorithms convert speech into text, providing automatic captioning and subtitles for audio and video content), etc.

PwC study [18] shows that AI brings the most substantial business value to the following areas: increased productivity through automation, improved decision-making, improved customer experience, innovative products and services, improved employee experience and skills acquisition and development of new, data-driven business models.

IV. CHALLENGES OF AI IMPLEMENTATION AND ADOPTION

Not all attempts at implementation of AI were successful. According to [2], the biggest challenges in AI adoption and implementation are lack of skilled people and difficulty in hiring (19%), data quality (18%), followed by finding appropriate use cases (17%) and company cultural problems (14%).

It's no surprise that the demand for AI expertise has exceeded the supply, but it's important to realize that it's now become one of the biggest obstacles to wider adoption. The biggest skills gaps were ML modelers and data scientists (52%), understanding business use cases (49%), and data engineering (42%). Organizations are starting to realize the importance of data quality and 18% consider this to be one of the major challenges to implementing AI.

According to a survey of global organizations [19] that are already using AI, a quarter of surveyed companies reported up to 50% failure rate of their AI projects. Most AI used by businesses today has been trained to maximize model accuracy (the percentage of "correct" answers),

which for many real-world business cases won't bring best business results (AI that provides only five good sales recommendations is of little value to a sales rep who needs to close 100 deals per month). AI will rarely be able to recognize business cost-benefit trade-offs (due to financial/business benefits, you might be willing to pursue one deal and lose other 99 deals; an AI that would see this as only 1 win in 100 would be very inaccurate). AI/ML accuracy is fully dependent on data sets and cases that have been used to train it, which often turns out to be one of the major reasons for AI failures. For example, Amazon had been building ML programs since 2014 to review job applicants' resumes, which turned out not to be gender neutral. Based on 10-years history of resumes and employment, ML algorithms started preferring male candidates over female ones.

Some of the obstacles to implementing AI are lack of skilled individuals and hiring shortages, unclear ROI metrics, complexity of AI systems, lack of governance, fear of job replacement. The biggest risks associated to AI adoption and digital trust, according to [1] seem to be cybersecurity, regulatory compliance, personal/individual privacy, ability to explain how AI models come to their decisions, organizational reputation, equity and fairness, workforce/labor displacement, physical safety, national security, political stability. Many are concerned with AI ethical principles, particularly in AI applications like legal advice, medical diagnosis, autonomous vehicles, etc.

V. CONCLUSION

Based on the current results of AI implementation in business, it can be concluded that organizations that have the capacity to implement AI in their business processes, either in development segments or with optimization intentions, frequently experience significant cost savings and revenue increase, but also can use AI as differentiator toward their competitors.

The role of AI in business optimization is very promising. AI has the potential to revolutionize the way businesses operate, it provides increased operational efficiency, improved decision-making, enhanced customer experience and loyalty, enhanced and new (AI-based) products and services. AI algorithms are becoming more capable of analyzing vast amounts of data to make predictions about future trends, expectations, and behaviors, helping business to make better decisions and respond more quickly to changes on the market.

AI is expected to play an increasingly significant role in shaping our future and will impact numerous industries, beyond those significantly impacted already today (retail, financial services, technology, telecommunications, healthcare, manufacturing, media). However, it's essential to ensure that the development and deployment of AI systems are guided by ethical principles and consideration for their impact on society as main priority.

REFERENCES

[1] Michael Chui, Bryce Hall, Helen Mayhew, Alex Singla, Alex Sukharevsky,.McKinsey and Company, "The state of AI in 2022", December 2022 https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

[2] Ashley Watters, CompTIA, "30+ Artificial Intelligence Statistics & Facts for 2022", February 2022 https://connect.comptia.org/blog/artificial-intelligence-statistics-facts

[3] Statista Research Department, "Increase in revenue of financial services companies worldwide because of the adoption of artificial intelligence (AI) as of 2020", May 2022

[4] Mike Loukides, „AI Adoption in the Enterprise", O'Reilly, Radar Report, 2021

[5] Camron Hashemi-Pour, OpenAI, Techtarget, 2023 https://www.techtarget.com/searchenterpriseai/definition/OpenAI

[6] Konstantin Sadekov, Mindtitan, "AI use cases in telecom relevant for 2023. with 8 examples", December 2022 https://mindtitan.com/resources/industry-use-cases/artificial-intelligence-in-telecom-business/

[7] Roboticsbiz, "Top use cases of Artificial Intelligence (AI) in telecom", 2019 https://roboticsbiz.com/top-use-cases-of-artificial-intelligence-in-telecom/

[8] McKinsey&Company, "How AI is helping revolutionize telco service operations", February 2022 https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-ai-is-helping-revolutionize-telco-service-operations#

[9] Sam Daley, "Artificial Intelligence in healthcare: 39 examples improving the future of medicine", 2022 https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare

[10] Mahya Sadeghi, Parmit Chilana, Jordan Yap, Philipp Tschandl, M. Stella Atkins, "Using content-based image retrieval of dermoscopic images for interpretation and education: A pilot study", Journal of Skin Research and Technology, 2019 https://onlinelibrary.wiley.com/doi/full/10.1111/srt.12822

[11] Nicholas J. Schork, Ph.D, "Artificial Intelligence and personalized medicine", The Translational Genomics Research Institute (TGen) 2020

[12] Cigniti, Top 5 AI Use Cases in Healthcare for Improving Patient Care & Hospital Efficiency, 2021 https://www.cigniti.com/blog/ai-use-cases-healthcare-patient-care-hospital-efficiency-ai-ml/

[13] Fleming, N., How artificial intelligence is changing drug discovery, Nature, *557*(7706), pp.S55-S55, 2018

[14] Thaifur, A.Y.B.R., Maidin, M.A., Sidin, A.I. and Razak, A "How to detect healthcare fraud? A systematic review". *Gaceta Sanitaria*, *35*,, 2021, pp.S441-S449

[15] Robert Hetu, "23Artigicial Intellicence Use Cases for Retail", Gartner, March 2021 https://blogs.gartner.com/robert-hetu/23-artificial-intelligence-use-cases-for-retail/

[16] Carrie Tharp, "Beyond COVID-19, retail looks to transform with AI/ML", Google Cloud, 2020. https://cloud.google.com/blog/products/ai-machine-learning/top-10-ai-use-cases-in-retail

[17] Mantra AI, "5 Artificial Intelligence (AI) use cases in retail", 2019 https://mantra.ai/blogs/5-artificial-intelligence-ai-use-cases-in-retail/

[18] Bret Greenstein, Anand Rao, PwC 2022 AI Business Survey, PwC, 2022 https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-business-survey.html

[19] R. Preston Mcafee, Arijit Sengupta, Johnatan Wray, "This much-hyped technology is failing business. Here is why", August 2020 https://www.fastcompany.com/90449015/this-much-hyped-technology-is-failing-businesses-heres-why

[20] CompTIA AI Advisory Council, Artificial intelligence in business: top consideration before implementing AI, 2022 https://comptiacdn.azureedge.net/webcontent/docs/default-source/research-reports/08497-ai-considerations-document-online-final.pdf?sfvrsn=789806ef_0

A Survey on Usage of Multimedia Databases for Emotion Elicitation: A Quantitative Report on How Content Diversity Can Improve Performance

M. Horvat* and P. Jerčić**
* University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Applied Computing, Zagreb, Croatia
** Graz University of Technology, Institute of Interactive Systems and Data Science, Graz, Austria
marko.horvat3@fer.hr, pjercic@tugraz.at

***Abstract* – Affective picture databases provide a standardized set of images to elicit controlled and consistent emotional responses in research participants. They are a valuable tool for studying various emotion-related phenomena across several research domains. These domains include emotion perception, emotion regulation, and the neural basis of emotion. However, affective picture databases have diverse schemas, structures, and content, making them difficult to use. Searching and retrieving optimal pictures relevant to affective stimulation may be challenging and time-consuming. In this context, we surveyed domain experts about their practices and experiences working with affective multimedia databases such as IAPS, NAPS, OASIS, GAPED, and others. The survey identified a need for novel data observatory software. This finding motivates the authors' intention to develop and validate such software platform that relies on AI. Such a platform would describe better, retrieve, and integrate various semi-structured affective multimedia datasets. The results prominently indicate the overwhelming dissatisfaction regarding stimuli content diversity and cultural bias, specifically regarding emotional and semantic context. The main driver of satisfaction from users of existing automated retrieval software is the quality of semantic descriptors available. This points to the direction AI should take in novel data observatory software. This survey follows up on a similar survey conducted ten years ago and explores the differences in researchers' opinions and experiences during that time. The complete aggregated results are publicly available at https://github.com/mhorvat/stimdbsurvey.**

Keywords - affective computing, emotion, databases, emotion stimulation, document retrieval

I. INTRODUCTION

Digital multimedia, apart from having a format, semantic content, and context, also can provoke emotions in observers. When people engage with various forms of digital media, such as movies, music, books, or virtual reality, they are inherently emotionally stimulated. The process of eliciting emotional responses through digital multimedia is complex and multifaceted, with underlying neuroanatomical mechanisms [1][2]. The type and intensity of the elicited emotional states can be modeled probabilistically for a given stimulus [3]. Affective multimedia databases, also known as multimedia stimuli databases, are a particular type of collection containing processed multimedia documents specifically designed to elicit emotions in a controlled laboratory environment [4]. Documents are often referred to as stimuli, depending on their purpose, while images and videos are typically referred to as visual stimuli [4][5]. These databases store the semantics, context, and affect of the stimuli and categorize them according to common emotion models [4][5][6]. Although traditionally considered non-interchangeable, recent research has challenged this notion [7][8]. Affective multimedia databases have practical applications not only in the study of human emotion mechanisms but also in perception, memory, attention, and reasoning [9].

Affective multimedia databases, for example, International Affective Picture System (IAPS) [10], Nencki Affective Picture System (NAPS) [11], Open Affective Standardized Image Set (OASIS) [12], The Geneva Affective PicturE Database (GAPED) [13], NimStim Face Stimulus Set [14], are designed to be used in research of emotion processing, attention, stress resilience, and mental health, yet a lot can be done to improve their models, facilitate their usage and expand prevalence in the field.

Our previous work showed that multimedia documents are sparsely annotated, making semantic retrieval difficult and resulting in low recall and precision [15]. Furthermore, high-level semantic content descriptors are informal and ambiguous due to the insufficient annotation methods which rely on unrestricted keywords. This limited meta-retrieval framework has no unified stimuli content description dictionary, underlying knowledge base, concept taxonomy, or terminology. These problems contribute to a prolonged and work-intensive construction of stimuli sequences, difficulties in finding optimal elicitation, and emphasize the need to thoroughly structure affective stimuli databases and develop new tools for searching and retrieving emotion stimuli [7][15].

Currently, stimuli are often manually extracted from multimedia databases through a time-consuming visual inspection of each stimulus and its associated database manuals. Complicating matters further, multimedia stimuli

databases are structured differently, may describe emotional and semantic data differently, and may contain different media formats. In addition, there is no agreement in the research community on an optimal structure or standard implementation of a multimedia stimulus database. For these reasons, a typical user in any database must be skilled in both emotion capture and technology-related tasks such as stimulus selection and extraction. Unfortunately, it is difficult and time-consuming to master such a fusion of different skills. Furthermore, because databases differ in structure, these skills must be learned separately for each database.

These difficulties suggest that a novel AI-based computer system is required to assist experts in finding the most appropriate stimulus and completing it quickly. A system of this type must be as universal as possible regarding database models and media formats and efficient and user-friendly. Furthermore, to aid the expert, the system should ideally include empirically derived rules for decision support and the automatic generation of stimuli sequences. Again, such a stimuli generator should produce personalized stimuli tailored to specific emotional and semantic parameters [7][15].

To determine the need for such a software tool in the professional community, we conducted an online survey to examine how researchers in psychology, neurology, cognitive science, and related fields use multimedia stimuli databases to elicit emotional responses. In addition, the survey identified current problems in the use of affective multimedia databases and areas where further improvement is needed. The study follows up on a similar survey on the usage patterns of stimuli databases undertaken ten years ago [16].

The remainder of this paper is organized as follows: Section 2 lists all the survey questions and systematically describes the methodology and the mixed-effects model used to explain the research questions. The aggregated survey results are presented in Section 3. Section 4 presents the statistical analysis of the survey results and shows which semantic descriptors (i.e., tags) influence users' satisfaction with affective multimedia databases. In addition, in Section 4, the results of this survey are compared to the previous survey's results, and the key differences and similarities between the two studies are highlighted. Finally, Section 5 summarizes the paper and suggests future directions for developing affective multimedia databases.

II. SURVEY METHOD

Between 1 December 2022, and 15 January 2023, an invitation to participate in an online survey was sent to 336 e-mail addresses of authors of published papers using at least one emotionally annotated database. Invitations were sent to each author twice within 2-4 weeks. The authors were demographically diverse from different higher education and research institutions. IAPS, NAPS, OASIS, NimStim, and GAPED were the databases addressed. In addition, some invitees used Karolinska Directed Emotional Faces (KDEF) [17], International Affective Digital Sounds (IADS) [18], and Pictures of Facial Affect (POFA) [19]. Relevant publications and author contact information were manually searched via Web of Knowledge, PubMed, IEEE Xplore, and Google Scholar by tracking citations in the corresponding databases. Of 336 invitations, 87 were bounced because of invalid emails, of which 54 were resent. In total, N=37 invited individuals completed the survey (12.13%). Individual responses were automatically recorded and aggregated using the Google Forms service for creating online forms and surveys. The anonymity of the participants was guaranteed by the service provider, as stated in the e-mail invitation. Participation in the survey was completely anonymous. Individuals participating in the survey were not identifiable, and no personal information was stored. Multiple responses were not permitted.

The survey consisted of 19 questions: 5 questions (1, 13, 15, 18, 19) were free text, and 14 others with predefined answers, including the Likert scale with a range of 1–5. Questions 15, 16, and 19 were optional. Likert scale questions also had the "Not Applicable/Prefer Not to Answer" option as one of the possible answers.

The survey questions were:

Q1. What best describes your research topic?

Q2. Please rate the image retrieval process from multimedia stimuli databases such as IAPS, NAPS, GAPED, OASIS, KDEF, POFA, JAFFE, and IADS?

Q3. How satisfied are you with the above-mentioned level of difficulty in the image retrieval process from the database?

Q4. How much time was necessary to effectively search the database and construct one picture sequence that was used in your research?

Q5. On a scale of 1 to 5, with one being very inadequate and 5 being very adequate, please rate the quality of the documentation (guide instructions and support)?

Q6. Have you at any time felt that the picture set you were using is missing images with a particular emotion that would be useful for your stimuli sequence?

Q7. Have you at any time felt that the picture set you were using is missing images with semantic content that would be useful for your stimuli sequence?

Q8. Have you felt that the picture set you were using addressed the cultural values in your target group?

Q9. On a scale of 1 to 5, with one being very insufficient and 5 being very sufficient, please rate the diversity of semantic and emotional content in the picture sets you used.

Q10. On a scale of 1 to 5, with one being very inadequate and 5 being very adequate, please rate how inadequate (insufficient) did you find the predefined semantic descriptors (e.g., keywords or tags) of the images you used?

Q11. On a scale 1 to 5, with one being extremely useless and 5 being extremely useful, please rate how helpful would a user-friendly software tool for intelligent retrieval of emotionally annotated images be to your research?

Q12. Have you at any time during your research wanted to find the most appropriate emotionally-annotated images faster and more efficiently?

Q13. Can you give a rough indication of the duration in minutes of one of your test sessions (the period in which one participant is tested continuously)?

Q14. Did you construct the sequence manually or with a help of a software tool?

Q15. Which software tool did you use (if any)?

Q16. Did your group actually develop the tool used in your experiment or acquired it elsewhere?

Q17. How useful or useless a stimuli database with realistic and immersive Virtual Reality (VR) images instead of just still ones, would be to your work?

Q18. Please list the names of multimedia stimuli databases you have used in your research (e.g., IAPS, NAPS, GAPED, OASIS, KDEF, POFA, JAFFE, IADS).

Q19. Do you wish anything to add that was not covered in the survey?

Addressing the ambiguity on what influences user ratings concerning affective multimedia databases, we wanted to investigate the impact of the following elements mentioned in the survey: 1) the quality of documentation (manuals, online help), 2) the diversity of available stimuli, and 3) the expressiveness of the content descriptions (keywords). Furthermore, considering the difference between users in terms of previous use of automatic software to create stimuli sequences and building it manually, we proposed the following research question:

Which factors impact user satisfaction with affective multimedia databases – the quality of documentation (manuals, online help), diversity of available stimuli, or expressiveness of the content descriptions (keywords)?

We have designed a linear mixed-effect model to investigate the drivers (i.e., factors) for user ratings of affective multimedia databases. More specifically, we considered the quality of the documentation (manuals, online help), diversity of available stimuli, and quality of semantic descriptors as independent variables, affecting satisfaction ratings with affective multimedia databases as dependent variables. The analysis aims to estimate the rating scores of the users of affective multimedia databases based on the survey data we have collected. First, we grouped the users based on the previous usage of automatic retrieval software and those retrieving it manually. Following this, we focused on the variables that significantly affect the satisfaction ratings while still considering the variability from other factors. We used a mixed-effect model because the rating scores from the same user groups regarding previous experience in using automated retrieval software were correlated. Moreover, it makes sense to structure the data given our interest in automated retrieval software.

Given the small number of users involved in this high-dimensional data, it was considered correct to use the restricted maximum likelihood as an estimation method since the bias to this method might over or underestimate the true variance.

III. Survey Results

Keywords best describing participants' research topics (Q1) are visualized as a tag cloud in Figure 1. Stop words were removed, and keywords were lemmatized.



Figure 1. Keywords that best describe survey participants' research topics. Normalized answers to Q1.

The names of multimedia stimuli databases the participants used in their research (Q18) are shown in Figure 2. The most frequent are IAPS (27 answers), NAPS (15), KDEF (11), GAPED (8), OASIS (8), IADS (5), NimStim (3), JACFEE (2), and MSFDE (2). Further, 32 different databases were mentioned one time each.



Figure 2. Keywords that best describe survey participants' research topics. Answers to Q18.

The aggregated survey results to Likert scale questions are displayed as individual charts in Figure 3 and Table 1 for better visibility. Because of the restrictions of the survey tool, the standard answer "Not Applicable/Prefer Not to Answer" is shown in a different color between the charts in Figure 3.

All survey results are freely available to the academic community and for research purposes on the GitHub online repository (https://github.com/mhorvat/stimdbsurvey) and can be downloaded in CSV, Excel, and other standard formats.

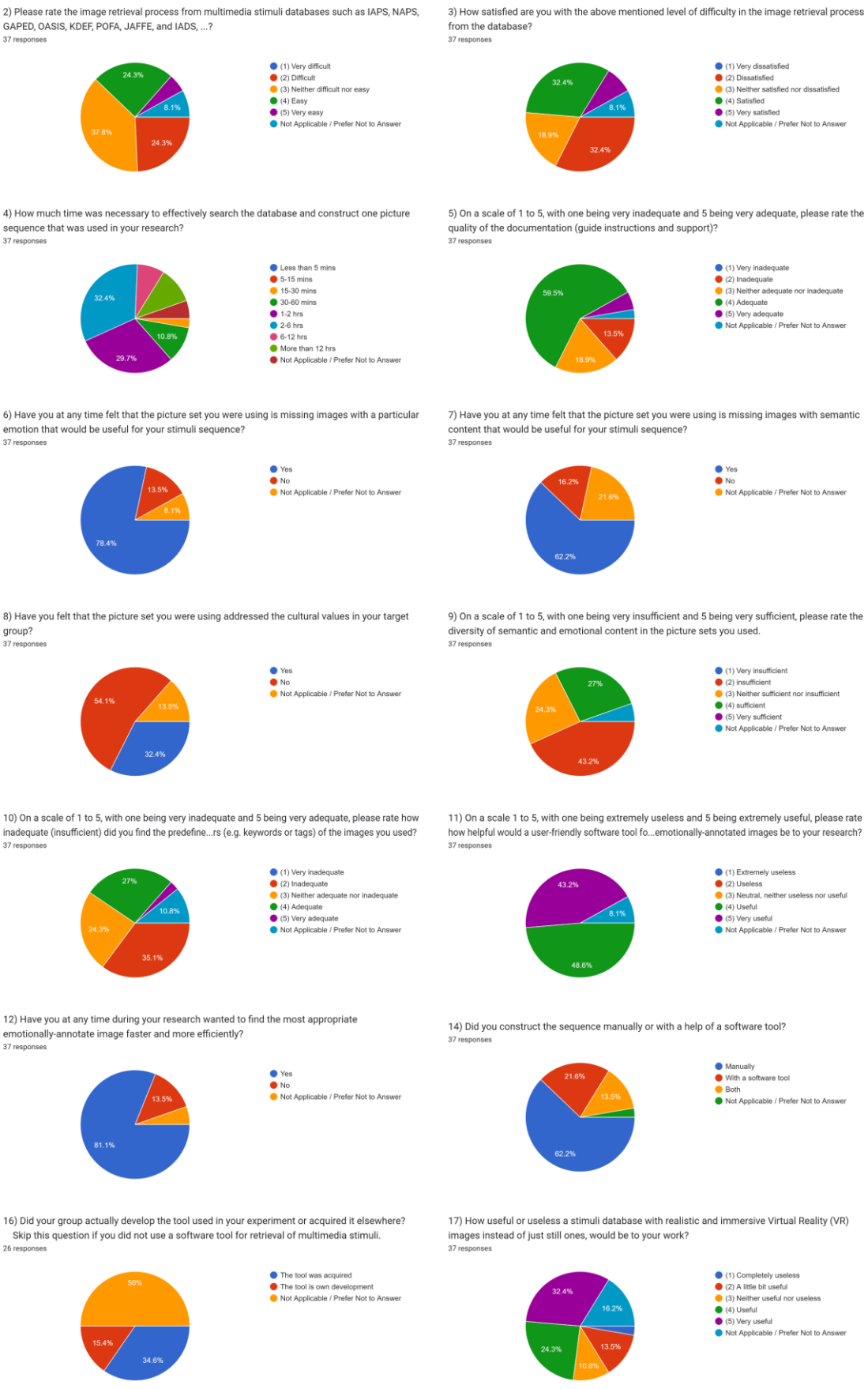


Figure 3. Overview of the aggregated survey results for Likert-scale questions.

TABLE I. TABLE OVERVIEW OF THE AGGREGATED SURVEY RESULTS FOR LIKERT-SCALE QUESTIONS.

Question	Answers
Q2	Very difficult = 0%, Difficult = 24.3%, Neither difficult nor easy = 37.8%, Easy =24.3%, Very easy = 5.4%, NA = 8.1%
Q3	Very dissatisfied = 8.1%, Dissatisfied = 32.4%, Neither satisfied nor dissatisfied = 18.9%, Satisfied = 32.4%, Very satisfied = 8.1%, NA = 8.1%
Q4	Less than 5 mins = 0%, 5-15 mins = 0%, 15-30 mins = 2.7%, 30-60 mins = 10.8%, 1-2 hrs = 29.7%, 2-6 hrs = 32.4%, 6-12 hrs = 8.1%, More than 12 hrs = 10.8%, NA = 5.4%
Q5	Very inadequate = 0%, Inadequate = 13.5%, Neither adequate nor inadequate = 18.9%, Adequate = 59.5%, Very adequate = 5.4%, NA = 2.7%
Q6	Yes = 78.4%, No = 13.5%, NA = 8.1%
Q7	Yes = 62.2%, No = 16.2%, NA = 21.6%
Q8	Yes = 32.4%, No = 54.1%, NA = 13.5%
Q9	Very insufficient= 0%, Insufficient= 43.2%, Neither sufficient nor insufficient = 24.3%, Sufficient = 27%, Very sufficient = 0%, NA = 5.4%
Q10	Very inadequate = 0%, Inadequate = 35.1%, Neither adequate nor inadequate = 24.3%, Adequate = 27%, Very adequate = 2.7%, NA = 10.8%
Q11	Extremely useless = 0%, Useless = 0%, Neutral, neither useless nor useful = 0%, Useful = 48.6%, Very useful = 43.2%, NA = 8.1%
Q12	Yes = 81.1%, No = 13.5%, NA = 5.4%
Q14	Manually = 62.2%, With a software tool = 21.6%, Both = 13.5%, NA = 2.7%
Q16	The tool was acquired = 34.6%, The tool is own development = 15.4%, NA = 50%
Q17	Completely useless = 2.7%, A little bit useful = 13.5%, Neither useful nor useless = 10.8%, Useful = 24.3%, Very useful = 32.4%, NA = 16.2%

Regarding the main drivers for user satisfaction with using these databases, semantic descriptors reached a statistical significance coefficient, while diversity and documentation did not (95% confidence interval). The results are presented in Figure 4 below, and detailed results are in Table 2 (Appendix).

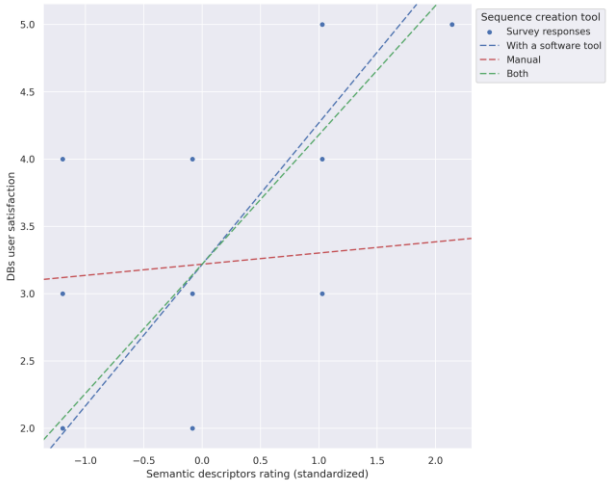


Figure 4. Keyword semantic descriptors were influencing user satisfaction with affective multimedia databases. Data is grouped based on experience with sequence creation tools (automatic retrieval software, manual creation, or both).

IV. Discussion and Results Analysis

While researchers are divided on the difficulty in retrieving images from stimuli datasets (Q2 and Q3), the survey clearly showed that sequences are still predominantly constructed manually, with experts visually searching multimedia repositories for prolonged periods (Q4). The process of stimuli construction takes 2-6 hours (32.4%), 1-2 hours (29.7%), more than 12 hours (10.8%), 30-60 minutes (10.8%), and 6-12 hours (8.1%). Therefore, 51.3% need 2 hours or more to construct an emotion elicitation sequence. Dataset documentation (Q5), when available, was primarily judged to be adequate (59.5%) and very adequate (5.4%), while a minority (13.5%) believed that the documentation was inadequate or neutral (18.9%).

A large majority of participants (78.4%) think that images with a specific emotion are missing from the dataset they use (Q6), and 62.2% believe that images with semantic content that could be useful for their work are missing (Q7). In addition, 54.1% feel that the imagery they use does not address the cultural values of the target group, and only 32.4% feel that it does (Q8). Finally, the diversity of semantic and emotional content was rated as insufficient by 43.2% of the participants (Q9).

Semantic information, keywords, or tags describing the stimuli were considered inadequate by 35.1% and neither adequate nor inadequate by another 24.3% of the researchers (Q10). Although no one (0%) thought that the existing stimuli labels were very inadequate; on the other hand, only 2.7% felt that they were very adequate. Clearly, the description of stimuli information should be improved.

An overwhelming percentage of participants (total of 91.8%) declared that a user-friendly software tool for intelligent retrieval of affective images would be very useful (43.2%) or useful (48.6%) for their research (Q11). Very importantly, none (0%) indicated that an intelligent software tool for stimuli retrieval would either be highly useless or had a neutral opinion, believing it would be neither useless nor useful. This (Q11) indicates that the professional community would welcome such a tool.

In addition, 81.1% of examinees stated that they would like to find the most appropriate images more quickly and efficiently. However, only 13.5% reported this was not the

case and were satisfied with the current stimuli search speed (Q12).

Responses to Q13 (Figure 5) show that one test session lasts 30 minutes for 29.7%, 60 minutes for 18.9%, and 20 minutes for 16.2% of most researchers. Outliers are 1 (min) and 120 minutes (max).

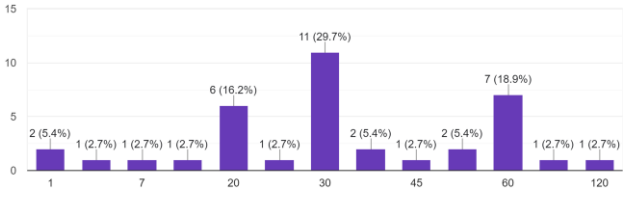


Figure 5. Duration of one test session in minutes. Answers to Q13.

As many as 62.2% of respondents constructed stimuli sequences manually, and only 21.6% with the aid of a software tool, while another 13.5% used both methods (Q14). The tools used for constructing stimuli sequences are (Q15): Psychopy, E-Prime, Psychtoolbox in Matlab, PicRotator (v1.0 and v2.0), Qualtrics, Testable, Presentation, and an online random sequence generator. In answering Q15, some survey respondents indicated tools intended only for presenting stimuli. As a result, 15.4% of researchers (*N*=26) developed the tool, and 34.6% acquired it elsewhere (Q16).

Only 2.7% of respondents consider immersive Virtual Reality (VR) completely useless (Q17). On the other hand, a large majority (70.2%) support introducing new computer-based methods in creating and delivering emotion-provoking content at any level, while 13.5% are neutral (Q17). Together with the previous answers, this once again underscores the need to continue producing fresh multimedia content and creating new databases with rich and personalized stimuli.

In comments on various topics (Q19), participants called for more consistency in documentation. They also pointed out that more racial diversity of images and more emotion types would be helpful (e.g., bored, disgusted). Some asked for neutral stimuli that are as semantically related as the emotional stimuli and positive arousing stimuli and emotional stimuli that do not occur in distant, exotic locations. The number of available images, as is the ability to assign images based on various features other than valence, is essential. It was also pointed out that there have yet to study children. Some were concerned about the quality of the ground truth data and how the dataset authors collected it. Some researchers found it easy to construct sequences based on published normative ratings but found it challenging to work with data dimensions beyond those covered by the normative ratings (i.e., beyond valence, arousal, and basic emotions).

Notably, regarding user satisfaction with these databases, we found that semantic descriptors have the most significant influence on predicting user satisfaction and significantly affect rating scores. Furthermore, the effect was strong for users with previous experience with automated retrieval software, while it was weaker for those who retrieved manually. This effect was significant even after controlling for the variance from other confounding factors in the survey.

A. Comparison to the previous survey results

In the earlier similar survey reported in 2013 [16], most respondents felt that the descriptions of the stimuli were inadequate and that it generally takes 1-2 hours or more than 24 hours to complete a single stimulation sequence. In the earlier survey, a large majority (84%) of participants indicated that real-life videos would be useful in their research. The experts recognized the need for an intelligent stimuli retrieval application to assist them in their experiments. Almost all experts then agreed that such software tools could be helpful in their work.

In light of the ten years between the two surveys, there were some shifts in the researchers' opinions. The image retrieval process has become more difficult in recent years. While at the same time, satisfaction with this difficulty increased. With more invested effort, possibly higher gains can be obtained using multimedia databases in recent years. This fact is crucial if researchers invest more time effectively searching the database and constructing one picture sequence. More specifically, there has been a 3-fold increase in spending 2 - 6 hours on this task. By segregating semantic or emotional content questions in the new survey, we observed that both are a problem, with emotional content being a more significant issue than the semantic one. At the same time, the adequacy of the descriptions of the images remained stable. Half the researchers responded that diversity or underrepresented cultural values might be responsible for this dissatisfaction.

On the other hand, researchers have positively consolidated their opinions on a user-friendly software tool for the intelligent retrieval of emotionally-annotated images. Increased difficulty in the image retrieval process might shift some respondents towards expressing a need for such a tool. However, the need to find the most appropriate emotional annotated image faster and more efficiently remained stable. Furthermore, researchers who used software tools in the past to accompany their manual construction of the sequence have shifted towards fully manual construction. This fact might indicate a need for a more intelligent software tool, as research slightly sifted towards acquiring software rather than developing their own.

V. CONCLUSION

The survey provided valuable insights into the opinion, motivation, and experience of experts who use affective multimedia databases to elicit emotional responses from subjects. Although the survey was not large in terms of the questions' complexity or the number of participants, it provided statistically significant results. Although email invitations certainly limited the number of responses received, it was the only method available to obtain the valuable opinions of a wide range of professionals.

Users creating the stimuli sequence manually had no significant effect of the investigated variables on their rating scores compared to the users of automatic software for stimuli sequence creation. Furthermore, their user ratings were remarkably stable over the investigation space, generally fixed on the middle rating.

The results indicate that automatic retrieval software for stimuli sequence creation has a clear advantage over manual creation. This is because users may give higher score ratings for the availability of good semantic descriptors. However, on the other hand, poor semantic descriptors in such automatic software drive their ratings very low, lower even than manual creation. Potentially, the non-availability of good semantic descriptors renders such automatic software unusable to create a satisfying sequence, which users with manual pipelines would still be able to operate in such adverse affective multimedia databases.

Considering these findings, an intelligent software platform for automatic sequence generation from affective multimedia databases should focus on providing quality semantic descriptors. This idea feeds well into our initial goal of developing more effective methods and procedures for generating semantic descriptors based on knowledge graphs to better describe, retrieve, and integrate different semi-structured stimuli datasets. We hope that such a tool will help all researchers in the field of affective computing, psychology, neurology, and cognitive science to find the most appropriate multimedia stimuli with minimal effort and in the shortest possible time. This should improve the quality of personalized emotion elicitation and reduce the experts' workload, allowing them to focus more on the subject and stimulation procedures.

The collected results indicate the need to develop new stimuli databases with fresh multimedia content. Generative AI graphics and VR delivery are good research directions, as they offer many advantages over acted or real visualizations. However, creating more databases further highlights the need for intelligent software to help manage multifaceted affective multimedia content.

The survey's authors would like to cordially thank all experts who have taken part in the survey.

REFERENCES

[1] K. R. Scherer, "What are emotions? And how can they be measured?", Social science information, vol. 44, no. 4, pp. 695–729, 2005.

[2] M. Lewis, J. M. Haviland-Jones, L. F. Barrett (Eds.), "Handbook of emotions," Guilford Press, 2010.

[3] R. W. Picard, "Affective computing: challenges", International Journal of Human-Computer Studies, vol. 59, no. 1, pp. 55−64, 2003.

[4] M. Horvat, "A Brief Overview of Affective Multimedia Databases," In Central European Conference on Information and Intelligent Systems (CECIIS 2017), pp. 3−11, 2017.

[5] Y. Wang, W. Song, W. Tao, A. Liotta, D. Yang, X. Li, G. Shuyong *et al.*, "A systematic review on affective computing: Emotion models, databases, and recent advances," Information Fusion, 2022.

[6] M. Horvat, A. Stojanović, Ž. Kovačević, "An overview of common emotion models in computer systems," In 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 1008–1013, 2022.

[7] M. Horvat, A. Jović, K. Burnik, "Investigation of Relationships between Discrete and Dimensional Emotion Models in Affective Picture Databases Using Unsupervised Machine Learning," Applied Sciences, vol. 12, no. 15, p. 7864, 2022.

[8] S. Zhao, G. Ding, Y. Gao, X. Zhao, Y. Tang, J. Han, ... Q. Huang, "Discrete probability distribution prediction of image emotions with shared sparse learning," IEEE Transactions on Affective Computing, vol. 11, no. 4, pp. 574-587, 2018.

[9] T. Brosch, K. Scherer, D. Grandjean, D. Sander, "The impact of emotion on perception, attention, memory, and decision-making," Swiss medical weekly, vol. 143, no. 1920, w13786−w13786, 2013.

[10] P. J. Lang, M. M. Bradley, B. N. Cuthbert, "International affective picture system (IAPS): Affective ratings of pictures and instruction manual," Technical Report A-8, University of Florida, Gainesville, FL, 2008.

[11] A. Marchewka, Ł. Żurawski, K. Jednoróg, A. Grabowska, "The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database," Behavior research methods, vol. 46, pp. 596−610, 2014.

[12] B. Kurdi, S. Lozano, M. R. Banaji, "Introducing the open affective standardized image set (OASIS)," Behavior research methods, vol. 49, pp. 457−470, 2017.

[13] E. S. Dan-Glauser, K. R. Scherer, "The Geneva Affective PicturE Database (GAPED): A new 730 picture database focusing on valence and normative significance," Behavior Research Methods, vol. 43(2), pp. 468–77, 2011.

[14] N. Tottenham, J. W. Tanaka, A. C. Leon, T, McCarry, M. Nurse, T. A. Hare, D. J. Marcus, A. Westerlund, B. J. Casey, C. Nelson, "The NimStim set of facial expressions: Judgments from untrained research participants," Psychiatry Research, vol. 168:3, pp. 242–249, 2009.

[15] M. Horvat, Generation of multimedia stimuli based on ontological, affective and semantic annotation (Doctoral thesis), University of Zagreb, 2013.

[16] M. Horvat, S. Popović, K. Ćosić, "Multimedia stimuli databases usage patterns: a survey report," in Proceedings of the 36nd International ICT Convention MIPRO, pp. 993–997, 2013.

[17] D. Lundqvist, A. Flykt, A. Öhman, "The Karolinska Directed Emotional Faces – KDEF, " CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 1998.

[18] P. J. Lang, M. M. Bradley, B. N. Culthbert, "International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings (Tech. Rep. No. B-2)," The Center for Research in Psychophysiology, University of Florida, U.S.A., 1999.

[19] The Paul Ekman Group, LLC, retrieved from: https://face.paulekman.com/face/productdetail.aspx?pid=1

VI. APPENDIX

TABLE II. Parameter of the linear mixed-effects model and the formula: "c2 ~ q10 + 1", groups= "q14", re_formula="~q10 + q5 + q9 + 1". The questions used in the model are listed in Section 2.

Parameter name	Beta	Std. Err.	Lower-95	Upper-95	Random effect (SD)
Intercept	3.221	0.177	2.873	3.568	
Q10	0.741	0.312	0.130	1.352	
Q5					0.191
Q9					0.021

AI Chatbot For Job Interview

N. Boudjani*, V. Colas*, C. Joubert†, D. Ben Amor*
* SogetiLabs, Paris, France
† Capgemini engineering, Paris, France
nadira.boudjani@sogeti.com

***Abstract*—In this paper, we propose an interactive AI chatbot in French language which able to ask questions to a potential candidate, to detect incomplete answers and to ask additional questions in order to obtain a complete answer to a given question. Additionally, the proposed chatbot allows an interactive environment where candidates can also ask questions during the interview. We also propose a system of interview based on an exhaustive behavioral diagram. All the chatbot's functionalities have been validated by experimentation. The results show that our chatbot is complete in terms of the questions asked and information to be collected during the interview. Moreover, experimentation has shown that the diagram covers all cases of scenarios between an interviewer and a candidate during a job interview.**

***Keywords*—NLP, interactive AI chatbot, job interview**

I. INTRODUCTION

Research on the application of AI in the HR domain has been increasing in recent years [1]–[3]. The aim is to improve management processes and make it possible to abandon administrative and repetitive activities in favor of tasks with high added value [4].

AI is used in several HR activities such as employees training, detection of employees ready to resign and particularly in hiring process at different stages : candidate outreach [1], CV filtering [5], profile matching [6], job interview [7] and on-boarding [8]. Authors in [9] define the job interview as a personally interactive process of one or more people asking questions orally to another person and evaluating the answers for the purpose of determining the qualifications of that person in order to make employment decisions. Most of the work proposed in the literature in job interview aim to evaluate the candidate: sentiment analysis [10], skills evaluation and personality traits extraction [11], [12] and [13].

We can group these works in two main categories: deferred video job interview [14]–[16] and job interview chatbot [17]–[19].

In the first category, the candidate connects to the platform and will be filmed with a camera during the interview. The candidate's answers are recorded to be analysed after the interview. The second category is based on chatbots, a textual conversational agent, which interviews the candidate. The chatbot asks the questions to be answered by the candidate.

According to our knowledge, the chatbot based are facing with two main limits, as follows:

1) the lack of interaction with the candidate during the interview: the interview questions are predefined. When the candidate answers a question, the answer is saved and the chatbot asks the following question regardless of candidate's answer. In order to acheive a more realstic interview, we propose an intraction between the chatbot and candidate.
2) the inability for the candidate to ask questions during the interview: the proposed recruitment chatbots do not allow the candidate to ask questions. We believe that giving the candidate this possibility is essential to make the interview becomes more close to a real interview.

To overcome these two limitations, we propose an interactive job interview chatbot based on AI techniques. This chatbot is able to interact with the candidate in real-time allowing the candidate to ask questions during the interview. We believe that it is important that the interviewer gives feedback to the candidate's answer and the latter can ask questions or request information from the interviewer.

The rest of the paper is organized as follows: Section 2 presents the proposed AI chatbot functionalities and diagram that structure the behavior of the chatbot during the job interview , Section 3 presents results of the experiments. Finally, Section 4 presents the conclusion and the perspective.

II. PROPOSITION

We propose an interactive chatbot that conducts a job interview with a potential candidate. The present paper does not aim to evaluate candidate's skills. Our goal is to collect as much information as possible about the candidate from his answers during the interview. We assume that nothing is known about the candidate prior to the interview. We do not take into account the information given on a CV.

During the interview, the chatbot must be able to ask questions to the candidate and answers the candidate's questions. The candidate answers the questions posed by the chatbot and can ask questions to the chatbot. Fig. 1 presents a comparison between a classic job interview chatbot and our proposed chatbot in this paper.

The design of our chatbot is based on 4 steps detailed below: definition of the questions to ask, definition of the intentions and entities to be extracted during the interview, candidate's actions during job interview and behavior of the chatbot during the job interview.

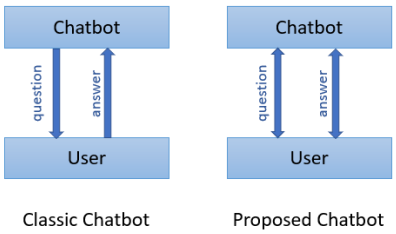


Fig. 1: Classic chatbot vs proposed chatbot.

A. Job Interview Questions

We structure the interview by "axis". There are 6 axes with 2 greetings (welcome message and goodbye message): marital status, educational career, experiences, skills, languages and hobbies. Each axis contains a list of default questions to ask in an interview (It is the case when the candidate answers perfectly all the questions asked).

In contrast, in a real interview, the candidate might not understand a question asked or might answer the question asked partially. In the first case, the interviewer reformulates his question and in the second case he asks additional questions to obtain the missing part of the candidate's answer.

For this reason, we define 3 types of questions to be asked by the chatbot:

- **Main questions** is a question asked by default to collect information from the candidate.
- **Reformulated questions** is a main question rephrased into another question when the candidate do not understand the main question
- **Additional questions** is a sub-question of a main question asked by the chatbot to recover the missing part of the answer.

B. Intentions and Entities

We propose a chatbot based on AI techniques. It relies on intent classification and entity extraction. For example, in Fig. 2 the candidate is answering a question. The chatbot can either move on to the next question, or ask a reformulated question or ask additional questions.

To match between questions, intents and entities, we defined for each question a list of intention which corresponds to a correct answer. If the classified intent is in this list the chatbot moves on to entity extraction. For each intent we have a list of entities to extract. Fig. 3 summarizes the match between questions, intentions and entities in the educational career axis.

C. Candidate Actions

Our goal is to offer the candidate a set of actions to allow him to interact freely with the chatbot. The interview begins with a greeting message sent by the chatbot. This message asks the candidate if they agree to be interviewed by the chatbot : "Hello, I'm your recruitment chatbot. Do you agree to start an interview with me?". Fig. 4 represents

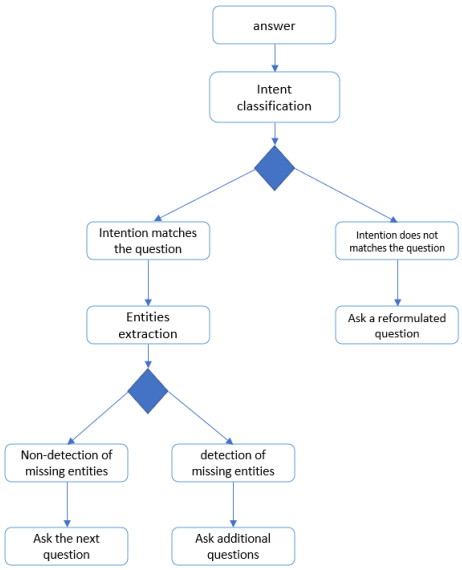


Fig. 2: NLP role in the conduct of the job interview.

the proposed actions the candidate can perform during the interview:

- Accept interview: allows candidate to accept the interview. The chatbot sends the first question.
- Refuse interview: allows candidate to refuse the interview. If so, the chatbot sends a goodbye message.
- Answer a question: allows candidate to answer the question asked by the chatbot
- Refuse to answer: allows the candidate to refuse to answer a question asked by the chatbot. The chatbot will move on to the next question. This action is necessary to not block the interview if the candidate does not want to answer a given question.
- Ask a question: allows candidate to ask a question to the chatbot. This action is important to make the interview close to a realistic interview.
- Stop interview: allows candidate to stop the interview

D. Chatbot Process Diagram

In this section, we propose a decision tree that represents the behavior of the chatbot during the recruitment interview. In order to cover all possible cases during an interview, we have generated effective scenarios. The idea is to vary as many scenarios as possible to take into account the different possible cases. Fig. 5 and Fig. 6 represent an example of generated scenario in French and English respectively.

From the generated scenarios and after 5 iterations, we propose the transition state diagram of the chatbot. The global process chatbot is given by Fig. 7. The job interview begins by sending the first question (greeting message) by the chatbot in B1. The chatbot receives the response and saves it in SD1. The intentions are processed in SD2, SD3 processes the control of the answer in relation to the question asked and the passage to the next question is processed in SD4. When the candidate answers the last

Axes	Main questions		Reformulated questions	intentions	Entities	Additionnal questions
axis 2 : educational career	diplômes	Q7 : Parlez-nous de votre dernier diplôme.	RQ7 : Quel est votre dernier diplôme obtenu ?	confirmer_diplome nier_diplome	intitulé_diplome spécialité_diplome organisme_diplome date_obtention	AQ7.1: Quel est l'intitulé de ce diplôme ? AQ7.2: Quelle est la spécialité de ce diplôme ? AQ7.3: Par quel organisme ce diplôme a été délivré ? AQ7.4: Quelle est la date d'obtention de ce diplôme ?
		Q8 : Avez-vous obtenu un autre diplôme ? Si oui, parlez nous de ce diplôme.	/	confirmer_diplome nier_diplome	intitulé_diplome spécialité_diplome organisme_diplome date_obtention	
	formations	Q9 : Parlez-nous de votre dernière formation.	RQ9 : Quelle est votre dernière formation obtenue ?	confirmer_formation nier_formation	intitulé_formation organisme_formation date_formation	AQ9.1: Quel est l'intitulé de cette formation ? AQ9.2: Par quel organisme cette formation est proposée ? AQ9.3: Quelle est la date de cette formation ?
		Q10 : Avez-vous fait une autre formation ? Si oui, parlez nous de cette formation.	/	confirmer_formation nier_formation refuser_answer	intitulé_formation organisme_formation date_formation	
	certifications	Q11 : Parlez-nous de votre dernière certification.	RQ11 : Quelle est votre dernière certification obtenue ?	confirmer_certification nier_certification	intitulé_certification organisme_certification date_certification	A11.1: Quel est l'intitulé de cette certification ? A11.1: Par quel organisme cette certification est proposée ? A11.1: Quelle est la date de cette certification ?
		Q12 : Avez-vous une autre certification ? Si oui, parlez nous de cette certification.	/	confirmer_certification nier_certification	intitulé_certification organisme_certification date_certification	

Fig. 3: Questions, intentions and entities defined in educational career

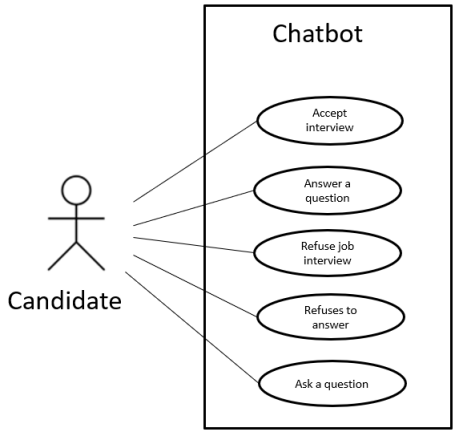


Fig. 4: Use case : candidate actions.

Q1 : Quel est le dernier diplôme obtenu ?
R1 : je n'ai pas de diplôme, j'ai une formation en A en novembre 2021
Q2 : Par quel organisme cette formation a été proposée ?
R2 : Vous voulez dire quoi par organisme ?
Q3 : Un organisme de formation ou un dispensateur de formation, parfois appelé prestataire de formation est une personne physique ou morale qui réalise des prestations de formation professionnelle continu
R3 : Ah ok, c'était proposée par X
Q4 : Bien ! Avez-vous fait une autre formation ?
R4 : Non
Q5 : Avez-vous des certifications ? Si oui, parlez nous de la dernière certification ?
R5 : non, je n'ai pas de certification. J'ai oublié de vous parler d'une formation que j'ai faite en octobre 2020. C'était une formation sur C proposée par Y

Fig. 5: Example of a generated scenario.

question, the chatbot sends a goodbye message and the interview ends.

The detailed process is given in Fig. 11. Sub-diagrams will be detailed below.

1) SD1: Receive and save the answer: A timer is activated when the chatbot is waiting for the candidate's answer. If the maximum duration is exceeded, the chatbot asks if the candidate wants a reformulation of the question (if he does not understand the question asked). If the answer is yes, connection to part SD3 is done to allow the chatbot to reformulate the question. Otherwise, the candidate's answer is saved and treated as a refusal answer.

Q1: What is the last degree you obtained?
R1: I don't have a degree, but I completed a training program in A in November 2021.
Q2: By which organization was this training program offered?
R2: What do you mean by "organization"?
Q3: An organization or provider of training is an individual or legal entity that provides professional and continuing training services.
R3: Oh, I see. It was offered by X.
Q4: Alright! Have you completed any other training programs?
R4: No, I haven't.
Q5: Do you have any certifications? If yes, can you tell me about your latest certification?
R5: No, I don't have any certifications. I forgot to mention that I completed a training program on C in October 2020. It was offered by Y.

Fig. 6: Example of a generated scenario translated in English.

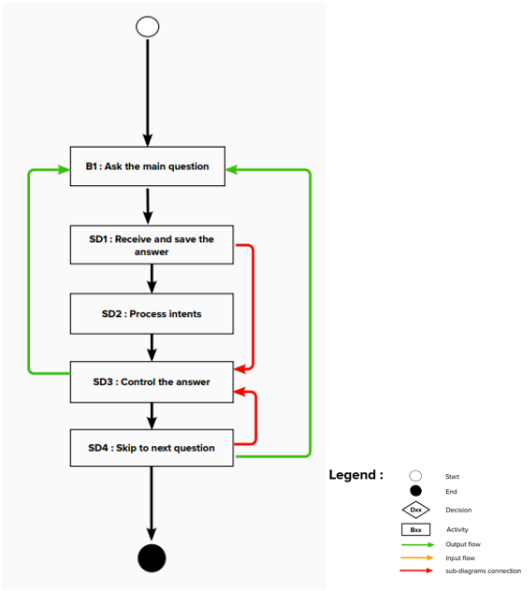


Fig. 7: Global process chatbot.

2) SD2: Process intents: The message intents sent by the candidate are classified by NLP intent classification model. A consistency check with the interview is carried out. If the intent is not adequate with the interview, the chatbot ignores it. Else, the intent can be of the 2 types:

- Informative: Answer to a question that the chatbot asked. Entities are extracted using NLP model.
- Information request: Question asked by the candidate. The chatbot saves the question, extracts the entities and prepares the answer.

3) SD3: Control the answer:

- If the candidate's message is an information request then the chatbot sends the information to the user.
- Otherwise, the chatbot checks if the answer corresponds to the question asked (response intent):
  - If yes : check response completeness
    - If no entity is missing:
      - if no information request is recorded: switch to SD4
      - else: the information is sent to the user and we connect to SD1.
    - else (some entities are missing): chatbot asks an additional question for each missing entity
  - Otherwise: chatbot finds that the answer does not correspond to the question asked, it reformulates the question

4) SD4: Skip to the next question: Chatbot checks that all the questions of the current axis are answered:

- If yes, the chatbot checks that all axes are processed:
  - If yes: the interview ends
  - Otherwise, the chatbot goes to the 1st question of the next axis and checks if there is already a complete answer to this question previously (this check is done in case the candidate has indirectly answered a question not yet asked, we believe that this question should not be asked again):
    - If yes, connect to D12
    - else, chatbot checks if part of the question has already been answered (if the candidate has partially answered a question not yet asked, the chatbot does not ask the main question again but he asks additional questions):
      - If yes, switch to SD2 to ask additional questions on missing entities
      - Otherwise, ask the next question (B1)
- otherwise, skip to the next question

III. EXPERIMENTS

The main of the experiments is to evaluate the completeness of our list of intents and entities to be detected in the interview and to check if our chatbot can conduct any type of interview by following the chatbot process diagram.

Population: 5 volunteers took part to the experiment, aged between 25 and 45 with various profiles (students, developers, project managers, etc.). The experiment took place in 3 stages:

Step 1: generation of interview scenarios. In this step, the participants generate a set of scenarios that will be used in the following steps. 16 job interviews were generated on the 6 axes of the interview.

Step 2: manual extraction of intents and entities for each candidate's answer in the interview.

The aim of this step is to verify the completion of the list of intents and entities the chatbot is going to detect. For this, we asked the participants to extract the intention and the entities of each answer. An example is given in Fig. 8.

Recruteur	Candidat	Intentions (entités)
Monsieur Dubois, vous postulez pour un poste de chef d'équipe de "poseurs de fermetures en bâtiment". Quel est votre niveau de formation.	Je n'ai pas de diplôme mais j'ai suivi un apprentissage chez un artisan pendant 3 ans.	nier diplôme Donner information sur apprentissage (employeur = artisan, durée = 3 ans)
Ou avez-vous fait cet apprentissage ?	Dans la région des Baux de Provence, car c'est une région ou il y a un grand savoir faire et beaucoup d'artisanat. Je l'ai fait chez un ami de mon père, lui-même dans la serrurerie.	Donner information sur lieu et contexte apprentissage (lieu = Baux de provence ; métier=serrurerie).
A très bien, et quelles spécialités avez-vous pu étudier ?	J'ai pu apprendre et m'exercer sur toutes les spécialités de pose de serrure, façonnage de clefs et de gonds, de conception de portes et volets en bois, en alu et en PVC.	Donner information sur spécialités étudiées (spécialités = pose de serrure, façonnage de clefs et de gonds, de conception de portes et volets en bois, en alu et en PVC).

Fig. 8: Manual intents and entities extraction Example.

Step 3: projection of the interview in the diagram

The objective is to verify that the diagram contains all possible cases and to detect blockages in the diagram projection with scenarios generated in Step 1. For this, the participants projected the interview on the diagram. We asked the participants to present their projections in a table as represented by Fig. 9.

Q/R	objet parcouru	label transition
Début		
Q1	B1	
R1	B2	
	B3	
	D1	Non
	D2	Oui
	D4	Informatif
	B5	
	D1	Oui
	D3	Non
	D5	Oui
	B10	
	D8	Aucune
	D7	Non
	B12	Non
	B14	
	D14	Non
	D15	Non
Q2	B1	
R2	B2	

Fig. 9: Diagram projection example.

The objective of the experiment is to test the completeness and validity of the intentions, the entities expressed in the candidate's answers in an interview, and the chatbot process diagram.

The results show that the proposed intentions and entities are very close and even identical to those found in the experiment. We have not detected any intent or entity not taken into account in the predefined list. This shows that the chatbot can cover any job interview with a potential

candidate. The results also show that the diagram can be used as a generic structure for job interviews. All the scenarios generated have been projected on the diagram without any blocking. This shows that the proposed diagram can represent the behavior of the chatbot during a job interview.

A first chatbot demonstrator on the civil status axis is developed using Rasa [20] plateform : DIETClassifier for intent classification and CRF (Conditional Random Fields) [21] model for entity extraction. The implementation of the other axes is in progress. Fig. 10 shows a screenshot of a user's conversation with the chatbot. In this example, the user gives an incomplete answer on the postal address where the postal code is missing. The chatbot asks him an additional question about the postal address.

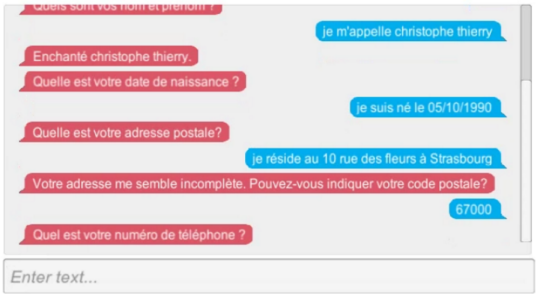


Fig. 10: Example of conversation with Job interview chatbot using RASA.

IV. CONCLUSION AND PERSPECTIVES

In this paper, we present an AI chatbot based on NLP techniques for job interview. We have defined the chatbot and represented its behavior during a job interview with a state transition diagram and presented also a demonstrator.

Diagram allows chatbot to:

- ask questions to candidates
- detect if the answer corresponds to the question asked
- detect if the answer given by the candidate is complete
- answer the questions asked by the candidates

The results show that the diagram, intents and entities cover thoroughly the different scenarios.

To the best of our knowledge, no job interview chatbot exists in the literature that has the defined functionalities of the chatbot proposed in this article. The only work that asks the questions according to the answer given by the candidate is [22]. Authors use follow up questions by detecting key words in the candidate's answer. However, it does not allow the candidate to ask questions during the interview. Further, in [22], authors use an android robot to implement their job interview and we are using a chatbot.

As a perspective, we have planed to finalize the implementation of our AI chatbot and test it in job interviews with potential candidates and evaluate the different functionalities. The final results will be presented in a future paper.

REFERENCES

[1] O. Ahmed *et al.*, "Artificial intelligence in hr," *International journal of research and analytical reviews*, vol. 5, no. 4, pp. 971–978, 2018.

[2] L. Menant, "L'acceptabilité de l'intelligence artificielle appliquée aux systèmes d'information des ressources humaines: cas d'un agent conversationnel," Ph.D. dissertation, Université Paul Valéry-Montpellier III, 2021.

[3] F. Chevalier and C. Dejoux, "Intelligence artificielle et management des ressources humaines: pratiques d'entreprises," *Le numérique et la refondation du système électrique*, p. 94, 2021.

[4] M. de La Rochefoucauld, "L'IA et la reconfiguration de la fonction rh," *Management & Datascience*, vol. 4, no. 1, 2020.

[5] P. Osipovs, "Classification tree applying for automated cv filtering in transport company," *Procedia Computer Science*, vol. 149, pp. 406–414, 2019.

[6] J. Martinez-Gil, A. L. Paoletti, and M. Pichler, "A novel approach for learning how to automatically match job offers and candidate profiles," *Information Systems Frontiers*, vol. 22, no. 6, pp. 1265–1274, 2020.

[7] C. de Ruijt and S. Bhulai, "Job recommender systems: A review," *arXiv preprint arXiv:2111.13576*, 2021.

[8] S. Balasundaram and S. Venkatagiri, "A structured approach to implementing robotic process automation in hr," in *Journal of Physics: Conference Series*, vol. 1427, no. 1. IOP Publishing, 2020, p. 012008.

[9] J. Levashina, C. J. Hartwell, F. P. Morgeson, and M. A. Campion, "The structured employment interview: Narrative and quantitative review of the research literature," *Personnel Psychology*, vol. 67, no. 1, pp. 241–293, 2014.

[10] S. Yakkundi, A. Vanjare, V. Wavhal, and S. Patankar, "Interactive interview chatbot," *International Research Journal of Engineering and Technology*, vol. 6, no. 4, pp. 2746–2748, 2019.

[11] I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated analysis and prediction of job interview performance," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 191–204, 2016.

[12] L. Chen, G. Feng, M. P. Martin-Raugh, C. W. Leong, C. Kitchen, S.-Y. Yoon, B. Lehman, H. Kell, and C. M. Lee, "Automatic scoring of monologue video interviews using multimodal cues." in *INTERSPEECH*, 2016, pp. 32–36.

[13] L. S. Nguyen, D. Frauendorfer, M. S. Mast, and D. Gatica-Perez, "Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior," *IEEE transactions on multimedia*, vol. 16, no. 4, pp. 1018–1031, 2014.

[14] L. Hemamou, G. Felhi, V. Vandenbussche, J.-C. Martin, and C. Clavel, "Hirenet: A hierarchical attention model for the automatic analysis of asynchronous video job interviews," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 573–581.

[15] L. Hemamou, "Analyse automatique des comportements multimodaux lors d'entretiens vidéo différés pour le recrutement," Ph.D. dissertation, Université Paris-Saclay, 2021.

[16] L. Hemamou, A. Guillon, J.-C. Martin, and C. Clavel, "Attention slices dans les entretiens d'embauche vidéo différés," in *Workshop sur les Affects, Compagnons artificiels et Interactions*, 2020.

[17] N. Nawaz and A. M. Gomes, "Artificial intelligence chatbots are new recruiters," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 10, no. 9, 2019.

[18] K. Kuksenok and N. Praß, "Transparency in maintenance of recruitment chatbots," *arXiv preprint arXiv:1905.03640*, 2019.

[19] K. Anitha and V. Shanthi, "A study on intervention of chatbots in recruitment," in *Innovations in Information and Communication Technologies (IICT-2020)*. Springer, 2021, pp. 67–74.

[20] "Rasa," https://rasa.com/, accessed: 2023-03-27.

[21] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," 2001.

[22] K. Inoue, K. Hara, D. Lala, K. Yamamoto, S. Nakamura, K. Takanashi, and T. Kawahara, "Job interviewer android with elaborate follow-up question generation," in *Proceedings of the 2020 International Conference on Multimodal Interaction*, 2020, pp. 324–332.

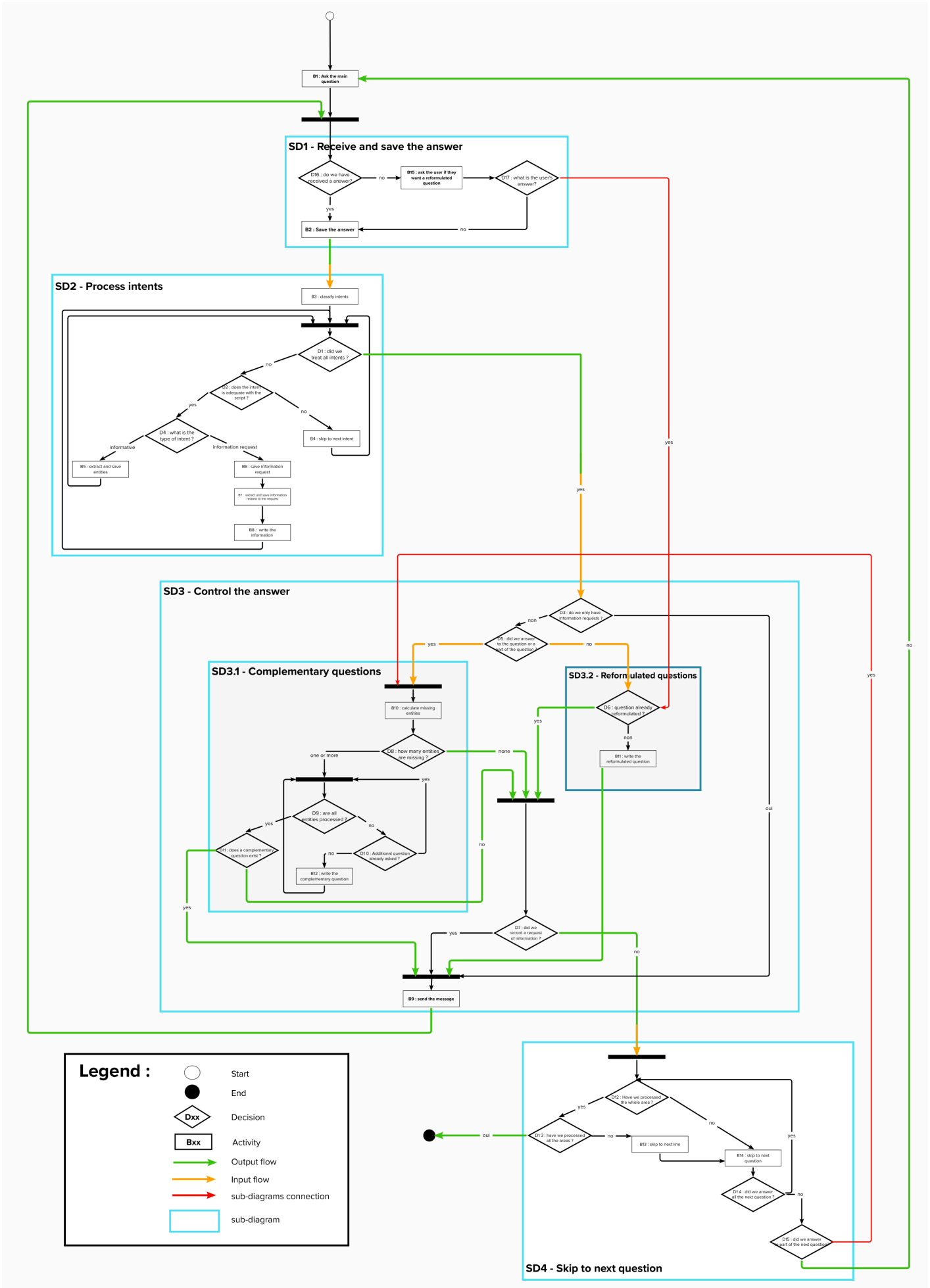


Fig. 11: Detailed process chatbot

Consumer Class Side Scanning Sonar Dataset for Human Detection

T. Aaltonen
Satakunta University of Applied Sciences, Pori, Finland
toni.aaltonen@samk.fi

***Abstract* - One of the leading problems for using modern deep neural networks with sonars is the lack of datasets, even rarer are datasets that are collected with consumer class sonars. This paper introduces a novel side scanning sonar dataset for humans under water. Data is collected with consumer class Garmin 8400 Xsv sonar with GT54UHD-TM transducer. Dataset is collected in shallow coastal water of the Baltic Sea, near Rauma Finland. The dataset contains 331 images of humans, and 364 images with other objects like tires and rocks. Dataset contains cropped images from objects, and full resolution images. Data is collected from two different locations, with different sonar settings. All images are from two rescue divers at the bottom of the sea. This Paper also introduces standard data split for collected dataset for training, validation, and test data for benchmarking different models with dataset, and the data collection system based on ROS.**

Keywords – side scanning sonar; dataset; rosbag; drowing; rescue diving

I. INTRODUCTION

Every year, more than 500,000 people drown in the world, most of them children and young adults [1]. It is of paramount importance for the survival of the victim of an underwater accident that the location is rapid, and that rescue, and possible resuscitation can begin immediately. On the other hand, in cold water, the victim may, under certain conditions, be able to recover more than half an hour after being submerged. Nowadays in Finland, the search continues for at least an hour after receiving an alarm.

In underwater rescue today, finding a victim is greatly influenced by the accuracy of the information about where the victim sank below the surface, if the place is well known and there are no significant currents in the place, there are a lot better chances of a successful rescue mission. The larger the area to be searched, the lower the chances of survival. It is especially difficult to find victim in murky waters, as conditions in Finland often are.

The process of underwater rescue in Finland has remained virtually unchanged through the decades. The safety diver is on the beach and handles a safety rope attached to the diver doing the dive. The searching diver starts from the side of the sector near the shore and begins to move along the bottom towards the other side of the sector, feeling the bottom with his hands until the safety diver detects that the other side of the sector has been reached, signaling the rescue diver to stop, and giving an additional half a meter of the rope. by which the diver moves away from the shore and begins to move toward the other side of the sector, while the bottom is felt by hand. This is continued until the victim is found, the sector has been searched, in which case the search area can be increased, or it is stated that the victim has been below the surface for at least an hour and there is virtually no hope of survival, and urgent search can be stopped.

Process has not changed much in recent years at least in Finland, even survey papers over 15 years ago were stating that development of autonomous underwater vehicles and cheaper consumer grade sonars where developed enough to start to deploy them in search and rescue mission [2].

Motivation behind this paper is to find the victim faster, with use relatively cheap consumer grade side scanning sonar with machine learning model to detect victim. These sonars targeted to consumers are priced from less than a thousand euros to couple of thousands of euros. Sonars used in research and in underwater robotics are usually lot more expensive from tens of thousands of euros hundreds of thousands of euros. Expensive sonars are usually more accurate, and in most cases, they are active scanning sonars, meaning that you get the full updated image even when sonar is still in place, opposite to cheaper side scanning sonars that will produce only one line of data, and sonar must move to get a full image.

The problem is that there are almost no public datasets about sonar images, and many studies are done with very limited amount of data, as stated also by other sonar data searchers [3]. For this work search about other sonar datasets was done in: kaggle databases, paperswithcode databases, datasetsearch provided by google.

kaggle datasets have 43 hits using sonar as keyword, most of these are replications of one old dataset that has sonar pings from rocks and mines [4]. Other search results are not related to underwater sonar but have for example twitter data where part of the user's name has sonar and similar nonrelated datasets, actual only sonar dataset is rock and mine dataset.

google over 100 hits with key word sonar, in this search there is also a lot of same rock or mine dataset

replications. Most of the datasets found with this search was from National Oceanic and Atmospheric Administration, which part of United states of America government. In this NOAA data there is a lot of raw side scanning sonar survey data from different oceans and lakes, another big part of this data is water column sonar data. Other USA government agencies also contributed with different water and ice column datasets. Other nations also have some other sonar survey datasets in these searches, for example Raw side scanning sonar data from Submarine structures in the mid-Irish Sea Area of Search, done by Irish government.

One dataset was a website hosted by North Carolinas government agency NCDMF that has map with sonar scanned artificial reefs, where you can zoom in the map, and save the images of the reefs that you are interested. The entirety of this side scan sonar image dataset was obtained using towfish that were built by Chesapeake Technology and operated at frequencies of 1200 kHz and 600 kHz. The subsequent data analysis was carried out by NCDMF Staff using SonarWeb and SonarWiz software. NCDMF staff members were responsible for gathering all field data from 2000 to 2015 [5].

There is also a dataset of sonar images of shoals of fish, this dataset contains camera and sonar images of fishes in an aquarium. In total dataset contains 1334 sonar images, sonar used to record images is not specified [6].

In papers with code website there are only two datasets with sonar keyword in them, database contains over 7900 datasets at the moment of the writing this paper. There are datasets about common garbage in the ocean [7], [8]. Another one is datasets about fishes, the CFC dataset, which comprises more than 1,500 videos from seven diverse sonar sources, With over 500,000 annotations of different fish species [9].

Outside of these dataset search engines, some datasets were found by reading papers related to AI methods in sonar data, and following references from those papers. An underwater observation dataset for fish classification and fishery assessment, 524 datapoints taken with DIDSON sonar, published 2013[10]. The ARACATI 2017 dataset contains both optical aerial and acoustic underwater images, enabling researchers to compare the two perspectives. By utilizing these diverse image types, dataset allows researchers to examine how differences between image modalities can impact the accuracy of underwater localization [11].

Recently there has also been released a big dataset consisting of over 9000 Multibeam Forward-Looking Sonar images captured using Tritech Gemini 1200ik sonar. Dataset provides raw data of sonar images with annotation of 10 categories of target objects and contains images with human like doll [12]. This dataset is the closest that was found to ours, but it uses human like doll and not a real human, and it is captured using expensive sonar.

Most of the datasets containing sonar images are collected with expensive active scanning imaging sonars, and it is still common to not express the exact sonar model used in the metadata of the dataset. In our best efforts we didn't find any public dataset that contains consumer class sonars with human images. Sonars that are used need to be cheap so that they can be deployed to small rescue departments, either used in boats or as part of an cheap underwater robot.

II. CATHERING THE DATA

The data was collected on a boat of the Satakunta rescue service, to which a Garmin 8400 Xsv sonar, and a GT54UHD-TM echo sounder were attached, as we can see in figure 1. The depth of the sensor could be adjusted by adjusting the height of the sensor mounting post, so data of different depths could be collected, even from the same places. When the boom was close to the maximum depth, it caused the sensor to wobble due to the water resistance, which could slightly reduce the accuracy of the data. However, this is likely to simulate the much faster movements of a much smaller underwater robot as it tries to maintain constant bearing and speed. The sonar image and other sensor data were collected in ROSBAG format on a laptop computer that acted as a ROS master. About 100 man-hours were used to collect the dataset, of which rescue divers accounted for about 30 hours.



Figure 1 Data collection devices installed into the boat of the rescue department. Power to the laptop and to the sonar is taken from the small gasoline generator. The system is designed to be easy to fit to most of the small boats without any holes or other permanent marks [6].

The data collection device shown in figure 2 had Garmin 8400 XSV. Main reason why this sonar plotter was selected was one of the few ones that had a HDMI output. So, it can be connected to laptop via HDMI to USB 3.0 capture card. 8400 XSV also has NMEA2000 data bus, that is also connected to same data stream with teensy 3.6 board. Teensy board was acting as ROS node with ROS-serial node connected with USB 2.0 to laptop. Data collected from NMEA2000 bus contained GPS position, Speed over ground, heading and depth. The device also had BNO055 IMU board connected with

serial to USB 2.0 connection to laptop. Speed over water propeller sensor was used, but data was incorrect to point to be unusable.

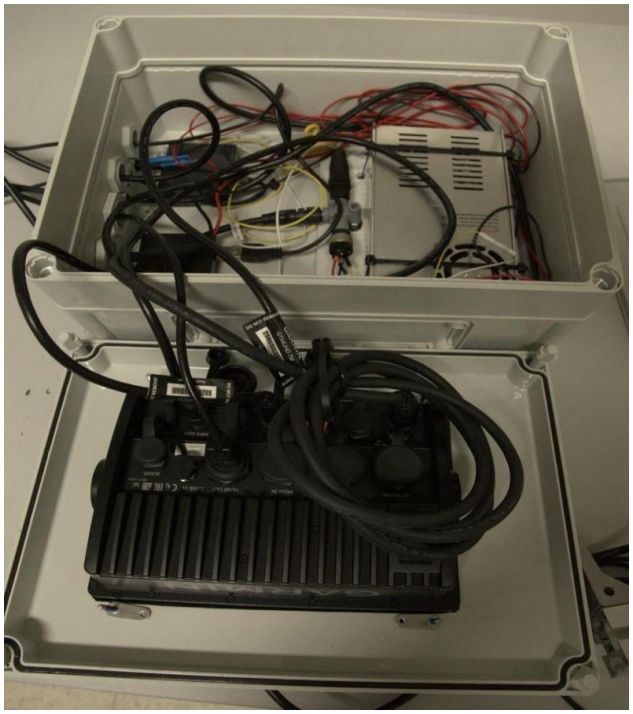


Figure 2 Data collection box, sonar plotter with touch screen is fitted to cover of the box. Meanwell power supply is fitted to the bottom of the box. IMU box, and NMEA2000 to ROS-serial microcontroller is clued to bottom in the left [7].

Sensor boom has 3 different sensors connected to it, as we can see from figure 4. First in the left is scanning sounder LVS12, next is GT54UHD-TM side scanning sonar, last sensor is a speed over water sensor. Sensor boom is made from aluminum. In figure 3 there is the connection diagram, where different connections, sensors and devices can be seen, all different dataflows are changed to USB signals before connecting to computer.

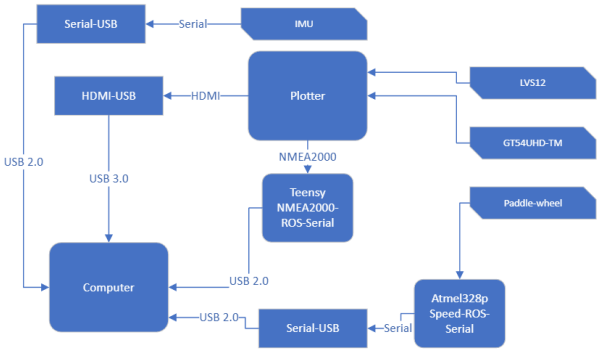


Figure 3 a connection scheme of the data gathering setup. Lines are connecting cables, and boxes are devices.

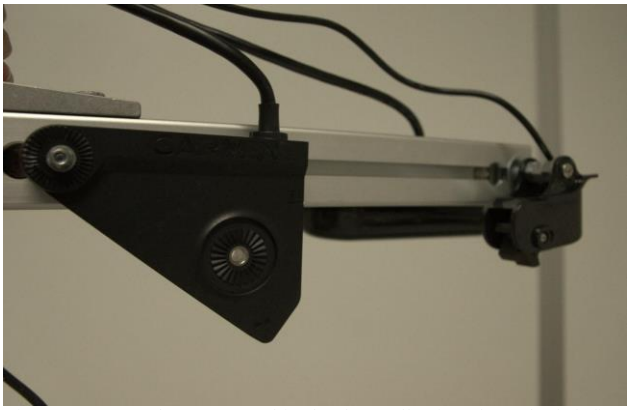


Figure 4 sensors that were used in the data collection. from the left first is real time sensor LVS12, next in the back side is GT54UHD-TM sonar, and the third sensor is speed over water paddle sensor [7].

In all the images, the diver is wearing a wet suit and a face mask. In most of the pictures, the diver's oxygen bottles are so that they are not visible in the picture. Divers are in various positions in different pictures. Data was collected over two days at sea off Rauma. The first day's data is from the Syväraumanlahti area, and the second day's data is from the Maanpäännnokka area. In the data of both days, there are pictures from different points of the bottom in different positions of both divers. Divers have a considerable height difference. In the photos from the first day, both divers have flippers. There are no flippers in the photos from the second day.

The dataset has also left all the worst quality images, which humans have great difficulty in identifying. The Ground Truth information was obtained from these images because it was known that there is a person in it based on the time stamp and location. Otherwise, some would certainly have gone unrecognized. Including such low-quality images in the dataset will certainly worsen the detection accuracy but gives a more realistic picture of the actual detection percentage when using inexpensive consumer grade echo sounders.

During the data preparation phase, the ROSBAG files were manually reviewed and checked, and at the same time stamps were recorded for the points where a person can be seen. After this, a python program was prepared that goes through the ROSBAG files and saves the images from the timestamp points where there is a person. In addition to this, the data was reviewed a second time. In this round, we searched for all those timestamp points where some other objects can be found, which can be used in teaching the neural network for teaching negative cases. Of course, many more such images could have been captured from the data, but we wanted to keep the ratio of positive and negative cases close to one. As we can see from picture figure 5, there are usually more than one object, a pile of rocks or another pattern in the bottom, that can be used. In figure 6 we can see an example of the image where there is a human in the image, some images are easier for humans to recognize than others, this one is between easy and average.

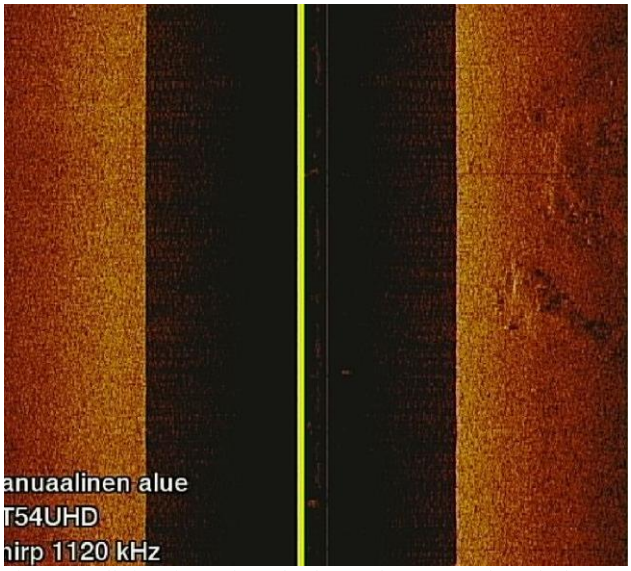


Figure 5 Full size picture from the dataset, boat moves from the bottom to the top, black area is the height of the water, and then finally the sea bead in left and right.

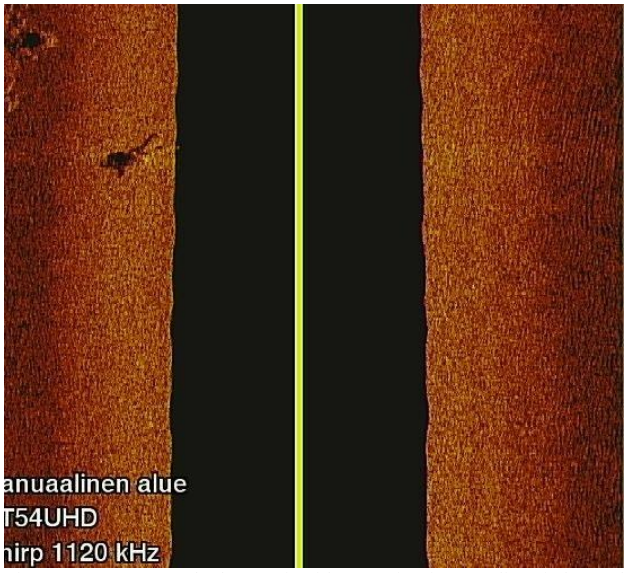


Figure 6 We can see the human in side scanning image, diver is located upper left side of the image, not all images in dataset are as easy find for human eye as this one, legs of the diver are easy to detect in this image [13].

After that, the pictures were in the form shown in figures 5 and 6. After that, the picture of the desired object was still cut from the larger picture. The final images have a resolution of 100x70 pixels and have three color channels. This cropped image size was used, because almost all images fit nicely on that. As we can see from figures 7, many patterns of the bottom can look a little like a human. In figure 8 we can see one example of the final cropped human image.


Figure 7 Cropped image taken from Figure 4, this is example of what kind of features the are in the negative dataset, from pictures where the are some features in the bottom what are not humans.


Figure 8 This image is the cropped version of figure 6. In the image, the diver is lying on his stomach at the bottom of the sea, legs are in upper right part of the picture, body of the diver is casting small shadow behind him in the upper body [13].

The reliability of the collected sonar data is good, and the internal optimal error is zero in light of the best information, but its external reliability is more difficult to define. This refers to how well the data represents all cases that can occur in transom echoes with a person on the waterbed. The only practical way to measure this is to use data from two different days as a separate data set in neural network training, one for training and validation and the other day's data for testing. If the results of the neural network trained with the training data generalize to the test data, it can be said that the dataset is comprehensive enough to generalize at least to a certain extent to the general situation with the sonar used. This can be said because there are many differences in first and second day data, depth is different, bottom of the sea is different, brightness settings are different, and the divers don’t wear flippers in second day, what makes the shape of the foot different. With small dataset like this, division to different days we can test the generalization of the model a lot better, and this is recommended way to use this dataset.

The final training/validation data set collected on the first day consists of 205 images of a person at the bottom of the sea, and 249 images with an object other than a person. Images of, for example, stones, plants, and car tires belong to this category. In addition to this, the test data set made from the data set recorded in the second data collection session contains a total of 125 images of people at the bottom of the ocean and 84 images where there are other objects or plain bottom of the ocean. All in all, there are a total of 695 pictures, of which there are 331 pictures of people, and 364 pictures without people.

Default dataset division is done in the way that would make train and validation dataset as separated from the test data as possible, and the second dataset is assigned to test date. Train and validation are made from the first dataset, by splitting it with random function. Full size images and cropped images of human are all with one-to-one naming, starting from 000000.jpg. Nonhuman images are not always one to one mapped from full size to cropped, because there are some pictures with multiple cropped pictures.

Default division for train and validation datasets, is provided with the dataset. Dataset is ready to use with neural networks without any preprocessing, with categorical detection networks like densenet, renesnets, and object detectors based on that kind of networks. This is main intended use of dataset, because in general that type of detectors needs less data to train, and the nature of the image produced by side scanning sonar is that is updated one horizontal line at the time, so sliding window type of object detection fits well to the problem. Extensive data augmentation is recommended, because of the small dataset size. If used with Yolo type object detection algorithms then use of full-size images are recommended, and annotation is needed, and is not currently provided with the dataset.

III. CONCLUSION AND DISCUSSION

There are many well-known limitations in this published dataset. The size of the dataset is still very limited, and it contains only data from the two different people, and all pictures are from adults, also in the images divers have wet suit, and that might give a different echo than human skin or other clothing. Still, it gives a novel dataset to understand what kind of quality can be obtained from relatively cheap side scanning sonars, and there is enough data to do preliminary tests with neural networks.

One important observation was that with consumer class sonars, the depth of the water needs to be shallow, best range seems to be from 1,5 – 3 meters and after 6 meters the human size object starts to be too small to detect for trained human eye. That means that sonar connected directly on the bottom of the boat limits the detection range to quite shallow water. Use of the sensor boom can extend this with some meters. Other options are to use depth controlled towed sonar, or underwater robot.

Dataset can also be used for training for rescue people for use side scanning sonars without any computer vision, with sonars mounted in to the boat. Observation after doing the work with this dataset is that after looking at sonar data with humans, detection comes a lot easier. Typical features start to clearer to see. In shallow murky coastal waters this could be helpful to detect victim faster than current methods.

Data is published in github repository, with following link: https://github.com/tonaalt/sonar_human_dataset

IV. FUTURE STUDIES

Next work is to publish results of the trained neural networks based on this data, using different neural network architectures. This study has been mostly done, but still no peer review papers have been published.

More data is needed to gather, to get more certain results, and it would be better to collect some images without wetsuit. An interesting option that we have been starting to investigate is the creation of synthetic or semi-synthetic data using different AI and traditional methods.

ACKNOWLEDGMENT

This work has been supported by Finnish Ministry of education project AI Forum and Satakunta Rescue Department Unit of Rauma.

REFERENCES

[1] D. Szpilman, J. L. M. Bierens, A. J. Handley, and J. P. Orlowski, “Current Concepts Drowning,” 2012.

[2] A. J. Murphy, M. J. Landamore, and R. W. Birmingham, “The role of autonomous underwater vehicles for marine search and rescue operations,” *Underwater Technology*, vol. 27, no. 4, pp. 195–205, 2008, doi: 10.3723/ut.27.195.

[3] M. Jegorova, “Generative Neural Data Synthesis for Autonomous Systems,” no. July, 2021, doi: 10.13140/RG.2.2.19876.60801.

[4] R. P. Gorman and T. J. Sejnowski, “Analysis of hidden units in a layered network trained to classify sonar targets,” *Neural Networks*, vol. 1, no. 1, pp. 75–89, Jan. 1988, doi: 10.1016/0893-6080(88)90023-8.

[5] “Artificial Reef Sonar Scans | Artificial Reef Sonar Scans | NC OneMap.”https://data-nconemap.opendata.arcgis.com/datasets/artificial-reef-sonar-scans/explore (accessed Mar. 29, 2023).

[6] T. Kei, K. Shin, M. Katsunori, and T. Koji, “A dataset of camera and sonar images.” 2018.

[7] M. Valdenegro-Toro, “Deep Neural Networks for Marine Debris Detection in Sonar Images,” no. April, 2019, [Online]. Available: http://arxiv.org/abs/1905.05241

[8] M. Valdenegro-Toro, A. Preciado-Grijalva, and B. Wehbe, “Pre-trained Models for Sonar Images,” 2021, [Online]. Available: http://arxiv.org/abs/2108.01111

[9] J. Kay *et al.*, “The Caltech Fish Counting Dataset: A Benchmark for Multiple-Object Tracking and Counting,” Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.09295

[10] E. McCann, L. Li, K. Pangle, N. Johnson, and J. Eickholt, “Data descriptor: An underwater observation dataset for fish classification and fishery assessment,” *Sci Data*, vol. 5, Oct. 2018, doi: 10.1038/sdata.2018.190.

[11] M. M. D. Santos, G. G. De Giacomo, P. L. J. Drews-Jr, and S. S. C. Botelho, "Cross-View and Cross-Domain Underwater Localization Based on Optical Aerial and Acoustic Underwater Images," *IEEE Robot Autom Lett*, vol. 7, no. 2, pp. 4969–4974, 2022, doi: 10.1109/LRA.2022.3154482.

[12] K. Xie, J. Yang, and K. Qiu, "A Dataset with Multibeam Forward-Looking Sonar for Underwater Object Detection," *Sci Data*, vol. 9, no. 1, p. 739, Dec. 2022, doi: 10.1038/s41597-022-01854-w.

[13] "tonaalt/sonar_human_dataset: Public sonar dataset dataset, two categories humans and other. Gathered with garmin consumer side scanning sonar." https://github.com/tonaalt/sonar_human_dataset (accessed Mar. 30, 2023).

Network Model of Multiagent Communication of Traffic Inspection for Supervision and Control of Passenger Transportation in Road and City Traffic

Aleksandar Đukić*, Radenka Bjelošević*, Mirko Stojčić*, Milorad K. Banjanin**
* University of East Sarajevo, Faculty of Transport and Traffic Engineering, Vojvode Mišića 52, 74000 Doboj, Bosnia and Herzegovina
** University of East Sarajevo, Faculty of Philosophy, Alekse Šantića 1, 71420 Pale, Bosnia and Herzegovina
** University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
a.djukic@inspektorat.vladars.net; radenka.bjelosevic@sf.ues.rs.ba; mirko.stojcic@sf.ues.rs.ba; milorad.banjanin@ff.ues.rs.ba

***Abstract* - As a rule, the supervision and control of the transportation of persons in different types of traffic is the responsibility of the Traffic Inspection of the state inspectorates of individual countries. This paper investigates the intelligent supervision and control of the transportation of persons in road and city traffic managed by the Traffic Inspection in the geo-space of the Republic of Srpska (RS). The aim of the research is to create a multiagent communication (MAC) network model of an ensemble of intelligent interactive teams of human agents (HA), software agents (SA) and cyber-physical agents (CPA) located in six distributed nodes of a star network topology. All nodes of this topology represent Intelligent communication agents, and a special team of intelligent management agents for Inspection Affairs is located in the sixth node of the MAC network model. In the proposed model, the communication platform for MAC is supported by process-adaptable applications of machine learning methods, and the ability of agents to learn from experience and predict based on digital data processing. The main goal of MAC model is optimization of solutions for timely planning, reliable functioning, continuous monitoring, preventive supervision and effective management of Traffic Inspection tasks in uncertain situations contexts.**

Keywords - multiagent communication, traffic inspection, agent, transport of persons, prediction, classification, categorization, situational contexts

I. INTRODUCTION

Logical and structural, traffic and transport systems are engineering systems whose structuring and behavior are inevitably influenced by innovative-technological, legal-regulatory, organizational-cultural, energy-ecological, social-contextual and other environmental factors. They are holonic structures characterized by the great complexity of the constellation of relations and relationships between numerous components of the system. The complexity of the configuration of states in the entire system, or some of its components as modules function is investigated in real situations. In the structure of the traffic system module of road and urban passenger transport, a special functional component is the *traffic inspection*. This is the subject of research in this paper, where the geo-space of the Republic of Srpska (RS), Bosnia and Herzegovina (BiH) was chosen as a case study. In this geo-space the system and process structuring and functioning of traffic is carried out according to technical-technological standards, norms, algorithms and protocols. This takes place in the organizational-management and social-contextual domains on the basis of prescribed legal norms and rules. Legal norms and consequent rules are prescribed by a set of Republican and state laws, international regulations, decrees, specifications and specialized functional and non-functional procedures. Traffic inspection works are planned, implemented, monitored and controlled according to the defined roles of intelligent agents in simple, complicated and complex situational contexts of the traffic system.

The hypothetical premise of this research is that the content-functional, technological-process and relational-contextual structure of traffic inspection works can function more accurately, precisely and reliably according to the network model of multiagent communication (MAC). The concept of the model is based on the interactive roles of human (HA), software (SA) and cyber-physical (CPA) communication agents and management agents. Agents function on knowledge and learning through the execution of tasks of supervision and control over objects of traffic flows and transport processes in unstable situational contexts. In the focus of communication agents are problems and solutions from the domain of safety of traffic participants, reliability of exploitation of mobile objects and stable infrastructure objects. Also, an important goal of MAC is information security of structured data and messages that are transmitted, exchanged and shared between singular agents, groups, teams and ensembles in the network architecture.

Intelligent management agents monitor, evaluate and control value attributes of structural dimensions of functional, equivalent and management processes, implementation of strategy theory in activities, legal

behavior of HA in singular and plural roles of traffic inspection. At the same time, points of interest are the functional and non-functional specifications of operational technologies, software applications and cyber-physical modules of learning and management in unstable contexts of road traffic flows. The interaction field of meaningful, dynamic, proactive, irreversible, interactive and contextual multimodal communication behavior of management agents in the ensemble is expanded by cooperation with non-inspection bodies and organizations. Their aim (goal) is solving specific tasks of prediction, classification, categorization and evaluation of inspection supervision and control works. Based on these conceptual settings, the proposal of the MAC network model of a real ensemble of interactive groups and teams HA, SA and CPA, which are located in six spatially distributed nodes of the star network topology, was created in the paper. The model is goal-oriented towards clearly defined functional tasks, stable problems and finding solutions through improving the way of working and decision-making in the work of the Traffic Inspection. The action platform of the MAC network model is supported by process-adaptive applications of machine learning methods and the ability of agents to learn from experience and make predictions based on digital data processing. Section 6 of this paper presents the results of training and testing of a classification model for solving one of the tasks within the intelligent and legal behavior of human agents in the singular and plural roles of the Traffic Inspection. The relevance of the model was tested in a case study in a selected geo-space of traffic area with the goal that the created solutions can have universal applicability.

II. Review of Relevant Published Research

So far, numerous studies have been published related to the application of the MAC model in the field of road and city traffic with different target orientations. In paper [1], an original model of the Next Road Rerouting (NRR) vehicle routing system based on the communication of intelligent agents was proposed. The goal is to assist drivers, based on the cost function, the driver's destination and local conditions, in finding the optimal route in order to avoid unexpected traffic jams. In [2], the authors presented a new adaptive multiagent system for road traffic management based on the ant colony algorithm, which is characterized by adaptability and the short time required to calculate new vehicle movement routes. The subject of research in [3] is a model for multiagent traffic signal control based on the decentralized and scalable Multi-Agent Reinforcement Learning (MARL) method. Considering the existence of a very large number of intersections in an urban traffic environment, the authors in [4] use Independent Reinforcement Learning (IRL) to solve the problem of cooperative traffic control. The aim of the research presented in [5] is to develop a multiagent traffic simulation methodology in order to evaluate potential road safety improvements using new technologies. In paper [6], the traffic signal control optimization problem is formulated as a distributed constraint optimization problem. Multiple agents are used to control road traffic based on traffic density at adjacent intersections. The study presented in [7] deals with the application of multi-agent learning algorithms with support for cooperative traffic signal control. A method based on Multi-agent Deep Deterministic Policy Gradient (MADDPG) was proposed in order to reduce the average waiting time of vehicles. Similar research was published in [8] where a traffic optimization system based on agent technology and fuzzy logic was presented.

III. The Roles of Intelligent Agents in the Multiagent Communication Model

In this paper, agents are "explicitly observed situated or positioned entities in the road and city traffic system that communicate" interactively in the contexts of HA, SA and CPA role performers in traffic inspection activities. The roles of groups, teams and ensembles of intelligent communication agents are oriented towards solving the tasks of monitoring and controlling actions, activities, events and processes in the domain of consistent application of traffic normative-legal regulations, representation of technical standards in traffic-infrastructure facilities, traffic-safe exploitation of dynamic facilities in public transport of people by buses, taxis and vehicles of rent-a-car agencies. In addition, agent roles in the MAC network model digitally treat data on the reliable operation of vehicle technical inspection stations, the safe operation of fuel filling stations and other external road and city traffic infrastructure facilities. In reality, communication is inseparable from some mode of information processing. The analogy established by the research between the computer and human way of information processing and the compatibility of physical data flows with the operations of software applications in cyber-physical systems enabled the development of multi-agent communication in the changing and adaptable situational contexts of road and city traffic and transport. Based on this, for the competent role-playing of the traffic inspection, the levels of knowledge and competence of the HA, the validity of the SA program and the functional ability of the CPA are important. Knowledge capacity and learning of all agents is a common characteristic on which the goal and methodology of creating a network model MAC in traffic inspection work is based.

Human agents (HA) - it is known that the human body functions as a neurophysiological system on three interactive communication platforms - cognitive, behavioral and conative. The human cognitive system has a powerful memory architecture whose components are: sensory register (SR) - closely connected with the modular sensory system; short-term, conscious, working memory (SM) and long-term memory (LM) in which experience and knowledge about phenomena, events and other knowledge are stored. Knowledge can be episodic (about specific events in spatio-temporal contexts) and semantic (abstracted from the spatio-temporal context). Semantic knowledge is declarative (know that) - general knowledge about the characteristics of a phenomenon or event and procedural (know how) - knowledge about procedures and actions which achieve certain goals. In the description of these components of the cognitive system, several communication dimensions are used, namely: source of information (speech, image, sound, computer data), capacity, function, length of retention of information, form in which information is stored, ways of which information

is lost. On the behavioral platform HA communications can function at the level of data. In simple situational contexts - at the level of information, and in complicated situational contexts - at the level of acquired knowledge and applied knowledge. But in complex situational contexts, their communication must function at the level of intelligence and wisdom or the level of the total capacity of the repertoire with all the mentioned types of cognitive continuum. On the conative platform, HAs function depending on the type, complexity and specificity of the tasks they perform in the traffic inspection business. Therefore, in HA management, the most important role is business process management (BPM). BMP is seen as a set of business process-oriented management approaches (BPO) instead of a classic functional organizational structure.

Software agents (SA) - although there are a large number of definitions of SA, papers [9,10] state that SA represents a software component capable of exchanging knowledge and information. Also, in the literature there are a large number of properties that can be attributed to software agents. In the paper [11], encapsulation, autonomy, interaction, persistence, activity and goal-oriented behavior are listed as strong properties, while mobility, truthfulness, benevolence and rationality are highlighted as additional weak properties. In the paper [12], the properties listed are autonomy, cooperation and learning of SA, on the basis of which their further classification can be made. Autonomy means the ability of agents to act without human intervention based on their internal states and goals. Proactivity is a key element of their autonomy and defines the ability to "take the initiative" [12]. Cooperation or collaboration with other agents is the ability to interact with other agents or humans using a language. Software agents learn while reacting and/or interacting with their environment, and learning is precisely one of the key features of every intelligent being [12]. Therefore, SA can be defined as an autonomous software entity located in an environment where it can proactively or reactively monitor and react to changes. This reaction is also possible through communication with other agents, in order to achieve a certain goal on behalf of the user or other agents [9,13,14]. Since SA autonomous units are capable of running on distributed computers, communication between agents is Machine-to-Machine (M2M) communication. Intelligent SAs are usually modeled and programmed with agent-oriented programming languages based on the principles of modal Belief-Desire-Intention (BDI) logic [15].

Cyber-physical agents (CPA) - represent the integration of components of the biological, physical and virtual world. Within the MAC model, they perform their role by acquiring data on physical variables using sensors, automatically controlling elements of the physical world using actuators, and promptly providing information for users and other agents. The virtual or cyber component of CPA is based on embedded computers that monitor and control physical processes in a network where many things happen in parallel, usually with feedback loops [16,17]. This means that physical processes also affect computer processes, which are defined by sequential steps. The feedback loop between physical and computational processes includes sensors, actuators, physical dynamics, calculations, software failures and network problems, communication delays. Thus, hardware elements - sensors, microcontrollers and actuators, as part of CPA function integrally with cyber elements such as computer algorithms for data processing, for monitoring and controlling physical processes, for making decisions, for managing storage locations and for data analytics [16,17].

The learning module and the decision-making module are integral parts of the CPA architecture. They play a very important role in the analysis and modeling of a large amount of collected data in physical streams that are fed into virtual sensor-network applications. The best solution lies in adaptive models based on artificial intelligence that have the ability to learn. Therefore, the most common techniques and algorithms implemented by the learning module are based on machine learning. Thus, virtual (software) sensors can also perform learning and decision-making functions based on various data obtained using physical sensors [17].

IV. Topology and Interactions in the Model of Multiagent Communication

In the proposed real network model of Traffic Inspection MAC, the agent teams are connected in a star network topology, as shown in Fig. 1. Each communication agent team has a point-to-point connection with the central management team. Intelligent management agents in the sixth-central team meaningfully, dynamically, proactively, irreversibly and interactively collaborate with all other agents of the ensemble. The management team node, which has the role of a server, controls all other network nodes in the roles of clients and is responsible for routing all data and information in the network. Management agents pursue goals respecting the environment and their defined space of action. Also, they create operational goals for the benefit of HA or other management agents with a certain degree of autonomy. Management agents proactively manage goals, take responsibility, make decisions, and create a model of their environment.

Communication agents distributed in six regional teams (Fig.1) enable agents to communicate with each other and generate the correct flow of relevant information at the exit point of the inspection process (Fig. 1(a)). They must possess the ability to reason in order to understand any information or actions in the inspection process. This also refers to the goals of management agents in generating the performance of traffic inspection works.

Within individual teams, as already mentioned, there are SA, CPA and HA that are connected in a Full Mesh topology, where each agent is connected to all other agents by direct links (Fig. 1(b)). In this way, the redundancy of the connection is increased, as well as the speed of data transmission and the robustness of the network. In this peer-to-peer network team model, each agent performs the role of server and client. The displayed network topology of six distributed teams of agents corresponds to the territorial organization of the observed space, which consists of six regions according to the

Spatial Plan until 2025. Therefore, each of the six teams belongs to one region: Team 1 - Prijedor, Team 2 - Doboj, Team 3 - Bijeljina, Team 4 - East Sarajevo, Team 5 - Trebinje and Team 6 - Banja Luka. The headquarters of the *Republic traffic inspection* is located in Banja Luka. That is why this city was chosen as the central node or the sixth team of agents, while each region includes several local self-government units.

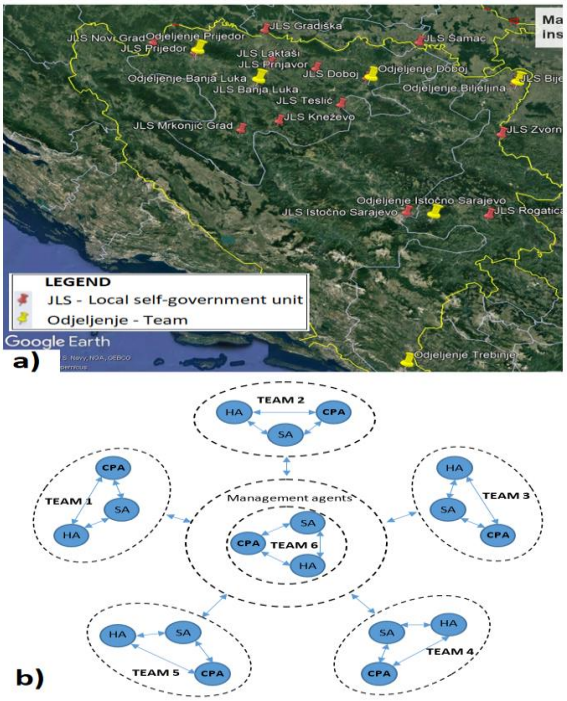


Figure 1. Communication agents distributed in six regional teams (RS, BiH): (a) Map of agents of the RS Traffic Inspection; (b) Network topology of the Traffic Inspection MAC model

Fig. 2 shows the interactions in the MAC Traffic Inspection model that imply on request or periodic exchange of data and information between agents in a team and between ensemble teams. Given that it was previously stated in section 3 that CPA consist of physical and virtual components, each of them has its own roles and tasks within the agent team and ensemble of the proposed MAC model. The physical component uses sensors to acquire data from the environment where objects and inspection control and monitoring processes are located (e.g. vehicles, drivers, roads, documentation, etc.).

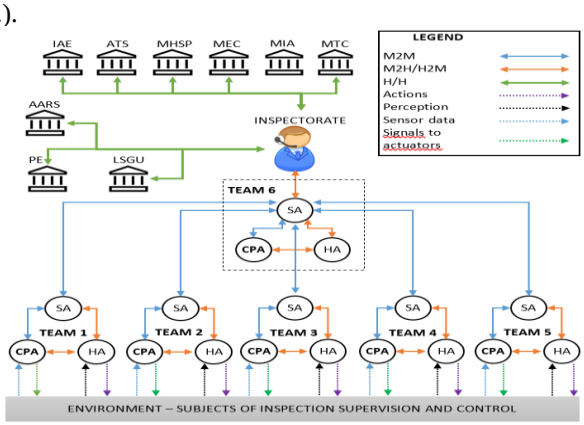


Figure 2. Interactions in the proposed MAC model

The virtual or cyber component of the CPA processes this data in the learning module and the decision module, and forwards the information to actuators or HA that act on the physical world (eg. by alerting inspectors or the police). In addition, CPA may forward information to SA and/or HA. HA or inspectors in the field receive data from the environment by perception, and after processing, perform actions on objects of control and supervision or forward data and information to SA and CPA. SA provide support to HA and CPA in data processing with machine learning methods and applications, which is shown by the results of the classification predictive model in Section 6 on the example of one of the Traffic Inspection's monitoring and control tasks.

Each agent communicates with all other agents from the team. Thus, the communication between HA and CPA, and between HA and SA is Human-to-Machine/Machine-to-Human (H2M/M2H), and between CPA and SA is M2M. The management team of the agent ensemble processes data from all teams and, if necessary, returns results, orders and tasks to individual teams. Important information is forwarded to the Chief Republic Inspector of the Traffic Inspection by H2M/M2H communication so that he has an insight into the work of the ensemble of agent teams. Furthermore, it is possible to direct them to the Ministry of Transport and Communications (MTC), the Ministry of Internal Affairs (MIA), the Ministry of Education and Culture (MEC), the Ministry of Health and Social Protection (MHSP), the Agency for Traffic Safety (ATS), the Institute for Adult Education (IAE), Automobile Association of RS (AARS), public enterprises (PE), entrusted management of the road network, bodies of local self-government units (LSGU). This way of forwarding data and information, from collection from the environment, through an ensemble of agents, all the way to institutions at the top of the jurisdiction, can be characterized as bottom-up. In addition, tasks, requests and inquiries can be sent from the top and passed down to agents in individual teams acting on the environment, which corresponds to a top-down flow of data and information.

This paper examines the intelligent and legal behavior of HA - traffic inspectors, expressed through numerical scores on a scale from 1 to 5 as an example of one of the tasks of supervision and control of the Traffic Inspectorate. Machine learning models for the classification of agents according to scores, the results of which are given in section 6, have been implemented in SA and allow the management team to evaluate the work of inspectors in the field, based on the input data obtained by contextual sensors or by other HA that have the role of control and supervision.

V. Machine Learning Models for Solving Inspection Monitoring and Control Tasks

In the proposed model, the MAC platform is supported by machine learning techniques, the most common of which are: 1) Artificial Neural Networks (ANN); 2) Decision Trees (Classification And Regression Trees - CART); 3) Support Vector Machines (SVM) [18]; and 4) k-Nearest Neighbors (k-NN).

In general, machine learning models usually represent very good solutions for prediction and classification problems in various fields. But, some research suggests that it is often necessary and desirable to transform the machine learning problem into a problem of cooperation between agents. In that way, the problem is broken down into several smaller problems that are solved by individual agents, thereby reducing its complexity and enabling the system to adapt to the evolution of individuals [18]. Our paper, proposed the MAC model consists of an ensemble of agents that can solve problems of control and supervision of the transportation of people in road and city traffic. That is a distributed manner using the Ensemble Learning method based on the idea of combining multiple individual machine learning models, commonly referred to as "weak learners". It can result in a model with better predictive performance than any single model. There are two main approaches to ensemble learning [18]: Bagging and Boosting [19]. Considering that one of the key steps for defining a model or an ensemble of ML models is the selection of independent/input and dependent/output variables. Table 1 shows input and output variables by individual models for solving the problem of monitoring and control in the tasks of traffic inspection.

TABLE I. INPUT AND OUTPUT VARIABLES OF THE MODEL FOR SOLVING TRAFFIC INSPECTION SUPERVISION AND CONTROL TASKS

Supervision and control tasks	Input variables	Output variables
1. Consistent application of traffic regulations	Type of regulation; Geographical location; Subject of control.	Number of offences
2. Representation of technical standards in traffic infrastructure facilities	Municipality/city; Control object: 1)Physically stable and some mobile objects of traffic and transport; 2)Legal-regulatory, energy-economic, security-protective and ecologically-sustainable facilities; Organization in charge of maintenance; Age of the object; Period of the year.	Safety assessment of facilities; Control result.
3. Traffic-safe exploitation of dynamic facilities in public transport of people by buses, taxis and vehicles of rent-a-car agencies	Number of regular controls of technical correctness; The number of extraordinary inspections of technical correctness; Prices of services; Regularity of departures in public transport;	Level of satisfaction of users of transport services [20]
4. Intelligent and legal behavior of human agents	Level of government; Competences; Role in the team; Location where it operates; Level of functional knowledge.	Agent grades
5. Functional implementation of operational software and cyber-physical technologies in the context of traffic flow security	Annual investments; The level of use of conventional business technologies; The level of use of Information Technologies (IT); Level of use of communication technologies (CT); Use of personal networks; The level of use of information and communication technologies (ICT); The level of use of software applications in CPA interactions.	Achieved results in terms of safety; Number of traffic accidents.

In accordance with the fourth task of supervision and control given in Table 1, the following are observed as input or independent variables of the machine learning model for the classification of HA according to grade: *Level of government (Level)*, which can be municipal or republican; *Competence*, expressed through attributes such as: experienced, communicative, meticulous, strict...; *Team role*, which can be: republican inspector, chief republican inspector, regional coordinator, city inspector or municipal inspector; *The location where it operates (Location)*, which means the city/municipality in the RS; *Level of functional knowledge*, which can be: data, information, knowledge, applied knowledge, intelligence, wisdom, capacity for knowledge sharing and exchange at complexity situations. The agent's grades, which can be a discrete value in the interval from 1 (unsatisfactory) to 5 (excellent), is viewed as the dependent or output variable.

VI. A MACHINE LEARNING MODEL FOR THE CLASSIFICATION OF AGENTS BY GRADES

This section presents the results of training and testing of machine learning models for solving one of the tasks within the intelligent and legal behavior of HA in singular and plural roles of traffic inspection. The goal is to classify the agents according to their ratings depending on the previously mentioned input/independent variables. The creation of the classification model was performed in the IBM SPSS Modeler software environment, whose main purpose is to work with predictive and classification models. The collected data for model training and testing are structured into 48 input/output vectors, 70% of which belong to the training set and 30% to the model testing set. Using the automatic classification method using the Auto Classifier option, models based on SVM, k-NN and CART techniques with default parameters were examined. After training and testing, SPSS Modeler ranked the mentioned models according to the overall classification accuracy, and the results are shown in Table 2.

TABLE II. RANKING OF THE TESTED MODELS FOR THE CLASSIFICATION OF AGENTS ACCORDING TO GRADES

Tested model	Overall classification accuracy
1. CART	84.615%
2. k-NN	69.231%
3. SVM	69.231%

As it can be seen from the results given in Table 2, the best classification performance is shown by the model based on CART machine learning technique with 84.615% accuracy of classified vectors. Models based on k-NN and SVM have equal classification accuracy values of 69.231%. As an significant result of the training and testing of the mentioned models using the Auto Classifier method, Fig. 3 shows the importance of the influence of each of the five independent variables (Predictor importance) on the classification results. As can be seen from the picture, the level of functional knowledge is the most influential variable on the classification results. But the level of action of the agents (municipal or republican) is the variable with the least influence on the classification performance of the model.

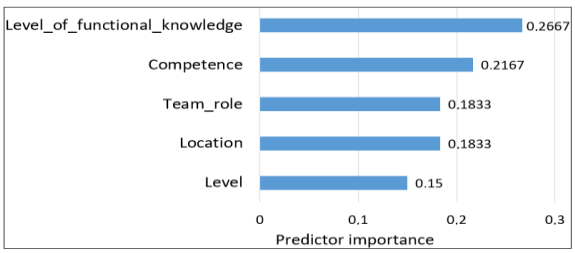


Figure 3. The importance of the influence of each of the five independent variables on the classification results

Fig. 4 shows the classification results for the best created model - CART according to Table 2. The x-axis of the diagram shows the grades of the agents from the data set for testing, while the y-axis shows the number of tested observations. It can be concluded that grades 2 (satisfactory) are correctly classified in all 6 observations, grades 3 (good) are correctly classified in 5 observations, and grades 4 (very good) are misclassified as well as grade 3 in two observations from the test set.

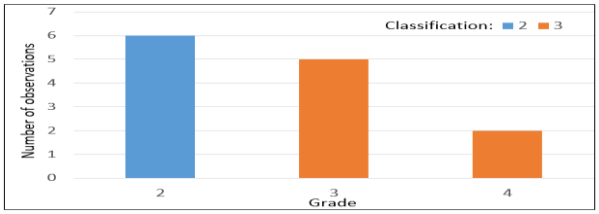


Figure 4. Agent classification results according to grades for the CART model

VII. CONCLUSION

The research in this paper is focused on the modeling of multiagent communication in the work of the Traffic Inspection which supervises and controls road and city traffic. The MAC model is presented in relational, descriptive, analytical, visual and classification modes based on the use of features of interactive HA, SA and CPA. Even though the relevance of the model was tested in selected geo-space of RS traffic area the aim of the created solution can have universal applicability. The MAC is presented in a star network topology with six distributed nodes according to the geo-locations shown on the map where intelligent communication agents are located. The sixth node also having intelligent management agents. A special diagram of interactions shows the possible state of functioning of the MAC.

In order to verify the functionality of the created MAC, machine learning models were specially created for the classification of HA categorical variables based on the collected grades of the agents. Data on selected independent and dependent categorical work variables in control and monitoring processes and tasks are structured in 48 input/output vectors, of which 70% of the volume is in the training set, and 30% in the model testing set. Using the automatic classification method by the Auto Classifier option, models based on SVM, k-NN and CART techniques with default parameters were examined. After training and testing, SPSS Modeler ranked the mentioned models according to the overall accuracy of the classification. The results indicate that the conceptual analysis, subject and methodology of the research justified the status of a scientific-innovative choice of the type of problem in the field of strengthening traffic safety. Actually that is a new topic of research in traffic and transport engineering systems. The main drawback and limitation of the paper is that the proposed model solves only one problem within the framework of traffic inspection supervision and control. Therefore, future research directions are oriented towards the modeling of other inspection and control tasks, as well as the application of the Ensemble Learning method to improve the accuracy performance of the model classification using the Boosting approach.

REFERENCES

[1] S. Wang, S. Djahel, Y. Zhang & J. McManis, "Next road rerouting: A multiagent system for mitigating unexpected urban traffic congestion", IEEE Transactions on Intelligent Transportation Systems, 17(10), 2888-2899, 2016

[2] H. Hamidi & A. Kamankesh, "An approach to intelligent traffic management system using a multi-agent system", International Journal of Intelligent Transportation Systems Research, 16, 112-124, 2018

[3] X. Wang, L. Ke, Z. Qiao, & X. Chai, "Large-scale traffic signal control using a novel multiagent reinforcement learning", IEEE transactions on cybernetics, 51(1), 174-187, 2020.

[4] C. Zhan, S. Jin, W. Xue, X. Xie, S. Chen & R. Chen, R., "Independent reinforcement learning for weakly cooperative multiagent traffic control problem", IEEE Transactions on Vehicular Technology, 70(8), 7426-7436, 2021

[5] S. Kitajima K. Shimono, J. Tajima, J. Antona-Makoshi & N. Uchida, "Multi-agent traffic simulations to estimate the impact of automated technologies on safety", Traffic injury prevention, 20 (sup1), S58-S64, 2019

[6] J. Withanawasam & A. Karunananda, "Multi-agent based road traffic control optimization", In 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC) (pp. 977-981). IEEE, 2017, October

[7] S. Li, "Multi-agent deep deterministic policy gradient for traffic signal control on urban road network", In 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA) (pp. 896-900). IEEE. 2020, August

[8] A. Ikidid, A. El Fazziki & M. Sadgal, "A Fuzzy Logic Supported Multi-Agent System For Urban Traffic And Priority Link Control", J. Univers. Comput. Sci., 27(10), 1026-1045, 2021

[9] G. K. Soon, G. K., On, C. K., Anthony, P., & Hamdan, A. R., "A review on agent communication language", Computational Science and Technology: 5th ICCST 2018, Kota Kinabalu, Malaysia, 29-30 August 2018, pp. 481-491.

[10] M. R. Genesereth, S. P. Ketchpel, "Software agents", Communications of the ACM, 37 (7), (1994) 48–53

[11] M. Juettner, S. Grabmaier, D. Vögeli, W. M. Rucker & P. Göhner, "Coupled multiphysics problems as market place for competing autonomous software agents", IEEE Transactions on Magnetics (IEEE Xplore), 53(6), 1-4, 2017

[12] H. S. Nwana, "Software agents: An overview" The knowledge engineering review, 11(3), 205-244, 1996

[13] M. Wooldridge, "An Introduction to Multiagent Systems", 2nd Edition. John Wiley & Sons, Inc., New York, 2009

[14] R. J. McCall, S. Franklin, U. Faghihi, J. Snaider, & S. Kugele, S., "Artificial motivation for cognitive software agents", Journal of Artificial General Intelligence, 11(1), 38-69, 2020

[15] N. Alechina, M. Baldoni & B. Logan B., "Engineering Multi-Agent Systems", 9th International Workshop, EMAS 2021 (Vol. 13190, pp. 1-381). Springer Cham. 2022

[16] M. K. Banjanin & M. Stojčić, "Conceptual Model of the Cyber-physical System in the Space of the M9J Road Section", In 2021 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS) (pp. 299-302). IEEE. 2021, October

[17] M. K. Banjanin, M. Stojčić & D. Drajić, "Software Networks in the Logical Architecture of the Cyber-Physical Traffic System", Science and Higher Education in Function of Sustainable Development – SED 2021, Užice, Serbia, 2021.

[18] T. Fourez, N. Verstaevel, F. Migeon, F. Schettini, & F. Amblard, "An ensemble Multi-Agent System for non-linear classification" arXiv preprint arXiv:2209.06824, 2022

[19] Projectpro. Retrieved February 3, 2023. from https://www.projectpro.io/article/bagging-vs-boosting-in-machine-learning/579

[20] M. K. Banjanin, M. Stojčić, D. Danilović, Z. Ćurguz, M. Vasiljević & G. Puzić, G., "Classification and Prediction of Sustainable Quality of Experience of Telecommunication Service Users Using Machine Learning Models", Sustainability, *14*(24), 17053, 2022.

Closed-loop Artificial Pancreas Development: A Review

Sabina Glumčević [1], Zerina Mašetić [1] and Benjamin Viteškić [2]
[1] International Burch University, Faculty of Engineering, Natural and Medical Sciences, Sarajevo, Bosnia and Herzegovina
[2] University in Tromsø, The Arctic University of Norway, Faculty of Health Sciences, Tromsø, Norway
sabina.glumcevic@ibu.edu.ba

***Abstract* - Diabetes is a widespread disease, suffered by millions, including children. Treatment of diabetes type 1 and sometimes even type 2, entails multiple blood glucose checks and insulin injections per day, and can thus be extremely exhausting, especially for very young children. Open-loop systems of insulin delivery, insulin pumps, used today commercially, require human interaction which can lead to low blood glucose control due to human mistakes. Fully automated closed-loop systems of artificial pancreas, as one-hormone as well as dual-hormone systems, are being developed. This paper is the literature survey of the latest research on the automated closed-loop artificial pancreas. The objective of this paper is to explore the development of devices and techniques to facilitate the daily life of diabetic patients with emphasis on the latest research on the topic. From so-called pens to open-loop systems of insulin pumps, closed-loop systems with user interaction - hybrid closed-loop, to the latest fully automatized closed-loop – artificial pancreas. In total 300 articles are reviewed from which 150 articles are retained for the literature survey and 50 are analysed in this literature review.**

Keywords – artificial pancreas, bionic pancreas, close-loop, prediction-based algorithm, detection-based algorithm

I. INTRODUCTION

Diabetes is one of the most prevalent diseases of today. Even children as young as 2 months can suffer from it. According to the WHO [1] more than 422 million people suffer from diabetes, of which 9 million suffer from type 1. In type 1 diabetes (T1D), a patient usually suffers from congenital flaws that prevent them from producing hormones that regulate blood glucose. Type 2 diabetes patients (T2D) produce those hormones, but their function in the body is diminished.

Currently, patients with diabetes are treated with either: 1. multiple daily blood glucose measurements and insulin injections, which can be very unpleasant, especially for young children (as well as for their parents and caregivers), or 2. continuous subcutaneous blood glucose measurements and insulin infusion delivered via an insulin pump. Various open-loop insulin pumps available in the market are programmable to deliver the required amount of insulin. However, all of them require user intervention. A fully automated closed-loop insulin infusion system that can deliver appropriate amounts of insulin to patients without any manual interference is being developed. Dual-hormone systems, insulin with glucagon, or pramlintide are also being developed. The closed-loop system contains three main components: continuous glucose monitoring (CGM), an intelligent controller, and an insulin pump.

In the last 20 years, several prediction-based and detection-based algorithms for insulin attenuation/suspension, as well as dual-hormone systems, have been proposed in the literature, with the aim of preventing or mitigating hyperglycaemia (high blood glucose level) and hypoglycaemia (low blood glucose level).

However, for artificial pancreas construction, there is a basic terminology to be comprehended:

- Basal insulin stabilizes blood glucose levels when a patient is in a fasting state, especially during periods of sleep. Patients that use conventional methods – needles or pens to inject insulin into the body, usually take basal insulin before sleep. Patients that use insulin pumps, either open-loop or hybrid close-loop, are administered basal insulin automatically, in small doses throughout the day.
- Bolus is a dose of insulin taken to handle a rise in blood glucose that happens after eating, especially peroral carbohydrate intake.
- Insulin on board (IOB) is defined as the amount of administered insulin that is still active in the body.
- Insulin pump is a system that consists of an insulin tank, pump controller, and in the latest systems that are on the market – a communications system with CGM via Bluetooth.
- CGM is the system that measures the level of blood glucose, usually in 5 minutes periods, and in some latest commercial systems it sends information to the pump controller via Bluetooth. In older systems, that are still in use, blood glucose level monitoring is done via NFC (near-field communication) and mobile phone applications.
- Open-loop systems consist of CGM and insulin pumps that are not interconnected, there is no communication between the two of them and the pump is completely user-controlled.
- The closed-loop system contains three main components: CGM, insulin pump, and intelligent controller that regulate insulin delivery.

Table I. Blood glucose target values

	healthy	diabetic
fasting (before a meal, during the night, …)	4–6 mmol/L 70–110 mg/dL	4–7 mmol/L 70–130 mg/dL
2h post-prandial (after meal)	<7,8 mmol/L < 140 mg/dL	<10 mmol/L <180 mg/dL

Information about the crucial parameter for close-loop artificial/bionic pancreas construction, target blood glucose levels for diabetic persons, compared with levels of a healthy non-diabetic person, is shown in table I.

Blood glucose levels for a diabetic patient:

4 – 7 mmol/L (70 – 130 mg/dL): euglycemia (normal level of blood glucose)

3 – 5 mmol/L (55 – 90 mg/dL): hypoglycaemia (low level of blood glucose)

< 2 mmol/L (40 mg/dL): life-threatening hypoglycaemia

7 – 15 mmol/L (130 – 270 mg/dL): hyperglycaemia (high level of blood glucose)

> 16 mmol/L (300 mg/dL): life-threatening hyperglycaemia

II. TECHNICAL BACKGROUND

Open-loop system of insulin delivery, insulin pumps used today commercially, require human interaction which can lead to low blood glucose control due to human mistakes.

Both single-hormone systems (delivering insulin only) and dual-hormone systems (delivering insulin and glucagon or another hormone) are being pursued clinically. The addition of glucagon has the potential to further alleviate the risk of hypoglycaemia but increases the system's complexity with separate drug reservoirs, infusion sets, and algorithms. From a patient perspective, the ideal closed-loop system requires minimal user interaction, device burden, and inconvenience while achieving optimal glucose control.

Fully closed-loop systems that detect and automatically dose insulin for meals have been attempted, but glucose control is compromised because of delays in the absorption of subcutaneous rapid-acting insulin analogue. Therefore, most closed-loop systems adopt a hybrid approach, requiring manual administration of insulin boluses for meals.

Closed-loop control overnight reduced time in hypoglycaemia and increased time in the target glucose range.

The efficacy and safety of closed-loop glucose control in the outpatient setting have been demonstrated in multiple studies using different closed-loop prototypes and in meta-analysis.

III. LITERATURE REVIEW

During the research, it has been found that diabetes and closed-loop artificial/bionic pancreas topic is very popular in the last 20 years, and there are many scientific papers on different research aspects. From about 300 papers retrieved from different scientific sites, such as Research Gate, Google Scholar, IEEE, Journal of Diabetes Science and Technology, and PubMed/Medline, around 150 met the inclusion criteria, technical aspect of closed-loop, and 50 of them are included in this literature review. A lot of papers explore the medical aspects and clinical research of available systems, so they are excluded from further review.

The function of basal insulin is successfully incorporated into artificial pancreas systems that are already in use, so papers that treat basal insulin control are excluded. CGM/pump communication papers are excluded, due to CGM and its communication with the pump are also successfully incorporated into commercially available systems.

Papers relating to mobile phone applications for monitoring and blood glucose regulation are also excluded, since research aim isn't mobile application.

Bolus insulin is the aspect that is hardest to automate, so the focus was placed on works that relate to algorithms to incorporate different factors that influence blood glucose dynamics. The majority of the technical papers treat the prediction of blood glucose dynamics.

Blood glucose is unpredictable. Each person is different and special, and there are a huge number of factors that influence blood glucose levels. Type of meal (number of carbohydrates, but also proteins and fats in a meal), physical activity (level and type), stress, hormonal fluctuation (during puberty, menstrual cycle, pregnancy, childbirth, menopause…), illness, high body temperature, sleep deprivation …. Everything mentioned above should be incorporated into artificial/bionic pancreas that emulates functions of the healthy human pancreas. But, human organisms, and every and each organ, are very complex systems that are almost impossible to emulate by mechanical means.

Children are the most delicate and unpredictable of all patients. Their bodies demand higher doses of insulin, which leads to a greater risk of low blood glucose hours after meals. So young children and their caregivers face many different obstacles. The majority of the research is for adult patients, and just a small number of research are led for the youngest diabetes patients. Papers [5], [9], [11], [14], [15], [27], [29] from the reference list address the use of the pump in the youngest population and their effects.

Teenagers with diabetes, on the other hand, provide another subset of challenges. Papers [10], [11], [15], [27],

[29] also includes teenagers and adolescent in their research.

All other papers consider only adult diabetic patients.

Only one paper [2] presents research on pregnant diabetes patients' blood glucose control.

Exercise has proven to be a particularly difficult-to-control metabolic disturbance. Exercise increases muscle demand for glucose and can increase insulin sensitivity for hours later. Given the long-time delays of insulin action, stopping insulin delivery for a rapidly falling glucose is unlikely to prevent hypoglycaemia. Much work has been done to detect exercise through the use of heart rate alone, but it alone isn't enough. Additional signals that detect physical activity faster than CGM alone, may improve blood glucose control and reduce exercise-related hypoglycaemia. In papers [16], [18], [25], the research was based on the inclusion of the detection of physical activity into the algorithm to perform corrections of insulin dosage.

One challenge of managing diabetes without significant hypoglycaemia relates to the wide fluctuation in insulin requirements to maintain euglycemia, between people and for the same person from day to day and in different situations. A second challenge is the time delay between subcutaneous insulin injection and insulin action to lower blood glucose levels, which can extend to more than an hour. As a result, it is exceedingly difficult for an individual to accurately predict how much insulin is needed in any particular situation, often resulting in hypoglycaemia or hyperglycaemia.

The majority of the research uses prediction-based algorithms. The most used are MPC (Model Predictive Control) and PID (Proportional-Integral-Derivative). Papers [2], [6], [7], [9], [17], [27], [31], [32] research closed-loop systems using the MPC algorithm, and papers [8], [17], [21], [24], [27] PID. Some of the papers use Machine Learning, Deep Learning, and Artificial Neural Networks [14], [15], [16], [18], [26], [32], [33], [37], [39] for decisions on insulin administrated.

Around 80% of the papers are about one-hormone pumps, while some, like[5], [10], [11], [13], [34], [40] test bi-hormonal cases, two separate pumps, and controllers. Also, the majority of the research is on hybrid close-loop systems that include user intervention in some way [2], [3], [7], [9], [16], [18], [19]. Almost all reviewed papers are about T1D, and just a few include T2D patients [22], [25], [35]. The reason is that T1D patients' bodies don't produce insulin (and some other hormones), so they are condemned to multiple injections day and night, unlike T2D patients whose body produces insulin in a certain amount and the need for a pump isn't that substantial. Even though, some scientists are considering constructing insulin pumps for T2D [22], [35].

All testing that is done in-silco is done using the UVa/Padova diabetes simulator.

As previously mentioned, the biggest challenge is blood glucose regulation around mealtime due to slow digestion and absorption of food (conversion to blood glucose), and on the other hand slow effect of injected insulin. In a healthy human body, a healthy pancreas starts insulin production and secretion at the moment a person starts a meal. In everyday life, diabetic patients, both T1D and T2D that use insulin for blood glucose regulation, calculate the number of carbohydrates to be eaten, calculate the responding bolus insulin dose, and inject it, via the conventional method – pens, or enter the number of insulin units into pump controller. How to detect or predict meals, how to detect or estimate the size of meals, number of carbohydrates consumed, and similar questions are discussed in papers [2], [3], [4], [6], [7], [16], [28], [30], [31], [32], [38].

In further literature review, the aim was to find researches that use artificial neural networks (ANN) for bolus insulin regulation, or for detecting some irregular events that affect blood glucose level like stress, physical activity, illness, life phases with hormonal changes, etc. There is a small number of papers that treat the topic of bolus and/or meal detection/prediction using ANN [14], [15], [16] and [37]. Similar is for physical activity using ANN [16], [18]. No research is done in a case of some irregular events like stress, illness, or hormonal changes during pregnancy, puberty, etc.

A Convolutional Neural Network (CNN) and a Long Short-Term Memory Recurrent Neural Network (LSTM RNN) are used in research [14] for blood glucose forecasting of 10 virtual paediatric patients. ReLu activation function was used for both training models. The models were implemented and trained on Google Colab using libraries Keras and TensorFlow. Both models take as input a 3×30 matrix of values, corresponding to the last 30 min of the 3 feature values. Dynamic range quantization, in *.tflite* format, is used for performing regression tasks on Rasperry, and the full integer quantization, *uint8*, for DevBoard. Results of the research showed that LSTM model achieves the best numeric and clinical accuracy when tested in the *.tflite* format, whereas the CNN achieves the best clinical accuracy in *uint8*. The models achieved numerical accuracy comparable to those reported in the literature for adult patients.

In their research [15] authors use offline Reinforcement Learning (RL) algorithms: batch constrained deep Q-learning (BCQ), conservative Q-learning (CQL) and twin delayed deep deterministic policy gradient with behavioural cloning (TD3-BC) for nine virtual diabetes patients, with UVa/Padova simulator, and compares results with PID controller's. Performance was evaluated by monitoring blood glucose levels over a simulated 10-day test period and aggregating the results over three test seeds per training seed. Offline RL algorithms had better results in time in the healthy glucose range (6.4±1.1)%, meal estimation, even in case of irregular meal patterns. The work showed that offline RL conduce more effective and safer insulin dosing from

smaller samples of data comparing with the current standard of glucose control algorithms.

In [16] authors described a deep reinforcement learning algorithm they developed using a T1DM dataset, data from wearable devices, with aim to automate insulin dosing. They built patient-specific computational models using systems biology informed neural networks (SBINN) with the Roy-Parker model, to mimic the glucose-insulin dynamics by simultaneously considering patient-specific carbohydrate intake and physical exercise intensity. Dataset that is used contained 8-week continuous glucose monitoring, insulin, physiological sensor, and self-reported life-event data for 12 patients. According to the authors conclusion model not only correctly predicts the hidden states that cannot be measured with current diabetes technology, but also accurately infers patient-specific parameters, different daily routine of physical activities and insulin injection, governing the patient-specific Roy Parker model by reconstructing all states of interest.

The study [18] investigates the classification of different types of acute psychological stress (APS) and physical activity (PA) via physiological variables measured by a wristband device. Random convolutional kernel transformation is used to extract a large number of feature maps from the bio signals measured by a wristband device (blood volume pulse, galvanic skin response, skin temperature, and 3D accelerometer signals). Three different feature selection techniques (principal component analysis, partial least squares–discriminant analysis (PLS-DA), and sequential forward selection) as well as four approaches for addressing imbalanced sizes of classes (up sampling, down sampling, adaptive synthetic sampling (ADASYN), and weighted training) are evaluated for maximizing detection and classification accuracy. A LSTM RNN model is trained to estimate PA (sedentary state, treadmill run, stationary bike) and APS (non-stress, emotional anxiety stress, mental stress) from wristband signals. The balanced accuracy scores for various combinations of data balancing and feature selection techniques range between 96.82% and 99.99%. The combination of PLS–DA for feature selection and ADASYN for data balancing provide the best overall performance.

A LSTM NN is used in [37] for glucose forecasting, to predict blood glucose levels up to 60 minutes in advance, using continuous glucose measurements and insulin data collected from 175 people with T1D. Authors introduce the glucose variability impact index (GVII) and the glucose prediction consistency index (GPCI) to assess the accuracy of prediction algorithms. The inputs to the model are glucose prediction up to 60 minutes, and glucose and insulin on board (IOB) 3 hours values. Activation function used is ReLu. During the network training phase, the mean-square-error (MSE) loss function was minimized and multiple passes over the entire training set were done. Weights were updated using batches of 64 training sequences. The model was compared with several naïve approaches for estimating glucose and with alternative machine learning algorithms. The LSTM had highest accuracy and best GVII and GPCI.

Research of a machine learning model developed for probabilistic prediction of hypoglycaemia is explained in the paper [39]. The model is developed for 30- and 60-minute time horizons, based on CGM datasets and obtained from 112 patients over a range of 90 days. The model predicted hypoglycaemia with >91% sensitivity while maintaining a specificity of >90%. Two approaches were considered for prediction: Logistic Regression (LR) and Random Forests (RF). Classifiers based on Decision Trees, Gradient Boosting, and Support Vector Machines were developed. From the aforementioned research, it is concluded that 30 minutes time horizon is sufficient for the results.

IV. CONCLUSION

Bolus insulin isn't completely regulated so far. A very limited amount of research has been done on children, and adolescents, as well as the effects of pregnancy and exercise. An even smaller amount of research exists for dual hormone pumps specifically. Recent researches on ML, DL and ANN algorithms show better results in blood glucose regulation, more secure, accurate and precise, than PID and MPC algorithms already used in commercially available systems. Further research should be directed to bolus insulin regulation, and dual hormone pumps, and specifically include children and adolescents in their sample. The system in use should work efficiently even in periods of frequent blood glucose variations due to hormonal changes, like growing up, puberty, pregnancy, breastfeeding, virus/bacterial diseases, stress, etc. The system should also adapt to varying carbohydrate intake. It is of uttermost importance to get a model with high degrees of both precision and accuracy.

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/diabetes retrieved 06/10/2022

[2] H. R. Murphy, D. Elleri, J. M. Allen, J. Harris, D. Simmons, G. Rayman, R. Temple, D. B. Dunger, A. Haidar, M. Nodale, M. E. Wilinska, and R. Hovorka, "Closed-Loop Insulin Delivery During Pregnancy Complicated by Type 1 Diabetes" Diabetes Care. 2011 Feb; 34(2): 406–411.

[3] K.Turksoy, S. Samadi, J. Feng, E. Littlejohn, L. Quinn, and A. Cinar, "Meal-Detection in Patients with Type 1 Diabetes: A New Module for The Multivariable Adaptive Artificial Pancreas Control System", IEEE Journal of Biomedical and Health Informatics, Volume: 20, Issue: 1, January 2016

[4] Jinyu Xie and Qian Wang, "A Variable State Dimension Approach to Meal Detection and Meal Size Estimation: In Silico Evaluation Through Basal-Bolus Insulin Therapy for Type 1 Diabetes", IEEE Transactions on Biomedical Engineering, Volume: 64, Issue: 6, June 2017

[5] S. J. Russell, M. A. Hillard, C. Balliro, K. L. Magyar, R. Selagamsetty, M. Sinha, K. Grennan, D. Mondesir, L. Ekhlaspour,

H. Zheng, E. R. Damiano, and F. H. El-Khatib, “Day and night glycaemic control with a bionic pancreas versus conventional insulin pump therapy in preadolescent children with type 1 diabetes: a randomised crossover trial” Lancet Diabetes and Endocrinology, 2016 Mar; 4(3):233-243
[6] J. P. Corbett, J. Garcia-Tirado, P. Colmegna, J. L. Diaz Castaneda, and M. D. Breton, “Using an Online Disturbance Rejection and Anticipation System to Reduce Hyperglycaemia in a Fully Closed-Loop Artificial Pancreas System” Journal of Diabetes Science and Technology, Volume 16, Issue 1, December 2021
[7] J. P. Corbett, P. Colmegna, J. Garcia-Tirado, and M. D. Breton, “Anticipating Meals with Behavioural Profiles in an Artificial Pancreas System - An Informed Multistage Model Predictive Control Approach”, IFAC PapersOnLine Volume 53, Issue 2, 2020, Pages 16305-16310
[8] Ruiqiang Hu and Chengwei Li, “An Improved PID Algorithm Based on Insulin-on-Board Estimate for Blood Glucose Control with Type 1 Diabetes”, Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine, Volume 2015, Article ID 281589, 8 pages
[9] J. Ware et al., “Randomized Trial of Closed-Loop Control in Very Young Children with Type 1 Diabetes” The new England journal of medicine, vol. 386 no. 3, Jan 2022
[10] Steven J Russell et al., "Outpatient Glycaemic Control with a Bionic Pancreas in Type 1 Diabetes", F. H. El-Khatib, N Engl J Med 2014;371:313-25
[11] Bionic Pancreas Research Group, “Multicenter, Randomized Trial of a Bionic Pancreas in Type 1 Diabetes”, N Engl J Med 2022; 387:1161-72
[12] S. Srinivasan, L. Ekhlaspour, and E. Cengiz, “Adjunctive Therapies to Optimize Closed-loop Glucose Control”, Journal of Diabetes Science and Technology 2021, Vol. 15(6) 1243–1251
[13] L. M. Wilson, P. G. Jacobs, J. R. Castle, “Role of glucagon in automated insulin delivery” Endocrinol Metab Clin North Am. 2020 March; 49(1): 179–202
[14] F. D’Antony. L. Petrosino, F. Sgarro, A. Pagano, L. Vollero, V. Piemonte, and M. Merone, “Prediction of Glucose Concentration in Children with Type 1 Diabetes Using Neural Networks: An Edge Computing Application” Bioengineering 2022, 9, 183
[15] H. Emerson, M. Guy, and R. McConville, “Offline Reinforcement Learning for Safer Blood Glucose Control in People with Type 1 Diabetes”, arXiv preprint arXiv:2204.03376. 2022 Apr 7.
[16] Y. Deng, K. Arao, C. S. Mantzoros, and G. E. Karniadakis, "Patient-specific deep offline artificial pancreas for blood glucose regulation in type 1 diabetes", preprint, Biorxiv, October 24, 2022
[17] M. F. Villa-Tamayo, M. García-Jaramillo, F. León-Vargas, and P. S. Rivadeneira, “Interval Safety Layer Coupled With an Impulsive MPC for Artificial Pancreas to Handle Intrapatient Variability”, Front Endocrinol, 2022 Feb 21
[18] M.R. Askari, M. Abdel-Latif, M.Rashid, M. Sevil, and A. Cinar, “Detection and Classification of Unannounced Physical Activities and Acute Psychological Stress Events for Interventions in Diabetes Treatment”, Algorithms 2022,15, 352.
[19] F. Golestani and M. S. Tavazoei, “Delay-Independent regulation of blood glucose for type-I diabetes mellitus patients via an observer-based predictor feedback approach by considering quantization constraints”, European Journal of Control, Volume 63, January 2022, Pages 240-252
[20] M.M. Rashid, M.R. Askari, C. Chen, Y. Liang, K. Shu, and A. Cinar, “Artificial Intelligence Algorithms for Treatment of Diabetes”, Algorithms 2022, 15(9), 299
[21] S. Taherinasab S. Soleimaniasl, and S. Taherinasab, “Application of model reference adaptive control and modified Smith predictor to control blood glucose in type 1 diabetic patients”, Mathematics and Computers in Simulation, Vol 194, 2022, Pg 198-209
[22] G. Freckmann et al, “Insulin Pump Therapy for Patients with Type 2 Diabetes Mellitus: Evidence, Current Barriers, and New Technologies”, Journal of Diabetes Science and Technology Volume 15, Issue 4, July 2021, Pages 901-915
[23] B. W. Alabdullah, H. R. Alowais, S. J. D'yab, M. S. Diab and A. Jarndal, “Health Care Device for Diabetic Patients”, 2020 Advances in Science and Engineering Technology International Conferences (ASET)
[24] P. J. Enright, “A classical control perspective on the design of automated insulin delivery systems”, Preprint · February 2022
[25] S. Samson et al, “Quantifying the Impact of Influenza Among Persons with Type 2 Diabetes Mellitus: A New Approach to Determine Medical and Physical Activity Impact”, Journal of Diabetes Science and Technology 2021, Vol. 15(1) 44 –52
[26] P. V. Astillo et al., “Federated intelligence of anomaly detection agent in IoTMD-enabled Diabetes Management Control System”, Future Generation Computer Systems, Volume 128, March 2022, Pages 395-405
[27] E. M. Aiello et al., “Review of Automated Insulin Delivery Systems for Individuals with Type 1 Diabetes: Tailored Solutions for Subpopulations”, Current Opinion in Biomedical Engineering, Volume 19, September 2021
[28] A. Haidar et al., “Reducing the need for carbohydrate counting in type 1 diabetes using closed-loop automated insulin delivery (artificial pancreas) and empagliflozin: A randomized, controlled, non-inferiority, crossover pilot trial” Diabetes, Obesity and Metabolism, Volume 23, Issue 6, p 1272-1281
[29] N. Magdelaine, P. S. Rivadeneira, L. Chaillous, A.-L. Fournier-Guilloux, M. Krempf, T. MohammadRidha, M. Ait-Ahmed, and C. H. Moog, “The Hypoglycemia-Free Artificial Pancreas Project”, February 2020, IET Systems Biology 14(1)
[30] A. B. Rakan, T. MohammadRidha, and S. A. Al-Samarraie, “Artificial Pancreas: Avoiding Hyperglycemia and Hypoglycemia for Type One Diabetes”, January 2022, International Journal on Advanced Science Engineering and Information Technology 12(1):194-202
[31] A. B. Rakan, T. MohammadRidha, and S. A. Al-Samarraie, “Automatic Glycemia Regulation: Avoiding Hypoglycemia and Hyperglycemia”, 2021 International Conference on Communication & Information Technology (ICICT), Basrah, Iraq, 2021, pp. 74-79
[32] S. A. Al-Samarraie, and S. D. Salman, “Backstepping Nonlinear Control for Blood Glucose Based on Sliding Mode Meal Observer”, Al-Nahrain Journal for Engineering Sciences 21(3):436-444
[33] M. Esfanjani, S. Safari, and F. Faress, “Backstepping Nonlinear Control by using Sliding Mode Observer for controlling Blood Glucose”, Technium Science, vol. 2(1), pages 119-130, Jan 21
[34] M. F. Tabassum et al., “Modelling and simulation of glucose insulin glucagon algorithm for artificial pancreas to control the diabetes mellitus” December 2021 Network Modelling Analysis in Health Informatics and Bioinformatics 10(1):42
[35] G. Freckmann, et al., Insulin Pump Therapy for Patients With Type 2 Diabetes Mellitus: Evidence, Current Barriers, and New Technologies”, New Technologies. Journal of Diabetes Science and Technology. 2021;15(4):901-915
[36] X. Sun, M, Rashid, N. Hobs, R. Brandt, M.R. Askari, and A. Cinar, “Incorporating Prior Information in Adaptive Model Predictive Control for Multivariable Artificial Pancreas Systems”, Journal of Diabetes Science and Technology, Volume 16, Issue 1
[37] C. Mosquera-Lopez and P. G. Jacobs, “Incorporating Glucose Variability into Glucose Forecasting Accuracy Assessment Using the New Glucose Variability Impact Index and the Prediction Consistency Index: An LSTM Case Example”, Journal of Diabetes Science and Technology 2022, Vol. 16(1) 7–18
[38] J. P. Corbett, M. D. Breton, and S. D. Patek, “A Multiple Hypothesis Approach to Estimating Meal Times in Individuals With Type 1 Diabetes”, Journal of Diabetes Science and Technology 2021, Vol. 15(1) 141 –146
[39] D. Dave, D. J. DeSalvo, B. Haridas, and S. McKay, “Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction”, Journal of Diabetes Science and Technology 2021, Vol. 15(4) 842 –855
[40] M. Infante, D. A. Baidal , M. R. Rickels , A. Fabbri , J. S. Skyler, R. Alejandro, and C. Ricordi, “Dual-hormone artificial pancreas for management of type 1 diabetes: Recent progress and future directions”, Artif Organs. 2021 Sep;45(9):968-986
[41] C. Boughton and R. Hovorka, “New closed-loop insulin systems” Diabetologia 64, 1007–1015 (2021)
[42] A. B. Rakan and T. Mohammadridha, “Review on Positive Control for Glycemia Regulation in Type One Diabetes”, Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), Vol. 21, No. 3, Sep 2021
[43] Bionic Pancreas Research Group, “Consensus Recommendations for the Use of Automated Insulin Delivery (AID) Technologies in Clinical Practice”, Endocr Rev. 2023 Mar 4;44(2):254-280

[44] A. Thomas and L. Heinemann, “Algorithms for Automated Insulin Delivery: An Overview”, Journal of Diabetes Science and Technology 2022, Vol. 16(5) 1228–1238
[45] E. Daskalaki et al, “The Potential of Current Noninvasive Wearable Technology for the Monitoring of Physiological Signals in the Management of Type 1 Diabetes: Literature Survey” J Med Internet Res 22 Apr: 24(4) e:28901
[46] R. M. Bergenstal, G. D. Simonson, and L. Heinemann “More Green, Less Red: How Color Standardization May Facilitate Effective Use of CGM Data”, Journal of Diabetes Science and Technology2022, Vol. 16(1) 3 –6
[47] M. A. Makroum, M. Adda, A. Bouzouane, and H. Ibrahim, “Machine Learning and Smart Devices for Diabetes Management: Systematic Review", Sensors (Basel). 2022 Feb 25;22(5):1843
[48] Jana Schmitzer, “Efficient Closed Loop Simulation of Do-It-Yourself Artificial Pancreas Systems” Journal of Diabetes Science and Technology2022, Vol. 16(1) 61 –69
[49] Bionic Pancreas Research Group, “Multicenter, Randomized Trial of a Bionic Pancreas in Type 1 Diabetes” N Engl J Med 2022;387:1161-72
[50] A. Thomas and L. Heinemann, “Diabetes Technology: Many Steps From an Idea to a Product”, Journal of Diabetes Science and Technology 1 –4, 2022
[51] T. Bomholt, D. Kofod, K. Nørgaard, P. Rossing, B. Feldt-Rasmussen, and M. Hornum, “Can the Use of Continuous Glucose Monitoring Improve Glycemic Control in Patients with Type 1 and 2 Diabetes Receiving Dialysis?” Nephron. 2023;147(2):91-96
[52] S. J. Moon, I. Jung, and C.-Y. Park, “Current Advances of Artificial Pancreas Systems: A Comprehensive Review of the Clinical Evidence”, Diabetes Metab J 2021;45:813-839
[53] A.Ahmed, S. Aziz, A. Abd-alrazaq, F. Farooq, M. Househ, J. Sheikh, “The Effectiveness of Wearable Devices Using Artificial Intelligence for Blood Glucose Level Forecasting or Prediction: Systematic Review”, J Med Internet Res 2023;25:e40259

ISS

International Conference on

INFORMATION SYSTEMS SECURITY

Steering Committee

Chairs: Stjepan Groš, University of Zagreb, Croatia
Tonimir Kišasondi, Apatura, Varaždin, Croatia

PAPERS

The Forensic Significance of Indexing Applications on the Windows Operating System

I. Špoljarić, D. Delija, G. Sirovatka
Zagreb University of Applied Sciences, Zagreb, Croatia
ispoljaric@tvz.hr, ddelija@tvz.hr, gsirovatka@tvz.hr

***Abstract* - When forensic analysis of the Windows operating system and the search for the existence of suspected files, applications, or artifacts of the operating system, the process of restoring deleted data is very often hard drives (SSD) SATA or NVMe interfaces in personal computers and taking into account properties such as wear leveling and garbage collection solid state hard drives, it is significantly difficult to recover deleted data as well as proving the start and presence of suspected files on the computers of the attacker or victim. This article analyzes the Windows Windows Search feature with a linked Windows.edb file as well as the 3rd Party application for indexing operating system files to find records of suspect files, metadata, applications, and their activities relevant to forensic analysis.**

Keywords – digital forensic; Windows indexing system; database; file recovery tools; record recovery tools.

I. INTRODUCTION

Digital forensics is defined as „The process used to acquire, preserve, analyze, and report on evidence using scientific methods that are demonstrably reliable, accurate, and repeatable such that it may be used in judicial proceedings" [1]. In such context, forensic artifacts are items that get left behind based upon the activities of the end user of the device, and digital evidence is any digital artifacts that can prove or disapprove forensic hypothesis.

Digital forensic analyses of Windows artifacts are a complex and difficult task in an all-time changing environment, especially artifacts related to proprietary databases and indexing systems, with very scarce documentation which often lags versions behind the current state of the analyzed Windows tools.

This article highlights the significance of different indexing bases and mechanisms of Windows operating system in terms of retrieving and extracting artifacts. Additionally, it also presents a comparison of the databases containing specific artifacts. The key element for an efficient approach with such artifacts is cooperation among practitioners and sharing of acquired knowledge through available online forums and workgroups.

There is a common path to how this knowledge is incorporated into digital forensics legally acceptable practice. This path covers a distinguished set of steps

- Initial recognition of new features in the forensic process;
- Feature analyses based on the malware analyses – identification of artifacts and importance of these artifacts;
- Development of forensic tools and procedures for acquiring and analyses of these new artifacts
- Testing and proving of forensic tools and procedures;
- Peer review and in-community proofing;
- Legal acceptance and proofing in court during some legal cases (US-specific);
- Incorporation of artifacts into commercial tools and education curriculums [7].

New features appear in Windows as a part of the new version of the operating system or through patch or service pack updates [6]. Documentation that exists is, from the forensic viewpoint, scarce and imprecise, presenting only a user or system administrator view on that feature, lacking specific implementation data, like internal application structure, database organization, and structure meaning of different fields. In practice, all implementation data which is often defined as a company intellectual property not available outside of the development department. It often takes about 6 months up to a year to get new feature support into commercially available software, leaving a period of‘uncertainty “ while practitioners are on their own in accessing and interpreting the „new“forensic artifacts.

II. FORENSIC VALUES OF INDEXING SUBSYSTEM

On the current version of the Windows operating system, several indexing applications are operational. The main goal of all these applications is to efficiently store metadata about files on the computer to simplify the search and organization of the file content.

Forensic analysis of indexing applications like Windows Search feature and 3td Party applications like Everything and Locate32 can provide many advantages for digital forensic experts during digital investigations. Windows Search feature save indexed data in Windows.edb file, Everything application save indexed files and folders in the proprietary database on the local machine and Locate32 application save indexed data to the PostgreSQL database on a local hard drive Mentioned applications can help experts and investigators to quickly find specific files on a Windows system, possibly prove the persistence of specific files, and provides additional information like file metadata such as dates, sizes, or type. In this paper, we analyzed three applications, Windows

Search (Windows.edb file), Everything application, and Locate32 application.

III. FORENSIC ANALYSIS

A. *Windows.edb file*

In forensic investigations, the "Windows.edb" [1] is located in „C:\ProgramData\Microsoft\Search\Data\Applications\Windows\" and the file is often analyzed in combination with other data sources, such as Windows Event Logs and users activity logs from Windows Registry hive "NTUSER.dat", to provide a comprehensive understanding of a subject's activities on a computer. However, due to the complexity and size of the file, its analysis typically requires specialized software, like ESEDatabaseView [2] from Nirsoft (https://www.nirsoft.net/utils/ese_database_view.html) or WinSearchDBAnalyzer [3] (https://github.com/moaistory/WinSearchDBAnalyzer). The "Windows.edb" store valuable information about the activities and data stored on a Windows computer [8]. The file can be used to:

-Identify recently accessed files, emails, and other content;

-Track user activities, such as email communication and file transfers;

-Retrieve deleted or lost files and email messages that may have been overlooked during the normal deletion process.

The frequency of re-indexing data in the "Windows.edb" file depends on the specific configuration and settings of the Windows Search service in the Windows Operating System. By default, the Windows Search service continuously indexes and updates the "Windows.edb" database in real-time, adding new files and folders and updating information as they are created, modified, or deleted. Despite this, users are not satisfied with the capabilities of the Windows Search feature, and new, 3rd party applications like Everything and Locate32 have been created.

To open the "Windows.edb" file for forensic analysis, we used a combination of specialized forensic software and native Windows tools, followed by these steps:

1. Mount the image of the hard drive containing the "Windows.edb" file.
2. Use a forensic software tool, such as EnCase, Xways Forensics, or FTK to create a bit-by-bit image of the drive, which will preserve the original file system structure and contents of the file.

Use a Nirosft „ESE Database Viewer“, version 1.71, to view the contents of the "Windows.edb" file.

In some situations recovery of "Windows.edb" is required, for such purposes extracted database can be accessed and recovered in a forensically sound way by the Microsoft tool esenutl.exe as it is shown in Figure 1. This process should be well documented and screenshots were taken to prove that the procedure was followed correctly and no unwanted changes were done to the data in the database.

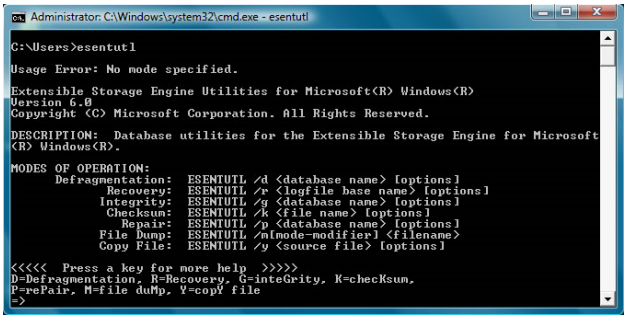


Figure 1: Recovery of the „Windows.edb“ through „esentutl.exe“ utility.

B. *Everything application*

The "Everything" application can be an important tool from a forensic perspective and it can provide valuable information about the user and system files and content stored on a computer. In addition, "Everything" stores a cache of its search results, which can be useful in reconstructing a timeline of activities and accessing information that may have been lost or deleted from the Windows file system. The Everything application stores its indexed data in a local database file named "everything.db". This file is typically stored in the same directory as the Everything executable file and is automatically created and updated when the application is run. The data stored in the "everything.db" file includes information about all the files and folders on a system, such as file names, paths, sizes, and dates. The database is designed to be fast and efficient, allowing the Everything application to quickly search for and display results based on a user's query.

The "Everything" database is a proprietary file format that is not designed to be opened or read directly. To access and search the "Everything" database, you need to use the "Everything" application.

In general, the data indexed by the Everything application will remain persistent as long as the underlying file system remains intact. The Everything application uses an NTFS-specific Master File Table (MFT) search algorithm, which is designed to be fast and efficient and to provide real-time results. However, if the file system is damaged or corrupted, or if files are deleted or otherwise changed, the indexed data in the Everything application may become outdated and may no longer accurately reflect the current state of the file system. In these cases, the Everything application will typically need to be re-indexed to update the indexed data and ensure that it remains accurate. The persistence of data in the Everything application depends on several factors, including the configuration of the application and the storage location of the indexed data. This means that the entries for deleted files and folders will remain in the database until the "Everything" application is re-indexed or until the database is manually cleared.

C. *Locate32 application*

Locate32 application version 3.1.11.7100, written by Janne Huttunen [3] is a free and open-source file indexing and searching application for Windows operating systems. It's designed to quickly locate files on a computer or network by indexing and storing the location of files and their metadata in a .dbs database file. Locate32 can be used to search for specific files or folders by name, date, size, or other metadata attributes. Locate32 support basic search functionality and advanced features, such as regular expression matching, support for networked file systems, and the ability to export search results to various file formats for further analysis.

Indexed data Locate32 save in a .dbs database on the local system (C:\Users\Username\Appdata\Roaming\Locate32\) where the application is installed. The default location for this database file is in the user's profile directory, in the folder named "locatedb". The database file is typically named "file.dbs" and contains the index of all the files and their metadata that have been indexed by Locate32. Some users may choose to store the database file on a network drive or other location to allow searching across multiple systems. Locate32 allows forensic analysts to export search results to various file formats, such as CSV or HTML, for further analysis or reporting.

D. *Indexing application comparison*

When comparing the "Windows.edb" file, Locate32, and Everything application in terms of forensic value, the following points can be considered:

•"Windows.edb" file: The "Windows.edb" file is a database file used by the Windows Search service to store information about indexed files and metadata. It can be valuable in forensic investigations as it contains information about files and their locations, which can be used to identify and analyze relevant data.

•Locate32: Locate32 is a search tool for Windows that creates and maintains a database of all files and folders on a system. This database can be searched quickly to locate specific files, making it useful in forensic investigations where time is a critical factor.

•Everything: Everything is a file search tool for Windows that uses a high-performance file search algorithm to quickly search and display files and folders. It can be useful in forensic investigations as it provides a fast and efficient way to search for specific files and folders on a system.

When running Everything, the application stores search data in the database "Everything.db", located in „C:\Users\<userprofile>\AppData\Local\Everything" and „Run History.csv" file located in the folder „C:\Users\ <userprofile>\AppData\Roaming\Everything\". In the "Everything.db" database we can find information about indexed files, while the "Run History.csv" file contains information about files searched using the Everything application, the number of times files were started or opened (Run Count), and the Last run time of applications, stored in Windows File Time format. If we compare the data about the launch of a single file from Run History.csv and the Windows Prefetch artifact, in our case we got additional information since Everything can store much older data, more precisely from the moment of Everything application installation, while the Windows Prefetch artifact store only 8 runs timestamp data.

TABLE I. INDEXING APPLICATIONS FEATURE COMPARISON

Indexing feature / Application name	**Windows.edb**	**Everything**	**Locate32**
Folders (System and users)	yes	yes	yes
Files (System and Users)	yes	yes	yes
Internet Links	yes	(partial – parsing from the browsers database)	(partial – parsing from the browsers database)
Network Drives	yes	yes	yes
File and Folders Timestamps	yes	(partial – only last write time)	(partial – only last write time)
LNK files	yes	yes	yes
Audio, Video and Pictures	yes	yes	yes
Hidden System Files (hiberfil.sys, $MFT, pagefile.sys, etc)	yes	yes	yes

In Table I, the relevance of the forensic artifacts stored in different indexing applications is presented. It is important to notice that databases present the state of the artifact or the content of the artifact in the moment of indexing even in the situation when a file or data was later removed from the system, like in the situation when a file was erased or email removed. Also, it is important to stress that indexing databases are also available through system restore points and backups if backups are done. System restore points are a much more reliable mechanism to access an earlier version of indexing databases on Windows systems. In this situation, the only additional step is to access the restore point through forensic software, which is well described the standard practice and extracts the previous version of index database files.

IV. FUTURE DEVELOPMENT

As the size of the disk grows up indexing applications become more and more important to accessing local and also remote data on networked or cloud environments. Also, these applications are performance issues In the situation where multiple indexing applications exist on the same system, with radically different infrastructures varying from property database systems like edb to open-source ones like PostgreSQL. It is reasonable to expect rationalization to simplify things and narrow down attack surfaces that different database systems provide.

It is uncertain in which way this effort will go, especially if the current development of AI services like chatGPT is taken into account, probably additional modifications to indexing systems will be developed with an integrated AI interface. This will again require a whole new forensic cycle probably much longer than the current one to get support for the whole new class of artifacts.

V. CONCLUSION

Forensic analysis of Windows.edb file, Everything application, and Locate32 application can provide valuable information and evidence in digital investigations. Windows.edb file is a database file that stores information about all the files and folders on a Windows operating system, which can help investigators track changes, recover deleted files, and gather evidence of illegal or unauthorized activity. The Everything application is a fast and efficient search tool that can quickly search for files and folders, even in large data sets, which records the total number and last run time of launching applications from the Everything application, and record persistence are not time-limited, only depends on the date of installation. While the Locate32 application is a search tool that can help investigators, after a mounting suspected hard drive, quickly find specific files based on various criteria such as date, size, or type and easily export findings to a human-readable file. Findings from these tools can help investigators to efficiently and effectively analyze and find files or file artifacts on the Windows Operating System, reconstruct timeline and user activities, and provide valuable information for investigation [5].

REFERENCES

[1] W. Oettinger: „Learn Computer Forensics“, Packt Publishing, 2020.

[2] NISoft: ESEDatabaseView v1.72 https://www.nirsoft.net/utils/ese_database_view.html (visited 05.01.2023.).

[3] Winsearchdbanalyzer https://github.com/moaistory/WinSearchDBAnalyzer (visited 05.01.2023.).

[4] Locate32 https://locate32.cogit.net/ (visited 05.01.2023.).

[5] O.Skulkin, S.Sourcier: “Windows Forensics Cookbook“, Packt Publishing, 2019

[6] Sans: „Digital Forensics and Incident Response“ https://www.sans.org/digital-forensics-incident-response/, / (visited 05.01.2023.)

[7] D.Delija,I.Špoljarić,G.Sirovatka: „Preparation and Planning of the Development of a Proficiency Test in the Field of Digital Forensics“, 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), ISSN: 2623-8764

[8] H.Chivers,C.Hargreaves: „Forensic data recovery from the Windows Search Database”, https://eprints.whiterose.ac.uk/75046/1/Forensic_Data_Recovery_From_The_Windows_Search_Database_preprint_DIIN328.pdf, (visited 05.01.2023.).

Implementation of Biometric Verification of a Fingerprint Whose Image is Taken from a Glass Surface

M. Knezović, M. Žagar, D. Delija, G. Sirovatka, D. Možnik
Zagreb University of Applied Sciences, Croatia
mknezovic@tvz.hr, mzagar@tvz.hr, ddelija@tvz.hr, gsirovatka@tvz.hr, dmoznik@tvz.hr

***Abstract* - Biometric systems often use the comparison of fingerprint characteristics for authentication and identification of persons, especially in mobile devices. This paper presents the procedures, devices, tools and algorithms that, with the use of non-forensic specialized devices, successfully capture a fingerprint from a curved glass surface using a mobile phone's digital camera and create a 3D model of the fingerprint that successfully unlocks a Samsung Android mobile device.**

Keywords - biometric system, fingerprint biometrics, fingerprint sensor, fingerprints, bypassing fingerprint biometric system

I. INTRODUCTION

The practical purpose of this paper is to create a fake fingerprint from a latent fingerprint left on a surface. There are papers available that prove that fake fingerprints capable of bypassing biometric systems can be created using commercially available tools and products, however, most of them collect fingerprint data directly from the finger which requires cooperation from the host. [1] Publicly available research shows various fingerprint image extraction techniques [2] as well as methods of creating a fake fingerprint on different types of sensors such as capacitive, optical, and ultrasonic sensors.[3] Available research in this area shows that it is possible to fabricate a fingerprint using the printing technique in 3D format by creating the so-called golden finger [4]. Extensive insights into the morphogenesis of the fingerprint give us a very good explanation of why this part of the human body is relevant for authentication mechanisms (methods) [5].

This paper will focus on indirect methods of collection, more precisely from the picture of a fingerprint left on a surface, and ways to improve and digitally alter the collected data to be able to create a realistic 3D model from the picture. After the 3D model with enough details is created focus will be on identifying the 3D printing technologies capable of printing fake fingerprints and utilizing the same. The practical part will be divided into four parts. Collection – where the collection of the fingerprint data will be conducted and explained. Digital processing – where software tools needed for the optimization of collected data and creation of 3D objects will be identified and used. Creation – where 3D printing technologies needed will be identified, selected, and used to create a realistic fake 3D fingerprint. Testing - where a short test on a smartphone device will be performed to make sure we have created a fake fingerprint capable of bypassing the fingerprint biometric system used for security.

The guiding idea is to use only off-the-shelf commercially available tools and products to make working fake fingerprints at „home“ conditions.

II. COLLECTION OF THE FINGERPRINT DATA

When it comes to the collection of fingerprint data as mentioned before it can be done using one of the different types of sensors, it can also be done by using ink and paper, the oldest technique available. However, for all those types of collection, the host or holder of fingerprint data must be compliant to give his fingerprint. This practical experiment focuses on indirect methods of collection without the knowledge of the host. This means latent fingerprint is collected and processed.

A. Size of the minutiae

The average size of the fingerprint is 1.27cm x 1.77cm it is of course different in males, females, and children. The size of the ridge differs from 100µm (0.1mm) in width for smaller fingerprints in children or women to 300µm (0.3mm) in width for larger fingerprints in grown-up males. Ridge and valley together are usually around 500µm (0.5mm) wide. The depth of the ridges varies from 20µm (0.02mm) to 50µm (0.05mm). In essence, we are working with really small sizes that need to be accurately collected. The FBI standard for fingerprint collection is a resolution of 500 dpi, which means there are 500 dots of data in a line along one inch or 2.54cm. We have to take into consideration that the FBI takes fingerprint data with a direct method using ink and paper which produces perfect contrast and ideal scanning conditions. In a matter of collecting latent fingerprints where is expected to collect a certain amount of visual noise inside the fingerprint pattern, we would need to raise the amount of data collected so visual noise could be digitally removed. Because of this, we would need a camera that can take a picture of a standard fingerprint size in a resolution of 2000x2000 pixels [6].

B. Different surfaces

Professional forensic investigators can lift latent fingerprints from a wide range of surfaces, however, surfaces that are smooth, non-porous, and reflective are recognized as the best source of quality fingerprint data. In the experiment, the focus will be exactly on these kinds of surfaces because forensic investigators have at their disposal special tools and equipment while the experiment

will be conducted using regular household items. Furthermore, the idea is to find the most suitable surfaces which are used every day. The three types of objects are selected from everyday life: wine glass, ceramic plate, and cutlery.

C. The collection

Initially, there were two ideas for the collection of fingerprints. The first was to take a photo of fingerprints left on the mentioned surfaces, the photo had to be excellent capturing all minutiae details with good contrast. The second idea was to dust the fingerprint with baby powder, collect it using transparent tape, paste it on dark background and try to scan it using the all-in-one printer. While the first method was successful the second method failed, because it was performed with a white baby powder which did not collect enough minutiae details when transported with transparent tape to the black surface. The experiment was continued using the first method.

Tools used for the collection are the iPhone 12 pro max - smartphone camera, as a camera for taking fingerprints, baby powder as a dusting powder, a make-up brush for dusting, and a simple standing lamp as a light source. The most important tool in this experiment, is the smartphone camera, with 12 megapixels and a resolution of 3024x4032 pixels. All mentioned tools can be found in a regular household.

Original idea was to collect the fingerprints from a wine glass, cutlery, and a ceramic plate. Collection from the wine glass and cutlery was successful while collecting from the white ceramic plate yielded no results because no matter what set up the light source, the fingerprint would be visible with the naked eye but not with the camera. In the case of a wine glass, the surface was transparent and allowed to focus of the source of light at an angle that made the fingerprint visible to the camera, in the case of cutlery the surface was reflective with a darker background which also made fingerprint details visible at specific lighting. In the case of the white ceramic plate, the surface was smooth and non-porous but not necessarily reflective with a white background which made the fingerprint not visible to the camera. The experiment was continued with wine glass and cutlery as surfaces to collect the fingerprint.

The first assumption was that fingerprints had to be dusted to be visible to the naked eye and camera (Figure 3, Figure 4). However, in a dark environment with the light source carefully positioned it is possible to make good photographs with a satisfying number of minutiae details without dusting the fingerprints (Figure 1, Figure 2).

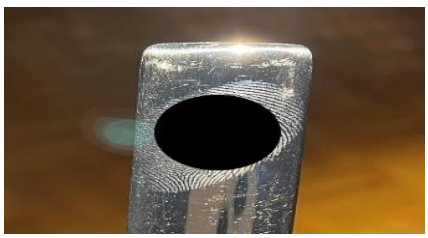


Figure 1. Fingerprint on cutlery without using the dusting powder

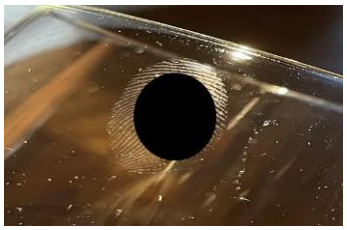


Figure 2. Fingerprint on wine glass without using the dusting powder

In the second round of fingerprint collection, baby powder was used with a makeup brush to highlight the ridges and valleys of the minutiae. The skin on the fingertips leaves small amounts of oil and sweat when pressed against smooths surfaces, fine powders like baby powder stick to the oils and sweat and enhance the fingerprint visually.



Figure 3. Fingerprint on cutlery using the dusting powder

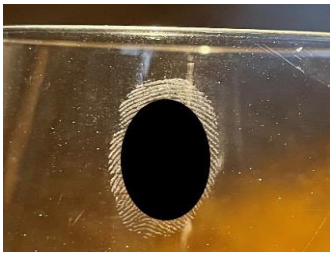


Figure 4. Fingerprint on a wine glass using the dusting powder

One thing that is important to do while taking a picture of a fingerprint is to measure the exact size of the fingerprint (Figure 5). The length and width of the fingerprint should be measured precisely because those are the only measurements that can be taken and based on that ratio the correct minutiae size will be replicated later in the process. It is not possible to measure the depth of the valleys or the height of the ridge.

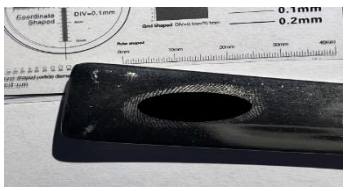


Figure 5. Taking fingerprint measurements

III. DIGITAL PROCESSING

After the fingerprint is collected with a satisfying number of minutiae details visible, the fingerprint picture will be optimized for crating a 3D fingerprint model.

A. Photograph optimisation

When fingerprints are collected by the police black or blue ink and white paper are used to produce an image of the fingerprint with great contrast so even the smallest details are distinguished. Considering this, the first thing to do in a latent fingerprint image is to enhance the contrast to make the ridge details more distinctive. For this task, photoshop software Adobe Photoshop

version 23.1.1. was used. In the case of latent fingerprint photographs, there is a lot of visual noise available and it should be reduced to the minimum. Because of this, the fingerprint image is converted to grayscale, and cropped for all excess image areas that are not representing information about fingerprint ridges or valleys (Figure 6, Figure 7, Figure 8). Also, Adobe Photoshop offers an option of image upscaling with a help of AI, this is making the picture clearer with a help of one of the offered resampling algorithms.

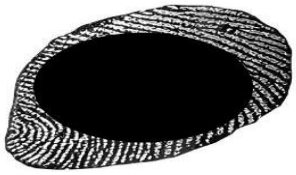


Figure 6. Latent fingerprint image in grayscale with enhanced contrast

It is really important not to bend the image while resizing because if there is a change in the proportions of the image, the most important data which makes fingerprints unique is lost. The scaling should be done by maintaining the aspect ratio regardless of the image size or shape.

Concerning digital optimization, it made little to no difference if a photograph taken using the baby powder or without it is used. This is mainly because photographed oils from skin ridges stand out in white color which when converted to grayscale produces the same contrast as photographs where the baby powder is used for enhancement.

Similar to this a fingerprint picture can be further enhanced using filters available in Photoshop which create a fingerprint picture with clear ridges and valleys without shadows, however, when this picture is used to mesh to a 3D model a lot of data about the depth of valleys is lost. This makes the 3D model have steep distinctions between ridges and valleys. Nevertheless, this kind of picture could be used to store and represent fingerprint data in 2D format.



Figure 7. Fingerprint picture with just minutiae data

B. Creating a digital 3D model

The next step in creating a fake fingerprint is to create a 3D model from a 2D photograph. There are existing research papers that describe a process of successful creation of the fake fingerprint when a real fingerprint is used for mold creation [7] and according to the Encyclopedia of Biometrics [8], there are three different methods and detection levels for fingerprint spoofing. Having that in mind it was decided to create a 3D mold for the fingerprint rather than printing the fingerprint itself. This is mainly because flexible materials like liquid latex or fabric glue can be used to create a fake fingerprint that can be attached to the real finger, for a more realistic representation of the fingerprint in a sense of temperature and conductivity.

Many different software solutions have been tested to create a satisfying 3D model of the molds. In Photoshop adobes 3D extrusion, the option works in a way that on a grayscale image, it makes ridges from white parts of the image and indentations from dark parts of the image. This in theory is exactly what we need but the final result was a model of the fingerprint which consisted of spikes where minutiae should be while valleys details were visible. Nevertheless, we have saved the file ready for 3D printing to see the final product when it is printed with a 3D printer.

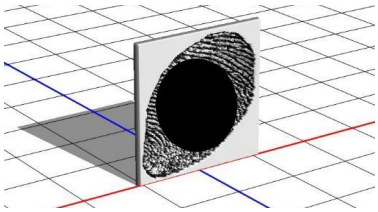


Figure 8. 3D model of the fingerprint made in Adobe Photoshop

After having experimented with other software like Blender and Meshroom it was decided to use a simple software initially developed to create a 3D relief on surfaces from a photograph (Figure 9). The software mentioned is “Photo ToMesh” by Ransen Software. Although its main purpose is to make personalized coffee mugs and medals it has served the purpose of creating a fingerprint mold perfectly. The software also offers an invert option which was used to invert ridges and valleys which made the creation of 3D printed mold from fingerprint pictures much easier. Photo ToMesh automatically converts the image to grayscale while preserving a lot of fine detail information about minutiae.

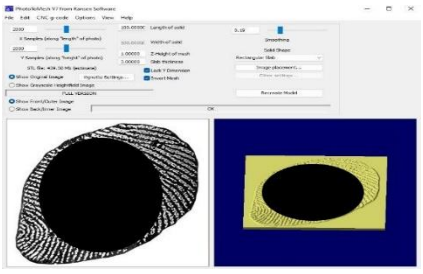


Figure 9. Photo ToMesh software - creating a 3D fingerprint mold

After the 3D object is created in Photo ToMesh it can be saved as an STL file which can be used by the majority of slicing software in a preparation for 3D printing.

Throughout the process of digitally optimizing the picture and creating a 3D model, it was important to

preserve the correct size ratios because the biggest obstacle to tackle in the process was creating a fake fingerprint of the same size as a real fingerprint. Even when the photograph was upscaled and the resolution was enlarged two times the ratios remained the same. All this was done to scale the fingerprint mold to the correct size just before printing to preserve the greatest amount of detail possible.

C. *Creating the fingerprint – 3D printing*

The final phase in fingerprint creation is 3D printing the fingerprint mold. To do this we had to identify the 3D printer with the capability of printing small minutiae size dimensions accurately.

IV. INTRODUCTION TO 3D PRINTERS

Generally, there are three biggest groups of 3D printers available commercially: FDM, SLA, and SLS printers. Each of these groups has its specific usage, advantages, and disadvantages according to what we have selected as the best printer for printing fingerprint mold. Layer thickness plays a big role in 3D printing precision, because of this every 3D object has to be sliced using the slicer software to a specific layer height just before printing.

Fuzed Deposition Modeling or FDM printers rely on heating a filament and creating a thin layer of material which immediately hardens to create layers. The filament is loaded in a form of a coil that goes through the nozzle which heats it and positions it carefully to create an object. These kinds of printers are affordable but are not precise with the best ones having a layer thickness of 0.1mm. Also, the surface of the objects printed with FDM printers is jagged and rough because of the way the layers of hot filament are placed on top of each other. Because of the low precision and rough final product surface, it was concluded that an FDM printer would not be able to print a fake fingerprint mold with a satisfying level of detail.

Selective Laser Sintering or SLS printers use high voltage laser to heat a powder-like material into solid layers. Each new layer is a new layer of powder glued to the previous one by applying heat with a laser. This enables SLS printers to work with harder materials like metal which is also the main advantage of SLS 3D printers. The surface of the final product from these kinds of printers is jagged and rough but can be polished after the material is completely cured. Precision is very good ranging from 0.08mm to 0.1mm layer thickness having in mind that products are cured after printing in a high-temperature oven which completely binds the layers. The main disadvantage here is price and size because of what SLS 3D printers are primarily used for industrial purposes. Rough final product surface and high costs are the main reasons SLS printers are not being used for printing fingerprint molds in this case.

Stereolithography or SLA 3D printers create the layers of a 3D object by curing the material in a form of liquid using UV light. These kinds of printers are the most precise with a layer height of 0.01m to 0.05mm. The main disadvantages of SLA printers are small working areas and the brittleness of printed objects while the main advantages are precision and price. Being that the width of the fingerprint ridges and valley range from 0.1mm to 0.3mm and the depth of valleys is from 0.02m to 0.05mm the SLA 3D printer is selected as the best option to print the fingerprint mold. [9]

A. *Selecting the SLA 3D printer and resin*

Stereolithographic printers are in the 3D printing community usually called resin printers because they are creating objects from a liquid similar to resin. Although this group of printers is known for precision one of the main characteristics by which the price of the printer is determined is the size of the working area. Because fingerprints are small in size, it was decided to select the printer with a small working area but with the best precision possible. For these reasons, Phrozen Sonic Mini 8K was selected as the printer of choice (Figure 10). The print volume of this printer is 16.5 x 7.2 x 18 cm which is more than enough to print multiple fingerprint molds. A field in which this printer is superior to all other commercially available SLA 3D printers is precision. With a layer thickness starting from 0.01mm and XY resolution of 0.022m m, this is one of the most precise 3D printers available. The precision is achieved by LCD type screen which is 16.5 x 7.2 cm in size and has a resolution of 7500 x 3240 pixels. The LCD screen emits UV light and cures the resin to create layers of an object.



Figure 10. Phrozen Sonic Mini 8K 3D printer [10]

When talking about 3D resin printing it is important to mention that there are many different kinds of resin, some are specially made to be flexible, others to be strong and there are a few resins that are made for high-precision printing. While it was possible to get the 3D printer with 8K resolution it was not possible to acquire an 8K resin for this experiment. For printing all of the fingerprint molds, the Creality Standard Rigid Resin Plus was used, this is an affordable resin capable of achieving standard levels of precision.

B. *3D printing the molds*

Just before the printing, a 3D object has to be prepared with slicer software, this type of software creates layers from the 3D object and saves it as instruction for the 3D printer. The software that has been used in the experiment is called Chitubox and it comes with a Phrozen Sonic Mini 8K printer. Many instruction properties like exposure time and layer height have to be set correctly for a successful 3D print. Among many others, two of the most important features that Chitubox offers are scaling and mirroring 3D objects.

With measurements taken while photographing the fingerprint, it is easy to scale the fingerprint to the correct size. In this example, the length measured on one of the fingerprints is 25mm. When loaded in the Chitubox, the program showed the size of the 3D model much larger when measured in millimeters. The length of the fingerprint

is presented as a Y axis, and scaling it to the correct measurement with a locked ratio provides the correct size of the fingerprint including the width of the minutiae (Figur 11).

Another important feature mentioned above is mirroring, to get the correct fingerprint shape after molding the initial shape has to be mirrored. This happens because a picture of the fingerprint is taken from the front and from that picture, the mold is printed rather than the fingerprint itself.

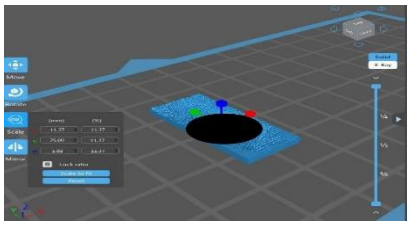


Figure 11. Scaling the model to the correct size in Chitubox

The one thing that was not possible to measure is the depth of the valleys, in the 3D model this measurement was adjusted by changing the mesh height. Experiments were done with different mesh heights to find appropriate valley depths which will define the height of the ridges once the molds are cast.

After a few failed attempts to print, mainly because of the wrong exposure time settings for a very small layer height, it was possible to successfully print the 3D model of the mold. Even though it is not visible from the pictures, the level of detail in these molds is very high (Figure 12, Figure 13). The layer height of the print was set to 0.02mm with an exposure time of 3.4 seconds per layer which made the print time 5 hours and 12 minutes long.



Figure 12. 3D printed fingerprint mold 1



Figure 13. 3D printed fingerprint mold 2

C. Casting the molds and adjusting

After the molds have been printed, support was removed, and cast the molds using liquid latex. Figure 22. shows four different fingerprint molds cast with liquid latex. Numbers one and two are molds of the same fingerprint with a difference in mesh height or valley depth, the same is true for numbers three and four. Mesh height in molds one and three is scaled to 1.0 being that it differs from minutiae to minutiae and it would not be correct to set a fixed value for all minutiae. Number 1.0 on a scale represents values software for 3D extrusion has made based on photograph shadows. Molds number two and four have mesh height scaled to 1.5 which means they have 50% higher ridges (Figure 14). This was done to detect perfect mesh height in future casts easier.

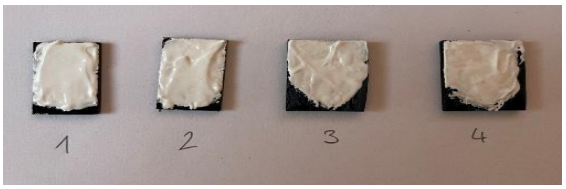


Figure 14. 3D printed fingerprint molds cast with liquid latex

Fake fingerprint minutiae were clearly visible, however, due to the sharp mold edges, the latex did not shape very well. Also, a mesh height of 1.5 or 50% bigger was a clear failure because the ridges were too big, and due to the sticky nature of latex, they ripped off while disassembling. A mesh height of 1.0 was good but the ridges were still too sharp to mold the latex perfectly.

To fix the problems mentioned above the smoothing option in Photo ToMesh software was used. This option created rounded edges and combined with the antialiasing option in Chitubox resulted in smoother layer transitions while printing.

After the second set of molds was printed, it was filled half of them with liquid latex and half of the with fabric glue. Drying the fingerprints lasted 12 hours after the molds were separated from the casts. There was no tearing and ripping while disassembling fake fingerprints from molds and the details on the fingerprints were of much better quality (Figure 15). Also, fabric glue molds have an overall better quality of detail for what reason they were selected for testing.

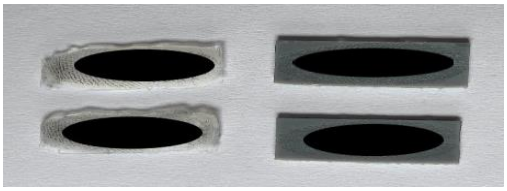


Figure 15. Molds and fabric glue fake fingerprints that were successful in unlocking the smartphone

D. Testing the fake fingerprints

Fake fingerprints were tested on a Samsung Galaxy A32 5G which uses a capacitive fingerprint scanner for fingerprint authorization. The success rate was 23% measured by 100 attempts to unlock the phone. It is important to mention that the success rate on a real fingerprint enrolled into the security system was only 54% which made fingerprint authentication with a real fingerprint fail almost every other attempt on this device. Also, during the tests, it was noticed that marginal parts of fake fingerprints work much better than the center of the fingerprint. We have figured that this is connected to the angle of taking the picture of the fingerprint. It happens because we cannot take picture of the fingerprint from the perfect 90-degree angle, we have to adjust the angle of taking a picture to the visibility of the latent fingerprint. As

a result, we usually take pictures from angles around 60 to 70 degrees which makes the final picture just like a fingerprint mold slightly out of proportion. That is why parts of the fingerprint whose picture has been taken from an angle closer to 90 degrees work better.

The whole process of creating a fake fingerprint lasted about 17 hours, where most of the time was consumed on drying the molded fingerprints (around 12 hours) and printing the molds (4 hours). Taking the picture of the fingerprint, digitally optimizing it, and creating a 3D object ready for printing takes less than an hour.

Extensive testing on all types of fingerprint sensors is needed to further adjust and enhance this method of creating fake fingerprints.

V. CONCLUSION

The experiment aimed at collecting fake fingerprints using commercially available tools and products was conducted. The focus of the experiment was to collect fingerprints from everyday objects such as a wine glass, ceramic plate, and cutlery. The experiment used a smartphone camera, baby powder, a make-up brush, and a standing lamp as a light source. The results showed that the collection of fingerprints from a wine glass and cutlery was successful, but from the ceramic plate was not. The optimization process involved using baby powder and Adobe Photoshop to enhance the details of the fingerprint and create a clear 2D representation.

To create a digital 3D model of a fingerprint, the experiment used Photo ToMesh software and an SLA 3D printer. The Phrozen Sonic Mini 8K was selected as the printer due to its small size and precision, while the Creality Standard Rigid Resin Plus was used for printing the molds. The fingerprint was scaled and mirrored in Chitubox slicer software to get the correct size and shape. However, the final molds had issues with sharp edges and high ridges, which were addressed using smoothing and antialiasing options in the software.

In conclusion, the experiment demonstrated that commercially available tools and products can be used to collect and create fake fingerprints, but preserving the correct size ratios and preserving detail was a challenge. The optimization process and 3D printing process required careful consideration of the tools and techniques used to achieve the best results.

The whole process of creating a fake fingerprint lasted about 17 hours, where most of the time was consumed on drying the molded fingerprints (around 12 hours) and printing the molds (4 hours). Taking the picture of the fingerprint, digitally optimizing it, and creating a 3D object ready for printing takes less than an hour. Extensive testing on all types of fingerprint sensors is needed to further adjust and enhance this method of creating fake fingerprints.

REFERENCE

[1] M Knezovic: "BYPASSING FINGERPRINT BIOMETRIC SYSTEM" Masters Thesis,TVZ 2022

[2] Stephen M. Bleay, Ruth S. Croxton, Marcel De Puit: „Fingerprint Development Techniques: Theory and Application, Willey 2018, Chapter 5, 6, ISBN: 978-1-119-18744-8

[3] Hugh Aver: „Faking fingerprints — doable, but hard",Kaspersky daily, 14.04.2020. https://www.kaspersky.com/blog/sas2020-fingerprint-cloning/34929/

[4] IEEE Transactions of Information Forensics and Security, Arora, S. S., Jain, A. K., & Paulter, N. G. (2017). Gold Fingers: 3D Targets for Evaluating Capacitive Readers. 12(9): 2067–2077. https://doi.org/10.1109/TIFS.2017.2695166

[5] M. Kücken and A. C. Newell, "Fingerprint formation," J. Theor. Biol., vol. 235, no. 1, pp. 71–83, 2005, doi: 10.1016/j.jtbi.2004.12.020.

[6] D. M. A. K. J. S. P. Davide Maltoni, Handbook of Fingerprint Recognition, Springer, 2002.

[7] M. Ž. Domagoj Tuličić, "ANALYSIS OF THE SECURITY OF SMARTPHONE FINGERPRINT," POLYTECHNIC & DESIGN, vol. 9, no. 4, pp. 277-286, 2021.

[8] A. K. Li, Stan Z. Jain, Ed., Encyclopedia of Biometrics. 2015., ISBN: 978-1-4899-7488-4, pp. 67.

[9] J. A. Z. Victoria Zukas, "The Basics of 3D printing," in An Introduction to 3D printing, Sarasota, First Edition Design Publishing, 2015., pp. 1-14.

[10] "Phrozen Sonic Mini 8K 3D Printer," Phrozen Technology, 2022. [Online]. Available: https://phrozen3d.com/products/phrozen-sonic-mini-8k-resin-3d-printer#specs. [Accessed 25 September 2022.].

The Latest Developments and Future Perspectives of Artificial Intelligence Systems for In-Vehicle Communication Intrusion Detection

Gabriel Marvin
dSPACE Engineering d.o.o., Zagreb, Croatia
GMarvin@dspace.hr

***Abstract* - CAN (Controller Area Network) is a message-based protocol that achieves communication by the exchange of packets of data between devices on the network. The protocol is widely used for in-vehicle communication (IVC) in the automotive industry as it is designed to be robust, able to handle a high rate of data transfer and tolerant of the electrical noise. However, the original CAN implementation lacks built in security mechanisms which makes it vulnerable to intrusion attacks that can be detrimental to the driver or the system itself. With the very rapid evolution of artificial intelligence (AI) various intrusion detection systems (IDS) have been developed to tackle the problem of detecting these attacks. This article aims to get a better insight of the current trend of development by surveying the latest approaches in the field of AI based IDSs. A comparison of various known attacks, detection techniques on available benchmark datasets, and a few advanced improvements of current security implementations and limitations are emphasized.**

Keywords* - *Intrusion Detection Systems, In-Vehicle Communication, Controller Area Network, Artificial Intelligence

I. INTRODUCTION

In-vehicle communication (IVC) refers to the interconnected communication between various electronic control systems within a vehicle also known as electronic control units (ECUs). With the advances in automotive technology and market demand, the modern vehicle now contains over 70 ECUs that control various functions within the vehicle [1]. The communication between various ECUs is facilitated by various standard in-vehicle communication protocols, such as CAN, FlexRay, Local Interconnect Network (LIN), Media Oriented System Transport (MOST), Ethernet, etc. In the automotive industry the CAN communication protocol is standardized and widely used because of its robustness, fast data transmission and reliability [2]. Various experimental attacks were performed to showcase and exploit the protocol's shortcomings. Hoppe et al. [3] conducted frame sniffing and replay attacks in a simulated environment to gain control over systems such as window lift, warning lights, and airbags. Similarly, Koscher et al [4] conducted various attacks on a real vehicle and successfully gained control over various modules such as the body control module, radio and engine. To counteract potential cyber threats significant efforts have been made to secure vehicles from security vulnerabilities. As a result, the research is focused on developing intrusion detection systems (IDS)s for in-vehicle networks as a reactive security measure. IDSs can be classified into two categories: signature-based detection and anomaly-based detection [5]. Signature-based IDSs report an intrusion when a match between the known attack and observed events is found. The anomaly-based approach on the other hand only knows what normal behavior looks like and considers deviations from the normal behavior as intrusions. Anomaly-based detection has received more attention due to its ability to detect novel attacks, whereas signature-based detection has limitations, such as the need for frequent database update in form of novel attacks [5]. This work focuses on the application of artificial intelligence (AI) to IVC IDSs and surveys several papers published from December 2022 to January 2023, examining attack types, detection methods, evaluation of available benchmark datasets and evaluation of performance. The paper is organized as follows: In section 2 existing surveys on the subject are discussed. Section 3 serves as a general introduction to the CAN protocol and common attacks on IVNs. Section 4 introduces the novel AI-based IDS methodologies for IVC, the used datasets as well as the training features. Section 5 discusses the limitations and future research directions. Section 6 concludes the paper.

II. RELATED WORK

This paper covers exclusively the state of the art (SoA) papers regarding IDS in CAN communication. Hafeez A. [2] reviewed five different fingerprint-based IDS methods for CAN communication. In the context of this work the term fingerprint-based is used to describe a signature-based method that does not necessarily include the use of machine learning (ML) or AI. This study provides a comprehensive introduction to various IDS techniques. The methods are presented in detail as well as their accuracy in detecting spoofing and impersonating attacks on CAN. However, it lacks a review on the used datasets on which these methods are tested. In the work by Karopoulos G. [1] a unified taxonomy for IVC IDSs is provided. The paper does not provide summaries on individual papers. However, a comprehensive description of the publicly available datasets is presented. The survey also discussed the available open-source simulation tools for communication data farming. In [5] the authors studied 44 articles presenting a division of IDSs by

categories. They also addressed research challenges and gaps in IVC IDSs. While authors in [6] compare current methods based on criteria such as real-time constraints, hardware types, changes in CAN bus behavior, attack mitigation types, and the software/hardware used to validate these approaches. They conclude with a discussion of the limitations of attack strategies and research challenges for the future. Loukas G. [7] reviewed 13 ML based IDSs for CAN, among other techniques that are not based only on the automotive sector. In [8] the authors introduce a detailed and systematic view of SoA AI-based IDSs for IVC. Focusing on the used methods to detect attacks, evaluation of used benchmark datasets, reviews on possible attacks and evaluation of performance. The AI-based IDSs in the papers are categorized by the features they are developed on. The methods within each category are presented, with emphasis on whether the learning conducted was supervised or unsupervised.

III. BACKGROUND AND TERMINOLOGY

A. CAN Communication Protocol

CAN is a communication protocol used for real-time control of message-based systems. It is robust, efficient, and uses a multi-master broadcast system [2]. The physical characteristics include the transmission of data which is done using a differential signal over a twisted pair of wires. The topology of the network can be either linear or starred, with each node connected to a single communication line. This allows for multiple nodes to transmit and receive messages on the same line.

The CAN bus data frame consists of several components. Some of the components are of great importance to the reviewed papers [8].

1. Start of Frame (SOF): Signifies the beginning of a new data frame and is used to synchronize the nodes on the network.
2. Identifier (ID): A 11-bit or 29-bit identifier used to identify the source and priority of the message.
3. Data Field (DATA): The payload of the data frame, containing up to 8 bytes of data.
4. End of Frame (EOF): Signifies the end of the data frame.
5. Intermission: A period of idle time between the end of one data frame and the start of the next. Used to allow the nodes on the network to process the received data and prepare for the next message.

An illustration of a CAN frame and all its components is shown in Figure 1.

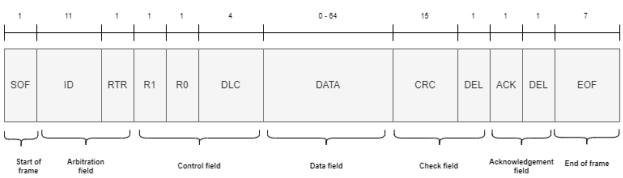


Figure 1: CAN frame

B. Known Attacks

An intruder can gain access to the CAN bus through various means such as wired mediums like the On-Board Diagnostics (OBD) port or USB port, wireless channels such as Global System for Mobile Communication (GSM) or Wi-Fi, and APIs such as SMS, web-interface, and mobile APIs. Once access is gained, the intruder can carry out an attack on a various number of vital processes [8, 9]. Some common attacks are listed below.

1. Denial of Service (DoS): Attacks attempt to disrupt communication services by overwhelming the system with a high volume of frames.
2. Frame Overwrite Attack: The attacker sends frames with a higher priority (lower CAN ID) to overwrite legitimate frames and manipulate the system's behavior.
3. Message Spoofing: Where the attacker sends fake frames that appear to be from a legitimate source to gain unauthorized access or manipulate the system.
4. “Man in the Middle” attack: The attacker intercepts and manipulates communication between nodes, disrupting the flow of data.
5. Replay Attack: where the attacker captures valid frames and retransmits them to the system to gain unauthorized access or manipulate its behavior.
6. Eavesdropping: Where the attacker intercepts and listens to communication between nodes to gain sensitive information.
7. Buffer Overflow Attack: The attacker sends a large number of frames that overflow the system's buffer and causes it to crash.

IV. CAN BUS AI-BASED IDSS

In this section a review of several papers is conducted. Each reviewed article is devised into five sections: general paper overview, data preprocessing, used method, used dataset, and achieved results.

Zhang H. et al. [10] proposed a novel approach for IDSs, which is based on a graph neural network (GNN) that can detect multiple different attacks: message injection, suspension, and falsification. The IDS is designed to work as a two-stage cascade. In the first stage, a one-class classification layer is used for anomaly detection and the GNN is trained only on normal CAN messages. The second stage determines the specific attack type or labels it as a new attack. The authors provide a detailed analysis of their proposed method by comparing the performance to several other techniques that tackle different attack types. In this way the authors prove the effectiveness of their solution by identifying different attack types. The paper also conducts research to validate the impact of vehicle states on the CAN message contents. The experiment is conducted by comparing two message sequences, with different time periods between

messages and then calculating their respective cosine similarities between consecutive messages. With this experiment the authors proved that there is an intrinsic connection between vehicle states and message content variations. A federated approach of training the model is proposed using two different federated learning optimizers. The results indicate that better performance can be achieved using the mentioned optimization techniques. Filtering out malicious participants and preserving privacy and security remains a challenge in federated learning.

Data preprocessing in [10] is implemented by using message graphs to describe CAN message streams. To ensure real-time analysis, CAN messages are analyzed in intervals that take a few milliseconds to process. If the interval is too short the message sequence becomes unstable due to the infrequent transmission of messages with high CAN IDs. This instability has been validated in the work through experiments using a dataset from [11]. The message graph comprises of nodes and edges, where the values of the nodes represent the CAN ID of the frame with an edge pointing to another node. This denotes the sequential transmission of CAN IDs. Each edge carries a bias value representing the number of corresponding message pairs appearing in the given message interval. In this way the authors can simulate the behavior of the connected system. The authors in [12] proposed a method for dividing the data in CAN messages to improve the accuracy of intrusion detection. The method called READ, divides the message contents based on the bit-flip rate, which is the rate at which bits change value. If a bit with a high bit-flip rate is followed by one with a low rate, it is likely that the two bits belong to two different signal behaviors or data blocks. To calculate the bit-flip rate, a certain number of CAN messages need to be collected for each CAN ID. Finally, to prepare the data for use in a GNN design, the authors need to describe all the CAN message contents using a node matrix. After applying the READ method, the message contents with different CAN IDs will be separated into different data blocks, each of which is then converted to a decimal number.

The IDS method uses graph learning as it increased in attention in recent years. The proposed CAN bus IDS is based on graph level algorithms. All three levels of models in graph learning start with graph convolution layers, and graph level schemes are used for graph classification tasks through the addition of pooling and readout layers. The proposed CAN bus IDS is based on the GNN model in [13]. Multiple convolution layers are stacked to capture graph substructure features at different scales. Pooling and readout layers are introduced for graph classification tasks and a sorting operation to sort nodes according to their structural roles is introduced. This helps solve node indexing issues caused by variations in CAN message graphs due to different driving regimes. The readout layers consist of 1-D convolution layers and dense layers. Due to the difficulty of acquiring intrusions to CAN buses in real vehicles, the first-stage classifier is designed as an anomaly detection IDS with a one-class classification layer or Support Vector Data Descriptor (SVDD) and only normal CAN bus data will be used for training. The whole architecture of the GNN, except the last one-class classification layer will be considered and the training process will be conducted using mini-batch stochastic gradient descent (SGD). After the anomalies have been filtered by the first-stage classifier, the second-stage classifier takes over the specific attack classification. To address new unknown anomalies the second-stage classifier includes a layer based on meta-recognition algorithms [14]. The paper also describes a novel federated learning strategy as a type of distributed ML where a cloud server works with multiple local users to train a model. During this process each local device trains a model using its own data and only sends the model parameters to the cloud server for optimization. Two federated learning schemes are evaluated [15, 16] in the proposed model. This allows vehicles in varying conditions to train their local model based on the optimized parameters on the cloud.

Two evaluation datasets are used to evaluate the performance of the proposed model with different attack types. The first is collected on a Ford Transit 500 and separated into three sets. The first set is composed of normal CAN messages while the remaining two sets include message injection attacks. In the second and third sets, the CAN IDs related to vehicle speed and revolutions per minute (rpm) are targeted and compromised CAN messages are randomly injected into these two datasets after specified time intervals. The second evaluation dataset is the CAN Signal Extraction and Translation Dataset provided by the Hacking and Countermeasure Research Lab (HCRL) and includes anomalies with five attack types: DoS, fuzzy, suspension, replay, and spoofing. The first dataset in the first baseline is used as the training set. Three baselines [16, 17, 18] are selected to test the GNN. The comparison between the proposed IDS and the first baseline is performed to demonstrate the former's effectiveness in detecting CAN message injection attacks. The second and third baselines are used to test the effectiveness of the GNN to identify other types of attacks with the metrics accuracy, precision, recall, and F1-score. The three intrusion detection strategies in the first baseline are thresholds for Cosine Similarity (CS), thresholds for Pearson Correlation (PC) and statistical CAN message sequence reconstruction using Long-Short Term Memory (LSTM). In the second and third baseline the detection strategies that are compared against the authors models are also LSTM and a CNN-LSTM model named CLAM.

The results are presented taking into consideration various perspectives of the model to correctly estimate its place with the mentioned baseline models. The GNN outperforms all first baseline models in all accuracy metrics. However, the

model is not so successful at predicting replay attacks as the CLAM model. A scalability evaluation is also performed where the model displayed negligible difference in accuracy in respect to the message interval length. In terms of the federated training approach the results show that using the scheme from [16] greatly improves accuracy in this scenario.

Zhao Y. et al. [19] introduces the Same Origin Method Execution (SOME) attack, an advanced variation of a masquerade attack. It can imitate the frequency of messages without changing the ID sequence on the CAN bus while sending attack information. At the same time a GAN-based Vehicle IDS (GVIDS) is proposed to counter the SOME attack. Tests on real vehicles show that GVIDS trained only on SOME can detect various attacks, including spoofing [25], bus-off [26], masquerade [27] and SOME with an average accuracy of 96.64% and a detection time of 0.18 ms. However, the novel method is only tested in controlled environments, but it's planned to evaluate its robustness in more complex environments.

Data preprocessing for GVIDS utilizes one-hot encoding to transform the data into CAN images, reducing the time cost for training and detection. It converts the data of 16 consecutive frames of the CAN bus into a single CAN image. To standardize the data, GVIDS supplements data segments of varying lengths to 8 bytes before encoding. The 16 CAN data frames are transformed into a 64 × 64 CAN image through one-hot encoding.

The IDS method used in this paper is GAN. A ML model composed of a generator (G) and a discriminator (D) that work together to generate data with the characteristics of the training set. G learns the characteristics of the training data to generate new data to further train D, while D distinguishes real data from false data generated by G. The GAN model adopts the GANomaly [20] network training model. The model consists of a G, D, and a reconstruction network (E). G is further composed of an encoder and decoder while E and D have similar structures to the encoder. During training, the input image is compressed into a vector, which is then regenerated into an output image. D discriminates between the output image and the generated image, based on a threshold value to determine an anomaly or normal data.

The model's training dataset is collected from two different vehicles (Luxgen U5 and Buick Regal). For each vehicle 1 million of normal CAN message images are collected. Several different evaluations are conducted to represent the results. The quality of the model was evaluated by a receiver operating characteristic (ROC) curve to plot the false positive rate (FPR) against the true positive rate (TPR) of the model. The best performance of the ROC curve is reached when the FPR is less than 0.1 and the TPR is greater than 0.9. The performance of the model is further evaluated using area under curve (AUC) metric. A higher AUC value, closer to 1, indicates a better performance of the model. In Luxgen, the AUC values for four different attacks are 0.9941 (spoofing), 0.9787 (bus-off), 0.9768 (masquerade), and 0.976 (SOME) respectively. Similarly, the AUC values for the same four attacks in Buick are 0.9982, 0.9778, 0.9665, and 0.9662 respectively. The intrusion detection accuracy for both individual vehicles was over 0.9600 for Luxgen and over 0.9312 for Buick. The real-time performance of GVIDS was analyzed by measuring the running time of intrusion detection in Luxgen and Buick. The average running time for both vehicles was 0.18 ms and it was enough to meet the requirements for the transmission of data frames in CAN bus networks.

The results show that the intrusion detection running time of GVIDS is minimal and can meet the demand for real-time detection.

Al-Jarrah O. Y. et al. [21] proposed a ML based IDS that creates high-level representations of CAN messages transmitted on the bus using the temporo-contextual dependencies between messages in a time frame and individual message data. These two representations are processed by two neural networks designed to detect novel intrusions by combining the views. The performance of the proposed IDS was evaluated and compared to SoA detection methods like Decision Trees (DT), Random Forests (RF) [22] and a deep learning-based IDS from [23]. The results show that the proposed IDS outperforms these methods in various evaluation metrics.

The data preprocessing was done by selecting a timespan that would result in the capture of 10 messages at a time. This enables the authors to construct recurrence plots (RP) that consist of nine CAN messages. The authors use this approach to capture the contextual information which resulted in the sending of the last message. The RP is the first input to the model while the second input is the payload of the last message in the set. Considering that not every CAN message contain exactly 8 bytes of data, the authors resort to padding the payload with values of -1 to create a uniform dataset.

The IDS method requires two views of the received CAN data, the first view from the data of one CAN message and the second from its temporal context represented in RP format. A LSTM neural network is trained using a stream of individual messages while a convolutional LSTM is trained using the generated RPs. The two LSTMs results are combined to form the input feature vector for a dense neural network that classifies each message as normal or intrusion. The performance of the model was evaluated using the "CAN-intrusion-dataset" [24]. The dataset contains 4 types of attacks, including DoS, rpm and gear spoofing attacks and fuzzy attacks. The DoS, rpm and gear spoofing attacks were used as the training dataset, while the fuzzy attack was used as the testing dataset to measure the performance of the different intrusion detection models. The proposed model was compared to other ML algorithms in

literature, including DT and RF, as well as a deep learning approach [23].

The results showed that the proposed model using both inputs achieved the highest accuracy, with a marginal drop in precision, while keeping a low FPR in comparison with the DT, RF [22] and the model used in [23]. The results also showed that using a combination of both views gives the best results in comparison with using one or the other. Table 1 aims to provide a concise and comprehensive summary of the key aspects of each reviewed paper.

TABLE I. COMPARISSON OF THE REVIEW ARTICLES

Reference	[10]	[19]	[21]
Algorithm	GNN	GAN	LSTM
Dataset	[11], [10]	N/A	[24]
Attacks	DoS, fuzzy, suspension, replay, spoofing	SOME [19], spoofing, bus-off, masquerade	DoS, fuzzy, spoofing
Advantages	Brings to attention that different vehicle states caused by variations in message graphs. Detailed analysis on run-time performance and comparison with relevant models.	Dataset collected from multiple vehicles, real time performance analysis, training model on advanced SOME attack shows generalization on other types of attacks.	Individual CAN frames are evaluated as well as their temporo-contextual dependencies.
Limitations	The evaluation of the federated learning approach was conducted on one dataset.	Lacks detailed comparison with more recent deep learning implementations. The driving conditions, number of ECUs and their functions are not mentioned.	Lacks detailed comparison with more recent deep learning implementations. Labeling each individual message may affect the run time performance of the model.

V. DISCUSSION

This paper focused on reviewing the most recent papers in AI-based IDS for IVC. The survey is conducted by summarizing the reviewed papers in several main sections that represent common aspects in all the articles. In this section a discussion is conducted based on the findings, limitations, and future improvements of the reviewed methodologies.

A. Findings

By analyzing the reviewed papers several findings have been made. All the papers display similar trends of research such as the concern for development of models that can detect unknown cyber-attacks, the common use of unsupervised detection techniques primarily based on benign data, feature set selection and used datasets. The biggest concern is to adapt the models to detect attacks based only on benign data. To tackle this requirement researchers in the reviewed papers have turned mostly to unsupervised learning algorithms, which result in better results in detecting previously unknown attacks in comparison with supervised learning methods. Used detection models are GAN, GNN and LSTM which all display great adaptability to variations in the environment and ability to distinguish anomalies from one class data. In the case of GAN and GNN with federated learning a great degree of accuracy is achieved in novel attack detection. This further solidifies the findings in [8]. The use of unsupervised learning models greatly eases the data collection process as it requires only benign data for training. Datasets containing malicious messages and attack tactics at this point are only required to test the models. The selection of the dataset feature sets have also a great impact on the generalization capabilities of the model as on the learning process. The vast amount of different vehicle models and car manufacturers lead to a great amount of diversity in message streams. This has led researchers to search for alternative ways to create feature sets that will lead to training of more generalized models. All the reviewed papers propose methods to extract features from the datasets by either new representations of the dataset like in the works [10, 19] or creating new features as in [21].

VI. LIMITATIONS AND FUTURE IMPROVEMENTS

The learning capabilities of AI methods are primarily impacted by the quality of the learning datasets. Due to the low quality of the publicly available datasets, it is often required to create new datasets which can lead to poor translation to real world environments. This process can be resource consuming as unsupervised learning algorithms required large amount of data to generalize properly. This issue is further amplified with the need for datasets that describe daily routines of individual vehicles. The paper [10] takes note of this drawback and proposes by considering the different variations of message streams due to driving conditions, which is also a viable point in the argument of poor datasets. The federated learning approach could be promising in the future coupled with a great amount of data gathered by individual vehicles in different driving conditions. Highly specialized software for ECU communication simulation could be used to farm a great amount of data under different driving conditions.

VII. CONCLUSION

This paper provides a comparative review of three of the most recent AI-based IDSs for the CAN protocol (December 2022 to January 2023). The reviewed papers are categorized based on the training and testing datasets used, the preprocessing methodology and the AI algorithms used. The reviewed papers are significant as they all highlight the importance of creating a system that is aware of the context from which the CAN frame is coming from, to be able to predict the anomalies in the

bus. It is important to mention that the impact on the field of research may not be clear as the papers are published in a short time range of only three months, but they suggest that the trend of research they represent is a current and active area. The findings of the paper also highlight the crucial role that high quality datasets play in the development of effective AI-based IDSs and propose several key improvements that can be made to existing systems.

REFERENCES

[1] G. Karopoulos, G. Kambourakis, E. Chatzoglou, JL. Hernández-Ramos, V. Kouliaridis, "Demystifying In-Vehicle Intrusion Detection Systems: A Survey of Surveys and a Meta-Taxonomy". Electronics. 2022; 11(7):1072. https://doi.org/10.3390/electronics11071072

[2] A. Hafeez, K. Rehman and H. Malik, "State of the Art Survey on Comparison of Physical Fingerprinting-Based Intrusion Detection Techniques for In-Vehicle Security," SAE Technical Paper 2020-01-0721, 2020, doi:10.4271/2020-01-0721.

[3] T. Hoppe, S. Kiltz, and J. Dittmann. "Security threats to automotive CAN networks practical examples and selected short-term countermeasures". In International Conference on Computer Safety, Reliability, and Security., Springer, 235-248., 2008.

[4] K. Koscher et al., "Experimental Security Analysis of a Modern Automobile," 2010 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 2010, pp. 447-462, doi: 10.1109/SP.2010.34.

[5] O. Y Al-Jarrah, C. Maple, M. Dianati, D. Oxtoby, and A. Mouzakitis. "Intrusion detection systems for intra-vehicle networks: A review". IEEE Access 7 (2019), 21266-21289.

[6] E. Aliwa, O. Rana, C. Perera, and P. Burnap. 2021. "Cyberattacks and Countermeasures for In-Vehicle Networks". ACM Comput. Surv. 54, 1, Article 21 (January 2022), 37 pages. https://doi.org/10.1145/3431233

[7] G. Loukas, E. Karapistoli, E. Panaousis, P. Sarigiannidis, A. Bezemskij, and T. Vuong. 2019. "A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles". Ad Hoc Networks 84 (2019), 124-147.

[8] S. Rajapaksha, H. Kalutarage, M. O. Al-Kadri, A. Petrovski, G. Madzudzo, and M. Cheah. "AI-based Intrusion Detection Systems for In-Vehicle Networks: A Survey". 2022. ACM Comput. Surv. https://doi.org/10.1145/3570954

[9] S. V. Kumar, G. A. A. Mary, P. Suresh and R. Uthirasamy, "Investigation On Cyber-Attacks Against In-Vehicle Network," 2021 7th International Conference on Electrical Energy Systems (ICEES), Chennai, India, 2021, pp. 305-311, doi: 10.1109/ICEES51510.2021.9383720.

[10] H. Zhang, K. Zeng and S. Lin, "Federated Graph Neural Network for Fast Anomaly Detection in Controller Area Networks," in IEEE Transactions on Information Forensics and Security, doi: 10.1109/TIFS.2023.3240291.

[11] L. b. Othmane, L. Dhulipala, M. Abdelkhalek, N. Multari and M. Govindarasu, "On the Performance of Detecting Injection of Fabricated Messages into the CAN Bus," in IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 1, pp. 468-481, 1 Jan.-Feb. 2022, doi: 10.1109/TDSC.2020.2990192.

[12] M. Marchetti and D. Stabili, "READ: Reverse Engineering of Automotive Data Frames," in IEEE Transactions on Information Forensics and Security, vol. 14, no. 4, pp. 1083-1097, April 2019, doi: 10.1109/TIFS.2018.2870826.

[13] M. Zhang, Z. Cui, M. Neumann, and Y. Chen. 2018. "An end-to-end deep learning architecture for graph classification". In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence (AAAI'18/IAAI'18/EAAI'18). AAAI Press, Article 544, 4438–4445.

[14] P. Zhang, J. Wang, A. Farhadi, M. Hebert and D. Parikh, "Predicting Failures of Vision Systems," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 3566-3573, doi: 10.1109/CVPR.2014.456.

[15] T. Li et al., "Federated optimization in heterogeneous networks," arXiv:1812.06127., 2018.

[16] M. Jedh, L. Ben Othmane, N. Ahmed and B. Bhargava, "Detection of Message Injection Attacks Onto the CAN Bus Using Similarities of Successive Messages-Sequence Graphs," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 4133-4146, 2021, doi: 10.1109/TIFS.2021.3098162.

[17] A. Taylor, S. Leblanc and N. Japkowicz, "Anomaly Detection in Automobile Control Network Data with Long Short-Term Memory Networks," 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 2016, pp. 130-139, doi: 10.1109/DSAA.2016.20

[18] H. Sun, M. Chen, J. Weng, Z. Liu and G. Geng, "Anomaly Detection for In-Vehicle Network Using CNN-LSTM With Attention Mechanism," in IEEE Transactions on Vehicular Technology, vol. 70, no. 10, pp. 10880-10893, Oct. 2021, doi: 10.1109/TVT.2021.3106940.

[19] Y. Zhao, Y. Xun, J. Liu and S. Ma, "GVIDS: A Reliable Vehicle Intrusion Detection System Based on Generative Adversarial Network," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 4310-4315, doi: 10.1109/GLOBECOM48099.2022.10001410

[20] S. Akcay, A. Atapour-Abarghouei, T.P. Breckon (2019). "GANomaly: Semi-supervised Anomaly Detection via Adversarial Training," In: Jawahar, C., Li, H., Mori, G., Schindler, K. (eds) Computer Vision – ACCV 2018. ACCV 2018. Lecture Notes in Computer Science(), vol 11363. Springer, Cham. https://doi.org/10.1007/978-3-030-20893-6_39

[21] O. Y. Al-Jarrah, K. E. Haloui, M. Dianati and C. Maple, "A Novel Intrusion Detection Method for Intra-Vehicle Networks Using Recurrence Plots and Neural Networks," in IEEE Open Journal of Vehicular Technology, doi: 10.1109/OJVT.2023.3237802.

[22] T. P. Vuong, G. Loukas, D. Gan and A. Bezemskij, "Decision tree-based detection of denial of service and command injection attacks on robotic vehicles," 2015 IEEE International Workshop on Information Forensics and Security (WIFS), Rome, Italy, 2015, pp. 1-6, doi: 10.1109/WIFS.2015.7368559.

[23] G. Loukas, T. Vuong, R. Heartfield, G. Sakellari, Y. Yoon and D. Gan, "Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning," in IEEE Access, vol. 6, pp. 3491-3508, 2018, doi: 10.1109/ACCESS.2017.2782159.

[24] E. Seo, H. M. Song and H. K. Kim, "GIDS: GAN based Intrusion Detection System for In-Vehicle Network," 2018 16th Annual Conference on Privacy, Security and Trust (PST), Belfast, Ireland, 2018, pp. 1-6, doi: 10.1109/PST.2018.8514157.

[25] E. Seo, H. M. Song and H. K. Kim, "IDS: GAN based Intrusion Detection System for In-Vehicle Network," 2018 16th Annual Conference on Privacy, Security and Trust (PST), Belfast, Ireland, 2018, pp. 1-6, doi: 10.1109/PST.2018.8514157.

[26] K. Cho and K. G. Shin,. "Error Handling of In-vehicle Networks Makes Them Vulnerable." In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16). Association for Computing Machinery, New York, NY, USA, 1044–1055. https://doi.org/10.1145/2976749.2978302

[27] K.-T. Cho and K. G. Shin, "Fingerprinting electronic control units for vehicle intrusion detection," in 25th USENIX Security Symposium (USENIX Security 16), 2016, pp. 911–927.

Deep Learning within the Web Application Security Scope – Literature Review

Matija Kaniški, Jasminka Dobša and Dragutin Kermek
Faculty of Ogranization and Informatics, University of Zagreb, Varaždin, Croatia
{matija,kaniski, jasminka.dobsa, dragutin.kermek}@foi.unizg.hr

***Abstract* - Over the last few years, several breakthroughs in deep learning have contributed to the development of new models. One of many areas they are applied to is the web application security scope. Web applications are still one of the biggest information and business security threats. Requests sent to the Web application are divided into normal and malicious. Malicious requests contain a payload that exploits a discovered vulnerability. Detection of Web attacks can be reduced to natural language processing classification problem. Lately, pre-trained models on Transformer neural networks showed promising results in the detection of Web attacks. In development of models, the preprocessing step of data preparation is crucial. After preparation of good datasets and application of powerful models it is very important to evaluate and compare performance of algorithms. The goal of this paper is to conduct an overview of the deep learning methods used for Web attack detection. The research is conducted by querying scientific databases, analyzing relevant articles within the security scope, and summarizing the proposed state-of-the-art approaches. Findings of reviewed papers were summarized based on implementation details and used performance metrics. Also, open problems will be emphasized, as well as challenges and possibly new opportunities for the future research.**

Keywords - deep learning; transformer architecture; natural language processing; web application security; web attack detection

I. INTRODUCTION

Web applications are software systems that generate web pages and documents written in one of the programming languages and are executed on the server. Web browsers are programs through which we access the Web application. Requests sent to the Web application are divided into normal and malicious payloads. Malicious requests contain a payload that exploits a discovered vulnerability. With the emergence of new Web technologies, modules, frameworks and platforms, Web applications open new attack surfaces and vectors. Attack surface is what is being attacked, whereas an attack vector is a path or means by which an attacker can gain access to a computer system or network to deliver a malicious payload. Such a payload is sent to the Web application, and the response is analyzed for possible vulnerabilities. Vulnerability is a known or unknown weakness (zero-day) that can be exploited. Further, exploits take advantage of that vulnerability to gain unauthorized access or execute malicious tasks. In the last decade neural networks have regained popularity [1] due to the rapid development of hardware.

Also, large amounts of data provide new opportunities in their use, including natural language processing (NLP). Deep learning (DL) is a part of machine learning (ML) where neural networks are essential components of the algorithm [2]. Determining an attack on a Web application with DL methods is a classification task of whether the payload sent to the Web application is malicious or not. Therefore, the detection of Web attacks can be reduced to an NLP classification problem. In development of models preprocessing step of data preparation is crucial. After preparation of good datasets and application of powerful models it is very important to evaluate and compare performance of algorithms. The motivation for the research originates from practical applications of training neural networks to improve Web attack detection. Existing literature reviews [1, 2] cover DL models for a specific Web attack like denial of service (DoS) or analyze multiple ML and DL techniques for the detection of unknown (known as zero-day) Web attacks. This study considers DL models for Web attack detection that are not limited to specific types of Web attacks or devices (IoT). In addition, payloads divided as normal or malicious (binary classification) will be observed. This paper is structured in the following way. Section I is the introduction. Section II focuses on defining research questions, strategy, and criteria. Section III brings up the analysis of selected papers. Section IV contains research findings and a discussion. Section V concludes the paper.

II. METHODOLOGY

This paper analyzes and summarizes DL methods for Web attack detection within the Web application security scope. It is crucial to identify primary studies and conduct an overview of different approaches for classifying malicious payloads to achieve this objective.

A. Defined research questions

Insufficiently described or omitted factors in the DL literature lead to difficulty reproducing or replicating a given approach. Major contributing factors to these issues are the lack of published open-source implementations and datasets, missing data filtering details and descriptions of hyperparameters used [1,2]. The research questions are: (i) What types of Web attacks have been addressed by DL-based approaches? (ii) How are artifacts being extracted, prepared, and used in DL-based approaches for Web attack detection? (iii) What DL models are used to support Web attack detection? (iv) How well does DL-based approach perform in supporting various Web attack detection?

B. *Search strategy*

The search was conducted on three scientific databases: Web of Science, Scopus, and IEEE Xplore. Title, abstract and keywords in the Computer Science research topic were searched. The search query was adjusted several times to find only those papers that are within the scope of the research. Defined search keywords were *deep learning*, *neural network*, *web attacks*, *web security*, *web application security* and *web vulnerabilities*.

Each database had different search queries according to their search engine. The following search queries yielded the most relevant results:

- **Web of Science**: TOPIC: (“deep learning” OR “neural networks”) AND (“web attacks” OR “web security” OR “web application security” OR “web vulnerabilities”)
- **Scopus**: TITLE-ABS-KEY ((“deep learning” OR “neural networks”) AND (“web attacks” OR “web security” OR “web application security” OR “web vulnerabilities”)) AND (EXCLUDE (PUBYEAR, 2023)) AND (LIMIT-TO (SUBJAREA, “COMP”)) AND (LIMIT-TO (LANGUAGE, “English”))
- **IEEE Xplore**: ALL: (“deep learning” OR “neural networks”) AND (“web attacks” OR “web security” OR “web application security” OR “web vulnerabilities”)

All queries aim to extract relevant articles about Web attack detection with DL-based approaches.

C. *Search criteria*

Selection criteria are presented in Table I. These inclusion criteria are used to determine which paper is relevant and which is not. The table also clearly states all exclusion criteria for further reducing the papers.

TABLE I. INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria	***Exclusion criteria***
Papers written	Non-English papers
Papers published between 2016 and 2022	Papers published before 2016 and after 2022
Papers published in scientific conferences or scientific journals or relevant Web articles	Duplicate papers
Papers on DL for Web attack detection	Out of scope papers
Papers which proposed new models or approaches	Papers who do not provide sufficient details about their models or approaches

III. OVERVIEW OF INCLUDED PAPERS

Applying the queries leads to filtering all papers considering the constraints. The section will show the number of articles per database, per year, and by paper type. Afterward, the analysis is presented. In this section is presented the number of papers selected in each iteration. There were four iterations whose results are displayed in Figure 1.

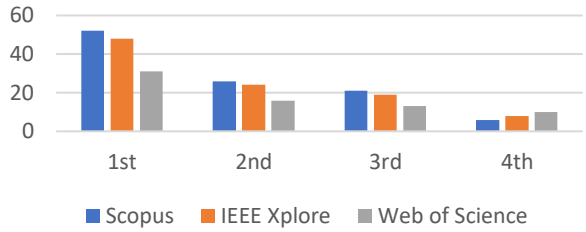


Figure 1. Number of papers per iteration and scientific database

In the first iteration queries are executed to get relevant articles without duplicates. In the second iteration it is done filtering based on reading titles and abstracts of the papers. The third iteration it is done filtering after reading full articles. The last iteration is based on narrowing the scope and focusing on answering all the research questions. After the process of selection, a total of 24 papers were analyzed and considered. Figure 2 shows the number of papers that used deep learning for detecting Web attacks in the observed period by years.

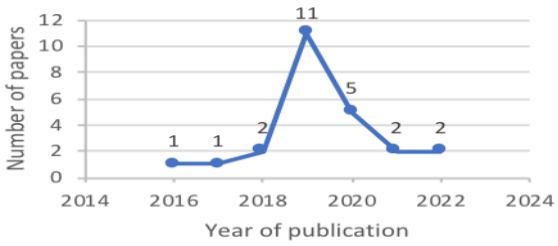


Figure 2. Number of selected papers per publication year

As one can see, most articles were published in 2019, and then the interest decreased. The reason for this decrease is probably the appearance of transformer models with self-attention mechanisms that yielded state-of-the-art performance in Web attack detection. The distribution of published research papers per year suggests that topics regarding Web attack detection with deep learning are an active area for research. We predict that it will still be a desirable area in the future since new and unknown Web attack variants are emerging which could be detected by new approaches using neural language models. Paper classification based on its type is shown in Figure 3.

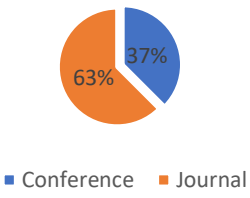


Figure 3. Precentage of papers per type of publication

Not surprisingly, the percentage of selected journal articles is higher than that of conference papers. Mainly because to answer all research questions, additional details of the proposed approach must be provided, which was not the case in most conference papers.

IV. RESULTS AND DISCUSSION

The existence of vulnerabilities doesn't mean it will be exploited. Only after a few years of initial discovery a particular attack technique is becoming widely misused. That's why we are exploring them right now. Table II summarizes the reviewed papers in six categories. Approaches were mainly supervised learning models (SLM) and unsupervised learning models (USLM).

A. Type of Web attacks

Injection Web attacks are considered to impact the most vulnerabilities in Web applications. Injection of HTTP payload [3-8] show different patterns hidden in the URL. Next, there are SQL injection (SQLi) and Cross-site scripting (XSS) [9-19]. SQLi attacks are server-side vulnerabilities targeting Web application databases, whereas XSS are client-side vulnerabilities targeting Web application users. Then there is XPath [20], a query language for XML with injection issues similar to SQL. Another Web attack is phishing [21-22]. Phishing is the malicious practice of luring users into disclosing their personal information. Fake websites may be created and employed by criminals to steal that information. Further, there are denial of service (DoS) attacks whose purpose is to make a system unavailable for access. In contrast to DoS attacks, distributed denial of service (DDoS) attacks [23-24] use several computers to execute a simultaneous DoS attack against one or more systems. Webshells (backdoors) [25] of the website are often implemented on the website servers to maintain the management authority. Web trojan [26] attacks are designed to damage, disrupt, and inflict harmful actions, like personal information leakage. Zero-day Web attacks are unknown attacks that have not been reported. They are very rare and hard to find in the HTTP requests. Web application firewalls (WAF) are a common defense against Web attacks, but they can be exploited to evade the filter detection pattern by leveraging the found vulnerability.

B. Datasets

The lack of public datasets in the Web application security scope represents an obstacle to training and testing DL algorithms. Most often publicly available datasets in the reviewed papers are:

- CISC2010 dataset [3-7, 10, 11, 15, 24] has automatically generated traffic with multiple Web attacks like buffer overflow, information gathering, SQLi, XSS, files disclosure, CRLF injection, parameter tempering and more.
- ECMLPKD dataset [5] contains more than 35000 normal requests and about 15000 malicious Web requests collected from real network traffic.
- httpParams dataset [7, 15] is created with several tools like sqlmap, xssya, Vega Scanner, FuzzDB.
- FWAF dataset [7, 15] is created by FSECURITY for detecting malicious Web queries with ML-driven WAF.
- XSSed dataset [16] contains more than 40000 different malicious XSS attacks.
- UCI phishing dataset [21, 22] consists of about 11000 samples which have 30 features pre-classified as phishing or normal.
- CICIDS2017 dataset [24] contains DDoS attacks and 83 features in the dataset, including both the network traffic and package information results.
- KDDCU99 dataset [24] contains samples with four types of anomalies DDoS, Remote-to-login, User-to-root, and probing.

Private data sets are also used. A dataset is considered private if it is not publicly available or if no procedure is described to replicate and reproduce it. Private datasets [3-5, 8-14, 17-20, 23, 25, 26] contain new examples, normal or malicious requests, on which the model was trained.

C. Preprocessing techniques

Original datasets must be transformed into numerical data by procedure called feature extraction. Not all features are equally important for the classification decision. Feature selection is an essential pre-processing task that aims to find a subset of relevant features from the original features. In examined papers, different data preprocessing techniques have been used, where tokenization and neural embedding are by far the most common. Also, encoding techniques such as HEX, URL, Unicode, HTML entity and generalization are used to reduce interference of redundant and irrelevant information. Tokenization [4, 11, 13, 15-17] is a broad technique that was applied to separate sequential data into tokens depending on the type of Web attack. Neural embeddings [8, 9, 22, 24, 26] use different DL architectures to represent the data. Two popular implementations of Wor2Vec [12, 14, 16] used for feature extraction are CBOW model [7, 10, 18, 25] and the Skip-Gram model [6]. Some papers did not explicitly mention the type of embedding they use: word-level [4, 5] or character-level embedding [3, 20]. Besides that, feature selection and weighting using Generic Algorithm (GA) [21] and statistical feature extraction [23] were present in the pre-processing task.

D. Classification model

Each model has its strengths and weakness in detecting certain types of Web attacks. CNN model [3, 8, 10, 16, 25] and its variants [6, 8, 11, 14, 24] have the most significant advantage in that its feature set is learned by itself. RNN [16, 20] is a DL model suitable for solving the sequential problem, identifying patterns in URLs. LSTM [3-5, 10, 20] is a special RNN to realize the long-term memory of characters. Compared with RNN, CNN training is shorter, whereas CNN compared with a deep neural network (DNN), its parameters are fewer, and the model is more concise. MLP [3, 10, 13, 15, 17] is a feed-forward neural network (FFNN) with multiple processing layers and at least one hidden layer, among which both the hidden layer and the output layer have the ability for Web attack detection. Autoencoder is an FFNN with one or more layers designed to minimize input and output differences. A stacked Autoencoder (SAE) [9, 26] is constructed by stacking the input and hidden layers of the Autoencoder layer by layer.

TABLE II. SUMMARY OF REVIEWED PAPERS

Ref	Implementation details of the proposed approach							Performance by most frequent performance metrics			
	Web attack	***Type***	***Framework***	***Dataset***	***Prep. technique***	***Over/under -fitting protection***	***Model***	***Acc. (%)***	***TPF (%)***	***F1-score (%)***	***FPR (%)***
[3]	Multiple Web attacks injected in HTTP request or payload	SLM	PyTorch	**CSIC2010***, Private	Character embedding	Data cleaning	MLP, CNN, LSTM	97.79	96.04	98.72	N/A
[4]		SLM	N/A	**CSIC2010***, Private	Tokenization, Word embedding	Data balancing	LSTM-MLP	98.42	97.56	N/A	N/A
[5]		SLM	TensorFlow / Keras	**CSIC2010***, ECMLPKD, Private	Word embedding	Data augmentation	Bi-LSTM	98.6	98.4	98.1	N/A
[6]		SLM	N/A	CSIC2010	Skip-gram	Boosting	Text CNN SVM	99.33	99.09	N/A	N/A
[7]		SLM	TensorFlow / Keras	CSIC2010, httpParams, FWAF	CBOW**,** FastText	Dropout	ResNet	99.41	99.55	N/A	N/A
[8]		SLM	TensorFlow / Keras	Private	Neural embedding	Early Stopping	CNN, GRU	99.61	99.58	99.61	N/A
[9]	SQLi, XSS, object deserial.	USLM	TensorFlow / Keras, WEKA	Private	Neural embedding	Denoising	SAE	N/A	92.8	91.8	N/A
[10]	SQLI, XSS	SLM	N/A	**CSIC2010***, Private	CBOW	Dropout, Data cleaning	Ensemble MLP, CNN, LSTM	99.47	99.70	99.0	0.33
[11]		SLM	N/A	**CSIC2010***, Private	Tokenization	Boosting, Data balancing	CNN with attention	99.845	98.29	99.1	0.0
[12]	SQLi	USLM	N/A	Private	Word2Vec	Data cleaning, Data balancing	Encoder-Decoder RNN	N/A	99.0	98.0	N/A
[13]		SLM	PyTorch	Private	Tokenization	Boosting	MLP	99.75	99.88	N/A	0.02
[14]		SLM	TensorFlow / Keras	Private	Word2Vec	Data cleaning	Elastic Pooling CNN	99.93	99.94	99.95	N/A
[15]		SLM	PyTorch	CSIC2010, httpParams, FWAF	BERT tokenizer	Data balancing	MLP	99.98	N/A	98.7	N/A
[16]	XSS	SLM	TensorFlow / Keras	XSSed	Tokenization Word2Vec	Cross-valid., Dropout	CNN-LSTIM	99.3	99.1	99.5	N/A
[17]		SLM	TensorFlow / Keras	Private	Tokenization	Dropout, Cross-valid.	MLP	99.32	98.0	98.7	0.31
[18]	XSS-DOM	SLM	TensorFlow / Keras	Private	CBOW	Cross-valid.	DFFN	N/A	95.0	N/A	N/A
[19]	XSS in **PHP*** and JS	SLM	N/A	Private (D1, **D2***)	Wod2Vec, **Code2Vec***	Data cleaning, Data balancing	DFFN with attention	95.38	99.90	91.80	N/A
[20]	XPath	SLM	PyBRAIN	Private	Character embedding	Data cleaning, Boosting	RNN, LSTM	N/A	84.2	N/A	16.2
[21]	Phishing	SLM	N/A	UCI	Feature selection and weighting using GA.	Dropout, Boosting, Cross-valid.	DNN	88.77	85.81	N/A	N/A
[22]		SLM	N/A	UCI	Neural embedding	Data balancing,	DNN	97.71	90.51	92.16	1.7
[23]	DDoS	USLM	WEKA	Private	Statistical feature extraction	Cross-valid.	SAE, LR	N/A	98.99	N/A	1.27
[24]		SLM	N/A	CICIDS2017, **KDDCU99***	Neural embedding	Data balancing	MC-CNN	99.18	N/A	N/A	N/A
[25]	Webshell in **PHP***, JSP and ASP	SLM	TensorFlow	Private	CBOW	Dropout	CNN	99.50	99.70	99.40	N/A
[26]	Trojan	USLM	N/A	Private	Neural embedding	Denoising	SAE	94.92	N/A	N/A	16.32

* best results

Logistic regression (LG) [23] was used with SAE for classification, as it can output a value corresponding to the probability of belonging to a given class. Transformer architecture like BERT [15] showed promising results to be used in Web attack detection. BERT can be considered as a stack of encoders and decoders. To combine all the strengths of all models, a typical approach is to use an ensemble. Voting ensembles are primarily selected based on the output of models. Hard voting predicts the majority class, whereas soft voting summed predicted probabilities from individual classifiers for each class. So, soft voting gives better insight better if all models provide probabilities [10].

E. Over/under-fitting protection

Classifying imbalanced datasets is a major challenge. DL algorithms solve these challenges and thus do not need a feature selection process performed during preprocessing [5]. However, if irrelevant or unnecessary features are selected, high variance (overfitting) or high bias (underfitting) may occur. Summary of reviewed papers showed that data cleaning and/or balancing was the most used method to combat over/under-fitting. Other techniques were dropout, cross-validation, boosting, denoising, early stopping and data augmentation. Data cleaning [3, 10, 12, 14, 19, 20] includes operation such as removing duplicates, modifying labels without values, etc. Data balancing [4, 11, 12, 15 19, 22, 24] involves random sampling and random parameter initialization in order to obtain balanced data set. Dropout [7, 10, 16, 17, 21, 25] is usually applied only to the neurons in the top one to three layers (excluding the output layer). Grid or random search [16, 17, 21, 23] has been used as a model hyperparameter optimization technique with a k-fold cross-validation approach. This approach randomly splits the training dataset into a set of k-folds of approximately equal sub-dataset sizes. Different boosting techniques for updating the weights of the DL models were used, inducing Adam [10, 11, 13], RMSprop [20], and AdaDelta [19]. Denoising [7,24] works by corrupting the original input with some noise. The Autoencoder then needs to reconstruct the input from that noise, which forces the hidden layer to capture the statistical dependencies between the inputs. Another way to regularize iterative learning algorithms, such as gradient descent, is early stopping [8]. The training is stopped as soon as the validation error reaches a desired minimum. Finally, data augmentation [5] artificially increases the size of the training set by generating many modified variants of the training instance. In this way, it compensates for a shortage in data.

F. Evaluation

In imbalanced classification, the number of examples in the training dataset for each class label is not balanced. Since the distribution of classes is skewed, e.g., they extend much farther to the left or right of the median. In this case, it is not sufficient to use accuracy alone as a metric of model performance. So, it is recommended to examine the confusion matrix. The idea of a confusion matrix is to count the times instances of class A are classified as class B for all A/B pairs. Elements of the confusion matrix in the context of Web security are the following: true positives (TP) or Web attack requests in the dataset that are correctly detected as Web attacks, true negative (TN) or normal requests correctly detected as normal request, false positives (FP) are those normal requests detected as Web attack requests and false negatives (FN) that are Web attack requests detected as normal requests. To compute the confusion matrix, we first need to have a set of predictions to compare them to the actual targets. Each row in a confusion matrix represents an actual class, while each column represents a predicted class. A perfect classifier would only have true positives and true negatives, so its confusion matrix would only have nonzero values on its main diagonal. Analyzing the confusion matrix often gives us insight into ways to improve your classifier. Below are just a few of the most common performance metrics [17]:

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (1)$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (2)$$

$$Sensitivity/Recall/TPR = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (3)$$

$$FPR = \frac{\mathrm{FP}}{\mathrm{TN+FP}} \quad (4)$$

$$F1 - score = 2\left(\frac{\mathrm{Recall} * \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}}\right) \quad (5)$$

$$AUC = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right) \quad (6)$$

The recall/precision trade-off should be noted. In this trade-off, the higher the recall is, the lower the precision becomes and vice versa. Because of this, the F1-score measure or the harmonic mean of precision and recall is usually applied [22]. AUC is an abbreviation for area under the ROC Curve, AUC is the probability that the model ranks a random positive example more highly than a random negative example. ROC stands for receive operating characteristic curve, and in a nutshell, it pots TPR vs FPR at different classification thresholds. A perfect classifier will have a ROC AUC equal to 1, whereas a purely random classifier will have a ROC AUC equal to 0.5. When a feature distribution has a heavy tail (i.e., when values from the mean are not exponentially rare), both min-max scaling and standardization will squash most values into a small range. So, before we scale the features, it is recommended first to transform it to shrink the heavy tail and, if possible, to make the distribution roughly symmetrical and bell-shaped distributions (e.g., by computing their logarithm or square root). To improve performance of the classifier, first we can gather more training data so that the classifier learns to distinguish them. Second, we could engineer new features that would help the classifier. Third, we can try to preprocess the data.

V. CONCLUSION

Client-side vulnerabilities exploit Web browser weaknesses. Most Web applications do not accurately filter the user's input data, so both the client and the server are invaded. The main difference between DL and traditional ML is that performance of DL algorithms benefits more when the dataset augments. DL models do not require complicated artificial feature engineering.

Retraining using online data opens the possibility of incorporating attack data into the normal dataset and avoiding detection. Class imbalance can cause DL models to be biased towards the negative class and negatively impact classification performance. Random sampling can help to balance it. The goal of detecting unseen attacks (zero-day) should be elaborated as detecting new attacks that share specific characteristics with the known attack used in training. Some models are better at limited dataset size and for specific Web attacks. Unsupervised DL models, like Autoencoders, can achieve high F1-scores in Web attack detection without domain knowledge and labeled training data. Publicly available datasets are outdated and not representative in most use cases. Therefore, we can conclude that data quality and feature extraction are critical components for the robustness and precision of these classifiers. Moreover, the proposed DL methods effectively detect Web attacks with FPR close to zero and the corresponding TPR (AUC-ROC curve) of almost 100%. Furthermore, codes and preprocessing details or datasets need to be publicly available for reproducibility and as a contribution to similar works in the future. New models should aim to achieve better detection accuracy, low computational cost, high flexibility, and high robustness. Some possible future works or improvements in this area are as follows: (i) finding and correcting misclassified labels in the original dataset, (ii) extracting more characteristic information of the Web attack besides the HTTP request or URL. (iii) creating a new standardized dataset of Web attacks (iv) improve existing approaches even further by exploring alternative combinations of DL models, (v) determine the frequency when DL models should be retrained and (vi) develop a methodology for creating new DL models and a framework for evaluating existing ones.

REFERENCES

[1] Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. applied sciences, 9(20), 4396.

[2] Ahmad, R., Alsmadi, I., Alhamdani, W., & Tawalbeh, L. A. (2023). Zero-day attack detection: a systematic literature review. Artificial Intelligence Review, 1-79.

[3] Gong, X., Lu, J., Wang, Y., Qiu, H., He, R., & Qiu, M. (2019, December). CECoR-Net: A character-level neural network model for web attack detection. In 2019 IEEE International Conference on Smart Cloud (SmartCloud) (pp. 98-103). IEEE.

[4] Liang, J., Zhao, W., & Ye, W. (2017, December). Anomaly-based web attack detection: a deep learning approach. In Proceedings of the 2017 VI International Conference on Network, Communication and Computing (pp. 80-85).

[5] Karacan, H., & Sevri, M. (2021). A novel data augmentation technique and deep learning model for Web application security. IEEE Access, 9, 150781-150797.

[6] Yu, L., Chen, L., Dong, J., Li, M., Liu, L., Zhao, B., & Zhang, C. (2020, July). Detecting malicious web requests using an enhanced textcnn. In 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 768-777). IEEE.

[7] Tian, Z., Luo, C., Qiu, J., Du, X., & Guizani, M. (2019). A distributed deep learning system for web attack detection on edge devices. IEEE Transactions on Industrial Informatics, 16(3), 1963-1971.

[8] Yang, W., Zuo, W., & Cui, B. (2019). Detecting malicious URLs via a keyword-based convolutional gated-recurrent-unit neural network. IEEE Access, 7, 29891-29900.

[9] Pan, Y., Sun, F., Teng, Z., White, J., Schmidt, D. C., Staples, J., & Krause, L. (2019). Detecting web attacks with end-to-end deep learning. Journal of Internet Services and Applications, 10(1), 1-22.

[10] Luo, C., Tan, Z., Min, G., Gan, J., Shi, W., & Tian, Z. (2020). A novel web attack detection system for internet of things via ensemble classification. IEEE Transactions on Industrial Informatics, 17(8), 5810-5818.

[11] Liu, T., Qi, Y., Shi, L., & Yan, J. (2019, August). Locate-Then-Detect: Real-time Web Attack Detection via Attention-based Deep Neural Networks. In IJCAI (pp. 4725-4731).

[12] Tang, R., Yang, Z., Li, Z., Meng, W., Wang, H., Li, Q., ... & Liu, Y. (2020, July). Zerowall: Detecting zero-day web attacks through encoder-decoder recurrent neural networks. In IEEE INFOCOM 2020-IEEE Conference on Computer Communications (pp. 2479-2488). IEEE.

[13] Tang, P., Qiu, W., Huang, Z., Lian, H., & Liu, G. (2020). Detection of SQL injection based on artificial neural network. Knowledge-Based Systems, 190, 105528.

[14] Xie, X., Ren, C., Fu, Y., Xu, J., & Guo, J. (2019). Sql injection detection for web applications based on elastic-pooling cnn. IEEE Access, 7, 151475-151481.

[15] Seyyar, Y. E., Yavuz, A. G., & Ünver, H. M. (2022). An attack detection framework based on BERT and deep learning. IEEE Access, 10, 68633-68644.

[16] Kadhim, R., & Gaata, M. (2020). A hybrid of CNN and LSTM methods for securing web application against cross-site scripting attack. Indones. J. Electr. Eng. Comput. Sci, 21, 1022-1029.

[17] Mokbal, F. M. M., Dan, W., Imran, A., Jiuchuan, L., Akhtar, F., & Xiaoxi, W. (2019). MLPXSS: an integrated XSS-based attack detection scheme in web applications using multilayer perceptron technique. IEEE Access, 7, 100567-100580.

[18] Melicher, W., Fung, C., Bauer, L., & Jia, L. (2021, April). Towards a lightweight, hybrid approach for detecting dom xss vulnerabilities with machine learning. In Proceedings of the Web Conference 2021 (pp. 2684-2695).

[19] Maurel, H., Vidal, S., & Rezk, T. (2022). Statically identifying XSS using deep learning. Science of Computer Programming, 219, 102810.

[20] Deshpande, G., & Kulkarni, S. (2019). Modeling and mitigation of XPath injection attacks for web services using modular neural networks. In Recent findings in intelligent computing techniques (pp. 301-310). Springer, Singapore.

[21] Ali, W., & Ahmed, A. A. (2019). Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm ‐ based feature selection and weighting. IET Information Security, 13(6), 659-669.

[22] Feng, F., Zhou, Q., Shen, Z., Yang, X., Han, L., & Wang, J. (2018). The application of a novel neural network in the detection of phishing websites. Journal of Ambient Intelligence and Humanized Computing, 1-15.

[23] Yadav, S., & Subramanian, S. (2016, March). Detection of Application Layer DDoS attack by feature learning using Stacked AutoEncoder. In 2016 international conference on computational techniques in information and communication technologies (icctict) (pp. 361-366). IEEE.

[24] Chen, J., Yang, Y. T., Hu, K. K., Zheng, H. B., & Wang, Z. (2019, February). DAD-MCNN: DDoS attack detection via multi-channel CNN. In Proceedings of the 2019 11th International Conference on Machine Learning and Computing (pp. 484-488).

[25] Lv, Z. H., Yan, H. B., & Mei, R. (2019). Automatic and accurate detection of webshell based on convolutional neural network. In Cyber Security: 15th International Annual Conference, CNCERT 2018, Beijing, China, August 14–16, 2018, Revised Selected Papers 15 (pp. 73-85). Springer Singapore.

[26] Xuan, S., Man, D., Wang, W., Qin, K., & Yang, W. (2018, November). Hybrid Classification of WEB Trojan Exploiting Small Volume of Labeled Data Vectors. In 2018 14th International Conference on Computational Intelligence and Security (CIS) (pp. 286-290). IEEE.

Social Engineering Aspects of email Phishing: an Overview and Taxonomy

Igor Tomičić
Faculty of Organization and Informatics, University of Zagreb
Artificial Intelligence Laboratory; Center for Forensics, Biometrics and Privacy
Pavlinska 2, 42000 Varaždin, Croatia
Email: igor.tomicic@foi.hr

Abstract— **Numerous online resources and reports are pointing to the growing effectiveness of email phishing techniques, with some indicating that 85% of IT breaches involve the human element, and that 96% or social attacks arrive via email. Phishing is a common occurrence, and a significantly successful one. While most of the available research on phishing involves phishing detection, prevention, filtering, anti-phishing tools, techniques and countermeasures, the remaining body of research is tackling phishing and social engineering in (too) generic and broad contexts. This paper will propose a focused effort to identify the specific groups of techniques that attackers are using in email phishing and the principles running "behind the scenes" that make these attacks successful. Thus, the goal of this paper is threefold: (1) to propose a taxonomy of the observed email phishing techniques, (2) to associate the principles and factors of influence with observed techniques and shed light behind their effectiveness, and (3) to raise awareness and lay the groundwork for working on the model of human resilience against these manipulative forms of cyber attacks.**

Keywords— social engineering; email; phishing; taxonomy

I. Introduction

According to the Verizon Data Breach Reports from 2020 and 2021 (DBIR), 85% of security breaches involved the human element, almost 100 % of social attacks in the public administration sector involved phishing, and 96% or social attacks arrive via email [1]. Some reports state that Internet users on average receive 117 emails per day and that 53% of such emails are spam [2]. Even if the automated spam filters are successful in blocking them 90% of the time, users could still experience few phishing emails in their inboxes every day - and it only takes one wrong click to become compromised. Related statistics can be found via numerous other online resources, emphasizing the importance of further research that needs to be conducted within this domain. Although phishing attacks can be simple and generic, the most successful ones are often highly targeted and elaborate, collecting the publicly available knowledge about the target using OSINT tools and techniques, and leveraging the principles of psychological influence for delivering highly effective campaigns. Technology is ever-developing at rapid speeds, enabling both more sophisticated attack and defense methods and techniques. However, within the same time frame, the human mind has remained relatively unchanged in its basic susceptibility to targeted manipulation. Thus, the main motivation of this work is to shed light on these manipulative techniques and to raise awareness of their existence and usage in phishing campaigns, so people could be trained to recognise indicators of manipulation and do so in a relatively automated fashion.

Phishing is a type of a complex socio-technical attack based on social engineering and interconnecting areas such as social psychology, information security, technologies in general, communication artefacts, even organisation processes. Phishing uses fraudulent communications crafted, for example, with the goal of tricking a person into revealing sensitive information to the attacker, or to deploy malware on the victim's infrastructure, often creating a gateway for future attacks. Phishing seems to be the most common type of attack in a cyber landscape. For example, FBI's Internet Crime Complaint Centre recorded 6 times more incidents of phishing than identity thefts in 2020 [3].

Phishing utilises a category of web threats called semantic attacks, where the focus is typically not on technical vulnerabilities, but on findings of how humans assign meanings to the message contents or interact with computers [4].

This paper is focused on email-related phishing attacks, as this vector is the most prevalent one.

II. Principles of Influence

The effectiveness of the phishing campaigns might be related to psychological principles of influence/persuasion presented by Cialdini [5], which are labeled as follows: reciprocity, commitment and consistency, consensus or social proof, authority, liking, scarcity and unity.

Reciprocity describes the inherent human need to give back to others the form of a behavior, gift, or service that they have received first.

Consistency manifests through people needing to be consistent with their previous attitudes and/or behaviours; it is usually triggered by asking for small initial commitments which could, after complying, lead to bigger ones. People have

aligned commitment with their self-image, so this mechanism is often exploited.

Social proof describes the state in which people tend to look to the actions and behaviors of others to determine their own, which more often happens in states of uncertainty. Social proof is often fabricated in order to trigger and exploit this principle.

Authority describes the state of mind in which people follow the lead of credible, knowledgeable experts - or the ones that are impersonating authority figures.

The liking/sympathy principle establishes that people prefer to comply with requests of those that they "like". Liking in this context is based on several factors, such as physical attractiveness, similarity, sincere compliments, contact and cooperation towards mutual goal. Liking is also triggered by showing vulnerability and using humor in establishing rapport.

Scarcity is not only restricted to physical resources, but to the limited time, limited stock, exclusivity, limited number of spots, etc. These limitations are often fabricated.

The unity principle states that more one feels like part of a group, the greater the chances to be influenced by that group. To exploit this, one has to provide the illusion of a shared identity, interests, goals, etc.

Throughout the years, these principles were exploited in a number of domains [5], and in modern times they are being widely used in the phishing and other forms of social engineering campaigns [6], [7], [8], [9].

Akbar [9] presented a quantitative analysis on suspected phishing emails and found that authority and scarcity were disproportionately the most exploited principles, followed by liking, consistency, reciprocation, and social proof.

III. Related Work

The lack of "a detailed and systematic discussion on the phishing techniques" is noticed by authors in [10], claiming that researchers are "more focused on the discussion of anti-phishing as opposed to phishing" itself. Similar observation was perceived in writing this paper, with observed discrepancies between categorisations of phishing attacks across different papers, inconsistent terminology and phishing domains. The review of related work presented within this section should shed a light on such cases, and later try to reconcile those diverse aspects. Moreover, most papers within the domain are tackling phishing in general, some even generalising to other forms of non-technical social engineering attacks. This paper however, is focusing on email-related phishing attacks because of their exceptional spread and effectiveness, making them often self-sufficient as an attack vector.

The authors in [11] describe phishing as a type of spam that employs two different techniques - social engineering schemes (spoofing a legitimate company, placing links to spoofed (fake) websites, etc.) and "technical" schemes, which rely on malware or security vulnerabilities found in user's system that would support target malicious activities. Such categorisation has its limitations - malware could be delivered by email and the incentive to open/click it could come from social engineering techniques. Attack on the vulnerable components of the user's system might not be related to phishing at all, if it does not include some form of user interaction - and if it does, than it also relies on some form of behavioural manipulation.

In [12] authors recognise two aspects of phishing; one with respect to target specialisation (such as spear-phishing which could lead to a business email compromise [13], [14], or "Fire-and-forget", the most common type according to [15]), and the other with respect to the attack vector (malicious attachment, phishing link). Authors also included several strategies for fighting phishing, such as user training, automated identification of phishing emails, domains, and websites, and aiding users in spotting suspected phishing emails or websites by providing warnings. The paper, however, does not dive deeper into the email phishing techniques themselves.

Authors in [16] are acknowledging several psychological factors within the phishing context. They are referring to manipulations with human curiosity, fear and empathy on top of "traditional phishing techniques" in order to trick the users into submission. Use-cases for all three aspects are provided within the paper. Authors also classify phishing attacks into a range of "techniques": spear phishing, clone phishing, malware-based phishing, search engine phishing. Such classification has its merits, although it may be considered incomplete and somewhat inconsistent, mixing both the target categories and attack vectors/deployment methods.

Authors [17] summarise phishing domain into two aspects: phishing categories (clone phishing, spear phishing, phone phishing, DNS-Based Phishing (Pharming), Man-in-the-middle-attack), and phishing attack techniques (email spoofing, web spoofing, DNS Cache Poisoning, malware). Clone phishing is a term which seems to be used throughout the academic literature and popular articles in two ways: (1) as a method of cloning a legitimate website, or (2) as a method of cloning and modifying a legitimate, previously sent email, all with the intent of utilising the Cialdini's "liking" principle of persuasion [5] for exploitation. DNS-Based Phishing (mapping the domain name of a genuine web site onto the IP address of a rogue website) and man-in-the-middle-attack are usually executed through technical vectors, where social engineering is rarely used - similarly as DNS Cache Poisoning, which authors included again within phishing attack techniques. Also, email/web spoofing and email/web cloning could easily refer to the same phishing methods, respectively.

Web phishing strategies are organized into three groups in [18]: (1) mimicking attack (effectively the same as clone phishing described before), (2) forward attack (a victim clicks on the link provided via phishing email, the link lands on the attacker's website where the victim leaves her personal information, after which it is forwarded to a legitimate site to minimise suspicion), and (3) pop-up attack (the victim is asked to leave her personal information in the attacker's pop-up window delivered by the MITM technique). From the perspective of our paper, where the focus rests on the email phishing, the analysis deals with the email contents - why would the victim be compelled to click on those links in the first place; which mechanisms are in play, which psychological

principles of influence are employed - all of which is not addressed within [18].

Phishing attack methods and defense techniques from five perspectives are described in [19]. Authors tackle the life-cycle and motivation of a phishing attack, distribution methods, taxonomy of various phishing-attacking techniques (Figure 1), provide phishing protection mechanisms and present some performance challenges faced by developers dealing with this domain. Although their taxonomy provides a basic orientation of the phishing context, it is important to note that smishing is susceptible to social engineering techniques in similar ways as email phishing is; and vishing is especially subjected to social engineering, as documented in numerous real-world examples such as presented in [6]. Therefore, there should be a clear link between social engineering and smsishing and vishing within the published taxonomy presented in Figure 1. In relation to the taxonomy proposed within this paper, the taxonomy by [19] does not provide an insight into the email phishing techniques.

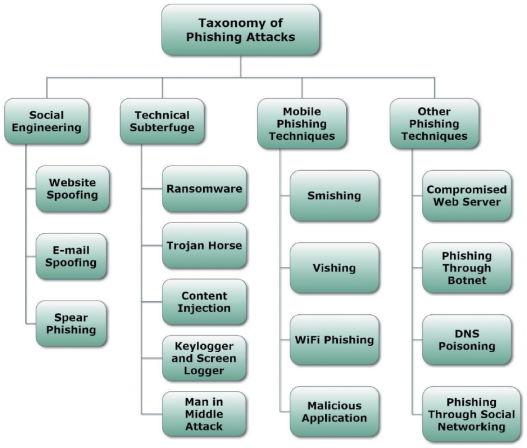


Fig. 1: Taxonomy of phishing attacking techniques by [19]

Authors in [20] have summarised several phishing trends, adding 2 of their own (smishing and vishing) on top of the work presented in [21], which includes: (1) the money scam (mainly targeting human greed in the context of offering some kind of money compensation), (2) information scam (targeting confidential data from the user with the aim of later using it for other malicious purposes - attacker usually present itself as an authority figure in order to exploit human susceptibility to authority, or as a recognisable/reputable institution like a bank, targeting the liking principle of influence, or fear, by using strong action words of threats, consequences and the lack of time to react (utilising the scarcity principle)); (3) malware distribution; (4) multiple file extensions (referring to hiding the real file extensions); (5) disguised links (including cloaked/obfuscated links, misleading named links, Homograph URLs); (6) spear phishing (requiring previous reconnaissance on the target for email customization); (7) Business Email Compromise (sending emails from the compromised or spoofed email accounts, often from the same organization, belonging to some high-ranking individual). Again, this summary of phishing trends cover some of the most important techniques and categories, but do so without apparent structure, which is a challenging task when including phishing across multiple vectors (email, web, voice, sms, even f2f), considering how the same techniques could be effectively used across several of such domains.

Interestingly, authors in [22] have recognised the importance of emotional triggers elicited by phishing emails and have created "a human-centric notification mechanism that extracts prospective psychological triggers, possible malicious intent, and a representative summary from emails." They present these findings to the user in a meaningful way to facilitate a better decision-making process. The work leverages transformer-based machine learning to "(i) analyze prospective psychological triggers, to (ii) detect possible malicious intent, and to (iii) create representative summaries of emails" [22].

Phishing emails often exploit the reciprocity principle (by engineering the context in which the victims feel like they have to return the favour), social proof (creating the impression that "other people" have already done the thing that is asked from the potential victim), and scarcity (creating the impression that the opportunity is limited, or that something has to be done as soon as possible, artificially limiting the time to react).

IV. Materials and Methods

After the systematic overview of the existing literature on the subject of email phishing techniques, hundreds of phishing emails were qualitatively analysed in order to find the indicators of manipulative techniques present in the email headers, subject lines and in the body of email messages. Phishing emails were collected, analysed or re-interpreted (1) from the existing literature review presented in Section III, (2) from the authors' personal collection, (3) from the Cornell University "Fish Bowl", a collection of phishing emails that have been spotted at Cornell [23], (4) from an archive of spoof email and phishing scams [24], and (5) from numerous web articles that have presented various phishing scenarios. During the analysis, various technique patterns were categorised according to their primary target (machine or human) and scope, creating a taxonomy of used techniques.

The paper associates identified email phishing techniques with the Cialdini's work on the principles of influence [5] in an effort to shed light on the evident effectiveness of the phishing email techniques. As those principles are already proven to work in other non-technical contexts, their universal nature is here applied to manipulative email communications. While not explicitly covered through these principles, other related emotional states are also considered, such as greed, fear and compassion/empathy, as significant factors of influence.

V. Email Phishing Techniques in the Wild

This section presents various phishing techniques and tricks discovered in the analysis of online sources and hundreds of different phishing emails. There are two major categories of

phishing techniques identified through the analysis - machine-oriented and people-oriented. Machine-oriented techniques are designed specifically to bypass the software-based email filters and scanners. People-oriented techniques are crafted with the goal to deceive humans through one or more of the influence principles and factors associated with a specific technique. As this paper is focused on the social engineering aspect of email phishing, machine-oriented techniques will be tackled in less detail. Naturally, there are hybrid approaches used with the capacity to deceive both human and software.

A. Machine Oriented Email Phishing Techniques

Some of the identified machine-oriented email phishing techniques are listed as follows:

1) The inclusion of legitimate links. If the email contains hyperlinks which lead to legitimate, well known websites, such an email would also look legitimate - both to software filters and humans [25].
2) Obfuscation of brand logos [25], [26]. Some email filters scan for signatures of HTML attributes that compose company logos, in an effort to detect misusing the logos for deceitful purposes (brand impersonation protection). Such attributes could be changed in ways that are invisible to human eyes, but the logo would become unrecognisable to a software filter.
3) Using less content and noise [25]. If there is no content to scan, the email filters could mark an email as safe. Therefore, the attackers could use images that contain text fused within them, instead of using the text format itself, which is usually not obvious to the average victim and would remain undetectable by some filters. Although, more sophisticated anti-phishing filters use optical character recognition to recover such text and identify potential phishing.
4) Mixing legitimate and malicious code [25]. This technique is used to avoid signature-based detection, and is used more often in phishing pages. A phishing page can for example include CSS and JS code from the legit corporate web pages [26]. The signature can also be obfuscated by adding random values, blank spaces, etc.
5) Linking most of the code in an external JavaScript file.
6) Inserting invisible Unicode characters to break up keywords. This technique is performed in a phishing email body or subject line, in order to bypass automated detection [26].
7) Zero-point font obfuscation [26]. Attackers insert hidden words with a font size of zero into the body of an email in order to avoid machine learning detections.
8) Using procedurally-generated graphics for brand impersonation [26]. There are reported cases of using HTML tables to imitate the logos and branding of legitimate and trusted organizations.

B. People Oriented Email Phishing Techniques

In people-oriented phishing techniques, the goal is to manipulate the perceived context and to deceive a human target. There are two main categories of such attacks: bulk, and personalised (spear) phishing. The bulk phishing techniques are described here as "generic" in the sense that they could be employed to any context or a victim profile.

The following people-oriented bulk email phishing techniques were identified:

1) Using URL shorteners. URL shorteners, as their name suggests, are services that are primarily and legitimately used for the reduction of long URLs. As such, they can completely mask the original (target) URL, as the two are mostly lexically unrelated and have no visual resemblance. The attacker could use this feature to mask any link needed for the attack to succeed. This technique does not directly utilise principles of influence, but can be used after a successful persuasion.
2) Using anchor texts. Links can also be hidden "below" anchor texts, which are usually designed to further motivate clicking on them (for example, "Click here now to renew your account", or "Microsoft Helpdesk", or "CLAIM YOUR ACCOUNT NOW"). They are also in many cases used in combination with URL shorteners to further avoid detection. Revealing links underneath the anchor text by "hovering the cursor" is especially challenging on mobile devices. Principles which are used via the anchor text depend on the message context, such as authority, liking, scarcity, fear, commitment.
3) Using redirections. There are two major techniques used here. Open redirects enable anyone to craft a URL that will redirect a user from the legitimate URL to the website of their choice; it is a vulnerability that allows attackers to use an established business reputation to perform phishing attacks, or in another words, to disguise malicious URLs with a trusted domain. Another type of technique is the post-exploitation redirection - after the user has been tricked into divulging his sensitive data on the phishing site, the victim is redirected to the legitimate site in order to remove potential doubt by exploiting the "liking" principle.
4) Using misspelled URLs. The misspelling is usually done subtly, by making the link appear to belong to the known/reputable organization (for example, facebok.com, mircosoft.com, etc.). Again, mostly the "liking" principle is exploited here.
5) Homograph attacks. Similarly to misspelled URLs, the goal is to craft the URL that is visually identical to a legitimate site by using similarly shaped characters (for a trivial example, micros0ft.com, FACEB00K.COM). More advanced attacks use homographs in internationalised domain names, which could be visually indistinguishable from each other (for example, in Latin the letter "a" is U+0061, and in Cyrillic the letter "a" is U+0430). Online article [27] showed an interesting example of masking the original domain by using the mathematical division operator ∕ (U+2215), which looks like a ASCII slash / (U+002F) in the following URL:

http://example.com/a-top-level-domain.com/ where the victim could falsely identify the string "a-top-level-domain.com" as a folder located in the root directory under the domain "example.com. Targeting the similarity and familiarity, this technique also mostly exploits the liking principle.

6) The use of subdomains. Attackers can use cloud services to create a malicious app that gets assigned a subdomain, and then host phishing pages. Several phishing URLs hosted on the domains of cloud computing services and platforms were identified in [28], one such company being appspotDOTcom, a Google cloud computing platform for developing and hosting web applications. The same article reports six appspotDOTcom subdomains that have been confirmed as phishing sites as of 2 October 2020. Many similar examples exist in the wild, exploiting the liking principle.
7) Hiding the malicious URL within the subdomain. For example, attackers can use an URL such as this one: legitimate.malicious.com, making it seem like the victim is directed to the legitimate URL, exploiting the liking/-familiarity principle.
8) Creating a sense of urgency, which exploits mostly scarcity and authority principles. By adding the social proof and unity to the mix, one can further increase effectiveness ('the rest of the group had already done it, and the time is running out').
9) Using overwhelmingly long emails "to deceive the recipients to overlook the phishing components and to focus on urgency and/or emotions"
[22], effectively creating cognitive overload and falling back on the emotional decision making process. Such emails might use the scarcity and authority principles, but it would depend on the context itself which ones would prove to be most effective.
10) Complex scenario-based emails. A most popular example would probably be a 419/Nigerian scam - a lengthy email from someone claiming to be a Nigerian prince, which either offers money to the victim (if the victim sends him a small amount first in order to claim it), triggering greed, or the attacker claims he is in a dire situation, and needs money to resolve it, triggering compassion/empathy. The number "419" refers to the section of the Nigerian Criminal Code dealing with fraud, charges and penalties for offenders. Numerous other scenarios are present in modern phishing landscape, targeting various emotions and principles of influence. If a person sees herself as a compassionate one, this would trigger a commitment/consistency principle. The belief that someone would gave a person some money could trigger the reciprocity principle. Although trivial in today's perspective, this technique seems to be quite complex from the persuasion point of view.
11) Attention-grabbing techniques within the subject lines, as detailed below.

Attention-grabbing techniques within the subject lines are using elicitation of emotions, personalised narrative, familiarity, special characters, absence of characters, immediate call for action, etc. Some of the most prevalent examples include:

- Using "RE:" or "FW:", implying the continuance of some past communication, which is in most cases fabricated, but in spear phishing campaign it may refer to some actual intercepted communication, targeting the familiarity/liking principle with possibilities to tackle commitment/consistency.
- Attention locks with special characters inside subject line, such as: ****High Severity Alert****, targeting mostly scarcity and authority principles.
- Call to action in subject line, such as "Action Required! ..." also targeting mostly scarcity and authority principles.
- Explicitly referring to some previous activity/communication which can be fabricated or real ("This is to notify you that your incident has been resolved"), targeting the familiarity/liking principle with possibilities to tackle commitment/consistency.
- Referring to some urgent event, usually fabricated ("High severity alert: Password Expiring Notice..."), targeting mostly scarcity and authority principles.
- Referring to some globally familiar event of interest ("COVID-19 Relief Fund"), targeting the liking principle.
- Empty subject line, eliciting curiosity.

Targeted, personalised (spear) phishing techniques identified in this section are referring to those techniques that are selectively pre-crafted and tailored to the specific context and/or victim profile. Where not explicitly mentioned, the individual technique is typically associated with the liking principle of persuasion, given the artificial context which is intentionally fabricated for the sole purpose of triggering the environmental familiarity with the victim.

1) Spoofing email headers. Attackers can use scripts to change the email header fields such as "from" address and the "reply-to" address, due to inherent vulnerability of SMTP which does not have a built-in method for authenticating email addresses by default. Attackers could thus exploit the liking principle should they use an email address which is known to the victim.
2) Appearing as if the email is coming from your own email (advanced customised email spoofing technique).
3) Cloning the previously sent legitimate email, where the attachment and/or link within the email might be replaced with the malicious version; the success rate of such an email can be increased by claiming to be a resend of the original, or an updated version to the original. We can call this "context hijacking".
4) Adapting the display name to impersonate a sender: for example, sending malicious emails from 'LegitimateName@gmail.com' rather than 'LegitimateName@legitimateCompany.com.'
5) The use of fake news articles. This technique is designed to elicit a strong emotion from the victim, causing him to

click a link without rationally considering where it might lead. The article is crafted with the victim's personal and/or professional interests in mind. It leverages the strong human desire to correct what obviously seems wrong from the personal perspective.

6) Any other personalization of the subject line and message body tailored to the victim's identity or general profile. Crafting such a spear-phishing email requires additional previous effort of performing the OSINT on the target. We can call this "context adaptation". Business email compromise could also fall into this category. Context adaptation could be crafted "real-time", by exploiting some immediate event, situation or circumstances (a business trip, a specific meeting, hardware maintenance or audit, etc.), or "all-time", exploiting victim's habits, hobbies, daily routines, etc. Context adaptation could target any of the Cialdini's principles - liking, authority, scarcity, reciprocity, commitment/consistency, social proof - and the decision on which one(s) to target should derive from the analyzed victim's profile.

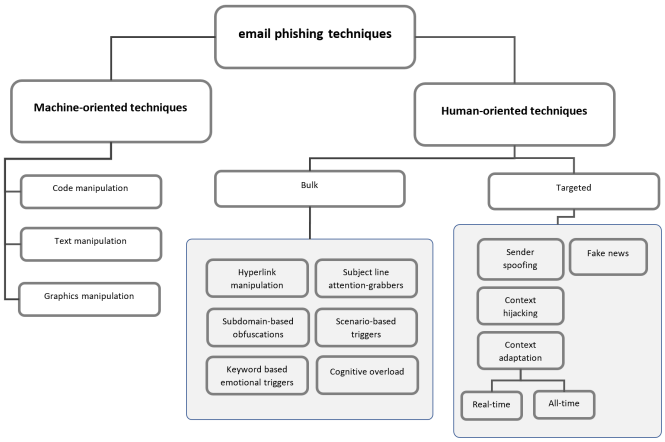


Fig. 2: The proposed taxonomy of email phishing techniques

VI. Conclusion, Limitations of the Study, Future Research

The study made a qualitative analysis of hundreds of email phishing messages from several sources, namely from the existing literature review presented in Section III, from the Cornell University "Fish Bowl" [23], an archive of spoof email and phishing scams [24], from the authors' personal collection and from numerous publicly available web articles.

The main contribution of the paper is the proposed taxonomy of email phishing techniques. Additional contributions include a comprehensive literature overview on the subject of email phishing, associations of phishing techniques with the principles and factors of influence, and raising awareness on the effectiveness of such techniques. The goal is to educate users not only on the technical aspects of phishing but also on the psychological triggers and social engineering techniques that may drive their actions towards unexpected and unwanted results.

Although a large number of messages were analysed in an attempt to identify the most prevalent types of phishing techniques, manual qualitative analysis has its limitations in terms of scope and thoroughness and may miss some of the significant techniques used in the wild. Processing a larger sets of data through automation could supplement such research with additional valuable information. Although there are some demonstrations in current research to automate such a process, there has been little effort to directly associate the specific phishing technique to principles of influence, to the best of author's knowledge.

The study made an observational association between the chosen principles and factors of influence and the phishing email techniques found within the described scope. While Cialdini's framework has been widely used in social engineering settings, it is important to acknowledge that other psychological factors, models, and frameworks beyond the scope of this study could provide additional insights into the effectiveness of phishing emails and social engineering techniques. Calculating specific correlation values might prove to provide deeper insight into the specific technique's effectiveness. By generating statistics that show the number of processed emails falling into different categories of the proposed taxonomy, it would be possible to prioritize a defense model based on the most prevalent attack techniques.

Thus for the future research there are three outstanding challenges: to enable a type of automation for processing larger amounts of phishing emails, to generate insightful statistics on the attack techniques regarding the proposed taxonomy, and to gain a deeper insight through more advanced correlation calculations.

VII. Acknowledgement

This research was funded by the project "Development of CSTI platform for retrieval and analysis of structured and unstructured data". The project received funding from the European Regional Development Fund through OP Competitiveness and Cohesion 2014–2020 within the Call for Proposals "Development of the products and services arising from research and development activities—Phase II", under grant number KK.01.2.1.02.0310.

References

[1] Verizon, "2021 data breach investigations report. technical report 11th edition." 2021. [Online]. Available: https://www.verizon.com/business/resources/reports/dbir/

[2] Symantec, "Symantec security report 2017," 2021. [Online]. Available: https://www.symantec.com/security-center/threat-report

[3] F. I. C. C. C. U.S. Federal Bureau of Investigation, "Internet crime report 2020," 2020. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf

[4] B. Schneier, "Inside risks: semantic network attacks," *Communications of the ACM*, vol. 43, no. 12, p. 168, 2000.

[5] R. B. Cialdini, "Influence: The psychology of persuasion," *New York: Quill/W. Morrow*, 1993.

[6] C. Hadnagy, "Social engineering: The science of human hacking. wiley publ," 2018.

[7] D. S. Oliveira, T. Lin, H. Rocha, D. Ellis, S. Dommaraju, H. Yang, D. Weir, S. Marin, and N. C. Ebner, "Empirical analysis of weapons of influence, life domains, and demographic-targeting in modern spam: an age-comparative perspective," *Crime science*, vol. 8, no. 1, pp. 1–14, 2019.
[8] S. Uebelacker and S. Quiel, "The social engineering personality framework," in *2014 Workshop on Socio-Technical Aspects in Security and Trust*. IEEE, 2014, pp. 24–30.
[9] N. Akbar, "Analysing persuasion principles in phishing emails," Master's thesis, University of Twente, 2014.
[10] K. L. Chiew, K. S. C. Yong, and C. L. Tan, "A survey of phishing attacks: Their types, vectors and technical approaches," *Expert Systems with Applications*, vol. 106, pp. 1–20, 2018.
[11] A. Almomani, B. B. Gupta, S. Atawneh, A. Meulenberg, and E. Almomani, "A survey of phishing email filtering techniques," *IEEE communications surveys & tutorials*, vol. 15, no. 4, pp. 2070–2090, 2013.
[12] J. Petelka, Y. Zou, and F. Schaub, "Put your warning where your link is: Improving and evaluating email phishing warnings," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–15.
[13] J. Hong, "The state of phishing attacks," *Communications of the ACM*, vol. 55, no. 1, pp. 74–81, 2012.
[14] IBM, "Ibm x-force threat intelligence index 2018," 2018. [Online]. Available: https://www.ibm.com/security/data-breach/threat-intelligence
[15] Verizon, "2018 data breach investigations report. technical report 11th edition." 2018.
[16] J. A. Chaudhry and R. G. Rittenhouse, "Phishing: classification and countermeasures," in *2015 7th International Conference on Multimedia, Computer Graphics and Broadcasting (MulGraB)*. IEEE, 2015, pp. 28–31.
[17] M. N. Banu and S. M. Banu, "A comprehensive study of phishing attacks," *International Journal of Computer Science and Information Technologies*, vol. 4, no. 6, pp. 783–786, 2013.
[18] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Tutorial and critical analysis of phishing websites methods," *Computer Science Review*, vol. 17, pp. 1–24, 2015.
[19] A. K. Jain and B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterprise Information Systems*, pp. 1–39, 2021.
[20] P. N. Mangut and K. A. Datukun, "The current phishing techniques–perspective of the nigerian environment," *World Journal of Innovative Research (WJIR)*, vol. 10, no. 1, pp. 34–44, 2021.
[21] C. Ross, "The latest attacks and how to stop them," *Computer Fraud & Security*, vol. 2018, no. 11, pp. 11–14, 2018.
[22] A. Kashapov, T. Wu, S. Abuadbba, and C. Rudolph, "Email summarization to assist users in phishing identification," in *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security*, 2022, pp. 1234–1236.
[23] C. University, "Phish bowl," 2020. [Online]. Available: https://it.cornell.edu/phish-bowl
[24] millersmiles.co.uk, "The archive of spoof email and phishing scams," 2020. [Online]. Available: http://www.millersmiles.co.uk/archives/331
[25] R. Basset, "5 common phishing techniques," 2019. [Online]. Available: https://www.vadesecure.com/en/blog/5-common-phishing-techniques
[26] M. D. T. I. Team, "Trend-spotting email techniques: How modern phishing emails hide in plain sight," 2021. [Online]. Available: https://www.microsoft.com/security/blog/2021/08/18/trend-spotting-email-techniques-how-modern-phishing-emails-hide-in-plain-sight/
[27] crypto it.net, "Homograph attack," 2021. [Online]. Available: http://www.crypto-it.net/eng/attacks/homograph-attack.html
[28] W. API, "Attack surface monitoring: Two ways to detect phishing subdomains," 2020. [Online]. Available: https://circleid.com/posts/20201009-attack-surface-monitoring-ways-to-detect-phishing-subdomains

Data leaks to third parties in web services for vulnerable groups

Robin Carlsson*, Sampsa Rauti*, Timi Heino*,
* University of Turku, Department of Computing, Turku, Finland
crcarl,sjprau,tdhein@utu.fi

***Abstract*—Third-party analytics services are increasingly being used to improve sales and usability of websites. While these services often have great value for companies and organizations using them, they also raise privacy issues. When information is gathered using third-party analytics, third parties also receive lots of personal data about users. This is often especially true for many vulnerable groups who may be forced to use online services instead of on-site services, such as the elderly, people with medical issues, or people living in remote locations. We conduct a comprehensive study of 15 Finnish web services often used by vulnerable groups by analyzing their network traffic. Our findings show most of these services use third-party analytics and, despite their delicate nature, send highly sensitive personal data to third parties. The study also discusses the implications of the found data leakages and offers recommendations on how to improve privacy of online services from a software engineering point of view.**

Keywords—privacy, web services, data leaks, vulnerable groups

I. Introduction

Digital technology provides a valuable means of delivering services to those who may have difficulties accessing them in person [1], [2]. As access to technology and internet connectivity has expanded, policymakers have recognized the importance of digital service delivery. For example, in 2019, Finland passed the Act on the Provision of Digital Services, which aims to improve the accessibility, quality, and security of digital services for all individuals. The COVID-19 pandemic has highlighted the need for online services, and it has further accelerated the shift towards digital service delivery.

In particular, many vulnerable groups can benefit from digital services. Vulnerable groups refer to individuals or communities that are at a greater risk of experiencing negative outcomes or facing challenges because of their socio-economic status, health, age, sexual orientation, or other factors [3], [4]. For example, elderly and older adults may become more vulnerable to health problems, isolation, and other challenges. Similarly, people with disabilities are often more susceptible to lack of access to services and opportunities, as are many immigrants and refugees.

Digital services, which can be easily accessed from anywhere with an internet connection, are particularly beneficial for individuals who may have difficulty leaving their homes or accessing transportation due to financial issues, remote location, disability, or health concerns [5]. Digital services also lower barriers to access, especially for people who may have language barriers, lack of documentation, or other challenges that may prevent them from accessing on-site services. It is also often argued that digital services can provide a higher level of confidentiality than on-site services, which can be beneficial for individuals who may be concerned about privacy or the potential for discrimination. However, online services may not be as confidential and private as expected, especially for many vulnerable groups [6].

This confidentiality aspect of digital services is the focus of the current study. We conduct an in-depth study of 15 Finnish web services often used by vulnerable groups by analyzing their network traffic while using the main functionality of the services. We study what third-party analytics these services use and explore what kind of personal data they transmit to third parties. Some previous studies have addressed third-party analytics services on websites used by vulnerable groups. For example, analytics in web-based governmental services [7], healthcare services [8] and pharmacies [9] have been studied. However, the current study discusses a wider selection of different online services frequented by vulnerable groups, while providing an in-depth analysis of found data leakages.

The rest of the paper is structured as follows. Section 2 outlines the study setting and methods used in the current study. Section 3 presents the results of our network traffic analysis and explores what kind of personal data the studied web services sent to third parties. Section 4 discusses the implications of our findings. Finally, Section 5 concludes the paper.

II. Research Setting and Method

The current study examined 15 web services targeted at or often used by vulnerable groups. Many of the selected services are offered by the Finnish government and cities, but also by different organizations and private companies. These services included among others immigration services, tax administration, public transport websites, mental health related web services, web portals for sexual minorities, as well as a private medical center and online pharmacy. The list of the services is shown in Table I. Aside from the websites of a couple of public

This research has been funded by Academy of Finland project 327397, IDA – Intimacy in Data-Driven Culture.

sector websites, we have opted not to identify the exact services. Instead, the type of service has been described.

The experiment involved testing web services by utilizing their main features and visiting key pages[1]. All cookies were accepted upon accessing the studied websites. Network traffic was recorded using Google Chrome's Developer Tools, which allows for examination of web page source code and network activity. When recording network traffic, caching was turned off and only the requests going to third parties were inspected. The cache was recorded traffic Figure 1 shows an example view of network traffic recorded with Chrome Developer Tools. The recorded web requests were saved as a log file and analyzed. Any data that could be used to identify the user and was sent to third-party analytics services was extracted from the log file. In particular, sensitive personal data was searched.

Finally, it is important to briefly define the term "personal data". In this context, we adopt the definition used by GDPR and the Finnish Office of the Data Protection Ombudsman: personal data is "all data related to an identified or identifiable person"[2] [3]. Examples include IP addresses, location data, and ID numbers used by analytics services to track users. It is important to note that data items that can be combined together to identify a person are also considered personal data. For example, screen size, while not directly an identifying piece of data, can be used in profiling a user. GDPR also defines certain "special categories of personal data" such as data on ethnic origin, health-related data and data concerning a person's sexual orientation[4]. The processing of this type of personal data is generally prohibited under GDPR, unless certain conditions are met, such as obtaining an explicit consent from the individual. Additionally, organizations must implement appropriate technical and organizational measures to protect this type of personal data from unauthorized access and processing.

III. Results

Figure 2 shows the third parties found in our network traffic analysis. Google and Facebook (Meta) are the most frequent third-party services, followed by LinkedIn and React&Share. Despite the legal concerns for the use of Google Analytics in EU [10], Google's services are still widespread even on websites frequently used by vulnerable groups, many of which are essential services. On the other hand, many public sector bodies such as Tax Administration and Digital and Population Data Services Agency have been careful to choose third-party companies based in Finland. On average, there were 2.6 third-party analytics services per website. This number can be considered quite high, as we are talking about web services handling sensitive data of vulnerable groups.

The data shared with third parties includes details like IP addresses, device and user identifiers, User-Agent headers with details about the OS and browser, and other technical data like screen resolution. The device's IP address, which is included in every web request, is a key piece of information used to identify the user [11], [12]. When contextual data, such as information about the page the user is visiting or the action they are taking, is connected with the previously discussed identifying pieces of information, privacy is compromised.

Table I lists the studied websites and presents the related main findings. The studied public sector websites did not contain a large number of third-party analytics services, but it is clear that they still cause some privacy issues. A *Finland information website* had 3 different third-party services such as Google Analytics. As the website consists of a large number of information pages about various topics and offers these pages in several languages (e.g. for refugees and migrants), Google can easily track the topics a user reads and the language they choose.

On the website of *Finnish Immigration Services*, there are 2 third-party analytics services that tracked the pages the user visits. This may be considered highly problematic, as delicate information such as the user accessing information about seeking asylum in Finland is revealed to third parties. Similarly, *Tax Administration* as well as *Digital and Population Data Services Agency* made use of third party analytics, and page visits are tracked. For example, the fact the user is accessing information about a specific tax is revealed to a third party.

Public transit service websites of two Finnish cities were also studied. One of these services contained 7 third party analytics services, and sent search terms input by the user to 5 different third parties. Search terms can include for example destinations the user is intending to reach. The second public transit service website only tracked page visits.

We also analyzed two websites frequently used by sexual minorities. The first one of these is a website of a *sexual minority rights organization*. The subpages a user visits were tracked by three third-party analytics services, including for example pages where the visitor can choose to join different associations related to sexual minorities. Also, the web form used to join organizations is located in an external service that has Google Analytics. This means the user's intention to join a specific association can be revealed. The second website was a *sexual minority web portal*. Visited pages are tracked on this website, including for example a page where users register to a discussion forum and threads on the discussion forum (a thread topic is a part of a URL which was sent to Google).

The most glaring privacy issues were found in web services provided by a private *medical center* and *online pharmacy*. When booking an appointment to a doctor

[1]We have thoroughly documented the testing sequence of each web service. This documentation, as well as a list of the tested web services, can be provided upon request.

[2]https://gdpr.eu/eu-gdpr-personal-data/

[3]https://tietosuoja.fi/en/what-is-personal-data

[4]https://tietosuoja.fi/en/processing-of-special-categories-of-personal-data

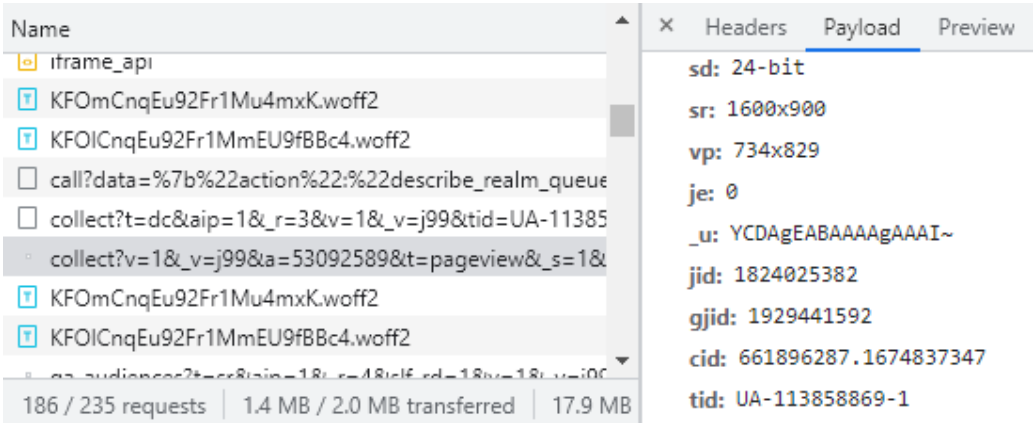


Fig. 1: A sample view of Chrome Developer Tools. Network traffic with HTTP requests and a payload with some pieces of identifying data is shown.

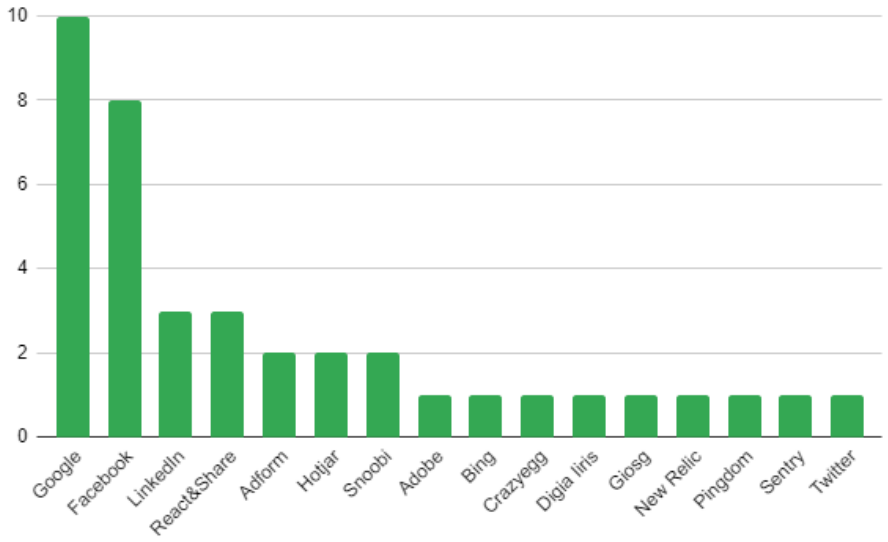


Fig. 2: The third parties found in the network traffic analysis.

on a website of the studied medical center, information about the intended appointment, the doctor's field of specialization (e.g. neurology), as well as the location and date of the appointment were leaked to 4 different third-party analytics services. There were also various sensitive information pages (such as a page about depression), where 5 different analytics services tracked user visits. Altogether, 8 third-party services were present on this medical center website. Possibly even worse, the online pharmacy we studied leaked the information about the prescription medicine the customer intended to order to 3 different third-party services! In addition to this, the medicines the user views were leaked as well. The pharmacy website made use of 4 third parties in total. In both of these cases, sensitive information about a patient's medical conditions can be deduced. It is very likely these privacy leaks are not intended by website developers. At the same time, the users of the websites also remain oblivious to their sensitive health data being leaked.

A *student health services* website had one analytics service tracking page visits. As in other cases in this study, this can be a problem with pages dealing with specific health related topics, such as a peer support group for depressed students. Similarly, a website of a *therapy center* had 2 third-party analytics services recording visited pages – for example, a page for a depression test. A *mental health chat* we studied used Google Analytics. The service was implemented as a single-page application, so it did not track separate page views, but a visit on the website is still reported. On the other hand, a *public sector social and healthcare service* portal was the only website with no third-party analytics in the current study.

We also studied an online service of a *violence victim shelter*, which contains information for violence victims. The website leaked visited pages to 3 different analytics services. Curiously, even when the user interacts with an information page by clicking it to open small sections on different topics, Google receives the information. This gives Google very fine-grained information on different issues and questions a potential victim of violence is

TABLE I: The studied websites and main findings

Description	# of 3rd parties	Main findings
Digital and Population Data Services Agency	2	Tracking of visited pages
Finland information website	3	Tracking of visited pages
Finnish Immigration Service	2	Tracking of visited pages (e.g. instructions on seeking asylum)
Medical center	8	Information about the appointment (date, location, doctor's specialization) leaks
Mental health chat	1	Tracking of visited pages
Online pharmacy	4	Information about the prescription medicine leaks
Public transport service 1	7	Information about searches (e.g. bus destinations) leaks
Public transport service 2	1	Tracking of visited pages
Sexual minority rights organization	3	Tracking of visited pages (e.g. registration pages and forms of different related associations)
Sexual minority web portal	2	Tracking of visited pages (e.g. forum registration page)
Social and healthcare service portal	0	(no third party tracking)
Student health services	1	Tracking of visited pages (e.g. depression peer support)
Tax Administration	1	Tracking of visited pages
Therapy center	2	Tracking of visited pages
Violence victim shelter website	3	Tracking of visited pages and opened of page sections

thinking. The website takes the user's safety into account e.g. by offering an "emergency exit" button for quickly leaving the page and providing instructions on how to clear cache and browsing history. This is good, but the privacy issues seem to be forgotten at least when it comes to third-party analytics. As violence victims may be agitated and in a hurry when visiting the website, it is worth asking whether the use of analytics services is appropriate on such a website. Although visitors may accept the cookies upon arrival, they probably are not in the right state of mind to think about analytics services or assume information on visited pages and opened text sections may leak.

IV. Discussion

We have seen that several websites used by vulnerable groups leak data to third parties, and in many cases this data is highly sensitive. The key findings of the current study are the presence of third-party analytics even on web services provided by the government and leaks of highly sensitive health data in web-based private healthcare services.

Several vulnerable groups may have insufficient digital skills and knowledge to make use of web services in an effective and secure way and understand the possible privacy risks involved [13]–[15]. Problems such as low literacy skills, cognitive decline or sensory impairments can further make using web services and understanding texts such as consent banners and privacy policies more difficult. The vulnerable groups that often have the greatest need for the digital services are also the ones who are likely to suffer most harm when their privacy is compromised [16]. While traditional face-to-face services will continue to be important to vulnerable groups, it is also essential to improve the existing digital services so that they are accessible and secure to use for everyone. Following the privacy-by-design approach when developing the systems and informing the users about third parties in the web service in a clear and transparent manner plays a big role in this development.

Web application developers should be aware of analytics services used in their websites, but different platforms and frameworks used in web development make it very easy to turn on web analytics without much consideration. Therefore, it is important for developers to better analyze what kind of data flows out of their application to third parties, and the use of each analytics service should be well justified. Network traffic analysis similar to this study can be used to observe the outgoing data transmissions.

Web developers may also not fully understand the types of personal data and implications of using analytics services. While the collected data can be used to improve the website's user experience, it is also used for profiling by large analytics companies. As more information is collected over time, analytics services are able to create comprehensive profiles on users and their behavior. In recent years, the use of browser and device information for user profiling on the web has become increasingly sophisticated [17].

It is unclear if the analytics providers actually store and utilize the health data they receive, but it is unacceptable for this information to be shared in the first place. However, the leaked sensitive data is only sent to analytics service providers who may not have a reason to use it, and the data is unlikely to enter the open data market. Utilizing the data effectively may also require manual effort and understanding of the specific website's implementation.

Our results, and the findings concerning healthcare

related web services in particular, give reason for further study of similar services in Finland and elsewhere in the world. Aside from the scientific aspirations, we also think that it is very important to quickly improve the privacy of the studied websites. To this end, we have informed the maintainers of the web services with the worst information leaks.

V. Conclusion

In this study, we investigated data breaches on Finnish websites used by vulnerable groups. Out of 15 websites studied, 14 leaked personal data to third parties. The most alarming cases were found in the healthcare sector, where the web services leaked details on intended appointment bookings and prescription medicine orders. Our results call for further investigation using a larger sample of websites used by vulnerable groups. In particular, our findings suggest the confidentiality of web-based services in the healthcare sector requires further study. Furthermore, examining similar web services in other countries would be beneficial. Additionally, future research could explore data breaches in more depth by experimenting and comparing different consent options on the studied websites.

Our results serve as a cautionary tale to software developers and data protection officers in charge of web services handling sensitive data. It is crucial for service providers to be aware of their responsibility in safeguarding customer privacy in areas where they are most susceptible, including increased awareness and control over the chosen third-party services. The use of multiple third-party analytics services on websites used by vulnerable groups is difficult to justify and undermines users' trust. Vulnerable groups should be able to trust the privacy of online services as much as the corresponding on-site services.

References

[1] A. L. Culén and M. Van Der Velden, "The digital life of vulnerable users: designing with children, patients, and elderly," in *Nordic Contributions in IS Research: 4th Scandinavian Conference on Information Systems, SCIS 2013, Oslo, Norway, August 11-14, 2013. Proceedings 4*. Springer, 2013, pp. 53–71.

[2] N. W. Eyrich, J. J. Andino, and D. P. Fessell, "Bridging the digital divide to avoid leaving the most vulnerable behind," *JAMA surgery*, vol. 156, no. 8, pp. 703–704, 2021.

[3] M.-A. Choukou, D. C. Sanchez-Ramirez, M. Pol, M. Uddin, C. Monnin, and S. Syed-Abdul, "Covid-19 infodemic and digital health literacy in vulnerable populations: A scoping review," *Digital Health*, vol. 8, p. 20552076221076927, 2022.

[4] B. L. Chang, S. Bakken, S. S. Brown, T. K. Houston, G. L. Kreps, R. Kukafka, C. Safran, and P. Z. Stavri, "Bridging the digital divide: reaching vulnerable populations," *Journal of the American Medical Informatics Association*, vol. 11, no. 6, pp. 448–457, 2004.

[5] S. O'Sullivan and C. Walker, "From the interpersonal to the internet: social service digitisation and the implications for vulnerable individuals and communities," *Australian Journal of Political Science*, vol. 53, no. 4, pp. 490–507, 2018.

[6] Y. M. Baek, E.-m. Kim, and Y. Bae, "My privacy is okay, but theirs is endangered: Why comparative optimism matters in online privacy concerns," *Computers in Human Behavior*, vol. 31, pp. 48–56, 2014.

[7] A. R. Zheutlin, J. D. Niforatos, and J. B. Sussman, "Data-tracking on government, non-profit, and commercial health-related websites," *Journal of general internal medicine*, pp. 1–3, 2021.

[8] M. Huo, M. Bland, and K. Levchenko, "All eyes on me: Inside third party trackers' exfiltration of phi from healthcare providers' online systems," in *Proceedings of the 21st Workshop on Privacy in the Electronic Society*, 2022, pp. 197–211.

[9] A. R. Zheutlin, J. D. Niforatos, and J. B. Sussman, "Data-tracking among digital pharmacies," *Annals of Pharmacotherapy*, vol. 56, no. 8, pp. 958–962, 2022.

[10] S. Winklbauer and R. Horner, "Austrian DPA Decides EU-US Data Transfer through the use of Google Analytics to Be Unlawful," *European Data Protection Law Review*, vol. 8, p. 78, 2022.

[11] V. Mishra, P. Laperdrix, A. Vastel, W. Rudametkin, R. Rouvoy, and M. Lopatka, "Don't count me out: On the relevance of ip address in the tracking ecosystem," in *Proceedings of The Web Conference 2020*, 2020, pp. 808–815.

[12] T. Heino, R. Carlsson, S. Rauti, and V. Leppänen, "Assessing discrepancies between network traffic and privacy policies of public sector web services," in *Proceedings of the 17th International Conference on Availability, Reliability and Security*, 2022, pp. 1–6.

[13] A.-M. Kaihlanen, L. Virtanen, U. Buchert, N. Safarov, P. Valkonen, L. Hietapakka, I. Hörhammer, S. Kujala, A. Kouvonen, and T. Heponiemi, "Towards digital health equity-a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era," *BMC health services research*, vol. 22, no. 1, p. 188, 2022.

[14] S. Ranchordás, "Connected but still excluded? Digital exclusion beyond internet access," *The Cambridge Handbook of Life Sciences, Informative Technology and Human Rights (Cambridge University Press, 2021, Forthcoming), University of Groningen Faculty of Law Research Paper*, no. 40, 2020.

[15] T. Heponiemi, K. Gluschkoff, L. Leemann, K. Manderbacka, A.-M. Aalto, and H. Hyppönen, "Digital inequality in finland: access, skills and attitudes as social impact mediators," *New Media & Society*, p. 14614448211023007, 2021.

[16] N. McDonald and A. Forte, "Privacy and vulnerable populations," in *Modern Socio-Technical Perspectives on Privacy*. Springer International Publishing Cham, 2022, pp. 337–363.

[17] N. Kaur, S. Azam, K. Kannoorpatti, K. C. Yeo, and B. Shanmugam, "Browser fingerprinting as user tracking technology," in *2017 11th International Conference on Intelligent Systems and Control (ISCO)*. IEEE, 2017, pp. 103–111.

Privacy in Popular Children's Mobile Applications: A Network Traffic Analysis

Robin Carlsson*, Sampsa Rauti*, Samuli Laato†, Timi Heino*, Ville Leppänen*
* University of Turku, Department of Computing, Turku, Finland
† Tampere University, Gamification Group, Tampere, Finland
crcarl,sjprau,sadala,tdhein,ville.leppanen@utu.fi

***Abstract*—Children increasingly download and use mobile applications from marketplaces such as Apple's App Store or the Google Play Store. One would expect that applications intended for children are free of third-party analytics, or at least make sure parents give their consent for collecting personal data from children. In this study, we performed an in-depth technical analysis of a representative snapshot of 15 applications from Google Play Store aimed at children (age group classifications 0–5, 6–8 and 9–12). We recorded the network traffic of these applications and compared it to the privacy policies of the applications. Across the applications, a significant number (13/15) were delivering more information about the users to various third parties than what was admitted in the respective privacy policies. We elaborate on details regarding the observed network traffic, and discuss implications of these findings on strategies for preserving user privacy, particularly for sensitive audiences such as children.**

Keywords*—privacy, web services, data leaks, vulnerable groups

I. Introduction

In recent years, the use of mobile devices such as smartphones and tablets has become widespread. In addition to adult audiences, there are tens of thousands of games and educational applications aimed at children on app stores such as Google Play Store and Apple App Store [1]. Due to phenomena such as the abstraction of development tools, online connectivity and the availability of online services, increasingly many applications, including applications aimed at children, now operate online - with multiple features such as advertisements, updates and multiplayer elements operating through online services [2]. These developments have led to a situation where there is a huge amount of network traffic being generated by mobile devices, and accurately understanding and classifying this traffic has become a challenge [2].

Several analytics companies such as Alphabet/Google, capture and record users' digital fingerprints [3], [4]. Personal data items such as device and user specific identifiers and gaming behavior are used to better understand the user base, more optimally target ads [5] or to improve the usability and performance of applications. However, in addition to this data being collected for developers, it is also leaked to third-party companies [6], [7]. Unfortunately, all of this tracking is usually not visible for users, and children in particular can be considered a group vulnerable to privacy violations [8].

This situation is concerning from the perspective of end user privacy. The GDPR recognizes the need for extra protection for children's personal data and sets specific requirements for its processing. The regulation requires that parents or guardians provide their consent for the processing of children's data, and that application developers implement appropriate measures to ensure that children's rights to privacy are respected. However, past work has showcased that parents and children alike may not understand digital privacy very well [9], and that application developers and app stores do not help by following privacy regulations poorly [4]. Therefore, it is critical to examine the network traffic of children's applications on app marketplaces for understanding the magnitude of the problem of privacy violations in children's mobile applications.

Former studies have explored the privacy of children's applications at a large scale, often by automatic methods. Binns et al. [10] used static analysis to study apps from Google Play stores. They discovered that apps aimed at children had a high number of third-party trackers. Reyes et al. [11] conducted a study of network traffic from the most popular free Android children's applications and discovered that most of the applications had potential violations of Children's Online Privacy Protection Act. The aim of this study is to take a closer look at the privacy of children's applications. Therefore, we carried out an in-depth investigation of a representative snapshot of mobile applications aimed for children. This included analyzing the applications' network traffic and studying their privacy policies and Google Play Store's data protection section to see how transparently the personal data leaked to third parties is reported. We also tested whether the applications request consent from parents before a child can use them.

In doing so, this study contributes important practical information on the status quo of end user privacy in available children's applications in popular app marketplaces. Our findings also have contributions to academic literature on privacy, as we discuss measures for improving user privacy from a software engineering point of view based on our findings.

The rest of the paper is structured as follows. Section II outlines our research approach. Section III presents the

This research has been funded by Academy of Finland project 327397, IDA – Intimacy in Data-Driven Culture.

results of the network traffic analysis, and compares the actual traffic to Google Play's data protection sections and privacy policy documents. Section IV summarizes our main findings, discusses implications for software development and addresses limitations of the current study. Finally, Section V concludes the paper.

II. Research approach

Methodologically our work followed a three-step approach. First, we searched for a representative snapshot of children's applications on the popular app marketplace of Google Play Store. Second, using an intermediary Linux computer between the mobile phone and the internet we recorded all network traffic passing between the surveyed applications and any third-party servers. Third, we analyzed the data from two perspectives: a descriptive look of what data is sent and where; and a comparison between the applications' Google Play data protection sections, privacy policies and observed data traffic. Next we elaborate on these three steps in more detail.

A. The application search process

We selected Google Play Store as the marketplace for searching for applications primarily for two reasons. First, it is a hugely popular marketplace with over 100 billion annual app downloads across the globe [12]. Second, it provides a specific section of applications aimed at children, and search tools for obtaining these applications. Previous research looking for applications in Google Play Store have used specific keywords and obtained the first results for analysis, with the justification that these applications are the ones average users are also most likely to find [1], and those studies looking at very specific applications have cut off the search when a new page of results did not yield any new matches [13]. In the case of this study, we had potential applications in the tens of thousands, meaning we also had to cut off the analysis at some point, and here we referred to the criteria of saturation: when the latest search results provided no additional insights, we could cut off the search. Hence, we commenced searching for children's applications on Play Store in summer 2022 from the "Kids" section. All apps that are accepted here have to go through a specific review process, and Google advertises that these are accepted by teachers [1].

In order to ensure we obtained a representative snapshot, we started reviewing applications from three age groups: 0–5, 6–8 and 9–12. Due to the authors' location, we focused on apps aimed at Finnish children, thus setting the preferred language to Finnish and English. Subsequently, without any further sorting, we downloaded the first applications from each of the three categories. We continued this process until the researchers were confident that we had reached a sufficient saturation in terms of the number of applications. Saturation here was measured so that once we had gone through two rounds of applications without any further significant findings from the network traffic analysis, we determined that observing any further applications would provide diminishingly small added value. Following this criteria, in the end we went through 15 applications, 5 for each age group category.

B. Recording network traffic

In order to record all mobile traffic, we set up an Android phone that was connected to the internet through a computer, which acted as a wireless access point. This Linux computer was used to record the phone's network traffic. Because there are some pre-installed applications related to the phone's operating system which generate network traffic, the system's background noise produced by these applications was recorded and removed from application-specific recordings.

On the Linux computer, mitmproxy and tcpdump tools were used to record the traffic. Mitmproxy is a free proxy tool that allows intercepting and inspecting network traffic by acting as a man-in-the-middle between the client and the server. Tcpdump was used to display and analyze contents of network packets transmitted through the access point. A traffic log file was generated for each studied mobile application, and the files were analyzed to find any personal data that can be used to identify a user. In addition to recording these data items, the third parties receiving personal data from applications were also listed. While recording the traffic, the main functionality of each children's application was actively tested for a few minutes[2]. In Figure 1, a sample snippet of captured mobile application traffic is shown.

C. Data analysis

Once we had the network traffic recorded we proceeded with the analysis. As we were interested in whether any personal data was being shared by these applications, it is important to clarify what it means. Both GDPR and Finnish office of the data protection ombudsman define personal data as "all data related to an identified or identifiable person"[3]. Therefore, data items such as a user's IP address, or accurate data revealing the user's location are chiefly considered personal data. However, it is also important to note that a combination of several pieces of data can often be used to identify a person and can constitute personal data. For instance, although a device's screen size as a technical data item does not directly identify a specific device or user, it can definitely be a very useful piece of information in profiling a user when combined with other technical data items.

First, we looked into whether any of such personal data was being shared, and where. We recorded all instances

[1]For more information, see Google's description of the Kids section here: https://play.google.com/store/apps/category/FAMILY?hl=en&gl=US &pli=1, visited February 1, 2023

[2]While the tested functionalities could not be included in this paper due to space constraints, detailed descriptions of how each application was tested are available upon request.

[3]See https://gdpr.eu/eu-gdpr-personal-data/ and https://tietosuoja.fi/en/what-is-personal-data, visited February 1, 2023



Fig. 1: A sample view from the mitmproxy tool. Network traffic with HTTP requests is shown.

of data being shared, and did some investigation on the domains to whom it was being sent, trying to identify the company behind the servers, as well as the probable cause for the traffic (e.g. Google Analytics, an elaborate data collection scheme or something else). Second, we read the privacy policies of all 15 applications and compared the observed network traffic to the data collection disclosed in these texts. Third, we then also compared these two to the Google Play Data Protection section, where developers must disclose whether any data in their application is being sent to third parties.

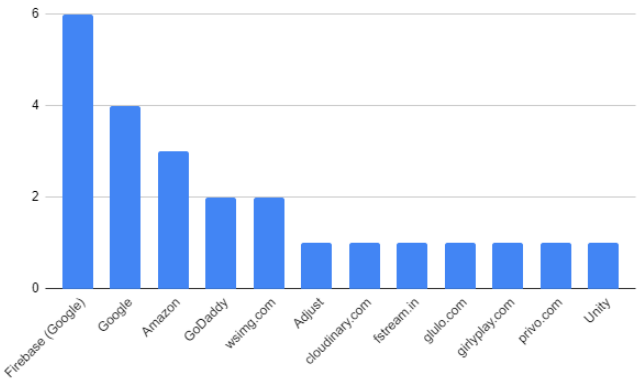


Fig. 2: The third-party analytics services found in the studied applications and number of occurrences (counted once per application).

III. Results

A. Findings from the network traffic analysis

The third parties receiving traffic from the analyzed applications are shown in Figure 2. Firebase, an app development platform provided by Google, is the most frequent receiver of personal data, followed by Google and Amazon.

The personal data items the studied applications sent to third parties are shown in Table I. The most frequently leaked pieces of personal data were the device's IP address and information on the phone's brand and OS. While the IP address is always sent out with every network connection an application forms, two of the analyzed applications did not connect to any third parties. Therefore,

TABLE I: The personal data items sent to third-party services by the studied mobile applications.

Sent data	Number of applications	Percentage
IP address	13	86.7
Phone brand and model	13	86.7
Phone OS	13	86.7
Phone OS version	13	86.7
Timestamp	5	33.3
Screen size	2	13.3
Processor	2	13.3
Is the phone rooted?	1	6.7
App store	1	6.7
Internet service provider	1	6.7
User identifier	1	6.7
Device identifier	2	13.3
Installation identifier	6	40.0
Instance identifier	4	26.7
Session identifier	1	6.7
Other unknown identifier	6	40.0
Timezone	2	13.3
Country	4	26.7
Language	4	26.7
Child's first name	1	6.7
Email address	1	6.7
Use of camera	1	6.7
Use of microphone	1	6.7

the IP address, as well as information about the phone and OS were leaked in 13 out of 15 cases.

While the device IP address is sent to all third parties a mobile application communicates with, many users may not know that a unique address, which can be used to identify them, is leaked to third parties and analytics services. Although IP addresses of home users are usually dynamic, it is common that the same IP address remains in the possession of the same device and user for time periods exceeding one month [14]. Moreover, large global analytics companies can be expected to be able to effectively connect even dynamic IP addresses to specific users.

Other frequently sent data items include numerous technical details such as screen size and processor, revealing additional information which can be used to profile users. In one case, even a sensitive piece of information on whether the phone has been rooted was leaked. We suspect that this is part of the functionality of the development platform and not a feature developers have intended to add. The findings also included leaked contextual information such as the app store the application was downloaded

from, user's country, language, timezone and internet service provider.

The technical and contextual data items mentioned before cannot be used to identify a person by themselves (except the IP address), but when combined together, they can be very useful when generating a profile for a specific user. This is especially the case when large analytics companies combine pieces of data from different mobile applications and websites.

Then again, several device or user specific identifiers shown in the list are usually very personal. Our findings also include many unknown identifiers. These are long strings that are probably used to identify users, devices or sessions. Unknown identifiers were sent out by 40% of the analyzed applications. One application also sent out the child's first name and email address, although this information was explicitly requested from the user on registration. Finally, one of the tested applications was also found to transmit data about the user's actions, revealing the use of camera and microphone.

B. Findings from the analysis of privacy policies and data protection sections

Table II shows whether the applications ask for parental consent and how they announce their target groups. We can see that none of the studied applications ask for a consent for collecting data from a parent or guardian. Still, 13 applications collected personal data, and almost all of the applications indicated (in the app name, on their product page, in the application itself or on a separate linked resource such as privacy policy) that the application is targeted for children. Only one of the applications, Indian Royal Wedding Game, did not clearly indicate that the application was targeted for children. In all cases, however, the target audience is quite clear from the Google Play age group as well as the contents of the application.

Only two of the studied applications, "Drawing Games: Draw and Color for Kids" and "Kids Carwash Service Auto Workshop" did not collect any data. It is quite clear that other applications, by collecting data on children while not requiring a consent, are violating GDPR.

Table III shows how the studied privacy policies and the data protection section in Google Play correspond to actual network traffic containing personal data. We can see that in several (5) cases, the data protection section in Google Play Store does not correspond to the actual network traffic. There are 4 applications that claim they do not collect personal data or deliver it to third parties despite doing so. Interestingly, there is also one application that reports collecting personal data but did not do so in our experiments.

On the other hand, we can see privacy policy documents usually mentioned that data was being collected and sent to third parties – although they did not always list the transmitted personal data items or receiving parties in sufficient detail. Also, several applications (4) did not have proper privacy policies at all, showing quite obvious disregard for privacy regulations. Privacy policies of two applications – "Tonka: Trucks Around Town" and "Color by Numbers: Cars" – claimed not to collect any user information at all, but at least IP addresses could still be collected by the developer and the associated third-party server.

Interestingly, developers of two applications, "Kids Art & Drawing Game" and "Indian Royal Wedding Game", seemingly try to wriggle out of their responsibilities by including statements such as "our services are not intended for use by children under the age of sixteen (16)" and "these services do not address anyone under the age of 13" in their privacy policies. It is obvious that this strongly contradicts the age group and the contents of the applications.

IV. Discussion

A. Key findings

The key findings of our study can be summarized as follows:

- Of the 15 studied applications, 13 sent out personal data to third parties, identifying the child using the application. This information can also be used to profile the user and their behavior, especially when other applications send information to the same third party, enabling them to build a more comprehensive picture of what applications the user uses and how. Because the user is a child, implications can be even more serious.
- None of the 13 applications sending personal data to third parties asked for parental consent. This violates the GDPR.
- When it comes to personal data delivered to third parties, there were significant discrepancies between 1) the studied application's data safety sections in Google Play, 2) the associated privacy policies, and 3) the recorded network traffic.
- Some application developers may try to escape from their responsibility by claiming in their privacy policy documents that their application is not intended for children. This blatantly contradicts the age range given for the applications in Play Store.

B. Implications for software development

It is apparent that the third-party libraries and platforms developers have chosen to use when building their application are a significant factor in what kind of personal data and technical details are sent out to third parties. Oftentimes, it seems developers and data protection officers do not pay sufficient attention to this side effect of using third-party libraries and analytics services, and the task of finding out how the used external services handle personal data is left to the user of the application.

Software companies and developers are, however, responsible for reporting the personal data items and the parties they are delivered to in the application's privacy

TABLE II: Consent and target groups of the applications.

			Does the developer indicate that the app is for children?			
Age group	Application	Parental consent	In app name	On product page	In the app	Linked file/website
0–5	Kids Art & Drawing Game	No	X	X		
	Coloring book - games for kids	No	X	X	X	X
	Drawing Games: Draw and Color for Kids	No	X	X		X
	Indian Royal Wedding Game	No				
	Paint for Kids	No	X	X		
6–8	Daddys Hair Salon	No				
	Jigsaw Puzzles: Games for Kids	No	X	X	X	X
	Kids Carwash Service Auto Workshop: Fun Game	No	X	X		
	Miga Town: My TV Shows	No				X
	Toca Hair Salon 4	No			X	X
9–12	ChatterPix Kids by Duck Duck Moose	No No	X		X	
	Color by Numbers: Cars	No		X	X	X
	Kids Coloring Book	No	X	X	X	X
	Montessori Nature	No		X	X	X
	Tonka: Trucks Around Town	No		X	X	X

TABLE III: Transparency of privacy policies and data protection section compared to actual network traffic.

Application	Sends personal data to 3rd parties	According to the Google Play Data Protection section, is data sent to 3rd parties	According to the privacy policy, is data sent to 3rd parties
Kids Art & Drawing Game	Yes	No	Yes
Coloring book - games for kids	Yes	No	N/A
Drawing games: Draw and Color for kids	No	No	No
Indian Royal Wedding Game	Yes	Yes	Yes
Paint for kids	Yes	No	Yes
Daddys hair salon	Yes	No	N/A
Jigsaw Puzzles: Games for Kids	Yes	Yes	Yes
Kids Carwash Service Auto Workshop: Fun Game	No	Yes	N/A
Miga Town: My TV Shows	Yes	Yes	Yes
Toca Hair Salon 4	Yes	Yes	Yes
Chatterpix Kids by Duck Duck Moose	Yes	Yes	Yes
Color by Numbers: Cars	(developer uses a 3rd party server)	N/A	No
Kids Coloring Book	Yes	Yes	Yes
Montessori Nature	Yes	N/A	N/A
Tonka: Trucks Around Town	(developer uses a 3rd party server)	N/A	No

policy document. That is why analyzing applications' network traffic should always be an integral part of the software development and testing process. This will help the developers to monitor what kind of personal data their mobile applications leak out to third parties. This knowledge is necessary to compile realistic and accurate data safety sections and privacy policy documents.

Privacy policy documents could also do much better when informing users about potential consequences of transferring information to a third party. This does not mean scaring users with unrealistic doomsday scenarios. It would be beneficial, however, to lay out some examples of how technical data items sent to analytics services can be used to profile children, for instance. In general, a potential method to improve privacy policy documents could be introducing a small number of different templates or structures that privacy policy documents would have to follow. Following a standardized format helps to include the necessary information (such as personal data items and their destinations) sufficient detail makes privacy policies easier to write and understand [15].

The discrepancies between network traffic, data safety sections and privacy policies may also speak of a gap in understanding between software developers and data officers. It is not only important to understand what kind of data the application sends out but also have a common appreciation of what constitutes personal data. With this renewed understanding, it is also worth reading the privacy policies of the used third-party libraries and reconsidering whether it is necessary to include analytics services in a mobile application aimed for children.

Google Play Store should also take steps to better ensure applications targeted for children are GDPR compliant. Applications that send personal data to third parties without consent should be disallowed. Also, when developers of applications clearly intended for kids declare their applications are designed for adults, Google Play Store seems to simply look the other way. Because the tech giants behind the app stores have the real power over the app economy, they should enforce stricter privacy.

C. Limitations

There are few limitations in the current study. First, the network traffic analysis we performed only reveals clearly detectable personal data mobile applications transmit. The analyzed traffic may contain some data items that have been obscured and purposefully made difficult to uncover to avoid detection or hide implementation details.

Second, while we can be certain that specific personal data items have been delivered to a specific destination, we cannot say what happens to this data (e.g. an IP address) once it reaches a third party server. The data can either be stored and possibly made use of, or immediately deleted. The network traffic analysis carried out in this study only covers the client side. Our analysis also is not exhaustive in terms of functionality available in the selected applications, despite our best efforts to cover the most essential functions.

Third, we only looked at applications in the Google Play Store, and it remains unclear whether similar findings would emerge in other Android app marketplaces such as Samsung Galaxy Store, and further in app stores for alternative operating systems, primarily the iOS App Store. Thus, so far our findings merely speak about the status quo of network traffic in children's apps in the Kids section of Google Play Store, and should be understood as such. However, the findings do raise broader and more critical questions regarding modern ecosystems and user privacy.

V. Conclusion

In this study, we demonstrated via network traffic analysis that several applications targeted for children, available in the world's most popular application marketplace, send data to third parties without parents' content, which violates the GDPR. We also found that there are discrepancies in what application developers disclose as part of the Google Play Data Protection section and the applications' privacy policies. Our network traffic analysis further showed that this information also does not align with what data is actually being sent and where, primarily in that more data was being shared than what was officially disclosed.

These findings have significant implications for both practice and research. On the practical side, users, and particularly the parents and guardians of children, should be aware that the applications they use and trust with their personal information, and the information of their offspring, may not be as privacy-friendly as they believe. Our work suggests that even if the parents and guardians read the data protection sections and privacy policies carefully, they may not get an accurate picture of what the applications really do.

On the research side, this study opens the door for further investigation into the extent of data leakage in applications and the impact it has on user privacy. It also highlights the need for further regulations and guidelines to ensure that application developers are transparent about their data collection practices and that user data is protected. Furthermore, as developer tools operate at higher and higher abstraction levels, it may be increasingly difficult for developers to keep track of whether there are some unintended functionalities embedded in some of the blocks of code they use, in our case, unintended network traffic. Overall, this study sheds light on the urgent need for greater awareness and action on the issue of data privacy in the age of digital applications, particularly for audiences in vulnerable positions, such as children.

References

[1] S. Laato, R. Lindberg, T. H. Laine, P. Bui, B. Brezovszky, L. Koivunen, O. De Troyer, and E. Lehtinen, "Evaluation of the pedagogical quality of mobile math games in app marketplaces," in *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE, 2020, pp. 1–8.

[2] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Multi-classification approaches for classifying mobile app traffic," *Journal of Network and Computer Applications*, vol. 103, pp. 131–145, 2018.

[3] I. Ullah, R. Boreli, and S. S. Kanhere, "Privacy in targeted advertising on mobile devices: a survey," *International Journal of Information Security*, pp. 1–32, 2022.

[4] F. Zhao, S. Egelman, H. M. Weeks, N. Kaciroti, A. L. Miller, and J. S. Radesky, "Data collection practices of mobile applications played by preschool-aged children," *JAMA pediatrics*, vol. 174, no. 12, pp. e203 345–e203 345, 2020.

[5] M. Shumanov, H. Cooper, and M. Ewing, "Using ai predicted personality to enhance advertising effectiveness," *European Journal of Marketing*, vol. 56, no. 6, pp. 1590–1609, 2022.

[6] T. Chen, I. Ullah, M. A. Kaafar, and R. Boreli, "Information leakage through mobile analytics services," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*, 2014, pp. 1–6.

[7] S. Rauti and S. Laato, "Location-based games as interfaces for collecting user data," in *Trends and Innovations in Information Systems and Technologies: Volume 2 8*. Springer, 2020, pp. 631–642.

[8] M. Liu, H. Wang, Y. Guo, and J. Hong, "Identifying and analyzing the privacy of apps for kids," in *Proceedings of the 17th international workshop on mobile computing systems and applications*, 2016, pp. 105–110.

[9] S. Livingstone, M. Stoilova, and R. Nandagiri, "Children's data and privacy online: growing up in a digital age: an evidence review," 2019.

[10] R. Binns, U. Lyngs, M. Van Kleek, J. Zhao, T. Libert, and N. Shadbolt, "Third party tracking in the mobile ecosystem," in *Proceedings of the 10th ACM Conference on Web Science*, 2018, pp. 23–31.

[11] I. Reyes, P. Wijesekera, J. Reardon, A. Elazari Bar On, A. Razaghpanah, N. Vallina-Rodriguez, S. Egelman *et al.*, ""won't somebody think of the children?" examining coppa compliance at scale," in *The 18th Privacy Enhancing Technologies Symposium (PETS 2018)*, 2018.

[12] Statista, "Annual number of app downloads from the google play store worldwide from 2016 to 2021," *ONLINE, available at: https://www.statista.com/statistics/734332/google-play-app-installs-per-year/*, vol. 1, 2021.

[13] D. Fernández Galeote and J. Hamari, "Game-based climate change engagement: analyzing the potential of entertainment and serious games," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CHI PLAY, pp. 1–21, 2021.

[14] V. Mishra, P. Laperdrix, A. Vastel, W. Rudametkin, R. Rouvoy, and M. Lopatka, "Don't count me out: On the relevance of ip address in the tracking ecosystem," in *Proceedings of The Web Conference 2020*, 2020, pp. 808–815.

[15] M. Rowan and J. Dehlinger, "A privacy policy comparison of health and fitness related mobile applications," *Procedia Computer Science*, vol. 37, pp. 348–355, 2014.

Analysis of DMARC Implementation in Republic of Croatia

Dražen Pranić
Atlantic Grupa d.d., Zagreb, Croatia
drazen.pranic@atlanticgrupa.com

***Abstract* - Many security reports and analyses continuously emphasize email service as the cause or starting point of numerous security incidents. The insecurity of email service results in a steady increase in financial fraud caused by compromising user accounts. The use of more advanced authentication protocols such as DMARC, significantly reduces the risk of these threats. The DMARC protocol and its role in improving email service authentication will be explained in detail. Numerous advantages of the DMARC protocol and implementation challenges will be highlighted. The extent to which DMARC is applied in Republic of Croatia will be shown by an analysis of DMARC implementation at most important companies per total income and relevant entities of public sector. Analysis will show how much are companies and public sector focused on security of email service and will give an insight into the DMARC complexity.**

Key words: email security, phishing, DMARC, authentication

I. INTRODUCTION

The insecurity of email services is a fact that is regularly pointed out by various security reports and analyses. Analyses of numerous security incidents indicate that one of the most common initial vectors of attacks is phishing email. Thus, according to one of the most respected security reports, the Verizon Data Breach Report, phishing email attacks are responsible for over 36% of the analysed incidents. Unfortunately, this devastating statistic has been unchanged for years. [1]

One of the world's largest insurance companies, AIG, said in its annual report: "The business email compromise of electronic mail service accounts (BEC) has overtaken ransomware as the main reason for claims." The report further explains that business compromise email service accounts caused by phishing attacks. [2]

The extent of this problem, i.e., its financial scale, is explained in more detail in the Report prepared by the US FBI on an annual basis. Namely, in the period from July 2016 to July 2019, the total reported damage is 26,2 billion $. The damage was caused by a total of 166,349 reported incidents. [3]

Analytics firm Gartner predicts that by 2023, e-mail service compromise attacks will double every year to more than $ 5 billion $ per year and will be a source of huge financial losses for many companies. Gartner also points out weaknesses in the authentication of e-mail services as one of the more important reasons why these simple attacks are still relevant. They recommend the implementation of advanced email authentication as one of the controls that significantly reduces the risk of such attacks. We will explain this in more detail in the following chapters of this paper. [4]

II. THE IMPORTANCE OF EMAIL AUTHENTICATION

Attacks aimed at compromising business email account are generally very simple. The most common form of such attacks is when an attacker changes, falsifies the "from" field of the e-mail message to deceive the recipient of the message.

A particularly effective method of attack is "CEO Fraud" in which attackers pretend to be the organization's management and persuade employees to enable payment to fake account or to transfer money from a business account without authorization. According to the Ministry of the Interior of the Republic of Croatia, this is one of the seven most used financial online frauds. [5]

Email authentication aims to reduce the risk of such and similar scams that exploit the insecurity of email services. The basic task of email authentication is to make the process of sending and receiving emails as secure as possible and significantly aggravate identity impersonation of the email sender.

Each sender’s email is analysed in detail by the incoming recipient email server. Criteria for analysing the message itself is based on the authentication method defined by the sender of the message. Based on this information, the recipient's email server defines whether to deliver, quarantine, or reject the received message.

This process is applied regardless of which type of authentication is used. When receiving the message, the recipient's server analyses certain data in the message and DNS records of the sender's domain. Based on the analysis of the authentication methods used, the message recipient's server decides on the authenticity of the sent message.

In the following chapters of this paper, we will explain the most important authentication methods:

- Sender Policy Framework (SPF),
- DomainKeys Identified Mail (DKIM),
- Domain Message Authentication Reporting and Conformance (DMARC).

III. SENDER POLICY FRAMEWORK

The Sender Policy Framework (SPF) is a method of email authentication whose purpose is detection of forgery of the sender's address during e-mail delivery. SPF allows authorized publishing hosts to send emails on behalf of a given domain. Publishing hosts IP addresses are defined in SPF DNS TXT record. SPF uses the Return-Path value to check the source email server.

Recipient email server is responsible for email validation. The decision is made by the check_host function described in RFC 7208 [6] that takes three arguments on input (IP address of the sender, the domain, the MAIL FROM or HELO identity) and returns one of the seven possible results: None, Neutral, Pass, Fail, SoftFail, TempError and PermError. Most important results are Pass and Fail. "Pass result means that the client is authorized to inject mail with the given identity while a Fail result is an explicit statement that the client is not authorized to use the domain in the given identity". [6] SPF email record check is performed as shown in Figure 1:

1. Receiving email server check a SPF record from the DNS server of the sending email server and analyses the list of IP addresses that are authorized to send emails.
2. If the sender's IP address satisfies the SPF check, then the receiving email server will continue to process the email.
3. If the SPF check is not successful, then the email will be rejected according to the receiving email server settings.

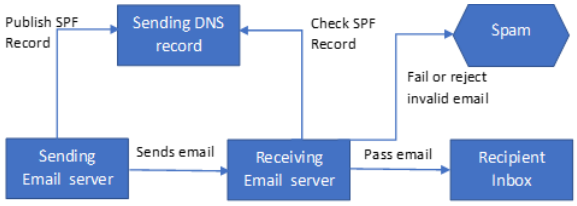


Figure 1. SPF record check

The biggest drawback of the SPF authentication method is that the From Field, which is only visible to the recipient of the email, is not checked. An attacker can specify his domain in the MAIL FROM or Return-Path value satisfy the SPF check and change the sender field and thus deceive the recipient of the email. Therefore, Sender Policy Framework can't help in email spoofing attacks. Unfortunately, many companies just rely on that basic level of email authentication.

IV. DOMAINKEYS IDENTIFIED EMAIL

DomainKeys Identified Mail (DKIM) permits a person, role, or organization that owns the signing domain to claim some responsibility for a message by associating the domain with the message. Assertion of responsibility is validated through a cryptographic signature and by querying the Signer's domain directly to retrieve the appropriate public key.[7]

The email server adds the digital signature of the message to the message itself using its private key. The sender's public key is available using the appropriate DKIM record in the sender's DNS domain. The DKIM DNS records are stored in the following format: selector._domainkey.domain. The domainkey is a fixed string, and the selector is a randomly chosen string by the domain owner. [8]

The recipient email server verifies the digital signature of the sender's public key message. It will detect the DKIM signature and look up the sending email server public DKIM key in DNS record. If the key is found, it can be used to decrypt the DKIM signature. This is then compared to the values retrieved from the received mail. If they match, the DKIM is valid. DKIM email record check is shown in Figure 2

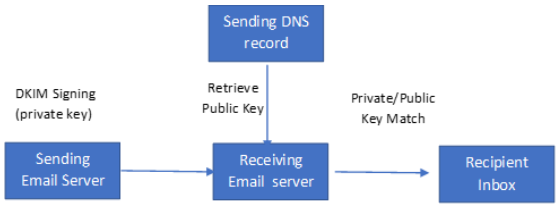


Figure 2. DKIM record check

Applying DKIM authentication prevents an attacker from intercepting an email message, changing it, and sending the changed email message to the recipient.

However, as with the SPF authentication method, DKIM does not check the From Field, which is only visible to the email recipient. An attacker can send a signed message from his DKIM domain and change the From Field of the message and thus deceive the recipient of the email.

V. DOMAIN MESSAGE AUTHENTICATION REPORTING AND CONFORMANCE

Domain-based Message Authentication, Reporting, and Conformance (DMARC) is a mechanism by which email operators leverage existing authentication and policy advertisement technologies to enable both message-stream feedback and enforcement of policies against unauthenticated email. [9]

It is based on the Sender Policy Framework and DomainKeys Identified Email protocols explained in the previous chapters. It solves the most important problem of email authentication, which is the unauthorized change of the From Field.

DMARC is based on the results of SPF and /or DKIM, and at least one of the above authentication methods is required for implementation. It is recommended to use both authentication methods to support the DMARC mechanism, due to the additional level of verification. In the case of using only one authentication method, the use of DKIM is recommended, since it provides a higher level of security and eliminates above mentioned security problems present with the SPF mechanism.

DMARC policies are published by Domain Owners and applied by Mail Receivers. A Domain Owner advertises DMARC participation of one or more of its domains by adding a DMARC DNS TXT record. [9]

DMARC records define how you handle messages that comply or not comply with this authentication method. There are three DMARC records / policies:

1. p = none
   Email traffic is only monitored. No additional activities are taken.
2. p = quarantine
   Unauthorized emails are quarantined or spammed by receiving email server.
3. p = reject
   Unauthorized messages are rejected by receiving email server.

When checking the compliance of an email message with the DMARC protocol, it is first checked whether the SPF and / or DKIM requirements are met. If they are then compliance with DMARC requirements is checked. Depending on the results of DMARC compliance and defined DMARC policies, emails are delivered to the recipient.

Consider an example DMARC TXT RR for the domain "sender.dmarcdomain.com" that reads "v=DMARC1;p=reject;pct=100;rua=mailto:postmaster@dmarcdomain.com". In this example, the sender requests that the receiver outright reject all non-aligned messages and send a report, in a specified aggregate format, about DMARC policy applied to a specified address. [10]

DMARC compliance addresses the biggest shortcoming of SPF and DKIM email authentication methods. Namely, DMARC connects the results of SPF and DKIM methods with the content of the e-mail message, i.e., with the content of the sender's field. The domain found in the email sender field is the part of the message that connects the entire DMARC processing.

Today, anyone can buy a domain and set up the appropriate SPF and DKIM records. Therefore, the result of the SPF and DKIM record analysis must be associated with the domain located in the email sender field. This concept is known as identifier alignment. [11] Identifier matching allows existing email authentication methods to become relevant to the content of the message itself. Matching all identifiers is the biggest challenge when implementing the DMARC protocol.

An important part of identifier matching relates to external service providers who communicate with other participants on behalf of the sender.

VI. DMARC REPORTING

DMARC protocol makes a major contribution in the field of reporting. Namely, without the DMARC protocol, it is not possible to answer the simple question: "Which email senders send messages on behalf of my organization"? Without the DMARC protocol, the answer to this question is not possible.

The DMARC protocol enables aggregate and detailed failure/forensic reports. Aggregate reports are sent automatically at a defined interval to the addresses specified in the "rua" parameter, by all email servers on the recipient's side that are set to the appropriate way.

The DMARC aggregate feedback report is designed to provide Domain Owners with precise insight into:

- authentication results,
- corrective action that needs to be taken by Domain Owners, and
- the effect of Domain Owner DMARC policy on email streams processed by Mail Receivers. [12]

Forensic reports defined in the "ruf" parameter, in addition to the above information, contain additional information such as the subject of the message, the header of the message and the URL of the links contained in the body of the message. Forensic reports are typically sent immediately upon receipt of a message that does not meet the DMARC authentication method.

Managing DMARC reports can be a challenge for many organizations. Namely, organizations with many users of email service can expect a huge number of received DMARC reports that need to be analysed. An additional challenge is that DMARC uses XML as a report format. This indicates the need to use ready-made software solutions or platforms for the analysis of DMARC reports.

As the DMARC protocol has grown in popularity, there are more and more such platforms on the market that offer several functionalities, the most important of which are:

- archiving DMARC reports,
- enrichment of DMARC status,
- check the status of SPF / DKIM / DMARC records,
- setting different notifications,
- integration with various security solutions.

VII. CHALLENGES WITH DMARC IMPLEMENTATION

Implementing the DMARC protocol is a complex and long process. This is especially true for large organizations. Such organizations often work with external service providers who need to send emails on behalf of the organization. Example of such service providers are marketing platforms like MailChimp or human capital management platforms like SAP SucessFactors. Prior to the full implementation of the DMARC protocol, the DMARC compatibility of the organization's email infrastructure and the external service providers used by the organization must be ensured. If DMARC is implemented hastily, there is a risk of rejecting many legitimate emails.

Best practice for the implementation of the DMARC protocol emphasizes the importance and need for a phased approach. In the first phase, it is necessary to set up the so-called DMARC auditing.

That is, it is necessary to set the appropriate DMARC record for example: *v = DMARC1; p = none; rua = mailto: dmarc@example.com*. With this record, the DMARC policy is placed in the supervisory or auditing mode, which means that electronic messages will not be rejected in the event of a failed DMARC check. The

"*rua"* parameter defines the email address to which the bulk DMARC reports will be sent.

Once DMARC is implemented in the auditing mode, the most demanding part of the implementation begins: the analysis of all external email servers that send messages on behalf of the organization. Once the DMARC policy has been put into auditing mode, at least few months must pass to collect enough bulk DMARC reports to identify all internal and external e-mail senders.

Not all email servers (belonging to the organization itself or external service providers) are at the same level of DMARC compliance. Some are at the very beginning and have, for example, only implemented SPF record and others can be almost compliant with the DMARC protocol. Depending on the number of senders and their level of compliance with the DMARC protocol, it depends on how long it takes for the full implementation of the DMARC protocol. This implies setting the DMARC record / policy *p = reject*. In our example, the DMARC record would look like this: *v = DMARC1; p = reject; rua = mailto: dmarc@example.com.*

Often these activities can be quite complex. Many organizations do not have enough knowledge, human resources for this implementation. As a result, many organizations see the risk of rejecting emails much higher than the risk of receiving unauthorized, spoofed email messages. Therefore, it is not surprising that in addition to the DMARC platform externalization service, the market also offers a service of full implementation of the DMARC protocol. Providers of such a service guarantee the application of p = reject for the agreed domain.

Perhaps it is the growing number of DMARC service providers that is improving the rather poor statistics on the use of the DMARC protocol. Namely, according to the security company Agari, only 34% of the world's largest companies (Fortune 500) have achieved full compliance with the DMARC protocol. In an extremely detailed analysis, Agari included almost 500 million Internet domains. Nearly 13 million domains have valid DMARC DNS record but nearly 4.8 million domains have p=reject DMARC DNS record. [13]

At the same time implementation of DMARC protocol is growing. Following graph is taken from Dmarc.org and shows the total of valid DMARC policies as published every six months from the end of 2016 through 2022. (The figures below are derived from analysis of data generously provided by Domain Tools). [14]

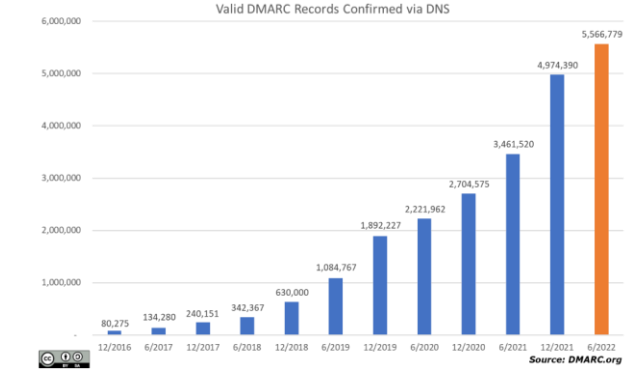


Figure 3. Valid DMARC records confirmed via DNS

These statistics can provide false sense of compliance since majority of DMARC policy records is p=none. This is shown on Figure 4.

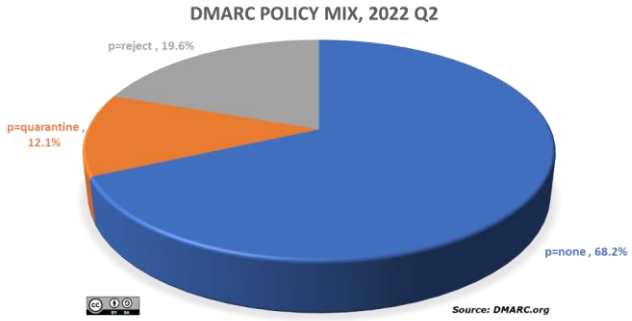


Figure 4. Dmarc policy mix

DMARC research analysis which is done on top one million domains in the three lists Alexa, Majestics, and Tranco between December 2018 and May 2020 show similar results. According to most recent scan, DMARC is only used by up to 11.5% of all examined domain. It is worth mentioning that DMARC policy p=none is implemented on 67.9% domains. [15]

While a "p=none" DMARC policy is not inherently a problem, it can leave a domain vulnerable to abuse by email spoofers and phishing attacks. It does not provide any direct benefit to the domain owner beyond receiving reports on DMARC activity. In scenario where p=none is not changed for a long period of time this means that DMARC is not implemented at all.

VIII. Analysis of DMARC implementation in Republic of Croatia

In this chapter will be given detail analysis of DMARC implementation in largest Croatian companies and compare it with world statistic. This will show where Croatia is compared to other countries. Is DMARC recognized as very important part of email security?

According to 2022 annual report of Republic of Croatia CERT email incidents related to phishing and phishing URL are responsible for 43% of total security incidents. This result is same compared to last year report and correlates with world statistics. It is obvious that email security is very important issue in Republic of Croatia as well. [16]

Analysis of DMARC implementation is given for most important Croatian companies based on total value created and relevant entities of public sector. List of top 200 most important companies is taken from business magazine Lider. Their methodology is ranking companies according to added value which is sum of gross wages and gross profit.[17] Some small companies in top 200 without registered internet domain are replaced with companies out of top 200. Added are additional companies from finance sector due to the importance on whole society. In total analysis included DMARC records for three hundred and thirteen internet domains.

DMARC records were analysed with MxToolbox Web page https://mxtoolbox.com/. On figure 5. and figure 6. are given examples of DMARC protocol implementation.

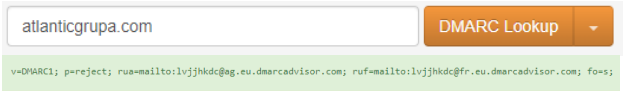


Figure 5. Example of implemented DMARC protocol

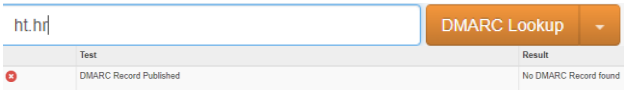


Figure 6. Example of not implemented DMARC protocol

DMARC implementation analysis of all entities in scope of this research on table 1. shows that 32% of them have some DMARC record and 68% don't have DMARC record.

DMARC policy	Count	Percentage
Not in use	213	68%
Policy none	50	16%
Policy quarantine	28	9%
Policy reject	22	7%

Table 1. DMARC analysis on all entities

DMARC implementation analysis of private sector on 234 domains in table 2. shows that 38% of analysed companies have some DMARC record and 62% don't have DMARC record.

DMARC policy	Count	Percentage
Not in use	145	62%
Policy none	48	20%
Policy quarantine	24	10%
Policy reject	20	8%

Table 2. DMARC analysis on private sector

DMARC implementation analysis on top 100 private companies in table 3. shows that 39% of analysed companies have some DMARC record and 61% don't have DMARC record.

DMARC policy	Count	Percentage
Not in use	61	61%
Policy none	21	21%
Policy quarantine	8	8%
Policy reject	10	10%

Table 3. DMARC analysis on top 100 private companies

DMARC implementation analysis on 79 most relevant public institutions (e.g., government bodies, parliament, public agencies, justice system etc.) shows that 14% of analysed entities have some DMARC record and 86% don't have DMARC record.

DMARC policy	Count	Percentage
Not in use	68	86%
Policy none	3	4%
Policy quarantine	5	6%
Policy reject	3	4%

Table 4. DMARC analysis on public sector

These analyses show that private sector in Croatia is far more advanced in email domain spoofing protection area than public sector. In private sector in total 18% of companies have at least some levels of protection (p=quarantine) or full protection (p=reject). Private sector also has potential for improvement because 20% of companies have DMARC policy p=none. This indicates some level of DMARC awareness.

Public sector is almost two times worse than private sector with only 10% of DMARC p=reject or p=quarantine. This indicates low level of awareness and lack of resources. Public sector is providing important digital services for whole Croatian citizenship. Poor DMARC practise is significantly increasing identity impersonation attacks based on email spoofing. Therefore, it is not surprising that the statistics of email attacks from the annual report of the Croatian CERT is worrying.

If we compare DMARC implementation in Croatia's top 100 private companies with world's largest companies situation is not so bright. [18] DMARC analysis done by email security company Agari in table 5. showed significantly largest DMARC adoption in US and UK companies than in Croatia. Australian and German companies are also at more advanced level.

DMARC policy	F 500	FTSE 100	ASX 100	HDAX 100
Not in use	19%	18%	16%	41%
Policy none	37%	35%	44%	28%
Policy quarantine	9%	9%	17%	10%
Policy reject	34%	38%	23%	21%

Table 5. DMARC Breakout: World's Largest Companies

One of the reasons why US and UK private companies have significantly better DMARC results is the fact that there are binding legal directives for their public sector. [19] [20] Directives like this also have tremendous impact on private sector through raising DMARC awareness and giving strong push to its implementation. Definitively strong regulative is always an efficient way to overcome neglecting important security controls.

Analysis of DMARC implementation in banking and pharmaceutical sector in Republic of Croatia is partially proving this thesis. Those sectors are at more advanced level compared to rest of private sector but still lacking comparing to world's largest companies.

DMARC policy	Count	Percentage
Not in use	8	40%
Policy none	6	30%
Policy quarantine	3	15%
Policy reject	3	15%

Table 6. DMARC analysis in banking sector

DMARC policy	Count	Percentage
Not in use	5	56%
Policy none	1	11%
Policy quarantine	1	11%
Policy reject	2	22%

Table 7. DMARC analysis in pharmaceutical sector

At the same time insurance sector in Republic of Croatia is completely neglecting DMARC importance. That statistic is even worse than public sector which is disappointing.

DMARC policy	Count	Percentage
Not in use	11	73%
Policy none	3	20%
Policy quarantine	1	7%
Policy reject	0	0%

Table 8. DMARC analysis in insurance sector

Better DMARC statistics in banking and pharmaceutical sector is highly related with sensitive type of business and strict regulation. Risk of email impersonation is quite high in those sectors and minimizing that risk with implementing DMARC is only reasonable option.

IX. CONCLUSION

The insecurity of email services is unfortunately always a hot topic due to significant financial loses. Attackers use anything that will grab the attention of the email recipients and make such messages as authentic as possible. It is the method of falsifying the From field that helps them a lot. Full implementation of the DMARC protocol significantly reduces the risk of such simple attack methods.

It is clear from the paper that this is not an easy task. However, the benefits of DMARC implementation far outweigh the risks that this implementation brings.

Analysis of DMARC implementation in largest Croatian companies and relevant entities of public sector has shown that this is neglected since only 7% have full DMARC compatibility and 9% have DMARC in quarantine mode. Situation in public sector is even worse. In private sector banking and pharmaceutical sector are leading in this area but still with significant space for improvement.

Probably the best way to push DMARC implementation in private and public sector in Republic of Croatia is binding legal DMARC directive. That kind of directives provided DMARC implementation momentum in USA and UK.

REFERENCES

[1] 2021 Data Breach Investigations Report, https://enterprise.verizon.com/resources/reports/dbir/, May 2021.

[2] Cyber Claims: GDPR and business email compromise drive greater frequencies, https://www.aig.ae/content/dam/aig/emea/uae/documents/aig-cyber-claims-report-2020.pdf,

[3] Business Email Compromise The $26 Billion Scam, https://www.ic3.gov/Media/Y2019/PSA190910, 10.9.2019.

[4] Protecting Against Business Email Compromise Phishing, https://www.gartner.com/doc/reprints?id=1-1Z0GPFC0&ct=200512&st=sg, 19.3.2020.

[5] Internet fraud, https://policija.gov.hr/prevencija/racunalna-sigurnost/internet-prijevare/456

[6] Scott Kitterman. 2014. Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1. RFC 7208. https://doi.org/10.17487/RFC7208

[7] Murray Kucherawy, Dave Crocker, and Tony Hansen. 2011. DomainKeys Identified Mail (DKIM) Signatures. RFC 6376. https://doi.org/10.17487/RFC6376

[8] Dennis Tatang, Florian Zettl and Thorsten Holz, The Evolution of DNS-based Email Authentication: Measuring Adoption and Finding Flaws; RAID '21: 24th International Symposium on Research in Attacks, Intrusions and Defenses, San Sebastian, Spain, October 2021, https://dl.acm.org/doi/fullHtml/10.1145/3471621.3471842

[9] Murray Kucherawy and Elizabeth Zwicky. 2015. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489. https://doi.org/10.17487/RFC7489

[10] Anatomy of a DMARC resource record in the DNS, https://dmarc.org/overview/

[11] Murray Kucherawy and Elizabeth Zwicky. 2015. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489. https://doi.org/10.17487/RFC7489

[12] Murray Kucherawy and Elizabeth Zwicky. 2015. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489. https://doi.org/10.17487/RFC7489

[13] 2022 Email Fraud & Identity Deception Trends, https://static.fortra.com/agari/pdfs/guide/ag-acid-2022-email-fraud-identity-deception-trends.pdf.

[14] DMARC policies increase 28% Through June 2021, https://dmarc.org/2021/10/dmarc-policies-increase-28-through-june-2021/, October 2021.

[15] Dennis Tatang, Florian Zettl and Thorsten Holz, The Evolution of DNS-based Email Authentication: Measuring Adoption and Finding Flaws; RAID '21: 24th International Symposium on Research in Attacks, Intrusions and Defenses, San Sebastian, Spain, October 2021, https://dl.acm.org/doi/fullHtml/10.1145/3471621.3471842

[16] Annual report of the national CERT for 2022., https://www.cert.hr/wp-content/uploads/2023/02/CERT-G.I.-2022..pdf, 8.2.2023.

[17] Liderovih 500 najboljih: Pogledajte kojih je deset najuspješnijih kompanija u Hrvatskoj, https://lidermedia.hr/biznis-i-politika/liderovih-500-najboljih-pogledajte-kojih-je-deset-najuspjesnijih-kompanija-u-hrvatskoj-145091, 21.9.2022.

[18] 2022 Email Fraud & Identity Deception Trends, https://static.fortra.com/agari/pdfs/guide/ag-acid-2022-email-fraud-identity-deception-trends.pdf.

[19] Binding Operational Directive 18-01: Enhance Email and Web Security, https://www.cisa.gov/binding-operational-directive-18-01

[20] Updating our security guidelines for digital services, https://technology.blog.gov.uk/2016/06/28/updating-our-security-guidelines-for-digital-services/

Towards Improving Online Security Awareness Skills with Phishing and Spoofing Labs

Alexander Kerr[1] and Timo Hynninen[1,2]
[1]South-Eastern Finland University of Applied Sciences, Mikkeli, Finland
[2]Laurea University of Applied Sciences, Espoo, Finland
alexander.kerr@xamk.fi; timo.hynninen@laurea.fi

***Abstract*—It can be especially hard for novices to examine their online browsing habits, or to be aware of the dangers related to online security. In this paper, we describe laboratory exercises designed to improve awareness of online security, in addition to teaching cybersecurity knowledge and skills. The lab exercises were used as a part of an information security fundamentals course for IT students with none or only novice experience in the field. These lessons are suitable for both delivery in the classroom and online in a virtual laboratory environment. The lab exercises were developed using the design science research methodology (DSR). Following DSR principles, the evaluation of the developed material was done by extracting observations from student reports. The thematic analysis method was used to process the data. The evaluation of the student learning reports revealed that the labs and the course were successful in the following ways: Improving security knowledge, improving critical thinking skills, and improving security awareness. Students also gained hands-on cyber awareness skills by analysing metadata in messages. Additionally, the material was perceived as useful for future working life.**

Keywords—cybersecurity;education;curriculum design

I. Introduction

Information and cybersecurity are recommended to be tightly integrated into computing education [1], [2]. This is because the need for cybersecurity knowledge and awareness is on the rise; Not only is the cybersecurity field in dire need of qualified professionals [3], [4] but security awareness is also paramount in other computing and IT-related jobs.

Cybersecurity is a multifaceted topic [5]. All computing students need to understand security to some degree, and they generally learn the necessary domain skills in their respective areas; For example, web application security can be covered in web development courses and infrastructure security in computer networking classes. However, all students should have the basic security thinking and awareness skills early on, as security cannot be an afterthought [6]. One important topic in security thinking and awareness is online phishing and spoofing scams.

Phishing attacks are common [7]. For example, the Anti-Phishing Working Group (APWG) reported having observed between 68 000 and 94 000 phishing attacks per month in 2022 [8]. According to APWG, the most targeted industries for phishing attacks are financial institutions, SaaS and Webmail systems, and social media providers [8]. Clearly, phishing attacks are common against services that everyone uses daily.

The current paper describes a case study in which laboratory exercises ('labs') are designed to illustrate phishing and online spoofing scams. These lab exercises include the investigation of online scams, the design and implementation of phishing and spoofing, and remedies to said dangers. The objective of the study is to improve awareness and security knowledge related to online browsing safety.

The rest of the paper is organized as follows. Section II presents related work on information and cybersecurity education. Section III describes the research methodology used in the study, particularly the thematic analysis method that was used to codify the data. Next, Section IV presents the design of the lab exercises and results from the data analysis. Finally, Section V concludes the paper by discussing the findings, and limitations of the work.

II. Related work

Cybersecurity is an interdisciplinary field that requires a strong foundation in computing technologies but also an understanding of the human and social aspects of computing [5].

The extant work in cybersecurity education is well established in the literature. The body of knowledge is detailed in several tertiary studies, e.g. [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. These studies can be found using the Google Scholar search engine with the keywords *("cybersecurity" OR "cyber security")+("teaching" OR "education")+("literature review" OR "mapping study" OR "meta analysis").*

According to the systematic review by Sağlam et al. [16] there are three main themes in the cybersecurity education literature: what to teach, how to teach, and who should teach. Cybersecurity courses should combine practical and theoretical education, with hands-on approaches being both popular and effective [17], [16], [29]. Roepke and Schroeder [15] point out that new cybersecurity tools and challenges arise constantly, and teaching should aim to foster the transfer of knowledge to solve new problems in a sustainable way. Usually students are able to gain an understanding of abstract security concepts early on [30].

With regards to best practices, technical cybersecurity should be taught using virtual laboratory environments

using an offensive approach [20]. According to the study by Wilson [31], teaching students to attack rather than to defend will increase learning results and student motivation. Virtual laboratory environments can also have the technical limitations which prevent them from portraying real-world situations true in every respect, however modern virtualization technologies can usually overcome these obstacles [25].

In addition, serious games can be an effective method to provide training [21]. Cybersecurity games have been shown to have a a positive effect on learning outcomes [28]. Games are especially useful for teaching security awareness and teaching students how to defend from cyberattacks [21].

Albeit the topic of cybersecurity education is abundant in literature, there is room for development. For example, the literature review by Chowdhury & Gkioulos [17] concludes that there is no existing consensus on the delivery methods and design of cybersecurity lab exercises. This suggests that evaluation of the cybersecurity lab exercises and their learning outcomes is a fruitful research avenue, and this paper attempts to fill the research gap.

III. Research method

The Design Science Research (DSR) method [32], [33], [34], [35] is suitable for various engineering work. DSR is an outcomes-based method that provides scaffolding for the creation and evaluation of useful artefacts [33], [34]. The artefacts can range from theories to designs and tangible products [36]. The novelty, quality, and applicability of the artefacts are evaluated based on their usefulness and capability to solve real problems [37], [33]. In this study, the Design Science Research approach will act as the guiding framework in the design and evaluation of the cybersecurity laboratory exercises.

Following the DSR method's requirement to evaluate designs in real-world context, the courseware will be analyzed using student reflections of their learning during the course. Therefore, the primary data source in this study is student submitted reports to a learning task at the end of the course. Turning in this reflection task was voluntary as such but it would accumulate points toward the course total, or students could opt to complete the reflection task instead of another laboratory assignment.

In this learning task, students were asked to name three things they learned during the course, and to explain their choices. Specifically, we asked the students to reflect on the following:

- name/choose three topics or things they learned during the course
- give a short explanation to justify the selection
- explain the importance of the chosen topics
- explain how the knowledge can be useful

The answers were open-ended. The *thematic analysis* research method was chosen to code the open-ended data. The thematic analysis method is a "qualitative research method for identifying, analyzing and reporting patterns (themes) within the data" [38]. In the data collection, processing, and analysis the ethical principles of research with human participants by the Finnish national board on research integrity were adhered to.

IV. Implementation and Results

A. Course design

In the autumn semester of 2022 we organized an introductory course in information and cybersecurity. The course focused on the following topics:

- cybersecurity concepts
- threats, attacks, and attackers
- security needs and security technologies
- computer security and access control
- network security, network defenses
- encryption, secured web communications, and certificates
- security processes and users in the process

The course's lectures followed the outline of the topics presented above. Weekly hands-on laboratory exercises were designed to implement the concepts in practice. While the course objectives were not designed to be particularly biased toward offensive or defensive approaches the labs turned out to follow a more offensive style. These lab exercises could be completed either in the virtual laboratory environment (the Virtual ICT Lab [39]), or in the classroom by installing the virtual machine images locally.

It was assumed that the majority of the students participating would have little to no knowledge of 1). Basic cybersecurity principles. and 2). The Linux operating environment. For this reason, detailed instructions were given for each lab. One week focused solely on getting familiar with the nested virtual environment, and Linux in general.

B. Virtual laboratory environment

The labs were implemented in a sandbox environment using virtual machines running in the Virtual ICT Lab. Figure 1 presents the topology of virtual machines (VMs) which were created for the students. The main platform for this was a Linux Ubuntu desktop 20.04, which was run as the main desktop that students used. Additionally, a standard Windows 10 desktop was available. The primary Linux VM was installed with all the required elements that students were expected to use over the semester. Several lab scenarios were available for the different weekly lab topics. The configuration included a Linux server running nginx (web server) and an IMAP (email) server (but students did not need to interact with the server directly).

The main VM was configured to use a lighter Xubuntu desktop manager and dynamic 25GB of storage. This made the VM capable of running on 1 virtual CPU, and it ran moderately well with 1GB of RAM. The initial installation would take up 9.25GB of storage.

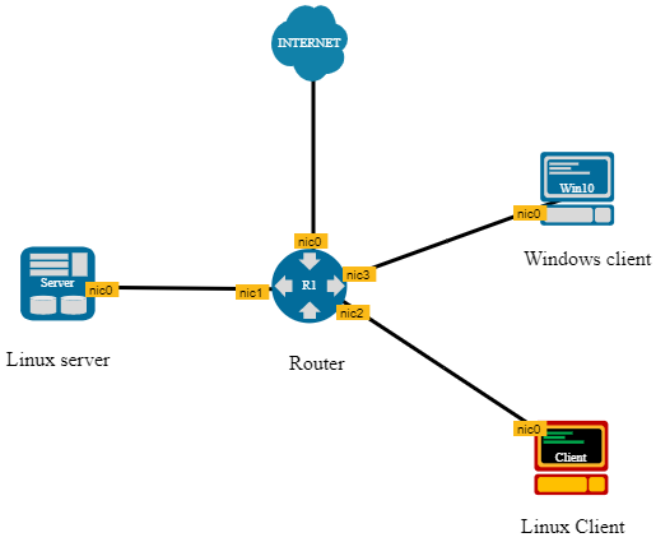


Fig. 1. Topology of virtual machines in the virtual laboratory environment

C. Lab exercise design

The Xubuntu VM was available as part of a virtual scenario where the students completed various cybersecurity hands-on labs. Table I presents the web security focused lab exercises.

The first lab demonstrated the ease of creating scam and phishing websites. First, the students were asked to design a scam and a phony website to accompany it. Then, the students were given some basic instructions on how they could create a targeted phishing site on the web.

The second lab exercise focused on manipulating e-mail headers. The students were tasked with sending emails from fabricated addresses using command-line tools. The objective was to demonstrate that on its own email is not inherently secure, and how easy phishing mail is to construct. In addition, the goal was to increase the awareness of other attack methods, giving the students a deeper understanding of how systems could quickly be compromised.

The third lab was designed to highlight awareness of phishing, spoofing, and password vulnerability. This was realized as the design and implementation of a credential harvesting portal. Students would copy a legitimate webpage for credential harvesting, and then point the victim's computer to the phony site by altering the Linux hosts file. This lab was also used to demonstrate browser features that can aid students in deciding if a website is legitimate or not.

D. Observations

In order to elicit information on how successful the phishing and spoofing laboratory exercises were, the responses to the reflection task were inspected by the two instructors (authors of this paper) of the course. Overall there were 39 students who submitted reports for the course. Out of these we selected 22 for closer inspection using the thematic analysis approach. This selection was based on whether student submitted reflection was related to the spoofing and phishing labs. The rest of the submissions described learning outcomes related to other course material, for example, unrelated labs or lecture material.

For the 22 selected reports an initial open coding process was done by both authors individually. This process involved reading the student reflection, and recording observations on *how the text described the spoofing and phishing labs*. Next, an axial coding process was conducted, both authors codified the observations into categories and concepts together. The resulting text categories are presented in Table II.

Most of the reflections (N=16) described how the tasks had increased the student's general awareness of phishing, spoofing, or online scams. Another common response was that the students found out how easy spoofing can be on the web (N=13). Many students also described gaining either generic or specific technical knowledge (N=12).

Next, a group of less common but equally important recurring themes were found. Some students reflected on the usefulness of the knowledge and skills in later working life (N=5). Other students reported on how they had gained new hands-on skills in analyzing email messages and their metadata (N=4). A few students described how the knowledge is useful given how common scams are on the web (N=3). Finally, the last students mentioned how the labs helped in developing critical thinking skills (N=2).

V. Discussion and Conclusions

This paper outlined our experiences in running hands-on laboratory exercises related to phishing and spoofing on an early information and cybersecurity course. The

TABLE I
ONLINE SECURITY AWARENESS RELATED LAB EXERCISES DURING COURSE

Name	Description
Desining convincing internet scams	Students used Google Sites (or other CMS editor) to design websites for malicious purposes. Students were instructed to design a convincing fraud or scam that could fool a user who lands on the website.
Email spoofing	Students used Linux command line tools to send email messages to an insecurely configured IMAP server. The message headers were altered to spoof the senders' email address.
Creating a phishing site	Students familiarized themselves with credentials harvesting. They studied the HTML and CSS code of a legitimate website and produced a convincing copy. Then the students would act as a malicious intruder and modify the Linux hosts file in an attempt to fool users to land on a credential harvesting site.

TABLE II
AXIAL CODES GENERATED DURING THE THEMATIC ANALYSIS PROCESS

#	Axial code	Number of observations	Detailed decription
1	Awareness	16	Descriptions of increased awareness of online security.
2	Critical thinking	2	Descriptions of increased critical thinking about messages on the web.
3	Gained knowledge	12	Free-form descriptions of different knowledge related to the phishing and spoofing exercises.
4	Useful in working life	5	Descriptions mentioning the awareness of phishing and spoofing when working in an organization.
5	Inevitability	3	Descriptions of spoofing being commonplace, and how everyone should be aware of the topic.
6	Analysing message metadata	4	Descriptions of gaining hands-on skills for analysing message metadata to verify the authenticity of the sender.
7	Spoofing ease	13	Descriptions of how easy it is or how little skill it takes to create spoofed messages.

labs should be suitable for a wide range of students - even though they were designed for IT students there was no prerequisite knowledge needed to take the course.

A. Findings

At the end of the course, we evaluated student reflections on what they perceived as the most important topics. It was found that the students described their experiences in seven different categories: Gaining awareness of online security; gaining critical thinking skills; gaining new knowledge; the course topics are useful for the future working life; scams are so ubiquitous that one must be wary of online dangers; gained hands-on computing skills to check email metadata; and learning that it takes little skill to concoct attacks online.

As a result, the data suggest, that the labs (and the course as a whole) helped students pick up both practical skills and knowledge about online security. Specifically, students reported now exercising more caution in online messaging / emails and they know some remedies for phishing attacks. Students generally have more awareness of phishing and spoofing online, and they are thus better prepared for professional working life. Additionally, students gained professional computing and IT knowledge about email systems and their operation.

B. Limitations and future work

The scope and limitations of the current study warrant some discussion. First, our sample of 22 student reflections is relatively low. However, the results are in line with previous findings (e.g. [20]). Additionally, the sample was taken from a larger pool of answers.

The research method employed a thematic analysis method. As thematic analysis uses an open coding approach it is possible that researcher bias affects the findings. For this reason both authors conducted the coding first individually, and then agreed on the resulted axial codes.

In future work we continue to evaluate our labs and improve the designs. We aim to answer the call given by Chowdhury & Gkioulos [17] to document and publish best practices in the cybersecurity education and course content.

REFERENCES

[1] S. Peltsverger and O. Karam, "Is teaching with security in mind working?" in *2010 Information Security Curriculum Development Conference*. ACM, 2010, pp. 15–20.

[2] The Joint Task Force on Computing Curricula, "Curriculum guidelines for undergraduate degree programs in software engineering," New York, NY, USA, Tech. Rep., 2015.

[3] (2017) This is what the future of cybersecurity will look like. [Online]. Available: https://www.weforum.org/agenda/2017/08/the-us-is-upping-its-game-against-cyber-attacks-but-the-security-industry-faces-a-huge-challenge

[4] F. Goupil, P. Laskov, I. Pekaric, M. Felderer, A. Dürr, and F. Thiesse, "Towards understanding the skill gap in cybersecurity," in *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1*, ser. ITiCSE '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 477–483. [Online]. Available: https://doi.org/10.1145/3502718.3524807

[5] H. Suryotrisongko and Y. Musashi, "Review of cybersecurity research topics, taxonomy and challenges: Interdisciplinary perspective," in *2019 IEEE 12th Conference on Service-Oriented Computing and Applications (SOCA)*. IEEE, 2019, pp. 162–167.

[6] J. McManus, "Security by design: teaching secure software design and development techniques," *Journal of Computing Sciences in Colleges*, vol. 33, no. 3, pp. 75–82, 2018.

[7] S. Bell and P. Komisarczuk, "An analysis of phishing blacklists: Google safe browsing, openphish, and phishtank," in *Proceedings of the Australasian Computer Science Week Multiconference*, 2020, pp. 1–11.

[8] The Anti-Phishing Working Group. (2022) The APWG phishing activity trends report. [Online]. Available: https://apwg.org/trendsreports/

[9] L. Zhang-Kennedy and S. Chiasson, "A systematic review of multimedia tools for cybersecurity awareness and education," *ACM Computing Surveys (CSUR)*, vol. 54, no. 1, pp. 1–39, 2021, publisher: ACM New York, NY, USA.

[10] V. Švábenskỳ, J. Vykopal, and P. Čeleda, "What are cybersecurity education papers about? a systematic literature review of sigcse and iticse conferences," in *Proceedings of the 51st ACM technical symposium on computer science education*, 2020, pp. 2–8.

[11] S. Laato, A. Farooq, H. Tenhunen, T. Pitkamaki, A. Hakkala, and A. Airola, "Ai in cybersecurity education-a systematic literature review of studies on cybersecurity moocs," in *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2020, pp. 6–10.

[12] X. Mountrouidou, D. Vosen, C. Kari, M. Q. Azhar, S. Bhatia, G. Gagne, J. Maguire, L. Tudor, and T. T. Yuen, "Securing the human: a review of literature on broadening diversity in cybersecurity education," *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education*, pp. 157–176, 2019.

[13] N. A. A. Rahman, I. Sairi, N. A. M. Zizi, and F. Khalid, "The importance of cybersecurity education in school," *International Journal of Information and Education Technology*, vol. 10, no. 5, pp. 378–382, 2020.

[14] H. Aldawood and G. Skinner, "Educating and raising awareness on cyber security social engineering: A literature review," in *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)*. IEEE, 2018, pp. 62–68.

[15] R. Roepke and U. Schroeder, "The Problem with Teaching Defence against the Dark Arts: A Review of Game-based Learning Applications and Serious Games for Cyber Security Education." in *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)*, 2019, pp. 58–66.

[16] R. B. Sağlam, V. Miller, and V. N. Franqueira, "A Systematic Literature Review on Cyber Security Education for Children," *IEEE Transactions on Education*, 2023, publisher: IEEE.

[17] N. Chowdhury and V. Gkioulos, "Cyber security training for critical infrastructure protection: A literature review," *Computer Science Review*, vol. 40, p. 100361, 2021, publisher: Elsevier.

[18] W. Chen, Y. He, X. Tian, and W. He, "Exploring Cybersecurity Education at the K-12 Level," in *SITE Interactive Conference*. Association for the Advancement of Computing in Education (AACE), 2021, pp. 108–114.

[19] M. Coenraad, A. Pellicone, D. J. Ketelhut, M. Cukier, J. Plane, and D. Weintrop, "Experiencing cybersecurity one game at a time: A systematic review of cybersecurity digital games," *Simulation & Gaming*, vol. 51, no. 5, pp. 586–611, 2020, publisher: SAGE Publications Sage CA: Los Angeles, CA.

[20] L. Riihelä, "Teaching information security: A systematic mapping study," Master's thesis, LUT University, Finland, 2019.

[21] F. Alotaibi, S. Furnell, I. Stengel, and M. Papadaki, "A review of using gaming technology for cyber-security awareness," *Int. J. Inf. Secur. Res.(IJISR)*, vol. 6, no. 2, pp. 660–666, 2016.

[22] J. Jeong, J. Mihelcic, G. Oliver, and C. Rudolph, "Towards an improved understanding of human factors in cybersecurity," in *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*. IEEE, 2019, pp. 338–345.

[23] R. Ramirez and N. Choucri, "Improving interdisciplinary communication with standardized cyber security terminology: A literature review," *IEEE Access*, vol. 4, pp. 2216–2243, 2016, publisher: IEEE.

[24] S. Das, "SoK: a proposal for incorporating accessible gamified cybersecurity awareness training informed by a systematic literature review," in *Proceedings of the workshop on usable security and privacy (USEC)*, 2022.

[25] J. Pittman, "Understanding system utilization as a limitation associated with cybersecurity laboratories–A literature analysis," *Journal of Information Technology Education. Research*, vol. 12, p. 363, 2013, publisher: Informing Science Institute.

[26] M. Lamond, K. Renaud, L. Wood, and S. Prior, "SOK: young children's cybersecurity knowledge, skills & practice: a systematic literature review," in *Proceedings of the 2022 European Symposium on Usable Security*, 2022, pp. 14–27.

[27] J. Hautamäki, M. Karjalainen, T. Hämäläinen, and P. Häkkinen, "Cyber security exercise: Literature review to pedagogical methodology," in *INTED Proceedings*. IATED Academy, 2019, issue: 2019.

[28] M. Hendrix, A. Al-Sherbaz, and B. Victoria, "Game based cyber security training: are serious games suitable for cyber security training?" *International Journal of Serious Games*, vol. 3, no. 1, pp. 53–61, 2016.

[29] L. Wang, J. Yang, and P.-J. Wan, "Educational modules and research surveys on critical cybersecurity topics," *International Journal of Distributed Sensor Networks*, vol. 16, no. 9, p. 1550147720954678, 2020.

[30] T. Hynninen, "On the learning activities and outcomes of an information security course," in *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, 2019.

[31] B. Wilson, "Teaching security defense through web-based hacking at the undergraduate level," 2017.

[32] S. Gregor and A. R. Hevner, "Positioning and presenting design science research for maximum impact." *MIS quarterly*, vol. 37, no. 2, pp. 337–355, 2013.

[33] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, Mar. 2004. [Online]. Available: http://dl.acm.org/citation.cfm?id=2017212.2017217

[34] A. R. Hevner, "A three cycle view of design science research," *Scandinavian journal of information systems*, vol. 19, no. 2, p. 4, 2007.

[35] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of management information systems*, vol. 24, no. 3, pp. 45–77, 2007, publisher: Taylor & Francis.

[36] B. Kuechler and V. Vaishnavi, "On theory development in design science research: anatomy of a research project," *European Journal of Information Systems*, vol. 17, no. 5, pp. 489–504, 2008.

[37] A. Hevner and S. Chatterjee, "Design science research frameworks," *Design Research in Information Systems: Theory and Practice*, pp. 23–31, 2010.

[38] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[39] J. Kasurinen and M. Kettunen, "Learning cyber security: Teaching technically challenging topics with games and virtual laboratories," *International Journal on Information Technologies and Security*, vol. 10, no. 3, pp. 103–114, 2018.

Secure, Accessible, Virtual Voting Infrastructure (SAVVI): Reducing Barriers for Disabled and Overseas Voters

Ted Selker and Justin Pelletier
Rochester Institute of Technology
Rochester, NY 14623
Email: jxpics@rit.edu

***Abstract*—We describe a way to deploy a secured ballot return for overseas, vision, or dexterity-impaired voters. SAVVI is a secure, accessible, virtual voting infrastructure. It uses multiple communication channels with synchronized multi–factor authentication, encryption, and hashing to preserve privacy, confidentiality, and vote integrity. Our usability goal is to supply voters easy access to a hardened system that will present secure instances of their familiar browser and email clients. A key to its viability is synchronizing authentication through two low-tech verifications such as phone calls. This strives to enhance security and usability for remote voting by mimicking best practices for in-person polling place procedures. Other more standard cryptographic measures include layered encryption and dynamically provisioning a secure virtual container–a virtual voting machine (VVM)–to process each ballot. In aggregate, the design of SAVVI seeks to allow remote voting while reducing difficulty for the voter, programming complexity for the election administrator, and material procurement for the voting authority.**

I. Introduction

Mail-in ballots are problematic for some classes of voters. Overseas Americans and Uniformed and Overseas Citizens Absentee Voting Act (UOCAVA) voters have had difficulties getting mail-in ballots to their home polling places in a timely way [49]. Blind voters, as well as people with limited manipulation, are not able to indicate selections and return a mail-in ballot privately and independently. For these classes of voters, some U.S. states such as North Carolina and Massachusetts allowed ballots to be delivered, marked, and returned electronically in 2020 [35]. This paper proposes an approach to assemble existing technologies to allow usable secure electronic return for these voters.

Previous work, such as Bruck, Jefferson, and Rivest's proposal for frog voting, considers architectures that separate vote generation from vote casting[17]. Our proposal also separates vote generation from casting. We also add multifactor authentication, dynamic provision of the vote generation machine, and provide some consideration of usability for disabled voters.

Our approach improves security for election officials who need to receive UOCAVA or disabled voter's ballots remotely. In doing this, we seek to minimize the need for an expensive procurement process, creating new software, or using security approaches that require deep expertise.

II. Background

In this section, we elaborate on the context for our goals to integrate Confidentiality, Integrity, and Accessibility for disabled and overseas voters. Confidentiality demands independent ballot marking and submission and protecting ballot confidentiality in transit. Integrity requires that voters can verify votes are cast as intended and counted as cast. We also strive to reduce the attack surface of voting infrastructures. Finally, we seek to integrate these security features while simultaneously improving Accessibility by requiring fewer and more familiar steps.

A. Voting Security Challenges

Many security problems result from the trade–offs that increase usability of voting systems, and a large number of reported problems arise from the user-interface. Though improved ballot–design practices might yield usability improvements, all practices need continual reassessment to balance security needs.

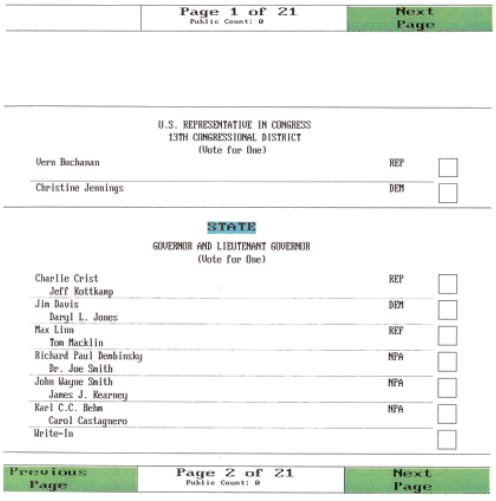


Fig. 1: 2006 Sarasota ballot where 13% of voters made no selections on Congressional CD13

Additional security concerns plague voting. People might rely on incorrect voter information that inhibits them from voting [33][31]. People might find it easier to vote for a selection that is first in a list [55]. Mechanical problems have at times kept votes from being counted, as in the incomplete selections from the hanging chad problems in the 2000 Florida election [56]. Tallying problems sometimes require recounts. Some voters notice that they have voted for the selection adjacent to the one they intended (called flipping). Sometimes data-record manipulation has been found, such as occurred in Volusia County in the 2000 election[24].

The question of how to securely deploy a system with complete privacy continues. Many people have stated that paper ballots are the best approach and they may be for now. Unfortunately, at times paper ballots have been compromised – even in wholesale ways – with reports of deliberate ballot access limitations, incorrectly printed ballots, chain voting to systematically coerce voters, lost ballot boxes, fraud in hand counting, and "new" ballots turning up in recounts [6], [29].

B. Mail-in Voting and Accessibility

During the COVID-19 pandemic, mail-in voting became critical for allowing vulnerable people to vote safely. Still, there are risks to mail-in voting [11]. Some ballots never get to the voter, get lost by the voter, or get spoiled by the voter. Some ballots do not get returned in time to be counted, are sent to the wrong addresses, or have incorrect use of the envelope and are thrown out. Various other simple system and user problems stop votes from being included in the vote. Electronic voting from home might still have some of these problems because voters might not set up the electronic voting system correctly. They might miss the emails. They might not use the software correctly to mark the ballot. They might not return the ballot correctly. Our architecture attempts to mitigate these risks.

There are even more challenges for disabled voters seeking to remotely return ballots. This remains a particularly important consideration because "individuals with disabilities also report voting by mail at much higher rates than do individuals not reporting a disability." [4]

1) Disabled Voters.: Section 508 under title 29 of the United States Code (https://section508.gov) defines functional requirements for technologies that accommodate nine classes of Americans with disabilities: (a) without vision, (b) with limited vision, (c) without perception of color, (d) without hearing, (e) with limited hearing, (f) without speech, (g) with limited manipulation, (h) with limited reach and strength, and (i) the need to minimize photosensitive seizure triggers.

We consider these requirements in light of specific requirements relating to absentee ballot preparation and submission. While deaf and mute voters should be able to use typical methods for voting, other disabilities present more challenges.

Voters with no/limited vision have used magnifiers, braille, and audio to vote. Although only some 10% of blind voters use braille [2] , Rhode Island for example, has made braille ballots available [7] Audio is a typical and accessible option made available to blind voters [36], [8], [5], [25].

Cognitively-disabled voters includes several populations. For example, more than 14% of voters have reading disabilities and more than 6% of voters have some short-term memory problems [47]. Structuring and the presentation of ballots for these voters can greatly reduce their errors. The practice of filling out a sample ballot to help them fill out the actual ballots reduces errors as well [26] [46].

Physically-disabled voters can have mobility or manipulation problems. Mobility problems might make it difficult for them to stand in line with their mobility solution. It has also been a challenge for polling places to make themselves adequately accessible to people with mobility problems. Manipulation problems are more difficult. People with manipulation problems can use prostheses to vote. Vote–marking machines are equipped with interfaces. Examples include special large buttons, mouth-sticks that can press buttons, or sip–and–puff devices to allow them to make selections with their mouth.

Mail-in voting presents privacy problems for blind and manipulation-challenged voters. We know of no workable prosthetic for allowing blind people to privately and independently vote on mail-in ballots. For people with manipulation problems, marking the ballot, putting it in the envelope, and mailing it without assistance can be impossible. Physically-disabled voters might also have difficulties assuring their paper ballots are received and delivered [35].

2) Overseas Voters.: Pernicious ballot delivery delays continue to impact military personnel and people living abroad who cannot be at home to vote. Some 3 million expatriates and military personnel are eligible to vote using UOCAVA. In 2006, only 26.5% of those requesting a ballot succeeded at having it counted [32], [41].

The voting security community is increasingly concerned with voting integrity verification. By letting voters know their vote is included in the count, the new verification approaches give much-needed evidence to increase confidence in outcomes[39]. Other projects such as Remotegrity [59] and VoteXX[19] concern technology that could be used for UOCAVA voters and some of those, such as Helios[9] and Election Guard[14], are even available as open-source software. However, these systems do not address how people with disabilities can vote privately and independently. Though some of these systems may have superior security properties, they retain notable usability limitations.

C. Usable Security in Electronic Voting

Concerns of electronic fraud have impacted the way we think about elections. Sadly, misinformation campaigns are common for reducing voting turnout. These are now being purveyed successfully online. For example, foreign messaging interference in the 2016 presidential election was impactful [12] and bad actors may turn or may have already turned to focusing on the voting systems themselves. To minimize risks, the Federal Election Commission and the Election

Assistance Commission have promulgated voting system certification recommendations. Every jurisdiction can choose how it will certify voting machines. Most states follow the EAC's Voluntary Voting Systems Guidelines (VVSG) for voting machine hardware and software [1] following rigorous testing and scrutiny defined by the National Institute of Standards and Technology (NIST) Handbook, pages 150-22. Still, each state has the right to control its approach to voting, so states mandate their specific certification procedures.

Due to past lost votes for the some 6 million expatriates, states also specify alternative approaches fulfilling UOCAVA [3]. The security standardization and the best practices in the Department of Defense (DoD) Inspector General's 2016 report on the department's policies, procedures, and practices for information security management [28] generally describe the same needs identified by the National Institute of Standards and Technology's (NIST) 2008 threat analysis of UOCAVA voting systems [42]. Though specific technologies have continued to change, both reports suggest three fundamental security control requirements: 1) strong authentication, 2) secure email, and 3) encrypted attachments.

Despite these controls, security concerns persist. For example, a body of research investigates a hypothetical user named Johnny and his use of encryption in email clients [54], [48], [44].

Researchers found that people can't encrypt because of usability concerns. They fail to properly implement pretty good privacy (PGP) because of the steps involved, or find it so difficult to use that they don't even try to encrypt. Most recently, Ruoti, Andersen, Zappala, and Seamons discovered that the usability of PGP-based email encryption has not advanced beyond a 10% success rate[44]. They attributed these failures to the perceived complexity of the encryption mechanism requiring both sender and receiver to materially participate in the PGP key exchange. The SAVVI system seeks to address this usability problem by automating the email encryption process, so Johnny doesn't have to encrypt: the secure container that processes his ballot does it for him.

While estimates vary, probably less than 0.3% of Americans have total blindness [30]. They have historically lower-than-average turnout [4], but weak security for them could still lead to fraudulent results and possibly even an overturned election. Layering best-practice protections can improve the security and access for UOCAVA and voters with other disabilities as well.

The component that remains missing is a secure way of transmitting electronic votes: an architecture that can allow a balance between accessibility, confidentiality, and integrity for remote voters while simultaneously minimizing the burden on the local elections board and the voter.

In the next section, we present our motivation and design rationale to integrate standard and non-centralized tools. Follow-on work will specifically test usability impacts of specialized CAPTCHAS, biometrics, and secure VMs in this voting infrastructure.

[1] https://www.eac.gov/voting-equipment/

III. Security Design Rationale

In the run up to the 2019 presidential election most people were concerned with blunting the pandemic by avoiding respiratory exposure to others. The authors of this paper met when a jurisdiction began asking: "What is possible for this election to allow people with disabilities to vote privately and independently?" We considered if there were inexpensive ways of giving people with disables a way to vote that was less prone to interference than what was currently in use. In exploring this, we looked at possibilities for improving the confidentiality, integrity, and accessibility for UOCAVA voters too. The challenge was posed by election officials who were concerned that there was no time to change anything. Could a safer, easier-to-administer system be organized inexpensively? SAVVI was the result of speaking to the above challenges.

The legitimizing question is which voting methods best reveal potential problems before they materialize and simultaneously lose the least votes. Questions of how to measure this are important too. Subsequently, the design of SAVVI focused on the following requirements:

1) **Prove to voters that their vote was cast as intended and recorded as cast.** Rationale: Trustworthy voting is the foundation of democracy. Giving voters a way to see their vote was tallied–while avoiding vote–buying–could be a way to increase trust in the electoral process.
2) **Resilient to technology change.** Rationale: Technologies deprecate and improve at a high rate. The infrastructure should be flexible enough to incorporate best-in-class security and virtualization tools as they become available. Also, replacing legacy (non-virtual) voting machinery can be tremendously expensive to elections districts.
3) **Allow remote voting from personal devices.** Rationale: Personal devices are more familiar to voters with specialized access technology implementations. Furthermore, using personal devices might reduce the anxiety of voting and may lead to less mistakes.
4) **Leverage process familiarity.** Rationale: The in-person voting process includes several steps that voters and election districts expect, such as checking in with a person at both the beginning and the end of each voting experience. These processes may increase both security and trust for voters. Process familiarity may also increase adoption rates among election districts.

In the next section, we present a design proposal that meets these requirements.

IV. Secure, Accessible, Virtual Voting Infrastructure (SAVVI)

The SAVVI system starts with the design rationale of analysing concerns in using private computers for voting. It then considers what can be done to address these concerns for disabled or UOCAVA voters. We propose defense–in–depth with synchronized multi–factor authentication (MFA) and ballot–integrity verification using Commercial Off–The–Shelf (COTS) security techniques in a 6–step protocol. These

steps are listed briefly here and described in detail in the remainder of this section.

1) Step One: Voter Registration. Voters register for electronic voting using their personal device.
2) Step Two: Check-in. Voters corroborate setting up their voting system with a synchronized telephone or video-conference connection with human volunteers at a check–in desk.
3) Step Three: Voting Machine. Voters log in to a secure, virtual voting machine (VVM) to mark, secure, and submit their ballot.
4) Step Four: Check-out. Voters establish a second synchronized telephone or video–conference connection with human volunteers at a check-out desk.
5) Step Five: Ballot Tally. The system decrypts the message and tallies the ballot.
6) Step Six: Vote Certification. The system publishes a record of the ballot in a way that does not disclose the voter's identity or the ballot's contents.

A. Step One: Voter Registration

Voters must pass a test of "humanness", such as by completing the recently proposed No Nonsense CAPCHA perceptual recognition and response puzzle or suitable alternative[38].

Voters have to share something about themselves, such as their voice audio or a video of their face, stating their name, address, and phone number. A critical part of this might be the simultaneity of interacting with a real person as part of the step.

The electronic voting registration database should also use a hash of a biometric signature such as a "voice–print" or "face–print". Instead of storing the biometric signature itself, a cryptographic hash is a pre–image resistant function that can be used to verify the integrity of a file and its contents. There are many off–the–shelf voice–print and face–print systems. There exist other promising techniques that may allow specific accommodations for specific disabilities, including mobile device-based hand recognition [57]. Also, several ear biometric methods may achieve high accuracy [53]. Biometric data can be practical, using off–the–shelf components that some researchers say can add `58` bits of security, even when "assuming the attacker knows one of the fingerprints of the user" [18]. While valuable as corroborating factors, privacy and public acceptance considerations direct us to not depend entirely on such biometric techniques [27], [22], [34]. We recommend simultaneity with a human verifier during the check–in process, described in Step Two. Furthermore, we consider privacy–preserving synchronization to assure that a voter's identity will not be revealed during this process because they are never directly identified alongside their ballot, which is described in Step Three.

At the conclusion of this step, the voter receives and retains their *voter registration identification number* (`PIN1`) through a separate communication channel, such as the U.S. Postal Service, as voters get registration verification today. The voter will use this in the voter check–in process. A signature of the device could be captured during the registration. The device signature might augment multi–factor authentication, which is described in Step Two.

B. Step Two: Check–in

As with requirements to show up with credentials at governmental offices the world over, proving you are who you claim to be and that you are authorized to vote remotely must include strong authentication. This step begins with another test of humanness and multi–factor authentication (MFA) at a virtual check–in desk. Such tests of humanness are central to the polling place experience and can be an important component in authorizing remote voting. Against practical threat models directed against off-the-shelf MFA, using two or more factors of MFA provides at least `58` bits of security [18]. The following MFA factors can help assure that an actual registered voter who has not previously voted in this election is obtaining a ballot and retains the ability to submit it for counting:

- **Something you have.** The check-in desk will verify the voter's *device signature* against the one captured during the initial voter registration in Step One. This corresponds with Alaca's findings on use of device signatures as an MFA augmentation technique [10]. If this fails, or as added proof, voters will demonstrate proof of residence in accordance with their jurisdiction's requirements.
- **Something you know.** Voters will provide their *voter registration number*, `PIN1`.
- **Something you are.** The check-in desk will verify the voter's *biometric signature*

Following successful authentication, the check-in volunteer follows instructions that automate deployment of a virtual container and whitelist the device and its IP address. The check-in volunteer then directs the voter to establish a secure connection to access the virtual container. The worker at the check-in desk then provides the voter a *ballot identification number* (`PIN2`), which the voter will use to authenticate with the voting machine in Step Three.

C. Step Three: Voting Machine

The system dynamically provisions a unique voting machine. This allows for the separation of personal registration information obtained in Step One and confirmed in Step Two from the private voting data generated in Step Three. The separation of personal identifiers from voting data is especially important in light of criticisms such as those directed against the accumulation of personally identifying information within the Voatz system [50].

1) Virtual Voting Machine.: Connecting to the internet and browsing the web without protection can compromise a device. The creators of secure software rely on computers to be clean of malware or cleaned before loading their software. A fresh browser 'voting system' install and updated security patches is a way to ensure there are no harmful extensions or plugins, but is not in itself adequate to protect a voting activity. There

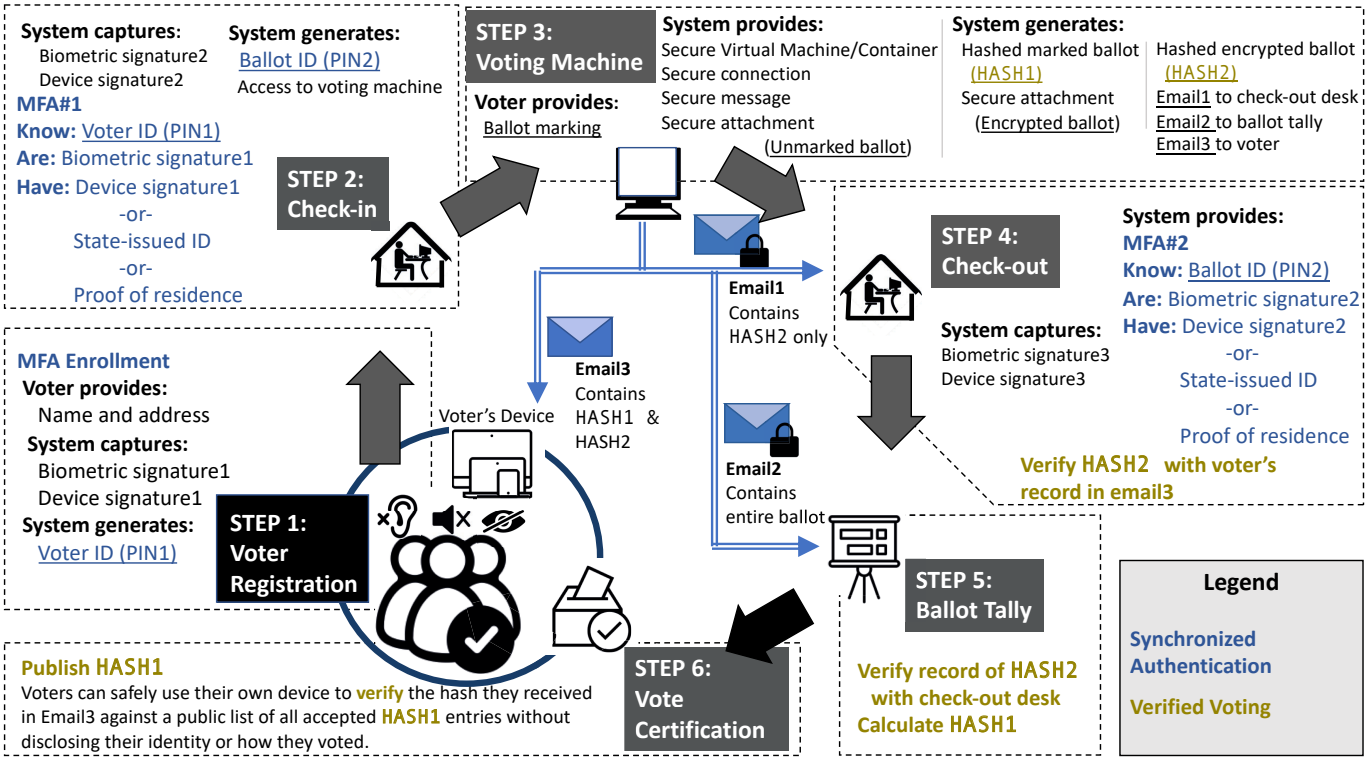


Fig. 2: Secure, Accessible, Virtual Voting Infrastructure (SAVVI)

may remain vulnerabilities in the underlying system that runs the browser.

Clean and secure implementations of browser software can solve some problems presented in today's insecure and inadequate vote-by–email procedures, such as in UOCAVA. However, SAVVI's use of a Virtual Voting Machine is intended to further secure these vote–by–email systems. It's important to note that the approach is made with corrupting possibilities in mind, as they are described in the 2018 National Academies of Science, Engineering, and Medicine (NASEM) report titled "Securing the Vote: Protecting American Democracy" [40]. The NASEM report calls out the possibility that a system can be corrupted at any layer and could flip votes. A protection against this in SAVVI is the dynamic provisioning of a Virtual Voting Machine upon voter check–in, and destruction of that machine upon voter check-out. The reduction of temporal attack surface makes it unlikely that any exploit chain originating on the user's device will be able to change the vote before check–out. Furthermore, different voters will use different systems on different networks. A system that flips ballots has to be aware of the correct ballot for the voter in the local machine used by the voter. As race choices change, some jurisdictions define their final ballot in the last day before providing it. For example, Los Angeles County provides something like 3000 different ballots to voters. It is hard to see wholesale power in creating specific code for changing particular voting choices.

Though it is not required for our system, and may not be immediately available as a COTS component, we envision a virtual voting machine (VVM) running on a secure virtual machine or virtual container that contains a popular, patched desktop. An example is a Security-Enhanced Linux container running a modern Windows Desktop. Clear step-by-step instructions could be displayed on the desktop background or in a simple graphical user interface. Also, to maximize compatibility with modern access technology for the blind and deaf, the secure VVM should automate the creation of a secure connection, a secure email, and presents a way to produce a secure, usable, and verifiable ballot attachment. Furthermore, the dynamic provision of the VVM will automate the correct installation and proper use of security software that requires administrative privilege on the VVM operating system, as in the VeraCrypt software recommended in the *Encrypt* section below.

A key advantage to using a virtual container is that each voting machine can be constructed in real time through common infrastructure automation tools like *Terraform*[2] and *Ansible*.[3] This makes provisioning a unique voting machine scalable, which greatly simplifies the verification of system security and usability for non–experts at the check–in desk. An open source repository of secure voting machine containers, virtual machines, and *Terraform*/*Ansible* scripts/playbooks will

[2] Terraform is an infrastructure compiler that translates program commands into pre–configured infrastructure. https://www.terraform.io/ [3] Ansible is another infrastructure compiler that specifically integrates IBM's *Red Hat* Security–Edition Linux containers. https://www.ansible.com/

simplify configuration. Furthermore, the dynamic creation and destruction of each machine in the span of only minutes presents reduced attack surface for a malicious actor seeking to compromise each voting device.

We recognize that the use of any phone or computer for secure communication does require testing, setting up, and rechecking. To address this concern, we recommend the voter accesses the VVM through a virtual private network (VPN). Several off-the-shelf VPN technologies exist and VPN applications are well-integrated with modern browsers [52]. If voters are unable to configure a VPN, there is still potential for secure connection using their favorite web browsers.

2) Secure connection.: On the secure machine, voters choose between secured ballot return channel options. This should include secured versions of popular email services running on popular browsers. A browser–based email service can enforce Transport Layer Security (TLS), a widely-adopted encryption scheme that adds security to websites and browser-based email. It uses public key encryption to secure the channel using a shared secret for each session of the connection. Several mechanisms can be used to secure the channel, and several are built into popular email–handling systems like *Gmail* and *Outlook*. The TLS 1.3, for example, "allows client/server applications to prevent eavesdropping, tampering, and message forgery over the internet" [43]. In 2017, Cremers, Horvat, Hoyland, et al. published a proof of the symbolic model of the TLS 1.3 [20]. The minimum effective key strength of TLS 1.3 implementations provide `128` bits of security [13]. The symbolic model of TLS 1.3 is proven, secure, and automated [15]. Still, there have been concerns about the implementation of the protocol, since insecure browser configurations can create susceptibility to TLS version–downgrade attacks, which then allow attackers to exploit vulnerabilities of previous TLS versions[16]. As of this writing, the National Vulnerability Database (NVD) indicates all the vulnerabilities specific to TLS 1.3 have been patched [1]. Ongoing attention to the NVD and maintenance of the secure container's contents will ensure that election managers avoid future security issues. Furthermore, we propose that a novel vulnerability discovery during the voting period could be obviated by our defense-in-depth strategy.

3) Secure message.: Pretty good privacy (PGP) and secure/multipurpose internet mail extensions (S/MIME) are widely–adopted for email encryption and digital–signature. These technologies add a layer of security to the message itself, which can preserve integrity and confidentiality even if the channel is compromised. S/MIME works with Gmail, Outlook, and other email systems. The current version of S/MIME is 4.0, which facilitates digital signature for authentication, integrity, and non-repudiation of the email. It also defines a standard for mail encryption exceeding `100` bits using public key infrastructure[45].

4) Secure attachment.: Pre-loaded scripts to hash, encrypt, then hash again will minimize user error and supplement vote integrity.

Hash. Hashing is a one–way function that transforms a message (or file) of any length into a string of fixed length. It can be used for integrity verification to check whether the contents of a message have been altered. SAVVI requires a voting jurisdiction to use hashing to help ensure a ballot is not tampered with and can also provide voters with a way to verify that their votes have been counted as cast. This concept of vote certification is discussed below in Section 3.6 (Step Six). The most common hashing functions are built into common utilities provided by all major device types, as it is with Microsoft's `certutil` tool.

```
H(marked_ballot}) =: HASH1 =:
   SHA512(marked_ballot)
E(marked_ballot) =:
   randomized cipher cascade
H(E(marked_ballot) =: HASH2 =:
   SHA512(E(marked_ballot))
```

Fig. 3: Pseudocode for a hash-encrypt-hash algorithm

Sensitive federal information must verify integrity with a minimum 256 bit Secure Hashing Algorithm (SHA 256) [28]. We recommend use of a 512 bit hashing algorithm (SHA 512), which presents complexities greater than 2^{158} even against methods such as the boomerang attack[58].

The system sends a copy of `HASH2` to the check–out desk, described in Step Four below. The system sends a copy of the first hash of the un–encrypted ballot `HASH1` to the voter, which allows the voter to independently verify that their electronic ballot was properly recorded after the ballot was decrypted. The handling of these hashes during the vote–tally and vote–certification steps are described below in Sections Five and Six, respectively.

Encrypt. There are diverse approaches to encrypt attachments and best practices are almost certain to evolve. For example, Meijer and van Gastel found vulnerabilities in hardware–based encryption and "strongly encourage users to instead use an open source... audited encryption software solution" and specifically recommend *VeraCrypt* as an exemplary software-based encryption tool [37]. *VeraCrypt* extends the popular cryptographic system, TrueCrypt (2004–2011). [4] A security evaluation in 2020 found that VeraCrypt protects confidentiality of data in an encrypted volume [23]. As of early 2021, the suite offers five encryption algorithms, each with a minimum practical security of `100` bits, and ten multi-layer encryption combinations (called cipher cascades). VeraCrypt allows encryption with any of its fifteen algorithms. This creates a randomization opportunity for added security: a script placed on the voting machine would frustrate an attacker's attempt to guess the encryption algorithm as a step toward decrypting the message.[5]

The security of this system can be calculated as $\exp(\kappa + \min\{\kappa(l' - 2)/2), n(l' - 2)/l'\})$, where l = the number of rounds in the multiple-encryption cipher cascade, κ = the

[4] https://www.veracrypt.fr/code/VeraCrypt/ [5] For an example implementation on the *Windows* operating system, see here: https://github.com/JP3L/SAVI

security-bit strength of each encryption (we assume here that $\kappa \geq 100$), and where $exp(t) = 2^t$, and where $l' = 2\lceil l/2 \rceil$ is the smallest even integer greater than or equal to l, for all $l \geq 1$. [21] For our system, where $l = 1$ to 3 depending on which cipher cascade is randomly selected, the security will range between κ and $\kappa + min\{\kappa, n/2\}$, which is between `100` and $100 + 100 =$ `200` bits. That said, a cipher cascade is not reliably independent and we should not assume a full 200 bits of security for the system. Instead we consider the cipher cascade as a hedge against the possibility of a cryptographic flaw in one of the algorithms. To calculate this security in practical application such as our proposal, we assume that E() and E()' are two different block ciphers each with a 100-bit key and propose the combined security for two general ciphers is effectively 101 bits, as per the person–in–the–middle attack [51].

D. Step Four: Check-out

Vote logging takes place at the virtual check–out desk. The desk confirms the voter received the email generated from Step Three. The desk also compares the voter's email receipt of `HASH2` with the desk's email receipt of `HASH2`. This biometric signature can be captured as a verified voice or image excerpt from the phone or video feed as part of a simultaneous second-channel support.

E. Step Five: Vote Tally

During this step, the system verifies the record of `HASH2` with the check-out desk, decrypts the *encrypted ballot*, calculates the hash of the unencrypted *marked ballot* as `HASH1`, and relays `HASH1` to the next step for vote certification.

F. Step Six: Vote Certification

Vote certification is an emergent security property of the SAVVI system. Once the vote has been decrypted and tallied, voters receive a certification notice confirming their ballot has been counted, along with public, anonymous access to the published `HASH1`. This notice includes instructions to verify their vote by matching hashes with the email they received at Step Five. One can verify their vote by searching for their ballot hash on a public website. A statistical verification can also assure integrity; non–matches strive to reveal issues before the tally is certified.

V. Security Considerations

Multiplying the probability of uncoupled things gives the probability that both will happen at the same time. It is the diversity of the sources of security and the use of them together–not only each of them individually–that increases security in a multi-factor authentication (MFA). The SAVVI protocol further synchronizes two MFA events, check–in, and check–out to achieve an enhanced security effect. We do not claim full independence of these security systems, but we propose that these individual components will work together in our architecture to defeat multiple types of threat.

The attack surface of our system includes many individual components, including the operating system on the virtual voting machine (VVM), the hosting platform for the VVM, the builders and operators of those machines and platforms, the local machine of the user, and so on. This enhances a standing web application by assuming that no single system will be completely secure for any longer than a few minutes. We instead propose a resilient architecture that presents a narrow temporal attack surface for the actual voting machine.

Technique	Bits of security	Threat Addressed
MFA1	58	Spoofing
TLS 1.3	128	Person-in-the-Middle (session)
S/MIME	100	Person-in-the-Middle (email)
Cipher cascade	101	Person-in-the-Middle (ballot)
MFA2	58	Spoofing

TABLE I: Each of the techniques proposed here provides protection against a specific security threat in the electronic marking and submission of a ballot. The authentication problem known as Spoofing remains the most vulnerable to attack, with 58 bits of security.

To further analyze the security of SAVVI, we consider five realistic and high–impact attack models: denial of service, spoofing, phishing, shoulder surfing, and person–in–the–middle. These attacks might prevent votes from being cast or recorded properly and we describe here the vulnerabilities relative to each type of attack and consider the potential for SAVVI to mitigate those vulnerabilities.

A. Denial of Service

Typical concerns surrounding electronic ballot returns include Denial of Service (DOS) attacks. These attacks usually bring down a service by overloading it. This was especially important when everyone voted on Election Day. Much of the country has early voting, with ballots sent to them as far ahead as 60 days so DOS attacks may be less of a concern. Nevertheless, communications could be disrupted by malicious parties. Denial of service attacks are a significant threat to email-based voting systems. Attackers could flood election email servers with large amounts of illegitimate traffic. This could prevent emails from arriving, or make it difficult for officials to distinguish valid ballots. SAVVI's email server could defeat such an email flood by white-listing only emails sent from the dynamically–provisioned Virtual Voting Machine, and including the white-list access control list entry as part of the container's provisioning process. Once the ballot is cast and the voter checks-out, the email server would remove the white–listed entry.

B. Spoofing

Spoofing attackers present the credentials of another to act like them. Multi–factor authentication in the voter registration phone call and again at the phone call for check-in can defend against such an attack. The voter identification log using biometric signatures facilitates detection and provides a potential remediation of such fraudulent voting attempts. In particular, it is plausible that this three-pronged system could provide adequate assurance that there is an actual person

interacting with the check-in desk and that it is the same person who previously registered to vote with PIN1. SAVVI requires synchronized authentication in voter registration and check–in/check–out, which is likely to detect a spoof attempt. Furthermore, the system might recognize the voice-print as other than the voter and could route the attacker to a honeypot where the attacker demonstrates their attempted fraud and thereby could contribute to future security improvements to the elections infrastructure. Geo–locating the fraudulent caller might also facilitate arrest.

C. Phishing

Phishing attacks use malicious links or attachments in an email to steal a person's credentials or compromise their device. The procedures described in this document create a unique voting email accessible only through a secure virtual container. These systems become live once the voter verifies with an election official for one-time use. Though an attacker might guess the set of all potential email addresses and spray phishing emails across them, the secure virtual container could purge all emails from the inbox upon loading the email client. This is built into the Application Programming Interface of popular email clients.[6] This provides a severely restricted attack surface within the email address and client provisioned within the Virtual Voting Machine that is probably not susceptible to phishing.

D. Shoulder Surfing

A person physically collocated with the voter could view or coerce actions. That said, SAVVI voters may be better able to control the timing, location, and potential for interlopers to reduce this concern in the safety of their own homes. SAVVI's process of check-in includes opportunities for uncovering and also discouraging coercion. The check in process can explicitly require affirmation that the voter is alone and voting privately.

E. Person-in-the-Middle

Person–in–the–Middle (PIM) and eavesdropping are potential threats in internet communications and especially threaten unencrypted communications. Marked ballots show how an individual voted, and may sometimes contain sensitive personal information about the voter. Anyone with access to the infrastructure could read or even modify email messages. In particular, email servers often store messages for a short period of time before passing them on to the next server or to the intended recipient. System operators for these servers could intercept or modify emailed ballots. It is difficult for election officials to identify ballots that have been modified in-transit. Also, emailed ballots are at risk before and after they are sent to election officials. The process of validating a voter, then giving them a new email or other ballot delivery channel is designed to detect bad actors and PIM attacks. Additionally, the provisioning of a secure virtual container would automate the encryption of the webmail channel (TLS 1.3), the marked ballot that is attached to the email channel (cipher cascade), and the email message itself (S/MIME). These three layers of security are complimentary and probably produce a cumulative security benefit. However, it is important to remember that a theoretical security bit-strength can be obviated by a side channel attack such as an endpoint compromise. The dynamic provisioning of a secure virtual container that contains a one-time generated voting machine, though, reduces this attack surface to a brief window of opportunity that makes it unlikely for an external attacker to penetrate. Also, the diversity of voter–owned end point devices makes it impracticable for attackers to create a wholesale disruption. Finally, the hash digests allow election officials to verify the vote was both cast as intended and recorded as cast. The NASEM authors' concerns about a single compromised layer in the application, operating system, BIOS, microprocessor, disk drive firmware, or across the internet reflect points of vulnerability to a PIM attack. The defense–in–depth approach described here provides multiple layers of encryption and dynamic provisioning of Virtual Voting Machines to reduce attack surface and reduce the potential exposure to PIM.

[6] For example, gmail, as in https://developers.google.com/gmail/api/guides.

VI. Conclusion

The difficulty of protecting privacy while establishing accuracy and integrity has made electronic voting a special concern for security experts. Integrity, accuracy, and privacy are all problematic with paper ballots too. Usable security for UOCAVA and disabled voters can be achieved without requiring them to become security experts themselves.

This paper came out of an exercise in contemplating what might be done for voters who could not present a paper ballot to their jurisdiction while preserving privacy and independence. Throughout 2020, several states were grappling with how to support ADA requirements. Disability advocates invited proposing a secure, accessible voting infrastructure. We started by considering blind voters, who generally cannot fill out paper ballots independently and privately. Some physically disabled voters have the same problem. Many UOCAVA voters were sending ballots by regular, unsecured email. These issues inspired devising the Secure, Accessible, Virtual Voting Infrastructure (SAVVI) framework to allow election officials to implement a secure ballot delivery system without significant new procurement or programming activity.

SAVVI uses available technology that can be implemented by a voting jurisdiction with relatively standard technical competence. Advanced Encryption Standard, Transport Layer Security, and Secure Multi–purpose internet Mail Extension represent available and open–source components that can work together to protect information in transit. SAVVI automates the preparation and orchestration of these tools to provide increased usability and security. It also verifies the voter and their ballot transfer using multi–factor authentication. The system represents a practical application of strong authentication, secure email, and encrypted attachments to achieve better usability than traditional mail–in paper ballots.

Potential implementations of this system–and future research–should consider usability impacts of specialized

CAPTCHAS, biometrics, and secure Virtual Voting Machines. These all require careful analysis and consideration that are beyond the scope of the current paper.

Detection is a critical part of judging security problems. A second idea lies in our suggestion of honeypot voting systems to discover and even entrap attempts to defraud the system.

Secure Accessible Virtual Voting Infrastructure is a flexible architecture that can accept evolving security tools that can be used for each part of the process described in Figure 2. Experiments testing these ideas through pilot voting demonstrations will provide deeper support for the approaches.

VII. Acknowledgements

This research was funded in part by ongoing activities in the Eaton Cybersecurity SAFE lab at Rochester Institute of Technology's ESL Global Cybersecurity Institute. We would like to thank the attorneys at Brown, Goldstein, and Levy for introducing us and encouraging us to help make voting safer, more secure, and more accessible to homebound voters with disabilities. J.M. Pelletier would also like to acknowledge the ongoing support he receives from the *Ordo Praedicatorum*.

References

[1] National vulnerability database. https://nvd.nist.gov/vuln/search/results?form_type=Basic, Accessed 15 October, 2020 with the exact-match search term *TLS 1.3*.

[2] The braille literacy crisis in america, Mar 2009.

[3] Schumer releases survey suggesting ballots of one in four troops deployed overseas went uncounted in '08 election, may 2009.

[4] *Recommendations to improve accessibility for absentee voting among recently injured service members*. Task 7 final technical report edition, 2012.

[5] Clemson university: Research alliance for accessible voting, 2014.

[6] Electoral fraud, Feb 2021.

[7] Rhode island general laws title 17. elections 17-19-8.1. ballots for voters who are blind, visually impaired or disabled, Feb 2021.

[8] Voting accessibility, Feb 2021. https://www.eac.gov/election-officials/voting-accessibility.

[9] Ben Adida. Helios: Web-based open-audit voting. In *USENIX security symposium*, volume 17, pages 335–348, 2008.

[10] Furkan Alaca and Paul C Van Oorschot. Device fingerprinting for augmenting web authentication: classification and analysis of methods. In *ACM ACSAC '16: Proceedings of the 32nd annual conference on computer security applications*, pages 289–301, 2016.

[11] Stephen Ansolabehere and Charles Stewart III. Residual votes attributable to technology. *The Journal of Politics*, 67(2):365–389, 2005.

[12] Adam Badawy, Emilio Ferrara, and Kristina Lerman. Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign. In *2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, pages 258–265. IEEE, 2018.

[13] Elaine Barker. Recommendation for key management: Part 1 – general, May 2020.

[14] J Benaloh. Electionguard preliminary specification v0. 85. *GitHub. https://github.com/microsoft/electionguard/wiki/Informal/ElectionGuardSpecificationV0*, 85:62–70.

[15] Karthikeyan Bhargavan, Bruno Blanchet, and Nadim Kobeissi. Verified models and reference implementations for the tls 1.3 standard candidate. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 483–502. IEEE, 2017.

[16] Karthikeyan Bhargavan and Gaëtan Leurent. Transcript collision attacks: Breaking authentication in tls, ike, and ssh. 2016.

[17] Shuki Bruck, David Jefferson, and Ronald L Rivest. A modular voting architecture (" frogs"). 2001.

[18] Rahul Chatterjee, M Sadegh Riazi, Tanmoy Chowdhury, Emanuela Marasco, Farinaz Koushanfar, and Ari Juels. Multisketches: Practical secure sketches using off-the-shelf biometric matching algorithms. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 1171–1186, 2019.

[19] David Chaum, Richard T Carback III, Jeremy Clark, Chao Liu, Mahdi Nejadgholi, Bart Preneel, Alan T Sherman, Mario Yaksetig, and Filip Zagórski. Votexx: A remote voting system that is coercion resistant. *UMBC Student Collection*, 2020.

[20] Cas Cremers, Marko Horvat, Jonathan Hoyland, Sam Scott, and Thyla van der Merwe. A comprehensive symbolic analysis of tls 1.3. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1773–1788, 2017.

[21] Yuanxi Dai, Jooyoung Lee, Bart Mennink, and John Steinberger. The security of multiple encryption in the ideal cipher model. In *Annual Cryptology Conference*, pages 20–38. Springer, 2014.

[22] Fiebig, Tobias, Krissler, and Jan. Security impact of high resolution smartphone cameras. *Workshop on Offensive Technologies (WOOT)*, Aug, 2014.

[23] Fraunhofer Institute for Secure Information Technology (SIT). Security evaluation of veracrypt, 2020.

[24] 24 October 2003 Friday. Diebold memos disclose florida 2000 e-voting fraud, 2003.

[25] Dan Gillette, Shama Hoque, and Edwin Selker. Improving write-in candidate text entry for audio-only voting interfaces. *Journal on Technology & Persons with Disabilities*, 2014.

[26] Jonathan Goler, Edwin Selker, and Lorin Wilde. Augmenting voting interfaces to improve accessibility and performance. *Proceedings CHI 06*, 38(4):1113–1136, Apr 2009.

[27] Marta Gomez-Barrero and Javier Galbally. Reversing the irreversible: A survey on inverse biometrics. *Computers & Security*, 90:101700, 2020.

[28] Carol N Gorman. Dods policies, procedures, and practices for information security management of covered systems (redacted). Technical report, Department of Defense Inspector General Alexandria United States, 2016.

[29] Cole J. Harvey. Changes in the menu of manipulation: Electoral fraud, ballot stuffing, and voter pressure in the 2011 russian election. *Electoral Studies*, 41:105–117, Mar 2016.

[30] Tagg Henderson. Blindness: total loss of vision is rare, Apr 2012.

[31] Astead W. Herndon. Georgia voting begins amid accusations of voter suppression, Oct 2018.

[32] Tom Intorcio. Absentee voting an 'obstacle course' for military and overseas citzens. In *National Conference of State Legislatures*. NCSL, 2009.

[33] Douglas W. Jones. A brief illustrated history of voting, 2003.

[34] Jan Krissler. Jan krissler the biometrics acker. *Workshop on Offensive Technologies (WOOT)*, Nov 12, 2018.

[35] Chris Larabe. Blind, visually impaired mass.voters have a variety of options. *The Daily Free Press*, Oct 29, 2020.

[36] Stephanie Lecci. Talking voting machines, other equipment ensure voters with disabilities can cast ballots, Mar 2016.

[37] Carlo Meijer and Bernard Van Gastel. Self-encrypting deception: weaknesses in the encryption of solid state drives. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 72–87. IEEE, 2019.

[38] Hendrik Meutzner, Santosh Gupta, Viet-Hung Nguyen, Thorsten Holz, and Dorothea Kolossa. Toward improved audio captchas based on auditory perception and language understanding. *ACM Transactions on Privacy and Security (TOPS)*, 19(4):1–31, 2016.

[39] Murat Moran, James Heather, and Steve Schneider. Automated anonymity verification of the threeballot voting system. In *International Conference on Integrated Formal Methods*, pages 94–108. Springer, 2013.

[40] Engineering National Academies of Sciences, Medicine, et al. *Securing the Vote: Protecting American Democracy*. National Academies Press, 2018.

[41] Federal Voting Assistance Program. The uniformed and overseas citizen absentee voting act overview, Feb 2021.

[42] Andrew Regenscheid and Nelson Hastings. *A Threat Analysis on UOCAVA Voting Systems*. US Department of Commerce, National Institute of Standards and Technology, 2008.

[43] Eric Rescorla and Tim Dierks. The transport layer security (tls) protocol version 1.3. 2018.

[44] Scott Ruoti, Jeff Andersen, Daniel Zappala, and Kent Seamons. Why johnny still, still can't encrypt: Evaluating the usability of a modern pgp client. *arXiv preprint arXiv:1510.08555*, 2015.
[45] J Schaad, B Ramsdell, and S Turner. Secure/multipurpose internet mail extensions (s/mime) version 4.0 message specification. Technical report, RFC 8551, April 2019. https://doi. org/10.17487/RFC8551. https://rfc-editor . . . , 2019.
[46] Ted Selker, Dan , Gillette, Shama Hoque, Kate Liu, Minhy Pham, and Mike Vroomen. 'research in accessible voting report, Jul 2014.
[47] Ted Selker. The technology of access: Allowing people of age to vote for themselves. *McGeorge Law Review*, 38, 2007.
[48] Steve Sheng, Levi Broderick, Colleen Alison Koranda, and Jeremy Hyland. Why johnny still can't encrypt: evaluating the usability of email encryption software. In *Symposium On Usable Privacy and Security*, pages 3–4. ACM, 2006.
[49] Captain Smith and Samuel F. Eleven california counties are late in sending absentee ballots for 2012 primary. *Law Review*, 1256, 2012.
[50] Michael A Specter, James Koppel, and Daniel Weitzner. The ballot is busted before the blockchain: A security analysis of voatz, the first internet voting application used in {US}. federal elections. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 1535–1553, 2020.
[51] Bing Sun. Provable security evaluation of block ciphers against demirci-selçuk's meet-in-the-middle attack. *IEEE Transactions on Information Theory*, 67(7):4838–4844, 2021.
[52] Haydar Teymourlouei and Vareva Harris. Effective methods to monitor it infrastructure security for small business. In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 7–13. IEEE, 2019.
[53] Zhaobin Wang, Jing Yang, and Ying Zhu. Review of ear biometrics. *Archives of Computational Methods in Engineering*, pages 1–32, 2019.
[54] Alma Whitten and J Doug Tygar. Why johnny can't encrypt: A usability evaluation of pgp 5.0. In *USENIX Security Symposium*, volume 348, pages 169–184, 1999.
[55] Matthew Wills. Ballot position: It matters, Aug 2016.
[56] Carter Yand. Presidency hinges on tiny bits of paper. Nov 12, 2000.
[57] Yilin Yang, Yan Wang, Yingying Chen, and Chen Wang. Echolock: Towards low-effort mobile user identification leveraging structure-borne echos. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, pages 772–783, 2020.
[58] Hongbo Yu, Yonglin Hao, and Dongxia Bai. Evaluate the security margins of sha-512, sha-256 and dha-256 against the boomerang attack. *Science China Information Sciences*, 59(5):052110, 2016.
[59] Filip Zagórski, Richard T Carback, David Chaum, Jeremy Clark, Aleksander Essex, and Poorvi L Vora. Remotegrity: Design and use of an end-to-end verifiable remote voting system. In *International Conference on Applied Cryptography and Network Security*, pages 441–457. Springer, 2013.

JavaScript Library Version Detection

Vilim Pagon*, Bruno Skendrović†, Ivan Kovačević‡, Stjepan Groš§
University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia

***Abstract*—There are more than 1.6 billion websites today, and almost every one of them uses JavaScript libraries. Knowing that, it's very important to show problems that occur as a result of not paying enough attention to security, such as using outdated versions of JavaScript libraries, insecure libraries, and so on. This paper proposes an algorithm for JavaScript library version detection. The algorithm detects version of JavaScript libraries based on differences between neighboring library versions. It's designed in such a way that it can be run periodically and automatically on a server. The paper also presents results and efficiency of the algorithm on a smaller set of data collected from the Croatian Web space. The success of the algorithm in detecting the correct version is about 50%, and the range of probable versions is an additional 25%. From these results, i.e. the detected versions, we found that the JavaScript libraries used on the websites of the Croatian web space are not regularly updated. Limitations and also possible potential improvements to the algorithm are listed at the end of the paper.**

Keywords*—*JavaScript, version detection, cybersecurity, Croatian web space

I. Introduction

Due to the simplicity of using the Internet, the concept of security has become an essential part of the development of software products that use the Internet as an access and management medium. Vulnerabilities can occur at all stages of the creation process of a software product. In most cases, they are introduced unintentionally because they are not given sufficient attention during development. It is necessary to know what errors can occur in order to prevent this problem. Likewise, it is useful in this area to know how much damage bugs can do if they are abused, what potential threats could exploit them, and how to protect against said threats. An introduced vulnerability must be eliminated as soon as it is discovered so that it cannot be abused.

Security patches are one method of eliminating such vulnerabilities. They are used to update systems, applications or software by inserting code to close or "patch" security vulnerabilities. It is necessary to regularly update the JavaScript libraries used by the website to reduce their vulnerability and thus ensure the website's security. The mentioned practice should be standard for every developer, but it turns out that this is not the case [1]. The article shows how the number of requests for older versions of the library decreases very little, if at all, after the new version is released. Another part of the problem occurs when a patch is released, but the developers do not update the existing one, so the vulnerability in the code persists.

The aim of this paper is to present the developed technique for detecting the vulnerability of the website, i.e. the version of JavaScript libraries used by the website.

The paper starts with an introduction to the topic of the paper, and then the next chapter titled Related Work contains other papers and projects similar to this one. In the third chapter, the main terms and techniques used in the paper are introduced to facilitate the understanding of the text. After that, the idea of the developed detection method and its implementation are described in detail in the chapters IV. and V. The Experiments chapter presents the results of this detection method. The problems and limitations of the implementation are presented in the chapter VII. The last chapter is dedicated to final conclusions and possible improvements in future papers.

II. Related Work

Numerous scientific papers address the issue of static vulnerability detection. Most papers use variations of machine learning, deep networks, or a combination of static and dynamic analysis for detection. Some examples of scientific papers on this topic are Wang et al. [2] and Russel et al. [3]. Unlike the aforementioned scientific papers that focus on source code analysis of various programming languages, the analysis in this paper is based on JavaScript code. Papers that use vulnerability detection methods in JavaScript are Tripp et al. [4], Guarnieri et al. [5] and Kluban et al. [6]. This paper describes vulnerability detection in the JavaScript library used based on libraries with known vulnerabilities. Specifically, the developed tool detects the version of the library and determines the presence of one or more vulnerabilities based on some known vulnerabilities in JavaScript libraries. This method is performed based on the differences between neighboring versions of the original libraries downloaded from known CDN servers. Since this is a new method for detecting vulnerabilities in JavaScript libraries, it is difficult to compare it to previous methods, as after an extensive search of published papers, not a single paper was found that performs detection using differences between neighboring versions. Vulnerability detection usually uses some form of artificial intelligence, machine learning, or dynamic analysis, which is not the case in this paper. However, the obtained results, which will be shown later, can be seen as successful and show the whole method as potentially successful. There are also Chrome extensions that analyze the JavaScript libraries used on the current website, such as Wappalyzer [7]. The JavaScript library vulnerability repository used

in this paper is the publicly available Retire.js [8]. This is a project that has developed a command line scanner that can be used to determine dependencies with known vulnerabilities in a web application.

III. Background

The detection method described in this paper is based on static analysis. Static source code analysis is a method for testing the source code of an application without executing the code [9]. Static analysis involves observing how the system, i.e., the solution, works.

The first two steps of JavaScript library version detection use hash values. The library hash value is a value calculated based on the source code of that library and passed to the function that calculates the cryptographic hash value. The value returned by this function is unique for each string. One of the functions used is the cryptographic hash function SHA-256 [10] which returns a unique value for each string.

The algorithm receives JavaScript code as input. It is the source code of a JavaScript library. At the beginning, the library and the version of the library were unknown, so this library is called an *unknown library*.

In contrast, *known libraries* are original JavaScript libraries downloaded from known CDN servers and used as a reference for a specific library, i.e. its source code.

Modified libraries are libraries from which redundant or non-functional parts of the code have been removed. These may be blank lines, spaces, new lines, comments, etc.

The third step uses *identifiers* to identify the library. Library identifiers are any keywords, variable names, function names, etc. that are characteristic of a particular library and do not appear in any other library. They are used to identify which library it is. They are selected from known versions of JavaScript libraries by analyzing their source codes and comparing them with each other to determine the identifiers that are unique to each library.

When detecting the JavaScript library version, the presence of a large number of minified libraries should be noted [11]. They pose a problem for detection because the code appears only in several and often only in one line. To solve this problem, reduction to a common form (caconic form), i.e., deminification, is used.

The reduction to canonical form - deminification - was performed using the Python library *JSBeautifier* [12]. Since minification detection was not performed, each library was reduced to a common form, which solved the problem of minimal differences between libraries, such as blank lines or spaces. For the rest of the program, it was assumed that all libraries were formatted the same and ready for comparison.

IV. Method of Version Detection

This paper is focused on static detection based on the source code of an unknown JavaScript library. The detection algorithm tries to find out which library and version that source code belongs to. The tool receives a JavaScript code as input for which the library version needs to be determined. The output, i.e. the detection result, is 10 versions of the library with the highest percentage of detection. Once the library version has been successfully determined, vulnerabilities can be searched in a publicly available repository with a list of vulnerabilities. It is a *Retire.js* [8] repository which contains a list of vulnerabilities for 26 of the most popular JavaScript libraries. Vulnerabilities are described as parameters *atOrAbove* and *below* that show the version in which the vulnerability was discovered, that is, which is the last version where this vulnerability is present. Detection was developed in steps where each step is performed separately and in a predetermined order. This paper discusses the fourth step of detection, which requires the first three steps to be completed for successful operation. These steps are shown in Figure 1.

1) The first step of detection uses a comparison of the hash value of each unknown library with the hash values of known versions of the libraries in the local database. The Library version is found if the hash values match, otherwise continue with the second step.
2) The comparison is done over modified libraries. The libraries have been modified in such a way that the removed lines of code do not affect the functionality of the rest of the code. Some of them are comments, empty lines, etc. The hash values of those modified libraries are recalculated and the first step procedure is repeated.
3) The third step is very important for detection. In this step, library name detection is done based on the identifiers. Identifiers are all keywords, variable names, function names, etc., which are characteristic of a certain library and do not appear in any other library. Each library has its own list of identifiers that are unique to it. The unknown library is compared to that list of identifiers and the results show which library it is with a certain percentage that indicates certainty. The obtained results are stored in the local database and used in the next step.
4) Previously obtained results are used in this step. For detection in the fourth step, a database of known versions of JavaScript library source codes was used. This data was collected from a well-known CDN server called cdnpkg [13] and saved in the database. Before detection itself, it is necessary for database of known JavaScript libraries to be supplemented with the differences between their neighboring versions. The differences are recorded locally in the *old* and *new* lists in the database. The list named *old* contains lines that were completely removed in the newer version, while the list *new* indicates the lines that were changed or completely added compared to the version before. These differences are obtained using the Python library difflib [14], which returns the

lines that are different between the two versions. In order to detect the version faster, a combination of the detected library name and the difference between neighboring versions of that library was used.

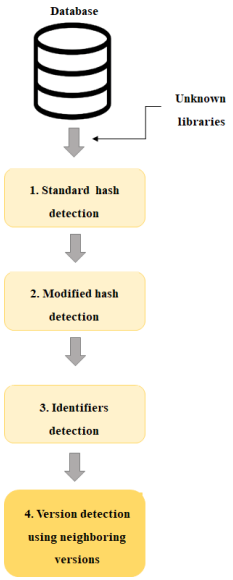


Fig. 1: Steps of detection

Each subsequent detection step is more detailed and determines the version of the unknown library more accurately. Detection with hashes gives the most accurate results but gives the least number of detections, while the fourth step is less accurate but gives much more results with the most accurate library version information.

A more detailed description of the fourth step is described in the next chapter. Previously detected library's name was compared with the new lists of each version of the library. The hit percentage is determined by the ratio of the number of found lines to the total number of lines in the *new* list. The detection results are also recorded in the local database with other results. Based on these results, it is possible to further continue with a more detailed detection.

V. Developing a System for Detection

A Python-based tool that automates the process was developed in order to detect the version of an unknown JavaScript library. The tool uses two MongoDB databases described in the next paragraph.

Two MongoDB databases were used in the detection process as can be seen in Figure 2. The first database contains the data on which the detection is carried out. This data was collected using a crawler in the laboratory and contains data from the websites of the Croatian Web space including HTML pages, JavaScript code used on those pages and some additional data [15]. In the following text, it will be called *Database with crawled data*. Second, the database was used to store results and auxiliary data in the process itself called *Results database*. Both databases are located on the private server where the detection is performed.

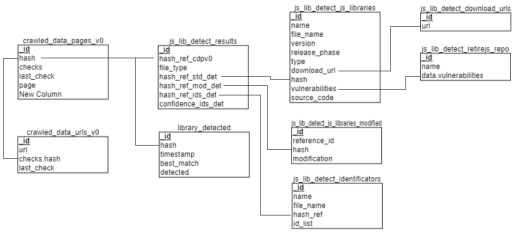


Fig. 2: Simplified structure of used databases

The first listed database, the *database with crawled data*, primarily serves as a database of collected data on which detection is performed, while the second one, *results database*, contains the data necessary for all steps of detection of an unknown library version. It also serves to store the results of individual detection steps for the purpose of further detection. The *database with crawled data* contains two essential collections. The first is a collection providing the source code that needs to be analyzed and the second is linked to the first in order to show which page it is about when the code is analyzed. That is shown in Figure 3.

In the *results database* there is a collection containing the results of all detections:

- standard detection with hash values
- modified detection with hash values
- detection by identifiers

A separate collection was created for the results of detections. It was made due to the different structure of the results.

There are also some auxiliary collections:

- URLs of known library versions
- codes of known versions of JavaScript libraries
- codes of known versions of JavaScript libraries - modified
- a copy of the Retire.js repository - vulnerabilities
- identificators



Fig. 3: Collections in the database with crawled data

Auxiliary collections show importance in various detection situations. Such as URLs of known library versions, so that subsequently the relevance of downloaded library versions is determined. Also, for detection by identifiers to be possible, it is necessary for those identifiers that are characteristic of a particular library to be stored somewhere. That collection contains lists of identifiers for each library and a hash value that represents the reference of the library to which the identifier detection result refers.

Each detection starts with a collection of known library versions and their source code. This collection is required at every step of discovery. It contains basic information needed for any detection, such as name, file name, version, release stage, type, hash, vulnerabilities, source code. The linked collection used in modified hash detection with hashes of known libraries is a collection containing the hashes of modified libraries.

To be able to continue with other more detailed detections, it's necessary to store the intermediate detection results in the database. All the results are in the aggregated collection of the results of all detections, which is linked to other collections as shown in Figure 2. It contains the results of standard and modified detection and detection using identifiers. The results are ordered according to hashes that represent references to the detected objects.

The created Python tool accesses the server's resources and uses the databases located on it at each detection step, which structure is shown in Figure 4.

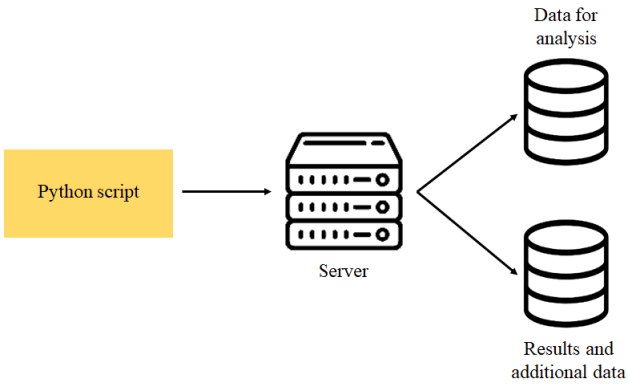


Fig. 4: System architecture

The version detection process starts by retrieving all JavaScript code from the collection with the results of previous detections. Filtering continues on the obtained records so that those libraries that were not detected by standard and modified detection using hash values remain in the process. Among the fetched records, those records detected using identifiers are added to the list of libraries to be analyzed. Furthermore, for each element from the previously created list, the source code is taken from the *database with crawled data* to which the record from the result collection refers. Each fetched source code of the fetched library is reduced to the canonical form using the tool Beautifier as well as the known libraries against which it is compared. Also, the library detected by the identifiers is retrieved to determine the name of the library with whose versions the algorithm must compare the unknown library. Once a library's name is determined, retrieve all versions of that library. And finally, the comparison process starts. Since the library's name is now known, it is necessary to determine the specific version. Lines of an unknown library reduced to canonical form are compared with a list of lines that did not exist in previous versions of each version of that library. The percentage of rows found determines the probability of a successful version hit. Based on match percentage, the top ten hits and the best hit are returned. After the comparison is complete, the results are saved in a separate collection created just for those results. Duplication is avoided by updating the existing document if it already exists in the library.

Detection results based on differences between neighboring library versions are saved in a separate collection. The reason for this is that the documents have a different structure than the previously described collection. This collection contains a hash field that points to the unknown library that was analyzed, a field that shows the most likely version of the unknown library, and a list of the top 10 possible versions. If the algorithm does not provide detection with 100% certainty, a more detailed analysis continues over the mentioned list of the next 10 most likely versions. If the library version is found, i.e. the algorithm is library version safe, using a collection that is a copy of the Retire.js repository, it is possible to determine if an already known library is vulnerable. A sample document from that collection can be seen in Figure 5.



Fig. 5: Sample document from the vulnerability collection

VI. Experiments

The first approach to version detection was to use hash values. The idea was to download as many source codes of libraries of different versions as possible and to calculate their hashes to store in the database. Retrieved hashes were used in comparison with hashes of unknown libraries. This approach did not succeed due to the small percentage of presence of original libraries. One of the reasons for weak

hash detection is the ratio of the number of plugins to the number of original libraries. The number of plugins used on websites is significantly higher than the number of original libraries. It is very important to emphasize that at that stage plugins were not part of the database of known libraries. Another reason is source code that has been modified by a third party in order to personalize or change the functionality of the libraries themselves. The percentage of detected versions of unknown libraries using this method was approximately 1%, so it was necessary to find a new method of detection.

Greater success was achieved by combined detection using library-specific identifiers and differences between neighboring versions of known JavaScript libraries. The procedure consists of filling the MongoDB database with *new* and *old* lists and then comparing the unknown library with the list *new*. The goal is to determine the percentage of matches with previously known libraries.

After describing the detection procedure, the results and statistics of the obtained results is presented. The algorithm was run on 16804 unknown libraries and each one was detected with a percentage of confidence. Results are divided into four important groups through a more detailed manual analysis:

1) Version match (100% match)
2) Probable version hit (>=50% match)
3) Probable range hit (>=10%, <50% match)
4) Version miss (<10% match)

The results are viewed as satisfactory, as can be seen from the percentages of each group. From the results, it can be seen that in the "Version match" group, i.e. one hundred percent detection was achieved for a quarter of unknown libraries. The second group “Probable version hit” is also important. It mainly contains correctly detected versions of the libraries, but due to changes, they have a lower hit percentage. If the first two groups are considered as correctly detected versions, a hit percentage of about 50% is achieved. This means that for half of the unknown libraries, more precisely the JavaScript codes that were analyzed, the correct version is found. The next group is "Probable Range Hit". In this group there are libraries that were detected with a certainty greater than or equal to 10% and less than 50%. Due to the very structure of the vulnerability repository, it is not necessary to know the exact version of the library to detect vulnerabilities, but it is enough to determine a smaller range in which the exact version could be located. In our case, that range is the 10 most likely versions of the libraries that the algorithm returns as the best 10. About 20% of the libraries are in the mentioned group which, together with the previous two groups makes up more than 70% of all given libraries. Libraries from this group require further analysis if the correct version is to be discovered, as they are too modified to be discovered this way. In the last group, i.e. "Version miss", there are libraries that the algorithm could not detect, i.e. no information is known about their version. Manual analysis found that the JavaScript codes in this group are often not even JavaScript libraries, but ordinary JavaScript codes found on a web page. Detection statistics and their graphic representation are shown in Figure 6 and Figure 7.

```
100% matched libraries: 4951/16804
[90%, 100%> matched libraries: 802/16804
[80%, 90%> matched libraries: 1211/16804
[70%, 80%> matched libraries: 250/16804
[60%, 70%> matched libraries: 1718/16804
[50%, 60%> matched libraries: 269/16804
[40%, 50%> matched libraries: 256/16804
[30%, 40%> matched libraries: 497/16804
[20%, 30%> matched libraries: 2386/16804
[10%, 20%> matched libraries: 1214/16804
[0%, 10%> matched libraries: 3250/16804
```

Fig. 6: Detection statistics

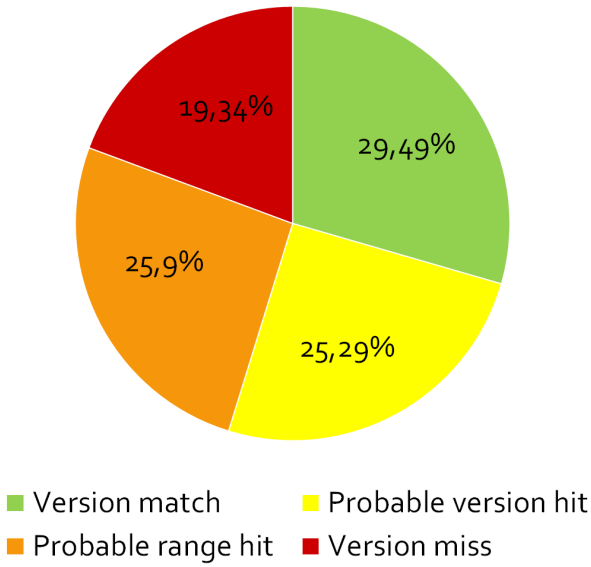


Fig. 7: Detection groups

VII. Discussion

Despite the relatively successful detection, the algorithm encountered some problems and limitations.

The algorithm itself is very complex. More precisely, its complexity is exponential $O(n^2)$. Even with additional optimization, the algorithm cannot be significantly speeded up on a large amount of data such as the database used in this paper. The greatest speedup is achieved by binding detection to identifiers, which reduces the range of libraries to compare. The range of libraries in the entire database, with slightly less than 300 different libraries and plugins, came down to exactly one library. Even after introducing this discovery step, execution time is still high. The algorithm took about 34 hours to discover 16804 unknown libraries, which gives us about 7 seconds per library.

One of the problems is the large number of unknown libraries that remain unknown due to the modification of the original libraries. They are made by developers in order to improve or adapt certain functionalities to their needs. To understand how much this affects the results, the fact that if the line differs by only one character, the line will not be detected is sufficient. Which results in a reduced hit percentage. Additionally, due to modification of the original library, it may happen that the vulnerability is fixed or a new one is introduced. If the vulnerability is fixed, detection will still report that the library is vulnerable. On the other hand, if the modification introduces a new vulnerability, due to developer carelessness or insufficient attention to security, detection will not occur. Some possible reasons for making changes are to delete parts of the code that will not be used and to add or rearrange code that complements and changes the functionality of the library to suit your needs. Also, one possible case is merging several libraries into one file in order to personalize it under its own name. An example of the modifications is the jQuery library [16], which contains more than 20,000 "forks" on its GitHub repository.

Another reason is that minified and obfuscated libraries are often used. As for minified, it's not that much of a problem because there are simple deminifiers. But if the library is obfuscated, such libraries will never be detected this way. There are very few of them, but they are present.

The next important problem that came after detecting a library version is determining vulnerabilities. There is no centralized database of known JavaScript library vulnerabilities. Even when the version is determined, if the library is not contained in the vulnerability repository it's not known whether it is vulnerable. This means that vulnerability detection is entirely dependent on the size and frequency of updating the vulnerability database. To identify vulnerabilities, the Retire.js repository was used but it contains only 26 of the most popular libraries. Therefore, a problem occurs when the detected library or plugin is not in the repository.

VIII. Conclusion & Future Work

When developing a software solution, it is very important to know potential vulnerabilities and to take care of the security of the program code at every step of development. Standard practice when developing any software solution is to use third-party libraries. The reason for this is to facilitate development if there is already a solution to some part of the problem and greater code security. That is why it is very important to monitor and update the versions of the libraries that the Web application uses in order to use the most secure version at all times.

The algorithm described in this paper automatically detects the version of an unknown library based on the library's source code. Furthermore, it can be used as a tool that analyzes large amounts of pages, looking for vulnerabilities that each page contains. Based on the detected vulnerabilities, warnings could be sent to Web site owners if they use a vulnerable version of the JavaScript library. Based on the detection results, it can be said that the algorithm detects the version of the unknown library very well. Nevertheless, this is only one technique of searching for vulnerabilities and should not be relied on alone but combined with already developed solutions.

Undoubtedly, there is a lot of room for improvement of the developed tool due to the limitations mentioned in the previous chapter. In order to improve detection, it is necessary to continue with new detection steps. Those steps should be even more detailed and continue detection of libraries that are not detected with 100% certainty. It is also possible to expand the database of known libraries with additional JavaScript libraries that are not currently in the database.

References

[1] Z. Bloom, "Javascript libraries are almost never updated once installed," Jan 2020. [Online]. Available: https://blog.cloudflare.com/javascript-libraries-are-almost-never-updated/

[2] Y. Wang, W.-d. Cai, and P.-c. Wei, "A deep learning approach for detecting malicious javascript code," *Security and Communication Networks*, vol. 9, no. 11, pp. 1520–1534, 2016.

[3] R. Russell, L. Kim, L. Hamilton, T. Lazovich, J. Harer, O. Ozdemir, P. Ellingwood, and M. McConley, "Automated vulnerability detection in source code using deep representation learning," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*. IEEE, 2018, pp. 757–762.

[4] O. Tripp, P. Ferrara, and M. Pistoia, "Hybrid security analysis of web javascript code via dynamic partial evaluation," in *Proceedings of the 2014 International Symposium on Software Testing and Analysis*, 2014, pp. 49–59.

[5] S. Guarnieri, M. Pistoia, O. Tripp, J. Dolby, S. Teilhet, and R. Berg, "Saving the world wide web from vulnerable javascript," in *Proceedings of the 2011 International Symposium on Software Testing and Analysis*, 2011, pp. 177–187.

[6] M. Kluban, M. Mannan, and A. Youssef, "On measuring vulnerable javascript functions in the wild," in *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security*, 2022, pp. 917–930.

[7] "Wappalyzer - identify technologies on websites," https://www.wappalyzer.com/, [Accessed 13-Jan-2023].

[8] "Retire.js - repository of known vulnerabilities." [Online]. Available: https://retirejs.github.io/retire.js/

[9] R. Bellairs, "What is static code analysis? static analysis overview." [Online]. Available: https://www.perforce.com/blog/sca/what-static-analysis

[10] J. Frankenfield, "Cryptographic hash functions: Definition and examples," Nov 2022. [Online]. Available: https://www.investopedia.com/news/cryptographic-hash-functions/

[11] "What is minification: Why minify js, html, css files: Cdn guide: Imperva," Dec 2019. [Online]. Available: https://www.imperva.com/learn/performance/minification/

[12] Beautify-Web, "Beautify-web/js-beautify: Beautifier for javascript." [Online]. Available: https://github.com/beautify-web/js-beautify

[13] [Online]. Available: https://www.cdnpkg.com/

[14] "Difflib - helpers for computing deltas." [Online]. Available: https://docs.python.org/3/library/difflib.html

[15] I. Kovacevic, M. Marovic, S. Gros, and M. Vukovic, "Predicting vulnerabilities in web applications based on website security model," in *2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 2022, pp. 1–6.

[16] "jquery," [Accessed 10-Feb-2023]. [Online]. Available: https://jquery.com/

Detecting JavaScript libraries using identifiers and hashes

Saša Lončarević, Bruno Skendrović, Ivan Kovačević, Stjepan Groš
University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia
{sasa.loncarevic, bruno.skendrovic, ivan.kovacevic, stjepan.gros}@fer.hr

***Abstract*—The web is a widespread platform for data exchange and service delivery on which many depend. Due to the high demand for website creation, methods are offered that provide more time-efficient design and programming. One of them is using JavaScript libraries, which can introduce vulnerabilities into web applications. Because of the sheer number of websites, as well as libraries and their vulnerabilities, it is difficult to maintain website security. There is a need to create tools that will automatically and proactively search for vulnerable libraries and enable timely remediation. This paper deals with the aforementioned problem and provides an insight into the implemented solution. The paper describes how JavaScript libraries are obtained and used, and proposes multiple methods for detecting them. It also presents the technical side of the implementation, as well as the results obtained by detecting libraries on a set of web pages from the Croatian web space. Finally, we discuss the observed distribution of popular libraries and their vulnerabilities, as well as the limitations of the proposed implementation, and offer potential solutions with the aim of improving the project.**

Keywords*—*JavaScript libraries detection, vulnerabilities, Croatian web space, cybersecurity

I. Introduction

The World Wide Web represents a large collection of websites that facilitate the exchange of information and provide important services to both companies and individuals. To make a website more dynamic and highly functional, developers use client-side JavaScript. Advanced features are commonly implemented through the use of JavaScript libraries. It is a modern solution that speeds up development of websites, but carelessness can give attackers an opportunity to abuse them.

For example, due to maintenance failures, websites can end up with out of date, and potentially vulnerable libraries. Vulnerabilities can be used to target both the websites and end users. Some of those vulnerabilities include cross-site scripting, input validation, stealing session data. Naturally, the need arises for a solution that will raise the level of security and prevent attackers from performing unwanted actions.

Therefore, the aim of this paper is to help and provide such tool, implemented in a way to detect JavaScript libraries using automated detection methods. Knowing the name and version of a library is valuable because it consequently gives the information about known vulnerabilities for the given library. Such tool would enable proactive action, so unsecure websites could be patched before potential attackers take advantage of them.

This paper is organized as follows. The introductory Section of this paper is followed by related research in Section II, where other papers and projects are discussed. Section III deals with the general understanding of using JavaScript libraries. In Section IV the theory of JavaScript library detection methods is presented. Section V describes the technical side of developing a system for detecting JavaScript libraries. In Section VI the conducted experiments and testing is described. Section VII contains the discussion about theoretical and practical limitations of implemented system. The paper ends with the conclusion of the topic and possible future work insights.

II. Related Work

While there are many prior studies about vulnerability detection, which focus on pure JavaScript code, such as Kluban et al. [1], Guarnieri et al. [2] and Shar et al. [3], there are not many that deal with detecting JavaScript libraries which are known to be vulnerable. Such approach could be of greater help in website security.

T. Lauinger et al. [4] conducted a comprehensive study of JavaScript library usage across popular web sites. While giving impressive insights about vulnerable libraries, some problems still arise during detection. Minimal modifications in library code, especially in minified versions, hinders detection. The source of the problem is attributed to many different, decentralized sources for obtaining JavaScript libraries. In comparison this paper demonstrates a way to handle minimal modifications in libraries and speed up detection.

Retire.js [5] is a project that deals with a similar matter. It is a tool that uses multiple static and dynamic methods to detect JavaScript libraries used in a website. It uses network traffic monitoring and regular expressions to detect libraries while this paper focuses on methods which are purely static. Another tool, in a form of a browser extension, is Library Detector For Chrome [6]. Strictly dynamic methods are used, detection is carried out by looking for known attributes of JavaScript libraries. This tool does not provide information about vulnerabilities. Content management system detector, Wappalyzer [7], can also detect JavaScript libraries by analyzing data from HTTP response or headers. Data gathered by afore-

mentioned tools, as well as our proposed solution, can be of great essence in developing website vulnerability prediction models such as [8].

III. Background

JavaScript is one of the most popular programming languages for creating web applications. Research from 2022, conducted on more than 80 thousand respondents, tells how over 65% of software engineers actively use it [9]. Frequent use of client-side JavaScript led to creation of libraries. Libraries represent a set of ready-made code that provides certain functionality, it is also sufficiently generalised so it can be used on larger set of problems. Some of them are built to access the HTML DOM (Document Object Model) [10] and introduce changes to it dynamically in specific situations. For example one of the most popular JavaScript libraries is jQuery [11], it allows developers to manipulate DOM and handle user-triggered events in a simple and fast way. Alongside jQuery, popular libraries are Bootstrap, Modernizr, React, Vue [12]. Apart from libraries, their plugins are also important because they can introduce vulnerabilities to websites as well. Plugins provide an additional functionality without requiring to modify the library. Due to flexibility they offer, they are widely used.

JavaScript libraries are not organized nor available in a centralized repository. The reason for this is that the libraries are mostly open source and made by different vendors, so owners publish them on different platforms. But many libraries are available on CDNs (Content delivery networks) [13]. CDNs are groups of servers, spread out over many locations, which store duplicate copies of libraries so requests could be answered as fast as possible. Common way to implement JavaScript libraries in websites is by using HTML tags like `<script src="">` and referencing a link to CDN resource.

Despite their popularity and widespread use, JavaScript libraries are susceptible to vulnerabilities. JavaScript is not necessarily an unsecure programming language, but with careless organization of code and disregard for good security practices the code may contain flaws that attackers can exploit. Research conducted on a set of popular websites has shown that more than 50% of websites contain at least one known vulnerability due to the use of unpatched JavaScript libraries [14]. Open Web Application Security Project (OWASP) publishes 10 most common web vulnerabilities each year, and some of them are also represented by JavaScript libraries [15]. A real example can be found in the jQuery library up to version 1.9.0 which provides `load()` function without properly validating input, which means executing JavaScript code is possible by passing it in a script tag [16].

IV. Detecting JavaScript Libraries

JavaScript libraries can be used in multiple ways, in different forms and thus different approaches can be used to detect them. Alongside fetching libraries from CDNs, they can also be stored on the same server as a website. The library itself can be in a certain form, depending on the need. Apart from the differences in versions there are differences in the way the library is packaged, specifically there are uncompressed, slim and minified types. The types are based on the idea of compressing code to make transfers over the network faster and reduce loading times. In production, minified code is preferred, i.e. code from which comments, empty characters and unused code are deleted in an automated way, and names of variables and functions are shortened [17].

A rough division in detection refers to static and dynamic methods. In general, static analysis refers to the study of source code without its execution, so in this case detection takes place without code execution. The method is often used in malicious code analysis because it is considered to be a safer approach than executing the code. On the other hand dynamic detection refers to recognition of code by its execution. Certain objects that the program creates and functions that it runs can be used to identify the library. What makes this type of detection possible is the absence of namespaces in JavaScript. Consequently, objects and functions are available globally in the code and can be called at runtime. Combination of recognized elements can be used to determine which library it is.

The paper implements two approaches of static JavaScript library detection. Those are detection by cryptographic hashes and detection by identifiers. A hash function is any function that takes an input string of arbitrary length and maps it to fixed length value. The idea of using hash functions comes from the need of truncating the strings being compared, specifically, instead of comparing the entire code of the JavaScript libraries, hash digests can be compared to see if they match. A prerequisite for performing detection by cryptographic hashes is a collection that contains all required JavaScript libraries and their calculated hashes. Then, for an unknown library we can compute a hash and look for a match in the collection. An assumption for using this type of detection is that the libraries are used in their original form, without any modifications. Reason for this is that hash functions don't tolerate even slight changes. This might be seen as a limitation, but procedures are implemented for reducing libraries to their canonical form, so slight modifications are not taken into account.

An alternative approach, in static detection, is detection by identifiers. It refers to extracting keywords from JavaScript code, where the keywords are identifier tokens obtained by lexical analysis. While collecting identifiers of a library it is important to take into account only those identifiers that do not appear in other libraries, because then they would not be unique, neither would they identify a given library. The assumption is that the libraries are relatively long scripts and that there are many identifiers, therefore the detection will be proportionally reliable. More identifiers means more diversity, which in turn suggest less chance of falsely detecting libraries.

The same can not be said for library plugins, because their code is often relatively short. For this reason, only libraries and plugins with significant amount of identifiers are capable of being reliably detected. To prevent false positive detections, libraries and plugins with short source codes are discarded.

This method, by its nature, is not resistant to situations where arbitrary JavaScript code that contains a large number of identifiers of some known library, is falsely detected as a known library. The disadvantage of detection by identifiers is the computational complexity. For larger library collections with many identifiers that must be compared to a multitude of unknown library identifiers a combinatorial explosion occurs. A large number of seemingly simple operations can drastically slow down detection, so it was necessary to consider more advanced measures during implementation.

V. Developing System for Detection

Developed software solution collects JavaScript libraries from CDN servers and builds a collection which serves as a basis for detecting unknown libraries. The same collection is supplemented with computed hashes and additional information about the libraries. The program implements three static detection mechanisms which are optimized and run on multiple processor threads.

The program is divided into units according to the functions performed. The system architecture and communication stakeholders overview is shown in Fig. 1.

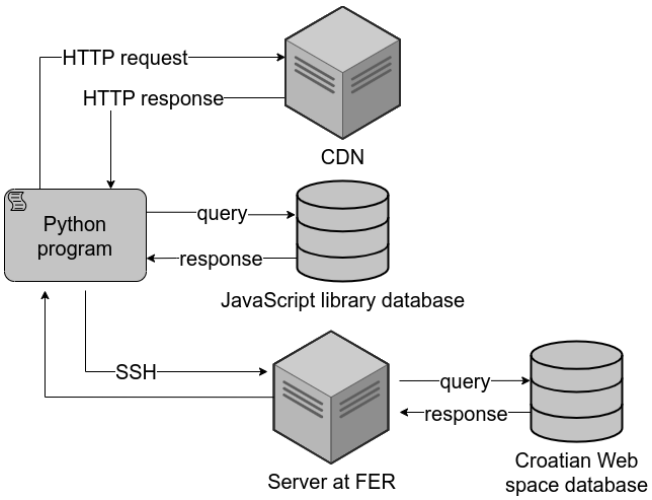


Fig. 1: System architecture components

While creating or updating the database, program communicates with CDN server and obtains library source codes, which are then stored and processed. The second part of communication is with FER server which provides access to a database of crawled web pages and used JavaScript libraries.

The database itself is based on 26 most used JavaScript libraries that have, or have had in past versions, publicly known vulnerabilities [5]. What enables automated library code fetching is a web scraper. It is programmed to retrieve all search results from CDNs for a given query, query being a library name. It collects all libraries and plugins related to the selected libraries and also all their versions and subtypes. All retrieved data is processed sequentially, for each request library name, type and version are extracted. Version semantics used is `major.minor.patch`, while types are organized by level of compression, so there are uncompressed, slim and minified types. Also, cryptographic hash digest is calcualted in advance to speed up the algorithm and remove redundant computation. If available, CVE references about library vulnerabilites are associated. This process generates a document as shown in Fig. 2.


```
_id: ObjectId('62532c1f4d7f5ed0619a31fe')
name: "jquery"
file_name: "jquery.js"
version: "3.3.1"
release_phase: "ga"
type: "uncompressed"
download_url1: "https://cdnjs.cloudflare.com/ajax/libs/jquery/3.3.1/jquery.js"
download_url2: "https://unpkg.com/jquery@3.3.1/dist/jquery.js"
hash: "d8aa24ecc6cecb1a60515bc093f1c9da38a0392612d9ab8ae0f7f36e6eee1fad"
timestamp: 1648991687.8593557
vulnerabilities: Array
  0: Object
    CVE: Array
      0: "CVE-2019-11358"
    summary: "jQuery before 3.4.0, as used in Drupal, Backdrop CMS, and other produc…"
  1: Object
    CVE: Array
      0: "CVE-2020-11022"
    summary: "Regex in its jQuery.htmlPrefilter sometimes may introduce XSS"
  2: Object
    CVE: Array
      0: "CVE-2020-11023"
    summary: "Regex in its jQuery.htmlPrefilter sometimes may introduce XSS"
source_code: "/*!
              * jQuery JavaScript Library v3.3.1
              * https://jquery.com/
              *
              * I…"
```


Fig. 2: Collection document example

The created collection is sufficient to perform a basic type of detection where only cryptographic hashes of libraries are compared, but for performing more advanced methods, it is necessary to further process the library data. To solve the problem of minimal changes in JavaScript libraries, they need to be reduced to a canonical form, i.e. parts of library code that are not crucial to the code itself are ignored. Precisely, all spaces, tabs, new lines and all comments are deleted. A new hash is computed from processed code and stored in the database. The same procedure is repeated on the unknown library during detection, for both sides to be equally reduced.

Furthermore, it is necessary to prepare data for detection by identifiers. The procedure includes tokenization of the library's source code, where identifiers are extracted by lexical analysis. These include all the names of variables, functions, classes and everything that identifies a certain object and is not specified by the syntax of the language itself. The resulting list of identifiers for each library must be compared to the lists of identifiers of other libraries in the database to remove all that appear in multiple different libraries. The goal is to keep only those identifiers that are truly unique to a given library. The entire described process of data collection and processing results in a database which contains 34 thousand documents. All available libraries, plugins, all their types, subtypes and version of multiple development stages are included.

The three static detection mechanisms extend each other and thus increase the number of detected libraries. The first method takes a cryptographic hash from an unknown library and searches for the same one in a collection of JavaScript libraries. Finding a match means that the unkown library was successfully identified, and has been used in its original form without any changes. This method guarantees absolutely accurate detection, that is, it guarantees that no false positive detections will occur. The amount of detected libraries depends on how complete the database is, meaning the more instances of libraries from different sources are contained the more libraries are going to be detected. The reason for this is that the various services which offer JavaScript libraries make slight changes, so the cryptographic hashes for virtually the same library will differ.

The second method refers to detecting libraries by using cryptographic hashes, but with implemented reduction to canonical form. It is an approach that alleviates the problem of minimal changes in the source code of the libraries. Unknown libraries go through the same procedure of reduction as libraries in the database. This method covers all the results generated by the first method, meaning, it will identify all the libraries as would the first method and more. The implementation and execution of this method comes with the price of complexity and longer runtime.

The third method is based on textual search with the aim of finding identifiers that determine individual libraries. By conducting lexical analysis on the unknown library identifiers are extracted and then the database is searched for matches. Whether the library is detected or not is dependent on the ratio of the number of detected identifiers to the total number of identifiers for a given library. Because of this, the method can not guarantee that false positive detection will not happen. In theory, it is possible that an unrelated JavaScript code contains all the identifiers of a library, and because of it, gets positively detected while it should not have been. During testing no such cases were observed. Results will show that this method works because developers follow good practices and assign identifiers that are meaningful and specific enough so false positives rarely occur. Again, due to the increase in complexity, this algorithm comes with longer runtime than the first two.

VI. Experiments

The experiment was conducted on 5 million documents collected by the web crawler from Croatian web space, which was part of previously conducted research [18]. The dataset consists of 93.03% of HTML documents, 3.41% JavaScript files, 1.98% CSS files, 1.58% other types of files. This file type distribution does not necessarily represent the actual distribution of file types on the web. It is crawler's fetching mechanism that causes it. Further testing is conducted only on JavaScript files, which there are over 170000 that make up the listed 3.41%.

While the developed tool is intended to analyze JavaScript libraries exclusively, besides the libraries, the dataset also contains a significant share of bespoke JavaScript code, i.e. code that is written for a specific website. There is currently no solution to differentiate libraries from other scripts, therefore the experiment is carried out over the entire dataset.

By performing the three detection methods, there were 966 libraries detected using the first method by comparing cryptographic hashes. Secondly, 2870 libraries were detected by reducing them to canonical form and then comparing cryptographic hashes. Thirdly, 16097 libraries were detected using detection by identifiers. As expected, the first method returns the least amount of results. Some of the reasons are minimal changes in the library code and the incompleteness of the database due to the large number of sources for obtaining libraries, where different providers make changes to the original libraries. On the other hand, second method with reduction to canonical form returns almost three times more detections, which proves that the assumed minimal changes are indeed present and frequent. A further assumption is that there are even more libraries with changes that could not be detected, however, there is no heruistic method which would provide information on their quantity. Lastly, the third method, relying on identifiers gives almost six times more positive detections than previous method. The advantage of this method is that minimal changes are irrelevant as long as identifiers are unchanged. The method has even higher potential for solving the problem of minimal changes. On the other hand, this method has the disadvantage that the exact versions of libraries can not be reliably determined, since not many identifiers are changed between versions.

The detection results show that certain libraries are more present than others. Fig. 3 shows the distribution of the 25 most represented libraries that were detected by using the second method with reduction to a common form.

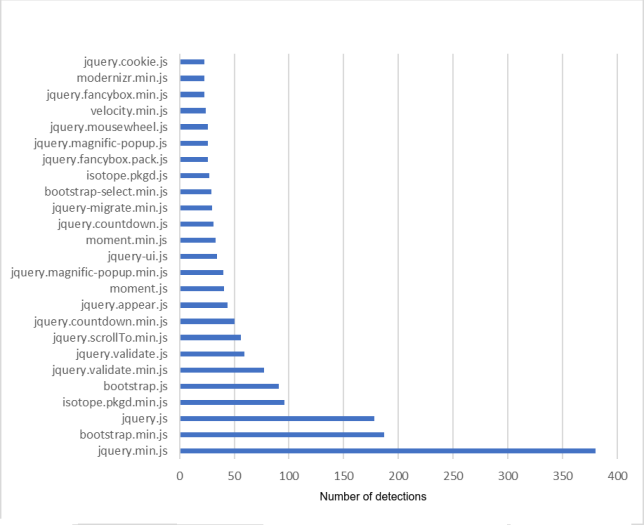


Fig. 3: Most represented libraries detected by reduction to common form and hash comparison

An important note is that minified libraries are used more often. This confirms the assumption from a paper on a similar topic [4] where it was the minified libraries that inhibited detection. Minification poses another problem in detection. Minification can be done by using different tools which will produce different minified versions of the same code, but even the same tool can, while shortening variable names, permute the names and thus brake detection.

Furthermore, detection by identifiers gives a slightly different distribution of detected libraries, the main reason for which is the limited set of data that serve as the basis for this type of detection. The method of extracting identifiers is suitable only for certain libraries, it is necessary that they have a considerable amount of the identifiers and that those are not generic names so that they can truly identify a given library. Fig. 4 shows a summary distribution of the most frequent libraries.

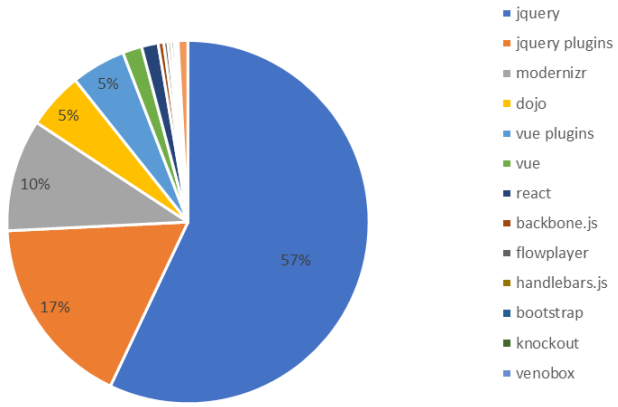


Fig. 4: Most frequent libraries detected by identifiers

The most recognized libraries throughout all methods of detection are jQuery and jQuery plugins. The accuracy of this distribution is also backed by research done on a selected set of websites [12]. It is confirmed that jQuery is certainly the most represented, followed by Bootstrap, Modernizr, React, Backbone. Unfortunately, past versions of these libraries have known vulnerabilities. Among the detected libraries, a non-negligible part of those has vulnerability records in the NVD database. Out of 966 libraries, detected using the first method, 165 (17.08%) of them are vulnerable. Out of 2870 libraries, detected using the second method, 636 (22.16%) is vulnerable. Third method does not provide reliable information about vulnerabilities because it can not determine the library version. Because of this the results rely on cryptographic hash detections. It turns out that jQuery is the most common vulnerable JavaScript library, specifically 80.82% of detected vulnerable libraries are versions of jQuery. A probable cause that lead to this result is that jQuery has over 80 publicly known vulnerabilities according to the NVD database [19], and only the last two versions of the library are considered secure, while all previous versions were proven to be vulnerable.

VII. Discussion

The implemented solution proves the validity of the theory in JavaScript library detection, but in practice certain limitations are imposed. To begin with, detection depends on how complete the database is, only those libraries that are contained in the database can be detected. The root of this limitation comes from the nature of hash functions that require identical inputs for detection to work. This would not be as big of a problem if JavaScript libraries were systematically organized on an official, centralized server. At the same time, all library instances would need to be identical, without minimal modifications. This is not the case, therefore the database should contain libraries collected from all platforms that are commonly used. Furthermore, the problem of minimal modifications introduced on the user side is algorithmically unsolvable. While some user based modifications follow certain patterns, there are plenty of unpredictable changes that are not easily extracted from library code. This problem passes on to detection by identifiers, even though this type of detection is more resilient to modifications, it still produced some false positive results. The solution was to declare libraries as detected only when they had a high percentage of matching identifiers. The percentage value was tuned until no false positive detections occurred, while giving the most results. Another limitation is the lack of organized information about vulnerabilities. Although many organizations invest a lot of effort to make vulnerabilities publicly known and available, the information is not as complete as such program would require. The implemented solution solves the mentioned limitations to a certain extent, but still leaves room for improvement.

VIII. Conclusion

The World Wide Web is of great importance in today's world and there is a need to raise its level of security. Vulnerabilities are proven to exist and are introduced in several ways, one of which is the use of vulnerable JavaScript libraries. A software solution that could detect libraries and determine whether they are vulnerable would be of great help. This paper deals with a static detection approach that detects JavaScript libraries based on comparing cryptographic hashes and identifiers. Experiments carried over the data from Croatian web space shows that many of the popular libraries are being used, with minified ones being preferred.

Detection results greatly improve when methods for resolving the problem of minimal modifications in JavaScript library code are implemented. First detection method is capable of detecting all the libraries that are used as-is. Second detection method implements reduction to a canonical form, and detects libraries despite minimal modifications. Third method is based on comparing identifiers and detects a larger share of libraries with greater amount of modifications. The implementation itself encounters certain limitations such as modifications in the library code and problems of obtaining all the necessary data, but to a

certain extent it successfully solves them. To improve the solution, combining different detection methods and using hybrid approaches is necessary.

REFERENCES

[1] M. Kluban, M. Mannan, and A. Youssef, "On measuring vulnerable javascript functions in the wild," in *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security*, 2022, pp. 917–930.

[2] S. Guarnieri, M. Pistoia, O. Tripp, J. Dolby, S. Teilhet, and R. Berg, "Saving the world wide web from vulnerable javascript," in *Proceedings of the 2011 International Symposium on Software Testing and Analysis*, 2011, pp. 177–187.

[3] L. K. Shar, L. C. Briand, and H. B. K. Tan, "Web application vulnerability prediction using hybrid program analysis and machine learning," *IEEE Transactions on dependable and secure computing*, vol. 12, no. 6, pp. 688–707, 2014.

[4] T. Lauinger, A. Chaabane, S. Arshad, W. Robertson, C. Wilson, and E. Kirda, "Thou shalt not depend on me: Analysing the use of outdated javascript libraries on the web," *Communications of the ACM*, vol. 61, no. 6, pp. 41–47, 2018.

[5] "GitHub - RetireJS/retire.js: scanner detecting the use of JavaScript libraries with known vulnerabilities." https://github.com/retirejs/retire.js/, [Accessed 13-Jan-2023].

[6] J. Michel, "Library detector for chrome," https://github.com/johnmichel/Library-Detector-for-Chrome, [Accessed 13-Jan-2023].

[7] "Wappalyzer - identify technologies on websites," https://www.wappalyzer.com/, [Accessed 13-Jan-2023].

[8] I. Kovacevic, M. Marovic, S. Gros, and M. Vukovic, "Predicting vulnerabilities in web applications based on website security model," in *2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 2022, pp. 1–6.

[9] L. S. Vailshery, "Most used languages among software developers globally 2022," Aug 2022. [Online]. Available: https://www.statista.com/statistics/793628/worldwide-developer-survey-most-used-languages/

[10] "Dom (document object model) - mdn web docs glossary," [Accessed 13-Jan-2023]. [Online]. Available: https://developer.mozilla.org/en-US/docs/Glossary/DOM

[11] "jquery," [Accessed 14-Jan-2023]. [Online]. Available: https://jquery.com/

[12] "Usage statistics of javascript libraries for websites," [Accessed 14-Jan-2023]. [Online]. Available: https://w3techs.com/technologies/overview/javascript_library

[13] "Cdn (content delivery network) - mdn web docs glossary," [Accessed 13-Jan-2023]. [Online]. Available: https://developer.mozilla.org/en-US/docs/Glossary/CDN

[14] T. Kadlec, "Open source security report - 77% of 433,000 sites use vulnerable javascript libraries," Nov 2017, [Accessed 13-Jan-2023]. [Online]. Available: https://snyk.io/blog/77-percent-of-sites-still-vulnerable/

[15] "Owasp top ten," [Accessed 14-Jan-2023]. [Online]. Available: https://owasp.org/www-project-top-ten/

[16] "Cve security vulnerability database - cve-2020-7656," [Accessed 13-Jan-2023]. [Online]. Available: https://www.cvedetails.com/cve/CVE-2020-7656/

[17] "Minification - mdn web docs glossary," [Accessed 13-Jan-2023]. [Online]. Available: https://developer.mozilla.org/en-US/docs/Glossary/minification

[18] E. Stambuk, S. Gros, and M. Vukovic, "Analyzing web security features using crawlers: Study of croatian web," in *2021 16th International Conference on Telecommunications (ConTEL)*. IEEE, 2021, pp. 142–145.

[19] NVD, "National vulnerability database jquery," [Accessed 17-Jan-2023]. [Online]. Available: https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=jquery&search_type=all&isCpeNameSearch=false

Detection and Analysis of Obfuscated and Minified JavaScript in the Croatian Web Space

Toni Dujmović, Bruno Skendrović, Ivan Kovačević, Stjepan Groš
University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia
{toni.dujmovic, bruno.skendrovic, ivan.kovacevic, stjepan.gros}@fer.hr

***Abstract*—JavaScript libraries allow for faster and easier programming of web content. In order to conceal know-how secrets and malicious code, obfuscation is used. Obfuscation is a deliberate act of reshaping something to make it harder to understand. Commonly used obfuscators provide many methods that produce different obfuscated versions of the same source code. Minification has similar techniques to obfuscation, but unlike obfuscation, minification reduces the size of the code, which speeds it up and is its primary functionality. This paper provides an overview of obfuscation and minification and the methods therein. The developed software tool uses regex, entropy and word size to detect and distinguish minified and obfuscated JavaScript libraries. The result of running the software tool on a database of pages in the Croatian web space is presented. The results show a high presence of minified and a small number of obfuscated JavaScript libraries. This automated detection has proven to be faster and in some cases more accurate than manual detection of obfuscation and minification. Observed problems with the tool implementation are commented on and potential improvements are discussed at the end of the paper.**

***Keywords*—Croatian web space, cybersecurity, obfuscation, minification, automated detection, JavaScript**

I. Introduction

Software security has become an increasingly important issue in recent years as more and more sensitive information is stored and transmitted online [1]. Obfuscation and minification are two techniques that can be used to increase the security of software by making the source code harder for attackers to understand or modify.

Obfuscation is a process that makes the source code more difficult to understand by transforming it into an equivalent but more complex form. The goal of obfuscation is to make it more difficult for attackers to reverse engineer the code, steal intellectual property, or find vulnerabilities. Minification, on the other hand, reduces the size of the source code, making it more difficult for attackers to find and exploit vulnerabilities.

Obfuscation and minification share similar techniques, which is why the results may look similar to the human eye or to detection software. Obfuscation can also be present in malicious programs, as it can hide commonly used functions or code patterns that an antivirus program would detect as malicious code. This paper describes a developed software tool that attempts to distinguish between source code, minified JavaScript, and obfuscated JavaScript.

In this paper, we make the following contributions:

- We introduce a developed software tool for detecting code minification and obfuscation in JavaScript libraries. The tool uses these parameters: entropy, word sizes, and regex.
- We demonstrate the effectiveness of the tool by applying it to a dataset of JavaScript libraries from the Croatian web space and analyzing the results.

The remainder of this paper is organized as follows. In Chapter 2, we provide a comprehensive overview of the existing literature on obfuscation and minification. Chapter 3 presents a detailed background on obfuscation and minification, describing some methods in the processes. Chapter 4 comments on methods for detecting obfuscated and minified JavaScript and discusses selected parameters for the detection. Chapter 5 presents the results of applying the implemented software tool to a set of JavaScript libraries used in the Croatian domain. Chapter 6 of this paper is a discussion chapter, where an analysis of the results of the study is made. Finally, Chapter 7 discusses the implications of our findings and discusses possible improvements for the software tool.

II. Related Work

Obfuscation and minification are techniques used to improve the security of software code. They make it more difficult for attackers to understand or modify the code, but can also be used by malicious actors to hide their malicious code. The detection and analysis of obfuscated and minified JavaScript has been the subject of increasing research in recent years [2], as JavaScript is one of the most commonly used programming languages in web applications [3].

Previous studies have focused on various aspects of obfuscation, its use in malware production, and the impact on detection of such techniques. W. Xu and coauthors [4] have conducted a comprehensive study on the use of JavaScript obfuscation in malware and have shown that obfuscation is a common technique used by attackers to evade detection by security tools. W. Xu's work highlights the importance of developing effective methods to detect and analyze obfuscated JavaScript, as this is a critical aspect of ensuring the security of web applications.

Other studies focused on the detection of such files, such as the work of YH. Choi [5]. The study evaluated the effectiveness of different detection methods using

parameters such as N-gram, entropy, and word size. The study found that the parameters were effective for obfuscation detection, but also indicated the need for additional parameters to improve classification accuracy.

One of the biggest challenges in obfuscated and minified JavaScript detection is that these techniques are constantly evolving. Obfuscation and minification tools and methods are constantly being improved and updated, making it difficult to develop a single, comprehensive detection method that is effective against all forms of obfuscation and minification.

Despite these challenges, previous studies have shown that it is possible to effectively detect obfuscated and minified JavaScript using a combination of lexical analysis, semantic analysis, and code similarity analysis. These methods have proven effective in identifying different types of code transformations that occur during obfuscation and minification.

There are also recent studies on the use of machine learning based approaches to detect obfuscated JavaScript, as in the comprehensive study by S. Aebersold [6]. The results of this study demonstrate that machine learning-based methods exhibit high accuracy in detecting code transformations. However, it is important to note that a reliable dataset is essential to achieve this goal. In addition, future research could explore the possibility of augmenting machine learning with other methods and using the parameters discussed in this work to provide more comprehensive code transformation detection.

In a recent study by A. Alazab [2], a machine learning method was developed to detect malicious JavaScript. The proposed solution used a total of 170 features, including statistical and lexical categories, to detect both obfuscated and unobfuscated malicious code with 98% accuracy. The malicious code samples used in the study often had one or more obfuscation layers. The main features used in the detection system were entropy, string length, number of occurrences of certain strings, and JavaScript keyword frequency. While the results of the study show the effectiveness of the selected parameters in accurately detecting malicious JavaScript, it is worth noting that further research with even better parameters or other machine learning approaches could potentially achieve even higher accuracy in detecting and mitigating security threats.

In summary, the literature on obfuscation and minification detection in JavaScript has shown the effectiveness of using various parameters such as N-gram, entropy, and word size in combination with and without machine learning-based approaches. These studies have provided valuable insights into the current state of the art in obfuscation and minification detection. However, further research in this area is essential to advance the understanding and development of robust detection techniques.

III. Background

Obfuscation is the deliberate act of reshaping something to make it harder to understand. In programming, obfuscation is the transformation of code into a version that retains the same functions but makes it extremely difficult or even impossible to understand, replicate, or modify without additional tools. Program code is often obfuscated to protect intellectual property or trade secrets. A notable disadvantage of obfuscation is its potential misuse to circumvent the detection of malicious programs by antivirus software [7]. Obfuscation serves as a defense mechanism against reverse engineering and is often used in the development of malicious code to intentionally obscure its understanding and increase its complexity. Antivirus programs typically scan code as part of their detection mechanisms for commonly used features and patterns that are often associated with malicious programs. With obfuscation, developers do not develop new malicious code, but place a layer of obfuscation over important functions to hide from detection.

Obfuscation uses various methods to achieve the desired result, including "Control flow flattening"," "Dead code injection" and "String Array." "Control flow flattening" is a technique that aims to streamline the code flow. For this purpose, all fundamental blocks of code,such as function bodies, loops, and conditional branches are broken down and reassembled in an infinite loop. A switch statement is then used to determine which part of the code will continue to execute. Such a structure makes it much more difficult to follow the execution of the program, since the usual structures that make the code easier to read are no longer present. Figure 1, titled "Control Flow Flattening," shows an abstract representation of the transformation of a simple code structure into a flattened version by manipulating the flow. The colored blocks represent the separated code blocks, while the new black block represents the main switch command that governs the flow.

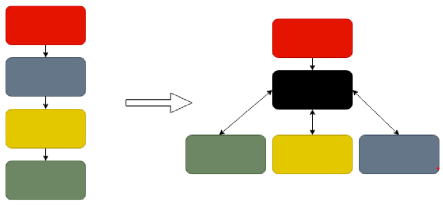


Fig. 1: Control flow flattening

The term "dead code" can be defined in two different ways. First, it refers to code that is never executed, such as a function enclosed in an "if" condition that is never satisfied. Second, it can refer to code whose output has no effect on subsequent code execution. Dead code can slow down execution time and consume system memory, making it difficult to understand the program. In the context of obfuscation, dead code is often inserted into source code to increase its complexity, complicate reverse engineering efforts, and increase the time required for deciphering.

Another method used to try to make the code more difficult to understand is to split strings into smaller parts and arrange them in arrays or array-like objects.

An example of this method would be splitting the word "classify" into "ssi", "fy" and "cla" and then storing the values in different parts of an array.

Minification is a technique used to reduce the size of code required to perform the same functionality. It is commonly used in the implementation of web pages and script files. Minification is an important method to optimize page loading times for users, as it speeds up page rendering. It also benefits users with limited internet connection as they consume less data compared to non-minified pages. The difference between the original and the minified jQuery JavaScript library version 3.1.1 is 176 kb [8]. Minification removes unnecessary or redundant elements from the source code, such as whitespace, comments, and naming conventions. This process results in more compact and efficient code that retains the same functionality as the original code.

After minification, the code becomes a less readable version that has the same functionality. For this reason, it can be said that minification is a type of obfuscation. Obfuscated and minified files are easy to distinguish from the files on which such operations were not performed, but in some cases it is difficult to distinguish minification from obfuscation. The goals of minification are different from those of obfuscation, but to minify the program code, some of the same methods are used. Some of these methods include shortening variable names, removing delimiters, and refactoring code. Shortening variable names reduces the size needed to store the code, refactoring code involves methods similar to "Control flow flattening".

IV. Detecting obfuscated and minified JavaScript

There are several ways to detect obfuscated JavaScript code. One way to detect obfuscated JavaScript code is to identify unusual or hard-to-read code patterns, such as long or overly complicated statements or unconventional naming conventions for variables and functions. Another approach to detecting obfuscated JavaScript code is to check for the presence of unnecessary or redundant code or statements, as this could be an indication of obfuscation techniques being used. In addition, developers can use a decompiler, a tool that can convert compiled code back into a more readable form, to understand and detect obfuscated code.

Minified code can also be difficult to read and understand since it removes variable names, comments, and block delimiters that serve to make the code readable but are not necessary for its execution. However, there are several tools that can help developers deal with minified code. One option is to use a code beautifier, which reformats and reorganizes the code to make it more readable. Another option is to use a source map [9], a file that maps the original, unminified code to the minified code, making it easier to debug and understand. Also, a debugger can be used to better understand the code and identify potential problems. In addition to the challenges posed by obfuscated and minified code, there are also security risks to consider. Malicious actors can use obfuscation and minification techniques to conceal malicious code, making it harder to detect.

In this work, the software tool initially categorizes JavaScript code into three categories: Minified Code, Obfuscated Code, and Source Code. This categorization is done through a separation process in which the tool identifies and distinguishes code that has been minified or obfuscated and code that remains in its original source code form. The initial separation of minified and obfuscated code from source code was performed using 5 parameters. The parameters increase the security index variable by 1 or 2, depending on the value given to the parameters from the given JavaScript. The specified parameters are:

- Regexes
- Entropy
- Size of the biggest word
- Average word size
- The ratio of the largest word to the number of characters of the entire file
- File size

After analyzing about 1000 obfuscated and minified JavaScript libraries crawled from the Croatian web space, regexes were written that could be used to detect the commonly used methods and patterns in their implementations. Some of the notable ones are regexes that detect parts of the code where variable names are used with one letter in a row and have been given values with one letter. This pattern could be used in minification to shorten variable names because the code no longer needs to be readable, or in obfuscation to shuffle values so that the code is harder to understand. Another regex was written for unusual return values that had been noticed in obfuscated code. Programmers would not normally overcomplicate code if it were possible to write it in a clearer way. For this reason, they would never write a return statement with only 3 Boolean values in it. In JavaScript, an empty array evaluates to false when computed as a boolean value, and the two exclamation points change the value to true and back to false, as seen in the second regex example. Since such code would never be written by a human, the regex was implemented to detect such patterns. Another regex was implemented to recognize a large number of "|" characters. In normal coding, such characters would not appear in the code, but in obfuscation, these characters are used in the string array method mentioned earlier, where strings in the code are broken into smaller pieces and stored in array-like structures. In this example, the large amount "|" simulates an array-like structure, where splitting the string by the "|" returns an array.

"[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,[a-zA-Z]=.,"

"return !!\[\]"

"\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|"

In computer science, entropy refers to a mathematical value that measures the degree of randomness in a set of characters. For JavaScript files, the standard entropy is usually 4.75, but after the process of minification or obfuscation, the entropy of a file tends to increase to around 5.1. By analyzing the entropy of a file using Shannon's formula [10], we can increase the probability of identifying obfuscated or minified code. The formula for calculating entropy is as follows:

$$H(X) = -\sum p(X) \log p(X)$$

The next three parameters analyze the word sizes in the code and compare them to word sizes normally found in source code files. After minification and obfuscation, many spaces, tabs, and other delimiters are removed from the code, resulting in larger words than in the source file. Following this logic also means that the average word sizes become larger as the word separators are removed. One problem that was encountered with this logic was that very small files can have large average word sizes, resulting in false positives. For this reason, the "largest word compared to the whole file" parameter was added. The last parameter contributes to the security index when the file size exceeds a certain threshold, as it was observed that larger files tend to be obfuscated, since JavaScript libraries do not usually exceed this threshold. For each parameter, an integer variable in the code would be incremented by 1 or 2, depending on the threshold value of that parameter. At the end of the classification, if the value is 4 or higher, the JavaScript is classified as minified or obfuscated. The threshold values for increasing the integer variable can be found at TABLE 1.

After separating the source code JavaScript from the obfuscated and minified JavaScript, the second round of classification begins, separating obfuscated and minified JavaScript. In the second round, a similar approach is taken as in the first round, but different regular expressions (regexes) are used and the thresholds for other parameters are adjusted to better distinguish between minified and obfuscated JavaScript code. This allows for a more accurate separation of the two types of code during analysis. Some types of obfuscation produce an exceptionally low entropy value, so a new threshold has been introduced when the entropy is below 3 to classify JavaScript as obfuscated. The regexes written for obfuscation detection are very specialized to the examples acquired during the research for the implementation of the software tool. In JavaScript, the eval() function is used to execute code written as a string. Some obfuscation tools may leave a signature such as "eval(function(p,a,c,k,e,r))" as a hint or indicator of deobfuscation. This signature can be used by analysts to identify and understand obfuscated code, as it often indicates the presence of code that is dynamically executed with eval(). Hexadecimal values can also be found in obfuscated code to change the values within the code without changing the values. This is not done in minification because it only increases the size of the code and does not increase the execution speed of the code. The regex for detecting hexadecimal values in the code can also be found below.

"eval\($function \ (p, a, c, k, e, r\)$"

"$\backslash b(0x[-fA - F]+) \ \backslash \ b$"

V. Experiments

The developed software tool was implemented in Python on a sample set of JavaScript libraries included in the Croatian web space. The source of the JavaScript libraries should not influence the written regexes because they were written according to commonly known methods of minification and obfuscation described in the previous two chapters. Only individual JavaScript libraries were used, as they were collected by their hash value. The tool is designed as a service that retrieves URL addresses from a database and retrieves their content using the Python request library. Once the content is retrieved, it is analyzed using six parameters as described in the previous chapter. The security index variable is then used to determine the type of the library based on the final index value.

Applying the software tool to a large number of JavaScript libraries revealed a significant flaw in the implemented solution. Despite its accuracy in detecting various forms of obfuscation, the solution suffered from a critical performance issue: catastrophic backtracking [11]. This problem occurs when a regular expression takes an exponential amount of time to match a string, resulting in a dramatic increase in processing time.

The cause of catastrophic backtracking was attributed to the use of complex regular expressions designed for a wide range of obfuscation techniques. However, these regular expressions also tended to match large amounts of irrelevant data, resulting in a large number of false positive matches. This in turn led to an exponential increase in the processing time required to match the regular expressions, resulting in catastrophic backtracking. One approach to solving this problem is to implement a timer during code processing or to skip very large files, since such files are more likely to be obfuscated. In such cases, only semantic and entropy analysis can be performed on these examples to optimize processing time. On an Intel(R) Core(TM) i5-1035G4 CPU @ 1.10GHz 1.50 GHz processor, the average processing speed for each sample was about half a second. However, for samples affected by the catastrophic backtracking problem, the processing time increased to five to ten minutes, resulting in slower classification.

The final test was performed on a set of 204842 JavaScript libraries crawled from the Croatian web space. The choice of the breakpoint for the security index had a large impact on the occurrence of false positives and false negatives when classifying libraries as obfuscated or minified. Selecting an appropriate breakpoint is critical to achieving accurate results and minimizing the occurrence of misclassification. In the first iterations of the tool, a value of 3 was used as the breakpoint for the classification security index, and there were no thresholds at which the security index would increase by 2. Using this logic, the

TABLE I: Security index increase table

	Increase by 1	Increase by 2
Regex	For each hit regex	
Entropy	Bigger than 5.05	Bigger than 5.3
Biggest word size	Bigger than 200	Bigger than 350
Average word size	Bigger than 11	Bigger than 45
Largest word compared to the whole file	Bigger than 10% of the whole file	
File size	Bigger than 1000000	

program contained a large number of false positives for source code JavaScripts that were classified as minified. If the analyzed library code contained an extremely long string, a function, or other built-in tools such as regex, these parts of the code contained very long amounts of text without spaces that would increase the security index by two in the case of false positives.

Within the libraries tested, there were also those that were so small that their largest word was 10 percent or more of the total size of the library, which increased the security index by one. Combining these two cases, the file is classified as obfuscated or minified, even though it is a source code library.

The tool proved to be better than manual classification in some cases. In certain cases where the code is so long that manual classification is difficult, the regexes were able to detect small functions that are obfuscated or minified and increase the security index, increasing the likelihood that the library will be classified as minified or obfuscated. These obfuscated functions in source code libraries usually contain know-how secrets or malicious functions that developers try to hide.

In this study, we present the results of a comprehensive analysis of a large sample set consisting of 204,842 randomly selected individual libraries. Our analysis revealed that a significant proportion of the samples were in either minified or source code format, while the proportion of obfuscated code was only 2% of the total set of examples. It should be noted, however, that the analyzed dataset was subject to potential bias because many identical files with the same hash value were treated as a single sample. Therefore, it is plausible that the actual number of minified files in the Croatian web space could be even higher than found in our analysis. This is especially true for commonly used libraries such as the popular JavaScript library jQuery, which may have the same hash value on different websites, leading to a possible discrepancy in the data. In summary, our results provide valuable insights into the prevalence of minification in the Croatian web as well as the prevalence of obfuscation in software libraries. This information could support the development of more effective tools for code analysis and optimization. The results are visually represented in the pie chart below.

After the experiment was completed, a study of the parameters of the classification system was performed. The entropy of the three file types was significantly different, with obfuscated code having the highest entropy of 5.51, followed by minified code at 5.23, and finally source

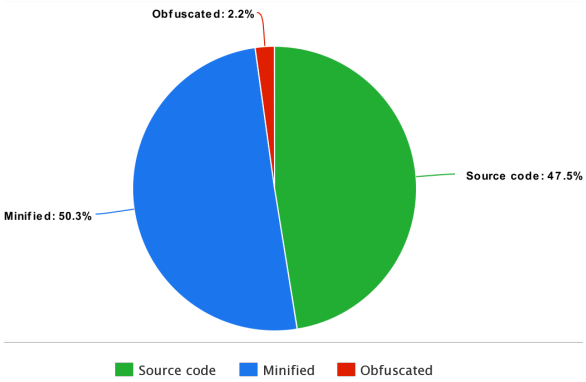


Fig. 2: Experiment results

code with an average entropy of 4.82. After manually analyzing some files that had an entropy of 5.3 or higher but were classified as source code files, it was clear that although the software had classified the files correctly, the files were written in such an ambiguous manner that the entropy was unusually high. The author of the code used few spaces, tabs, and new lines, and did not write the code in a way that made it easy to read, which increased its entropy. The code was also clearly not minified or obfuscated, as it followed other coding practices, such as clear and precise variable names, that are not followed in obfuscation or minification. When looking at the length of the longest word, it was found that the average length was 156 characters for source code, 2810 characters for minified code, and 18428 characters for obfuscated code. However, when analyzing the average word lengths, it was found that source code files had a length of 10 characters, minified files 116, and obfuscated files 69 characters. Although obfuscated files have a much longer largest word, the average word is shorter than the average minified word. One reason for this could be some functionality that creates very long words that are used in obfuscation but not in minification.

Some of the possible 3-pair parameters were examined for all file types. For source code files with entropy greater than 5.3, longest word size greater than 350, and average word size greater than 45, 2229 files were found. The best combination for minified files was entropy between 5.05 and 5.3, biggest word size greater than 350, and average word size greater than 45 with 20762 file matches. The combination with the most samples for obfuscated files was the same as for source code files, but there were 1170 obfuscated files.

Overall, the study provides valuable insight into the

characteristics of source code, minified code, and obfuscated code, as well as parameters that can be used for effective classification. These insights can be used to improve the performance and accuracy of code classification systems, leading to better detection of such methods.

VI. Discussion

The results of our study show that the implemented obfuscated JavaScript detection software tool effectively identifies the different types of code transformations that occur in software libraries. The results show that 45% of the libraries analyzed in the study were source code libraries, 53% were minified, and 2% were obfuscated.

The results are consistent with the expected trend that a significant portion of software libraries are minified, as minification is a widely used technique to improve the performance and security of software code. Nevertheless, the relatively low incidence of obfuscated libraries is unexpected, considering that obfuscation is also an important technique for bolstering software security and can be used for both legitimate and malicious purposes. It is important to conclude that even if obfuscation does not occupy a large portion of the JavaScript library ecosystem, even the 2% of these libraries can do great harm if they are malicious and will not be detected if not protected against.

It is important to note that the results of our study are specific to the dataset and software libraries analyzed, and may not be representative of the overall distribution of minified and obfuscated code in real-world applications. Nevertheless, these results provide valuable insights into the state of software security and the prevalence of various code transformation techniques.

The results of our study highlight the importance of using code transformation detection tools, as this can help identify vulnerabilities and improve the security of software applications. Our study also highlights the need to continue research on minification and obfuscation detection as these techniques continue to evolve and new methods emerge.

In summary, the results of our study demonstrate the effectiveness of the software tool for detecting minified and obfuscated JavaScript and provide valuable insight into the state of software security and the prevalence of various code transformation techniques. These results can inform future research on minification and obfuscation detection and be used to improve the security of software applications.

VII. Conclusion & Future Work

Obfuscation is the deliberate act of reshaping something so that it is harder to understand. It is used to keep secrets such as important code functions and to prevent the detection of malicious code. Minification involves reducing the amount of code, which is often used when implementing web pages and script files. Obfuscation and minification are two very similar processes that use a large amount of the same methods in their implementations.

Classifying obfuscated, minified, and source code libraries by hand is a straightforward but time-consuming process. This is because obfuscated libraries are on average more complex and therefore larger. It has been shown that using methods that examine word size and similar parameters in addition to regex and entropy provides a favorable result. The developed tool also speeds up the classification of libraries and confirms the success of the detection with the chosen combination of parameters. Further development of the tool would allow even more accurate detection of obfuscated and minified libraries by introducing new and better detection parameters. A limitation of using regex-based approaches is that they are inflexible and can only detect a fixed set of minification and obfuscation patterns. One possible way to mitigate this limitation is to use a machine learning approach, which could be more flexible in detecting a wider range of minification and obfuscation patterns.

The developed software tool has demonstrated its potential in accelerating the classification of source code, minified and obfuscated JavaScript. However, there is still room for improvement, such as exploring the use of machine learning-based approaches with additional detection parameters or better breakpoint adaptations for the existing ones to further increase the accuracy and efficiency of the tool.

References

[1] S. Shea, "Cybersecurity," https://www.techtarget.com/searchsecurity/definition/cybersecurity, accessed on April 2., 2023.

[2] A. Alazab, A. Khraisat, M. Alazab, and S. Singh, "Detection of obfuscated malicious javascript code," *Future Internet*, vol. 14, no. 8, p. 217, 2022.

[3] "Stack overflow developer survey 2022," https://survey.stackoverflow.co/2022/#technology, accessed on April 2., 2023.

[4] W. Xu, F. Zhang, and S. Zhu, "The power of obfuscation techniques in malicious javascript code: A measurement study," in *2012 7th International Conference on Malicious and Unwanted Software*. IEEE, 2012, pp. 9–16.

[5] Y. Choi, T. Kim, S. Choi, and C. Lee, "Automatic detection for javascript obfuscation attacks in web pages through string pattern analysis," in *Future Generation Information Technology: First International Conference, FGIT 2009, Jeju Island, Korea, December 10-12, 2009. Proceedings 1*. Springer, 2009, pp. 160–172.

[6] S. Aebersold, K. Kryszczuk, S. Paganoni, B. Tellenbach, and T. Trowbridge, "Detecting obfuscated javascripts using machine learning," in *ICIMP 2016 the Eleventh International Conference on Internet Monitoring and Protection, Valencia, Spain, 22-26 May 2016*, vol. 1. Curran Associates, 2016, pp. 11–17.

[7] N. Malviya, "Simple malware obfuscation techniques," https://resources.infosecinstitute.com/topic/simple-malware-obfuscation-techniques/, accessed March 14. 2023.

[8] Imperva, "Minification," https://www.imperva.com/learn/performance/minification/.

[9] "Use a source map," https://firefox-source-docs.mozilla.org/devtools-user/debugger/how_to/use_a_source_map/index.html, [Accessed: March 3. 2023.].

[10] E. Chiu, J. Lin, B. McFerron, N. Petigara, and S. Seshasai, "Mathematical theory of claude shannon. a study of the style and context of his work up to the genesis of information theory," *submitted for The Structure of Engineering Revolutions (MIT course 6.933 J/STS. 420J), nd*, 2018.

[11] J. Goyvaerts, "Catastrophic backtracking," https://www.regular-expressions.info/catastrophic.html, accessed March 12. 2023.

Forensic Implication of a Cyber-Enabled Fraud Taking Advantage of an Offline Adversary-in-the-Middle (AiTM) Attack

D.O. Lawal, D.W. Gresty, D.E. Gan, and T.C. Durojaiye
University of Greenwich/School of Computing and Mathematical Sciences, London, United Kingdom
D.O.Lawal@greenwich.ac.uk
D.W.Gresty@greenwich.ac.uk
D.Gan@greenwich.ac.uk
TD8962Q@greenwich.ac.uk

***Abstract* - Many computer users utilise the High-Definition Multimedia Interface (HDMI) for connecting external displays as this interface is common on modern computers. This work investigates the feasibility of performing an offline adversary-in-the-middle attack with a portable programmable device such as the Screen Crab which leverages the HDMI interface to covertly capture information being sent to the external display. This work also addresses the possibility of such attacks being carried out as the reconnaissance phase of a wider attack or being carried out as a standalone attack for data exfiltration, data theft, or espionage. Among the operational observations of the Screen Crab, while it was exfiltrating data, include its property of being storage and process efficient. In addition, there were no indicators on the external display (e.g., quality drop, lag/latency) to suggest to the target user that any form of tampering had been done to their machine. This paper also shows how it might be difficult for forensic analysts to detect the use of this device which poses a risk of the target user (victim) being falsely accused or wrongly prosecuted for divulging sensitive or classified information in this kind of situation.**

Keywords – Digital forensics, Adversary-in-the-middle, Cyber-enabled fraud, Screen Crab, Portable programmable devices, Miscariage of justice.

I. INTRODUCTION

Man-in-the-middle attacks (MiTM) or adversary-in-the-middle (AiTM) attacks are used to achieve a number of different attacks. A common use of an adversary-in-the-middle attack is to sniff or eavesdrop on communications that the attacker is not authorised to receive. There are a number of tools that help attackers achieve this, such as Wireshark, and Tshark, which could also be used by security administrators to monitor the flow of traffic for anomalies. ARP poisoning is a common technique applied by attackers to become the man-in-the-middle [1]. ARP poisoning tricks both ends of the communication into believing they are communicating with the intended recipient, when in fact they are communicating with an unauthorised third party (e.g., a client thinks it is communicating with the gateway router while communicating with the attacker machine, and the router also thinks it is receiving or forwarding packets to the intended client, also while talking to the man-in-the-middle). This is possible because the ARP protocol employs no form of authentication during communication which makes it easy for an attacker to send spoofed ARP responses and become the man or adversary-in-the-middle [1]. An AiTM attack through ARP poisoning can be detected while it is active using the 'arp -a' command to investigate the machine suspected of compromise. Investigation can also be done using other tools highlighted in [2].

An attacker could become the adversary-in-the-middle for the sole purpose of sniffing packets without modification or plans to use the captured information in a future attack. An attacker could also become the adversary-in-the-middle to capture information that would be useful to launch a larger attack where the information acquired is key to the success of the larger attack [1]. In this work, the information captured from the target is to be used to plan a larger attack of illegally modifying a record of accounts and whoever is considered responsible, could be charged with fraud if caught.

One common property of the highlighted AiTM methods is that they leave traces that can be detected during incident response or a criminal investigation [3]. This work investigates the effectiveness of the current digital forensic process in combatting the use of a portable programmable device like the Screen Crab when it has been used to commit or facilitate a crime; in this case, a crime of fraud by illegally modifying a record of accounts. The attacker would usually require adequate information about the target environment to successfully plan and execute a file tampering attack. Such information is acquired through reconnaissance which is usually the first phase of a cyber-attack [4]. The reconnaissance phase provides information about the target such as the operating system, applications, location of the system, location of the target file, and several other information that would be useful to the attacker.

This work focuses on the reconnaissance phase of a file tampering attack where the Screen Crab will be used for information gathering on the target. The experiment considered in this work is part of a larger file tampering

experiment involving the use of another portable programmable device (the O.MG Cable) for altering the file after the attacker has gathered information with the Screen Crab; however, this paper does not cover the file tampering aspect of the attack.

II. Related work

This section explores existing work and literature involving AiTM attacks and the use of portable programmable devices. There are many portable programmable devices that can be used to perform operations or automate tasks on a computer, and they include devices like the Bash Bunny [5], Rubber Ducky [6], LAN Turtle [7], and Shark Jack [8] to mention a few. Most of these devices are pocket-sized, thereby making them easily portable and not easily noticed or suspected. Some of these programmable devices like the Rubber Ducky and Bash Bunny have a USB interface. Some like the Shark Jack have LAN interfaces. The O.MG cable is able to operate both via USB and wirelessly [9]. The Screen Crab has HDMI (High-Definition Multimedia Interface) interfaces to capture screen data and store on the local storage or send it to a remote command and control server.

The harm caused by portable programmable devices does not end with disrupting system operations or stealing files. The implication of such devices includes a potential for miscarriages of justice where an employee is fired or an individual is punished for the actions of someone else (or in this case, actions of something else). Lawal, et al., [6] demonstrated how a portable programmable device could be used to plant false evidence like false web history and false file downloads on a target machine to implicate the user. This in turn could lead to innocent users being prosecuted for a crime they did not commit as the forensic analyst may mistake the actions of the programmable devices for the target user's. Many cyber-attacks can be achieved through plugin attacks (e.g., Stuxnet [4]), screen monitoring, malware installation, phishing, keystroke injection attack, etc. Many attacks like sniffing, data exfiltration, and data manipulation could be achieved or facilitated by becoming the adversary or man in the middle.

An AiTM attack or MiTM attack is a type of cyberattack where the attacker positions themselves between a communication source and its destination for ill intent [10]. In a network environment, this may be done with the intention of capturing packets to acquire sensitive information [1] from data in transit. It could also be done with the intention of modifying the packets in transit thereby compromising their integrity [11].

An attacker can also gather information about a target through the use of backdoors to create a reverse TCP connection [12] giving them unauthorised remote access to the target machine. The reverse TCP connection could be leveraged to take screenshots at specified intervals. Other types of information that can be stolen through a reverse TCP connection include web cookies and browser history. This approach often requires a clever delivery of the backdoor to lure the victim into clicking so the payload can be executed or installed on the target machine. The implication of this is that there is direct evidence that can be forensically examined or reverse-engineered to study its behaviour and gather other information. This would be to the advantage of forensic analysts; however, to the disadvantage of an attacker who would prefer to leave no trace.

Keystroke logging or monitoring is another means through which sensitive data can be exfiltrated [13]. Keystroke logging is the process of capturing keystrokes on a target computer and can be done using hardware or software [14]. Malicious software keyloggers could be planted on a target by attackers through phishing links, and trojans [15]. Programmable USBs could also be used to capture keystrokes or to install a software keylogger [16].

In the case of information theft or data exfiltration through a reverse shell connection, forensic analysts may still find tangible evidence such as an installed backdoor e.g., rootkits [17] which could be reverse-engineered to get further evidence as to the source and behaviour of the malware [18]. Tools like Wireshark could also be used for packet analysis to investigate a network MiTM attack [19]. Moreover, security solutions such as firewalls, antiviruses, and intrusion detection and prevention systems (IDPS) may be able to detect the presence of a trojan or network MiTM activity [20]. Security policies to prevent data exfiltration via USB may have also been put in place but the Screen Crab is not limited by these security solutions. It is still able to achieve data exfiltration and information gathering covertly without the interference or detection of security solutions or software security policies. The Screen Crab does not use the USB interface, nor does it need a network connection to capture information from the target machine. It acts as an adversary in the middle by abusing the HDMI connection between the target machine and the connected external display.

The Screen Crab does have some limitations, however, which will be discussed in detail in the results section of this paper.

In this work, the possibility of using the Screen Crab to facilitate a crime of fraud is considered, to determine the response of forensics to the use of such a device. Can a forensic analyst easily detect the presence or use of the Screen Crab during an investigation in situations where it might have been used to facilitate a crime or commit a crime?

III. Experimental setup

The experiment discussed in this work forms part of a larger file tampering experiment. The overall aim of this research is to determine the response of the current digital forensics process to the malicious use of portable programmable devices. However, this paper addresses the threats that devices like the Screen Crab pose; particularly their potential of causing a miscarriage of justice if they are not detected during a forensic investigation where they have been used to facilitate or commit a crime. There are three main phases of the attack simulated before the forensic investigation is carried out. The three phases are Reconnaissance, Arming, and Attack phases (see Figure

1). This work focuses on the reconnaissance phase of the wider file tampering attack. The attack phases for the experiment could also be viewed from the Cyber Kill Chain [21] perspective which breaks down the attack into seven steps including Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control, and Action on objectives.



Figure 1 Attack Phases

The Reconnaissance Phase: this is where information is gathered about the target machine, and the environment. This includes information like the physical location of the machine, the operating system of the machine, the kind of physical and logical access needed (e.g., doors, passwords), the available ports on the machine (e.g., USB, HDMI), and the arrangement or format of the user interface. This phase also involves planting the recon device if a hardware device is used for reconnaissance. The tool used in this work as a recon device is the Screen Crab. The gathered information can then be used to more effectively plan and execute the attack. Figure 2 shows the experimental setup for the reconnaissance phase of the attack which is the focus of this paper. Figure 2 also shows the adversary planting the Screen Crab and using the information collected to write a targeted payload which would then be loaded onto the hacking device shown in the wider experiment setup.

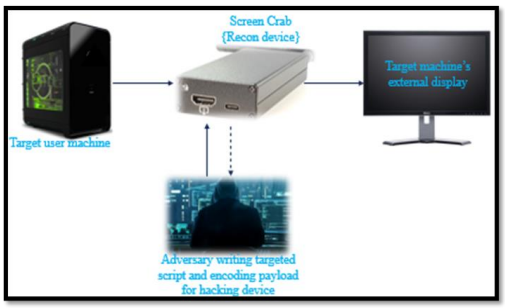


Figure 2 Setup of Reconnaissance Phase

The Reconnaissance phase is the focus of this work, to determine the response of forensics to the recon device used. The Screen Crab was set to video capture mode with a video interval of thirty seconds. The Screen Crab is then covertly positioned between the target system unit and its external display monitor and left to capture the target user's screen for three days simulating a near realistic amount of time needed to gather adequate information for the nature of the larger attack planned. The target system is connected to the Screen Crab's HDMI input interface while the external display is connected to the Screen Crab's HDMI output interface. Relating this experiment to the Cyber Kill Chain, this phase would also be considered as Reconnaissance. The Screen Crab is considered an offline AiTM tool in this work as it does not require network access to capture or exfiltrate data, unlike other methods of becoming the AiTM or MiTM discussed earlier in the literature review.

Arming/Exploit Development: this is the stage where the adversary or attacker uses the information gathered from the reconnaissance phase to develop and test a targeted payload. There is a need to test the payload on a clone of the target machine and environment to achieve a higher level of precision and accuracy when the payload is delivered. The payload needs to be tailored towards the Operating system, and user environment of the target machine as it might not operate as expected if programmed randomly or for a different machine. After the payload has been written and tested, there might be a need to encode the payload into a format the programmable device can process. The Rubber Ducky is an example of a programmable device that requires the payload to be encoded into binary before it can carry out the programmed instructions. In this phase, the attacker encodes the payload if necessary to do so and loads it onto the programmable device to carry out programmed instructions. Relating this to the Cyber Kill Chain, this phase would fall under Weaponization.

Exploit/Attack Phase: this is the final phase of the attack where the payload is delivered, and the target machine is exploited. The attacker plants the programmable device on the target machine when it is safe to do so, for the programmable device to execute the instructions programmed on it. The attacker could also utilise social engineering tactics to get the programmable device to the target machine. Relating this to the Cyber Kill Chain, this phase would cover Delivery, Installation, Command & Control, and Actions on Objectives.

IV. TOOLS AND TECHNOLOGIES

Screen Crab: this is the recon device used for information gathering on the target machine including its operating system, location of files, etc.

Target user machine: this is the target of the attack. The recon device is used to gather information about the user's activities and the location of the target file.

External Display with HDMI support: the external display needs to support HDMI as the Screen Crab has only HDMI interfaces for connection. The external display is connected to the output HDMI interface of the Screen Crab so it can receive data transmission from the target machine.

HDMI Cable: a minimum of two HDMI cables are needed for this kind of attack. One of the HDMI cables connects the target machine to the Screen Crab while the second connects the Screen Crab to the external display as seen in Figure 2.

Micro-SD Card: the Screen Crab requires a micro-SD card for local storage. A minimum micro-SD size of 16GB is recommended for the attack simulated in this work.

Micro-SD Card Reader: this is needed to view or modify the content of the micro-SD card on a computer. Among the content that can be viewed or modified include

the captured images, captured videos, and the Screen Crab's configuration file.

V. EXPERIMENT RESULTS

The results of the experiments are presented here. In this work, the Screen Crab was used for reconnaissance as part of a wider attack; however, it could also be used for standalone data exfiltration or espionage attacks. This raises the question of whether forensics can detect the use or presence of this device, especially when used as an offline AiTM.

A. Result Overview

The forensic investigation was conducted using industry-standard digital forensic tools Autopsy and FTK Imager and Registry Viewer. The results of the investigation were unable to detect any trace of the Screen Crab. This makes it near impossible for a forensics analyst to detect its presence.. Some risks associated with this include facilitating data exfiltration, or espionage which could lead to accusing an innocent person of divulging sensitive information, as the Screen Crab leaves very little to no forensics evidence. This may yield unreliable digital evidence leading to a miscarriage of justice (i.e. an innocent individual being punished or a guilty person walking free).

While the Screen Crab was operational, there were no signs such as lag or latency on the external display to raise suspicions to the target user that their computer connection had been tampered with. In addition, there was no drop in image quality on the external display which may have also been a cause for alarm.

Moreover, a stand-alone analysis was carried out on two different laptops including a Microsoft Surface Pro and a Toshiba laptop with Windows 10 installed. The computers detected only the monitor on the other end and not the Screen Crab in between. This suggests the Screen Crab works just like an HDMI splitter and requires no drivers to function. It relies solely on the signal going through the HDMI and so leave no trace on the registry.

In situations where the Screen Crab has been used against a target user who is the custodian of top-secret information related to national security, they may be falsely accused of disclosing classified information and charged with treason [22] [23].

The use of USB devices to extract data or steal files may still leave some traces like the product ID (PID) and Vendor ID (VID) [9]. However, the Screen Crab does not require any drivers and leaves no obvious Hardware ID that may have been an indicator of its use or presence to the forensic analyst during an investigation. The Screen Crab uses the HDMI interface and as mentioned, works like an HDMI splitter.

The Screen Crab can also use a Wi-Fi connection to gain access to the internet and communicate with a command and control (C2) server. However, it was used in an offline mode in this work. With Internet access, the Screen Crab can receive instructions from the remote C2 server, can store the captured files on the remote server, and can provide a live feed of the target screen to the adversary via the C2 server. If the wireless connection was used, it may be possible to use a tool like Wireshark to capture the traffic and see what kind of packets are being transmitted as seen in [19].

B. Operational Observations of the Screen Crab

The captured images are high resolution and are clear enough to pick out every detail on the screen (e.g., see figure 4). Picture quality is sharp and there is no evident effect of the screen crab on the functionality of the user's display. There was no drag, no delay, no break in transmission, and no cracks or shakes that would have led to the user attempting to check the connections.

The captured file sizes are not large which means the capture process for images is not storage intensive. The captured videos had varying file sizes with an average file size of 2MB. The Screen Crab can also detect when the computer screen is inactive (or off) and stops capturing. The Screen Crab resumes screen capture once it detects the computer screen being active again. This feature further increases the Screen Crab's storage efficiency as it prevents the capture of blank videos that would take up unnecessary space. There are two options for storage; either ROTATE or FILL. The FILL option continues the capture until the Screen Crab's storage is filled up whereas the ROTATE option continues the process while replacing the oldest files with the newest when the Screen Crab's local storage is full.

During the simulated attack, the Screen Crab was also not able to transmit audio despite its connection via HDMI which should be able to transmit audio. This may lead to the target user discovering the Screen Crab while attempting to troubleshoot.

C. Using the Screen Crab Against Touchscreen Devices

Figure 4 below is a screenshot from a video capture of the Screen Crab while the system user was entering their Windows logon password. The target machine used here was a Microsoft Surface Pro PC running Windows 10 operating system with its firewalls and the Microsoft



Figure 4 Screenshot from captured video during user logon

Windows Defender antivirus program active. Neither the firewall nor the MS Defender detected nor flagged up the Screen Crab while it was operational. As could be seen in figure 4, every keystroke or screen tap is picked including the user login password as the keys are pressed.

D. Limitations of the Screen Crab

The Screen Crab has only HDMI interfaces which means it might be useless against a computer and/or external display that has no HDMI ports. This would only be a limitation to the attackers who have no knowledge of alternative ways to use HDMI connections on devices that have no HDMI ports. This limitation could be overcome through the use of converters (e.g., VGA to HDMI converters or Display port to HDMI converters, etc.). It might be a slightly more complex improvisation, but it will get the job done on many systems.

The Screen Crab also requires physical access to be planted and to be removed which could be a setback for the attacker in certain situations. However, this limitation could be overcome through well-executed social engineering as seen in [24]. This would not be a limitation for personnel or individuals who would normally have access to the target machine's location e.g., office colleague, IT support, system administrator, janitors, security personnel or facility managers.

The Screen Crab requires a power supply through a USB Type C cable to function. No capturing can be done while it is turned off. Hence, proximity to a power source or a covert USB port for power is paramount for the successful operation of the Screen Crab.

Another limitation of the Screen Crab is that it does not transmit audio. HDMI is capable of transmitting both audio and video and in some situations where the target user expects audio to be transmitted, the absence of audio on the external display may raise concern e.g., if the target machine is connected to a TV that would normally play sound. This may prompt the user to troubleshoot the problem and may then lead to the discovery of the Screen Crab.

VI. Conclusion

This work demonstrates how a portable programmable device like the Screen Crab can be used to steal information anonymously from a target computer without much evidence left for investigators or forensic analysts to find. The investigation was conducted following an industry-standard forensic process and the use or presence of the Screen Crab on the victim's computer was not at any point flagged during the investigation. This increases the risk of a miscarriage of justice because an innocent person (the user of the target machine in this case), could be punished or prosecuted for an offence they did not commit. Possible repercussions for the victim include termination of employment, or in cases where non-disclosure agreements have been signed, they could be sued for a huge sum of money, or even worse, they could be prosecuted for divulging classified or top-secret information without authorisation. The user of the target computer could be charged with treason for such offences [22].

There are also security implications for individuals, organisations, or countries whose sensitive information might have been leaked through the use of devices like the Screen Crab. Such implications include a threat to national security, and facilitating espionage of organisations and countries. The possible applications of portable programmable devices are only limited by the creativity of the attacker.

In addition, there were no indicators on the external display (e.g., quality drop, lag/latency) to suggest to the target user that any form of tampering was taking place on their machine. This paper also shows how it might be difficult for forensic analysts to detect the use of this device which poses a risk of the target user (victim) being falsely accused or wrongly prosecuted for divulging sensitive or classified information in this kind of situation.

The experiment results indicate that the current digital forensics process may not be effective enough in combatting the use of portable programmable devices like the Screen Crab in situations where it has been used maliciously to commit a standalone crime or facilitate a larger attack. This means the evidence could be missed in the investigative stage by the forensic analyst and might also be missed in the expert presentation in court. This could potentially lead to a faulty conclusion in the court of law, as it looks like an open and shut case. Forensic investigators need to be aware of the potential of devices like the Screen Crab for spying or stealing confidential data and the subsequent potential for a miscarriage of justice.

In addition, this shows the need for a methodology for digital forensic investigations that could guide forensic analysts during an investigation to reduce the chances of missing evidence like this or at least consider its possibility in scenarios where its use is possible.

Investigations into this are still ongoing and future recommendations include investigating the Screen Crab's remote connection to the C2 server using Wireshark and other network investigation tools to determine the kind of information being transmitted. In addition, this may also provide information as regards the Screen Crab's indicators of compromise (IoCs).

References

[1] Nayak, G.N. and Samaddar, S.G., 2010, July. Different flavours of man-in-the-middle attack, consequences and feasible solutions. In 2010 3rd International Conference on Computer Science and Information Technology (Vol. 5, pp. 491-495). IEEE.

[2] Tuli, R., 2020. Packet Sniffing and Sniffing Detection. International Journal of Innovations in Engineering and Technology, 16(1).

[3] Saputra, D. and Riadi, I., 2019. Network forensics analysis of man in the middle attack using live forensics method. International Journal of Cyber-Security and Digital Forensics, 8(1), pp.66-73.

[4] Shaikh, R.A., Khan, M.S., Rashid, I., Abbas, H., Naeem, F. and Siddiqi, M.H., 2022, May. A Framework for Human Error, Weaknesses, Threats & Mitigation Measures in an Airgapped Network. In 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2) (pp. 1-8). IEEE.

[5] Zhao, S. and Wang, X.A., 2020. A survey of malicious HID devices. In Advances on Broad-Band Wireless Computing, Communication and Applications: Proceedings of the 14th International Conference on Broad-Band Wireless Computing, Communication and Applications (BWCCA-2019) 14 (pp. 777-786). Springer International Publishing.

[6] Lawal, D., Gresty, D., Gan, D., and Hewitt, L., 2021. Have You Been Framed and Can You Prove it? In 2021 44th International Convention on Information and Communication Technology, Information System Security (MIPRO). IEEE.

[7] Slunjski, M., Sumina, D., Groš, S. and Erceg, I., 2022. Off-the-Shelf Solutions as Potential Cyber Threats to Industrial Environments and Simple-To-Implement Protection Methodology. IEEE Access, 10, pp.114735-114748.

[8] Hak5 (n.d) Shark Jack, Hak5. Available at: https://shop.hak5.org/products/shark-jack (Accessed: February 6, 2023).

[9] Lawal, D., Gresty, D., Gan, D., Hewitt, L. (2022), 'Facilitating a Cyber-Enabled Fraud Using the O.MG Cable to Incriminate the Victim', World Academy of Science, Engineering and Technology, Open Science Index 189, International Journal of Computer and Systems Engineering, 16(9), 367 - 372.

[10] Mallik, A., 2019. Man-in-the-middle-attack: Understanding in simple words. Cyberspace: Jurnal Pendidikan Teknologi Informasi, 2(2), pp.109-134.

[11] Bhushan, B., Sahoo, G. and Rai, A.K., 2017, September. Man-in-the-middle attack in wireless and computer networking—A review. In 2017 3rd International Conference on Advances in Computing, Communication & Automation (ICACCA)(Fall) (pp. 1-6). IEEE.

[12] Tigner, M., Wimmer, H. and Rebman, C.M., 2021, October. Windows Reverse TCP Attack: The Threat of Out-of-Date Machinery. In 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 555-560). IEEE.

[13] Akhtar, Z., 2021. Malware detection and analysis: Challenges and research opportunities. arXiv preprint arXiv:2101.08429.

[14] Singh, A. and Choudhary, P., 2021, August. Keylogger detection and prevention. In Journal of Physics: Conference Series (Vol. 2007, No. 1, p. 012005). IOP Publishing.

[15] Ayap, N.R.D., Dabu, A.C.Y., Reyes, A.M., Ventura, A.B.K., Trinidad, G.P. and Blancaflor, E.B., 2022, Beat the Bait: A Case Study on Phishing Attack.

[16] Charan, B.V., Kulkarni, L. and Karad, V., 2023. Survey On Micro-Controller Based Bad USB Attacks. Journal of Positive School Psychology, pp.965-974.

[17] Abuzaid, A.M., Saudi, M.M., Taib, B.M. and Abdullah, Z.H., 2013. An efficient trojan horse classification (ETC). International Journal of Computer Science Issues.

[18] Bao, C., Forte, D. and Srivastava, A., 2015. On reverse engineering-based hardware Trojan detection. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 35(1), pp.49-57.

[19] Lan, H., Zhu, X., Sun, J. and Li, S., 2020, January. Traffic data classification to detect man-in-the-middle attacks in industrial control system. In 2019 6th International Conference on Dependable Systems and Their Applications (DSA) (pp. 430-434). IEEE.

[20] Aliyu, F., Sheltami, T. and Shakshuki, E.M., 2018. A detection and prevention technique for man in the middle attack in fog computing. Procedia computer science, 141, pp.24-31.

[21] Cyber kill chain® (2022) Lockheed Martin. Available at: https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html.

[22] Fabre, C., 2020. The Morality of Treason. Law and Philosophy, 39(4), pp.427-461.

[23] Dobrovidova, O., 2022. Russian scientist facing treason charges dies in custody. Science, 377(6604), pp.355-356.

[24] Bakhshi, T., 2017, December. Social engineering: revisiting end-user awareness and susceptibility to classic attack vectors. In 2017 13th International Conference on Emerging Technologies (ICET) (pp. 1-6). IEEE.

miproBIS

International Conference on

BUSINESS INTELLIGENCE SYSTEMS

Steering Committee

Chairs: Boris Vrdoljak, University of Zagreb, Croatia
Mirta Baranović, University of Zagreb, Croatia
Matteo Golfarelli, University of Bologna, Cesena, Italy
Mihaela Vranić, University of Zagreb, Croatia

Program Committee

Members: Marko Banek, University of Zagreb, Croatia
Mirta Baranović, University of Zagreb, Croatia
Ladjel Bellatreche, University of Poitiers, France
Ljiljana Brkić, University of Zagreb, Croatia
Alfredo Cuzzocrea, University of Calabria, Italy
Todd Eavis, University of Montreal, Canada
Matteo Francia, University of Bologna, Cesena, Italy
Enrico Gallinucci, University of Bologna, Cesena, Italy
Dragan Gamberger, Ruđer Bošković Institute, Zagreb, Croatia
Matteo Golfarelli, University of Bologna, Cesena, Italy
Marko Gulić, University of Rijeka, Croatia
Igor Mekterović, University of Zagreb, Croatia
Damir Pintar, University of Zagreb, Croatia
Sandro Skansi, University of Zagreb, Croatia
Zoran Skočir, University of Zagreb, Croatia
A Min Tjoa, Vienna University of Technology, Austria
Mladen Varga, University of Zagreb, Croatia
Mihaela Vranić, University of Zagreb, Croatia
Boris Vrdoljak, University of Zagreb, Croatia

PAPERS

Metrics for Estimating Accuracy, Reliability, and Bias in Peer Assessment

M. Fertalj, LJ. Brkić and I. Mekterović
Faculty of Electrical Engineering and Computing/Department of Applied Computing, Zagreb, Croatia
melita.fertalj@fer.hr ljiljana.brkic@fer.hr igor.mekterovic@fer.hr

***Abstract* - The challenge of teaching staff being overworked and not having enough time to provide individualized feedback and guidance to students is a significant problem. Open-ended assignments foster creative thinking but add to the workload. The abundance of solutions makes individualized assessment necessary. Peer assessment, where students evaluate their peers' assignments based on their motivation, knowledge, and resourcefulness, has proven effective in higher education courses.**

This paper provides an overview of existing metrics used in peer assessment, defines what constitutes correct and consistent evaluation, and suggests ways to choose appropriate metrics for learning objectives. The paper also highlights limitations of measures presented in the published papers, shows the advantages and disadvantages of different ways of calculating the selected measures through examples.

Keywords - peer assessment; metrics; accuracy; reliability; bias; grade calculation; criteria of assessment; open-ended works

I. INTRODUCTION

Nowadays, with platforms like Coursera, edX and Udacity online learning is considered mainstream. Online teaching has created new opportunities for improving knowledge and online courses often have large groups. Whether it's online or traditional courses, students like any other type of consumer are increasingly raising their expectation of the learning service. Therefore, it is important to define an assessment method that is suitable for large groups of students which is also useful when the number of students is considerably greater than the teaching staff.

Solving open-ended assignments encourages students to tackle the problem analytically and apply their own understanding of the given task. For example, writing an essay or an engineering design doesn't have an unambiguous solution like closed-ended assignments. Tasks where the answer is a simple yes or no or an exact number can be graded automatically. In open-ended assignments, the solutions can be in audio or video format, thus rendering automatic grading difficult or impossible.

A possible solution to this problem is the concept of peer assessment (PA) where students review their peers' work while providing feedback on the quality of the work they submitted. The given feedback affects the learning process and the improvement of solving assigned tasks. Teaching staff of the department of Computer Science at North Carolina State University have compared 18 machine learning models to analyze student comments [1] to their peer's work and concluded that comments that identify problems and offer suggestions that encourage students to improve their solutions.

Students might exhibit "rogue" and "non-rogue" [2] behavior when grading. "Rogue" students grade their peers' work with the highest, lowest, average grade or simply rate randomly. Fair assessment of students' work means reviewing their assignments according to the expected quality standard. It is important to assign a fair grade to maintain the motivation of students already interested in the course material and to increase the motivation of others. Grades assigned by graders influence the final grades of their peers, therefore it is necessary to ensure that they are graded fairly. Students are awarded for the quality of their assignment solutions as well as the quality of their grading peers' solutions. Calculating both the grade for the student's solution and the grade for student's evaluation in peer assessment should take bias, reliability, and credibility into account. Peer assessment can be implemented with calibration, where students are provided with assessment criteria called rubrics [3][4]. Using rubrics, students evaluate calibration tests of different qualities. Calibration tests are assignments created and graded by the teaching staff and their grades are considered correct. Students should evaluate as similar as possible to teachers.

II. FINAL GRADE CALCULATION

We will analyze separately different methods for calculating the final grade on a student assignment using peer assessment. For some methods (e.g. Mean and Median) a calibration test is not required as it cannot improve the result. Examples of assigned grades and the final grades are presented and compared to understand which method is better in a particular case.

A. Mean

The grades assigned by the student graders are aggregated and the sum is divided by the number of grades. We assume that the student graders did not exhibit "rogue" behavior and do not require additional calculations to improve the accuracy of the obtained grades. For student i and grader j, the final calculated grade is marked as g_i. The grades are calculated as $g_i^{(x)} = \frac{\sum_j A_{i,j}^{(x)}}{|A_i^{(x)}|}$. Set of all grades assigned to student i is marked as A_i, while $A_{i,j}$ is the grade assigned to student i by grader j. Let us grade each assignment task individually, rather than the whole assignment. The index x in the formula above marks the task in question. E.g., $g_i^{(1)}$ is the grade for student i for the

1st task, whereas $A_{i,j}^{(1)}$ is the grade assigned by grader *j* to student *i* for the 1st task.

B. Median

Median is an easy way to calculate the final grade instead of calculating the average. Median determines the central value of the distribution. This method is used in peer assessment applications like Coursera and Coursebank. For set of numbers X with n elements ordered by size from smallest to largest, the median is calculated depending on if n is an even or odd number. When *n* is even is defined as $median(X) = \frac{x_{(n/2)} + x_{(n/2)+1}}{2}$. Whereas n is odd, the formula is as follows: $median(X) = x_{(n+1)/2}$. We are interested in calculating the median for an assignment with multiple tasks. For student *i* and task *x*, grade g_i is calculated using the formula below, where A_i is the set of all grades that have been assigned to student *i* for task *x:* $g_i^{(x)} = median(A_i^{(x)})$.

C. Calibrated peer assessment

The "Calibrated Peer Review" (CPR) system was developed at the University of California [5] and was also used at the Texas University [6]. The CPR method for calculating the final grade is to evaluate the mean grade with the previously calculated weight. The students who evaluated closest to the teachers' grading are awarded the greater weight. Since their grading is considered most accurate among their peers, their grading has the most impact on the final grade. The weights are determined with assignments called calibration tests. These tests are created and graded by teachers and their grades are considered as a reference to compare how a student has graded in relation to the teacher. During peer assessment, students also grade one or more calibration tests along with their classmates' assignments. The expected grades for the calibration tests are teachers' grades who previously graded them in preparation for the peer assessment process.

The formula for calculating the mean is modified so that the weight w_j for the grader *j* is used and the final calculated grade is defined as $g_i^{(x)} = \frac{\sum_j w_j^{(x)} A_{i,j}^{(x)}}{\sum_j w_j^{(x)}}$. The weight for each student is determined from the difference between the grade they have assigned and the teacher's grade. For *N* calibration tests (assignments) each assignment consists of multiple tasks. The correct grade that should be assigned $t_n^{(x)}$ is defined for the assignment task *x* on calibration test *n*. The difference between the grade assigned by student grader j and the correct grade assigned by the teacher for task *x* is determined as $\Delta g_j^{(x)} = \frac{\sum n \left(A_{n,j}^{(x)} - t_n^{(x)}\right)^2}{N}$. If the student hasn't graded a task on the calibration test, that test is ignored in the formula above. Only the differences for the other calibration tests are calculated, where *N* represents the number of calibration tests that are not disregarded. The greater the difference $\Delta g_j^{(x)}$, the smaller the credibility of the student grader. Therefore, the weight is defined as $w_j^{(x)} = \begin{cases} w_{max,} \; if \; \Delta g_j^{(x)} = 0 \\ \frac{1}{\Delta g_j^{(x)}}, else \end{cases}$. If the student has graded the calibrated test exactly like the teacher, the difference $\Delta g_j^{(x)}$ is 0. Therefore, the weight of that student will have the maximum weight value calculated for other students.

D. Comparing mean, median, and calibrated PA calculations

Let us examine the methods for calculating the final grade mentioned in the previous paragraphs. Example of assigned grades are presented in TABLE I. The correct grades for students 1, 2 and 3 will be determined by using the grades given in the peer assessment process. The teacher has evaluated the calibration test and the assigned grades are considered as correct. A Lickert scale of 1-5 is often used for grading [7].

TABLE I. ASSIGNED GRADES

Grader *j*	Student 1			Student 2			Student 3			Calibration		
Grader 1	1	1	1	1	1	1	1	1	1	2	4	1
Grader 2	5	1	2	3	4	2	5	3	1	5	3	4
Grader 3	5	2	1	4	3	4	3	5	1	4	1	3
Correct grade $t_n^{(x)}$	/	/		/	/		/	/		**5**	**2**	**3**

Before calculating the final student grades with mean, median and calibration, the CPR method has an additional step where the weights for each student are calculated. Using the formulas mentioned in the *Calibrated peer assessment* paragraph, we determine that there is $N = 1$ calibration test with $x = [1, 2, 3]$ tasks. It is also necessary to set an upper limit for weights, because if the grader's grades correspond perfectly with the teacher's grades, the weight will be very high and would prevail over all other weights. We set the upper limit for the maximum weight value to 0.9. The differences between student and teacher grades as well as the calculated weights are as follows:

TABLE II. GRADE DIFFERENCES AND WEIGHTS

Grader *j*	$\Delta g_j^{(1)}$	$\Delta g_j^{(2)}$	$\Delta g_j^{(3)}$	$w_j^{(1)}$	$w_j^{(2)}$	$w_j^{(3)}$
Grader 1	$(2-5)^2 = 9$	$(4-2)^2=4$	$(1-3)^2 =4$	$\frac{1}{9} = 0.11$	$\frac{1}{4} = 0.25$	$\frac{1}{4} = 0.25$
Grader 2	0	1	1	0.9	0.9	0.9
Grader 3	1	1	0	0.9	0.9	0.9

Mean grade is calculated as a grade for the assignment, while grade for each task can also be calculated separately. The same is valid for calculating the median grade. As shown in TABLE III., the mean grades differ from the median grades. Using one method instead of the other makes more sense depending on the expected grading result. E.g., student graders can assign grades in range [1, 5] and the calculated grade should be whole or half of the grade. In this case, calculating the median results in integer or partial values.

TABLE III. COMPARISON OF MEAN, MEDIAN, AND CALIBRATED PA CALCULATIONS

		Student 1	Student 2	Student 3
Mean	$g_i^{(1)}$	3.67	2.67	3.00
	$g_i^{(2)}$	1.33	2.67	3.00
	$g_i^{(3)}$	1.33	2.33	1.00
	$\overline{g_i}$	2.11	2.56	2.33
Median	$g_i^{(1)}$	5.00	3.00	3.00
	$g_i^{(2)}$	1.00	3.00	3.00
	$g_i^{(3)}$	1.00	2.00	1.00
	$\overline{g_i}$	2.33	2.67	2.33
CPR	$g_i^{(1)}$	$\frac{0.11*1+0.9*5+0.9*5}{0.11+0.9+0.9}$ = 4.77	3.35	3.82
	$g_i^{(2)}$	$\frac{0.25*1+0.9*1+0.9*2}{0.25+0.9+0.9}$=1.43	3.19	3.63
	$g_i^{(3)}$	$\frac{0.25*1+0.9*2+0.9*1}{0.25+0.9+0.9}$=1.44	2.75	1.00

Calculating both mean and median grades is equally represented in the literature for determining assignment and task grades. The results also show that the student 1 deserved a higher grade for the 1st task than the grader 1 evaluated. Grader 1 assigned the lowest grade and has graded the calibration test poorly, so his assessment has the least weight in calculating the final grade.

III. METRICS

Metrics include measures used to monitor student activity and evaluate the success of teaching programs. Students learn through homework, quizzes, exams, or essays where the final grade is a metric comprised of the mentioned measures.

TABLE IV. contains different kind of measures found in research for bias, reliability, credibility, and accuracy. The terms are somewhat exchangeable i.e., reliability and credibility are both defined by consistency, Pearson's correlation coefficient is used for calculating reliability or accuracy depending on the literature. When calculating the final grade, it is possible to give less or more importance to the graders' evaluation, in which case the mentioned credibility index is used. Intra-rater reliability will not be analyzed in this paper because we are interested only in students reviewing the same assignment once.

TABLE IV. MEASURES FOR CALCULATING BIAS, RELIABILITY AND ACCURACY

Statement of meaning	Terminology detected in literature	Corresponding measures
Student evaluation is subjective because it is influenced by their perspective [8].	Bias	assessment median [8], average distance to diagonal [8], probabilistic models for estimating biases [9][10]
Inter-rater reliability is defined as consistent grading of different students on the same assignment. Intra-rater reliability is defined as consistent grading of the same student on the same assignment [11].	Reliability	agreement (exact, adjacent) and consistency (Pearson's and Spearman correlation coefficient) [11], probabilistic models for estimating reliabilities[9], Intraclass Correlation Coefficient [12], Krippendorff's alpha [13]
Grader credibility is defined by the dimensions [14]: Accuracy determines how close the assigned grade is to the grade we know to be correct (teacher's grade), consistency represents stability of assigned grades within the same test, and transferability provides information on how grader accuracy changes across multiple works.	Credibility	Credibility Index (CI)[14] [15], Reviewer Competency Index (RCI) [16], Competency index [2]
The teacher determines what the correct solution is and the associated grades for different qualities of the student solution depending on the deviation from the correct one.	Accuracy	Distance between gradings [17], Pearson's correlation coefficient [11], probabilistic models for improving accuracy [18] [19]

A. Accuracy

1) Distance between gradings

Distance between gradings published in the *Learning Analytics for Peer assessment: (Dis)advantages, Reliability and Implementation* article [17] considers each set of assigned grades for each student separately. From the set, the grades are compared pairwise. For example, for the student i, the grades of the student graders x and y are observed. The grades are organized into a vector and the distance between the vectors is compared. Grade vector for each task for student i that graded grader x is defined as $A_{i,x}$ = ($A_{i,x}^{(1)}$, $A_{i,x}^{(2)}$, … , $A_{i,x}^{(n)}$), and for grader y the vector is defined as $A_{i,y}$ = ($A_{i,y}^{(1)}$, $A_{i,y}^{(2)}$, … , $A_{i,y}^{(n)}$) where n is the total number of tasks in the assignment.

The authors of the mentioned article have opted for the modified Manhattan distance to determine the distance between these two vectors. Manhattan distance is calculated as the sum of the absolute differences of the Cartesian coordinates of the two vectors. The Manhattan distance between two points can be visually represented on a rectangular grid that is parallel to the coordinate axes where space between the gridlines is equal to 1. Distance between vectors $A_{i,x}$ and $A_{i,y}$ is calculated using the following expression, where A_{max} is the maximum possible grade on the assignment and X is the number of tasks:

$$d(A_{i,x}, A_{i,y}) = \frac{1}{A_{max}} \left(\frac{|A_{i,x}^{(1)} - A_{i,y}^{(1)}| + \cdots + |A_{i,x}^{(n)} - A_{i,y}^{(n)}|}{|X|}\right) \in [0, 1]$$

Vectors $A_{i,x}$ and $A_{i,y}$ are considered as set *A*. The diameter of *A* is defined as the maximum distance between two vectors in the set as $diam(A) = \max_{A_{i,x}, A_{i,y} \in A} d(A_{i,x}, A_{i,y})$. If the student graded the other students' work exactly or similar as the teacher, the difference between assigned grades is small. Therefore, the diameter is smaller, and the accuracy is bigger. To demonstrate, let us consider that we have conducted peer assessment for an assignment thar consists of one task. Six students have participated in assignment and graded three of their peers. For student *i*, the grades are as follows:

TABLE V. ASSIGNED GRADES EXAMPLE

Assign-ments	Stud 1	Stud 2	Stud 3	Stud 4	Stud 5	Stud 6
$A_{t,1}^{(1)}$	4	5	4	4	2	3
$A_{t,2}^{(1)}$	5	4	4	5	2	3
$A_{t,3}^{(1)}$	5	4	4	5	2	4

The calculated distances are presented in TABLE VI.:

TABLE VI. DISTANCES

Assign-ment distances	Stud 1	Stud 2	Stud 3	Stud 4	Stud 5	Stud 6
$d(A_{t,1}, A_{t,2})$	0,2	0,2	0,0	0,2	0,0	0,0
$d(A_{t,2}, A_{t,3})$	0,0	0,0	0,0	0,0	0,0	0,2
$d(A_{t,3}, A_{t,1})$	0,2	0,2	0,0	0,2	0,0	0,2

The maximum distance $diam(A)$ is 0,2. Therefore, peer assessment accuracy is calculated and presented as a percentage: $acc(A) = 1 - diam(A) = 1 - 0{,}2 = 0{,}8 * 100 = 80\%$. As shown, the accuracy is high concluding that the students greatly agreed on the quality of the solutions they graded. The main disadvantage of the ***distance between gradings*** method is that the grades used in the described calculations are individual grades assigned by the student graders. Final grades are calculated independently and are not verified with this method. The overall agreement in grading between students with the advantage of not requiring teachers grades for calculation.

2) Pearson correlation coefficient

Pearson correlation coefficient [11] is used to measure of the linear correlation between each student grader's grade and the teacher's grade on the same assignment. The use of the Pearson's correlation coefficient is widely represented in the literature for various purposes is defined as

$$pearson(G_i, T_i) = \frac{\sum_{x \in X}(g_i^{(x)} - \overline{G_i})(T_i^{(x)} - \overline{T_i})}{\sqrt{\sum_{x \in X}\left(g_i^{(x)} - \overline{G_i}\right)^2 * \sum_{x \in X}\left(T_i^{(x)} - \overline{T_i}\right)^2}}$$

Where G_i is a set of grades calculated for student *i*, g_i is the calculated grade for the student *i* for task *x*, while T_i is the set of correct grades (defined by the teacher) for student *i*. The similarity obtaines values from the interval [−1,1] where 1 represents complete similarity while -1 represents minimal similarity. The calculation is done for every student. The average value of all students is taken as the final result and cannot be expressed as a percentage. Positive Pearson's coefficient indicates that if the value of the first set of data increases, the values in the second set will also increase. Negative pearson coefficient indicates that if the value of the first set decreases, the value of the second set increases.

Let us examine an example for the Pearson's correlation coefficient so that we can inspect its properties. Vector $\overrightarrow{g_1} = (5, 4, 5, 3, 4, 2, 4, 5, 1, 5)$ is the vector of calculated grades. Vector $\overrightarrow{t_1} = (3, 1, 2, 2, 3, 1, 4, 2, 4)$ is the vector of teachers grades that are considered correct. The average grades are $\bar{g} = 3.8$ and $\bar{t} = 2.5$. To simplify the subsequent calculation, the respective averages are deducted from the vector values so now the vectors contain values as following:

$\overrightarrow{g_1} = (1.2, 0.2, 1.2, -0.8, 0.2, -1.8, 0.2, 1.2, -2.8, 1.2)$,

$\overrightarrow{t_1} = (0.5, -1.5, -0.5, -0.5, 0.5, -1.5, 0.5, 1.5, -0.5, 1.5)$. The Pearson's coefficient is calculated according to the formula mentioned above: $pearson(\overrightarrow{g_1}, \overrightarrow{t_1})$ = 0.5887. The value indicates that the correlation is relatively low which makes sense since the vectors contain distinctly different values.

For the 2nd example the $\overrightarrow{g_1}$ vector remains the same and the teacher vector is $\overrightarrow{t_2} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$, with calculated average $\bar{t} = 1$. When average teacher grade $\bar{t}$ is deducted from the $\overrightarrow{t_2}$ vector, the values are now all 0, i.e. $\overrightarrow{t_2} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$. The Pearson's coefficient is $pearson(\overrightarrow{g_1}, \overrightarrow{t_2})$ = 0. As seen in the example, Pearson's correlation coefficient gives a result of 0 in case when one of the vectors contains all 0 values. More precisely, the result is 0 if any of the vector components are set to the same number (e.g. the whole vector is 3). Of course, the student can solve the assignment 100% correct and will be assigned the same grade for all of the tasks but we cannot know if that is the case for certain.

The 3rd example demonstrates a Pearson's coefficient where the value cannot be interpreted as we expect. Again, the $\overrightarrow{g_1}$ vector remains the same and the teacher vector is $\overrightarrow{t_3} = (1, 1, 1, 1, 1, 1, 1, 2, 1)$, with calculated average $\bar{t} = 1.1$. When average teacher grade $\bar{t}$ is deducted from the $\overrightarrow{t_3}$ vector, the new values are $\overrightarrow{t_3} = (-0.1, -0.1, -0.1, -0.1, -0.1, -0.1, -0.1, -0.1, 0.9, -0.1)$. The Pearson's coefficient is $pearson(\overrightarrow{g_1}, \overrightarrow{t_3})$ = -0.7075. In this case, the correlation coefficient is drastically different from the example with $\overrightarrow{t_2}$ vector, although the $\overrightarrow{t_3}$ vector changed only in one value. The ideal conditions would be to always expect minimal changes in the calculated results when the input g or t vectors contain slightest change. Accuracy is required to be 0 in the case with maximum error when grading, i.e. when the calculated grades of the student graders are farthest from the expected teacher's grade. For example, the teacher gave the highest grade (5) for all tasks in the assignment, but calculated grades which gave the students are the lowest grade (1). We also expect accuracy of 1 when the g and t vectors are equal.

From the previous examples it is clear that Pearson's correlation coefficient does not satisfy these properties.

B. Reliability

Spearman's rank correlation coefficient is a measure for determining peer-grading reliability [11] between each student grader's grade and the **median grade** on the same assignment. Pearson's correlation coefficient is used to measure the linear correlation between student and teacher grades (paragraph *Accuracy*), while Spearman's correlation measures monotonic relationships. If one value increases (or decreases), the other value also increases (or decreases). It is defined as $\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$, where n is the number of grades (X_i, Y_i) and $d_i = R(X_i) - R(Y_i)$ is the difference between the two ranks for each grade comparison. If there are no ranks that are tied i.e., there are no equal grades, then the grades are ordered from greatest to smallest with rank 1 assigned to the highest grade, rank 2 to the next highest and so on. When ranks are not tied, the above formula for ρ can be applied. When ranks are tied to two grades of the same category (e.g., two X_i grades), both grades are assigned a mean rank calculated from the two ranks in question.

Spearman's coefficient returns a value from -1 to 1, where +1 indicates a perfect positive correlation between ranks, -1 value indicates a perfect negative correlation between ranks and 0 presents no correlation between ranks. Spearman's coefficient presents the inter-rater reliability i.e., how consistently did different students grade the same assignment. Assigned grades are in TABLE VII., where the usual mark for grade g_i is now replaced as X_i and the median is marked as Y_i. The example in TABLE VII. presents an assignment consisting of two tasks, where two student graders evaluated three of their classmates.

TABLE VII. ASSIGNED GRADES, RANKS AND THEIR DIFFERENCES

	Student 1	Student 2	Student 3
$X_i^{(1)}$	1	3	5
$R(X_i^{(1)})$	3	2	1
$X_i^{(2)}$	5	1	2
$R(X_i^{(2)})$	1	3	2
$\overline{Y_i}$	3	2	3.5
$R(\overline{Y_i})$	2	3	1
$(d_i^{(1)})^2$	1	1	0
$(d_i^{(2)})^2$	1	0	1

$$\rho^{(1)} = \rho^{(2)} = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6*2}{3*(9-1)} = 0.5$$

In this case, we calculated how consistently did two student graders evaluate a task. For both tasks the reliability ρ is 0.5 i.e., grading is dispersed around the median. Therefore, the graders are strongly consistent in their evaluation. If the students' grading is compared to the correct grades defined by teacher instead of the median, the Spearman's correlation shows how consistently students grade in relation to the teacher. The **positive ρ** (closer to +1) indicates that students give higher grades for better solved assignments i.e., assignments that the teachers also grade higher. The **negative ρ** (closer to -1) indicates that students grade consistently but the exact opposite of how the teacher will grade the same assignment. The **closer ρ is to 0**, the relationship between the student grade and median is weaker, and the less consistent the grading is among the students. In this case, students assign high grades for an assignment, but is equally possible they will give lower grades for the same assignment. Therefore, their grading is not reliable.

C. Bias

To guarantee reliability in peer assessment, it is necessary to detect bias and provide objective grading of student assignments. In the previous paragraph, we examined the students' lack of consistency in grading their peers. Student bias is displayed as the tendency to evaluate every assignment with the same grade or to consistently give the assignments higher or consistently lower grade then deserved. As mentioned in the *Introduction*, these students exhibit "rogue" behavior and their impact on the calculation of the final grade should be minimized.

There are different types of biases presented in the paper on metrics written by the teaching staff at University of Alicante [8]: restriction of range bias is the tendency to evaluate every assignment with the same grade; central tendency bias is the tendency to evaluate every assignment with the median or mean grade; leniency bias is the tendency to overrate, and harshness bias is the tendency to underrate assignments. We will examine the average distance to diagonal and its standard deviation measures to determine leniency and harshness biases.

For student i, the grader j assigns the grade a_{ij}. The grade assigned by the teacher (correct grade) is marked as c_i and n_j presents the number of grades assigned by the grader j. The average distance to diagonal is defined as $\overline{DD_j} = \frac{\sum_{\forall i} a_{ij} - c_i}{n_j}$ and presents the mean of differences between the grade assigned by the student grader and the correct grade assigned by the teacher. Every distance is positive if the student graded higher than the teacher, otherwise it is negative when the student graded lower. The standard deviation for the student grader j is defined as $S_{DDj} = \sqrt{\frac{\sum_{\forall i}(a_{ij} - DD_j)^2}{n_j - 1}}$. Considering both the DD_j and S_{DDj} values, it is possible to confirm the presence of leniency and harshness bias. If DD_j obtains a high positive value, the grader overrates his peers' assignments which indicates leniency bias. Whereas low negative values of DD_j indicate that the grader underrates the assignments which represents harshness bias. The low values of standard deviation S_{DDj} shows that the graders bias is more pronounced.

For the assigned grades in TABLE VIII., the grader 1 has a high positive value for DD_j, therefore leniency is present i.e., grader 1 overrated the three tasks of the student 1 assignment. This makes sense for the given data since we know that grader 1 rated all tasks with the highest grade. Grader 2 has a negative DD_j value i.e., underrates the student 1 which concludes harshness bias. For grader 3 the DD_j is 0, thus indicating perfect evaluation equal to the teacher's grades. All three graders have a high standard deviation value, so their grading habits are highly noticeable.

TABLE VIII. ASSIGNED GRADES, AVERAGE DISTANCE TO DIAGONAL AND STANDARD DEVIATION

Grader j	Student 1			DD_j	S_{DDj}
Grader 1	5	5	5	2	3.67
Grader 2	2	3	1	-1	3.8
Grader 3	3	4	2	0	3.8
Correct grade $t_n^{(x)}$	3	4	2	/	/

IV. CONCLUSION

Using peer assessment for calculating grades of assignments decreases time and effort pressure on the teaching staff. It is important to consider student behavior and ensure that the students are graded according to their grading habits. Students should be awarded for their assignment solutions as well as their grading of other students' work. The research shown in this paper has pointed out the interconnection of accuracy, reliability and bias when calculating the final grade in peer assessment process. "Non-rogue" students are reliable graders, and their grade should carry more weight when calculating the final grade. "Rogue" students tend to underrate or overrate. Therefore, they are considered unreliable and biased which can negatively affect the final grade. Their grade should carry less weight for the final grade calculation depending on the difference degree from the correct expected grading.

Analysis of different measures for accuracy indicates that Pearson's correlation coefficient isn't suitable in all grading cases. Choosing the right measure is critical for estimating accuracy, reliability, and bias because it affects the final grade calculation. Looking into calibrated PA, it is worth considering that if only one calibration test is used, which is for most cases, we are not able to evaluate its weight and the calculated evaluations will not be used. In other words, the weight is 0, which is one of the disadvantages of calibrated evaluation.

Peer assessment process is conducted using the Edgar system [20] for the *Databases* course at Faculty of Electrical Engineering and Computing. The PA process generates student data that can provide insight into student evaluation. Since every option that student selects in the Edgar system is saved, there is several years of log event data available for analysis. We plan to expand current analysis of metrics for estimating the best peer assessment grade and use the mentioned metrics on real student data collected from the Edgar system to improve the final grade calculation.

REFERENCES

[1] Y. Xiao, Q. Jia, and J. Cui, "What kind of peer-assessment comments help improve learning outcomes? Evidence from a programming course Machine learning in peer assessment View project Understanding Butterfly Effects and Predictability in Lorenz Models and Global Models View project," 2021. [Online]. Available: www.aaai.org

[2] J. Hamer, K. T. K. Ma, and H. H. F. Kwong, "A Method of Automatic Grade Calibration in Peer Assessment."

[3] Y. Song, E. F. Gehringer, Z. Hu, J. Morris, J. Kidd, and S. Ringleb, "Toward Better Training in Peer Assessment: Does Calibration Help? Paper presented at the CSPRED 2016: Computer-Supported Peer Review in Education," 2016. [Online]. Available: https://digitalcommons.odu.edu/teachinglearning_fac_pubs

[4] "An Analysis of Calibrated Peer Review (CPR) in a Science Lecture Classroom," 2008.

[5] H. K. Suen, "Peer Assessment for Massive Open Online Courses (MOOCs) l sis (SNA) in OnlineCourses."

[6] Y. Hartberg, B. Gunersel, N. J. Simspon, and V. Balester, "Development of Student Writing in Biochemistry Using Calibrated Peer Review," 2008. [Online]. Available: http://cpr.molsci.ucla.edu/

[7] J. R. Rico-Juan, A. J. Gallego, J. J. Valero-Mas, and J. Calvo-Zaragoza, "Statistical semi-supervised system for grading multiple peer-reviewed open-ended works," *Comput Educ*, vol. 126, pp. 264–282, Nov. 2018, doi: 10.1016/j.compedu.2018.07.017.

[8] R. Molina-Carmona, R. Satorre-Cuerda, P. Compañ-Compañ, C.-Rosique, F. Farao´, and F. Llorens-Largo, "Metrics for Estimating Validity, Reliability and Bias in Peer Assessment*."

[9] C. Piech, J. Huang, Z. Chen, C. Do, A. Ng, and D. Koller, "Tuned Models of Peer Assessment in MOOCs," Jul. 2013, [Online]. Available: http://arxiv.org/abs/1307.2579

[10] A. A. Namanloo, J. Thorpe, and A. Salehi-Abari, "Improving Peer Assessment with Graph Neural Networks." [Online]. Available: http://www.tml.cs.

[11] Y. Song, Z. Hu, and E. F. Gehringer, "Closing the circle: Use of students' responses for peer-assessment rubric improvement," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9412, pp. 27–36. doi: 10.1007/978-3-319-25515-6_3.

[12] Y. Tong, C. D. Schunn, and H. Wang, "Why increasing the number of raters only helps sometimes: Reliability and validity of peer assessment across tasks of different complexity," *Studies in Educational Evaluation*, vol. 76, Mar. 2023, doi: 10.1016/j.stueduc.2022.101233.

[13] F. Garcia-Loro, S. Martin, J. A. Ruipérez-Valiente, E. Sancristobal, and M. Castro, "Reviewing and analyzing peer review Inter-Rater Reliability in a MOOC platform," *Comput Educ*, vol. 154, Sep. 2020, doi: 10.1016/j.compedu.2020.103894.

[14] Y. Xiong and H. K. Suen, "A proposed Credibility Index (CI) in peer assessment The TRACE Center for Excellence In Early Childhood Assessment View project Cognitive Diagnostic Models (CDM) View project," 2014. [Online]. Available: https://www.researchgate.net/publication/281783815

[15] "What is 'Calibrated Peer Review' (CPR TM) and Why Are We Doing It?" [Online]. Available: https://cpr.tamu.edu/cpr/cpr/login.asp

[16] J. Russell, S. van Horne, A. S. Ward, E. A. Bettis, and J. Gikonyo, "Variability in students' evaluating processes in peer assessment with calibrated peer review," *J Comput Assist Learn*, vol. 33, no. 2, pp. 178–190, Apr. 2017, doi: 10.1111/jcal.12176.

[17] B. Divjak and M. M. Mareti´c, "Learning Analytics for Peer-assessment: (Dis)advantages, Reliability and Implementation," 2017.

[18] J. Xu, J. Liu, P. Lv, and P. Yang, "Improving Peer Assessment Accuracy by Incorporating Grading Behaviors," Dec. 2021, pp. 1162–1169. doi: 10.1109/ictai52525.2021.00184.

[19] A. Darvishi, H. Khosravi, S. Sadiq, and D. Gašević, "Incorporating AI and learning analytics to build trustworthy peer assessment systems," *British Journal of Educational Technology*, vol. 53, no. 4, pp. 844–875, Jul. 2022, doi: 10.1111/bjet.13233.

[20] I. Mekterovic, L. Brkic, B. Milasinovic, and M. Baranovic, "Building a comprehensive automated programming assessment system," *IEEE Access*, vol. 8, pp. 81154–81172, 2020, doi: 10.1109/ACCESS.2020.2990980.

Exploring the Relationship between Indoor Playrooms and Population in Skopje

A. Naumoski, G. Mirceva, K. Mitreski
Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje
Skopje, Republic of North Macedonia
e-mail: {andreja.naumoski, georgina.mirceva,kosta.mitreski}@finki.ukim.mk

***Abstract* – Play has a crucial role in childrens' growth. It teaches them new skills, strengthens their self-confidence, and promotes creativity. Both indoors and outdoors offer play opportunities, but a nearby playground is crucial as it provides a safe and fun place for children to exercise and explore. This can lead to a healthier lifestyle and improved social skills through interacting with other children. In this paper, we analyze the spatial relationship between indoor playrooms and the population in the study area in Skopje. Our study is based on publicly available data on indoor playrooms, including their location, user satisfaction ratings from customers, and data on population density in the area. Our goal is to find potential locations for new indoor playrooms and improve existing indoor playroom offerings through interpolation, hot spot analysis, and spatial data analysis with multiple ring buffers. Our analysis reveals spatial areas of high population density that offer opportunities to improve indoor playroom products and services.**

Keywords – GIS; indoor playrooms; population; Interpolation; Mutltiple Ring Buffers; Hot Spot Analysis

I. INTRODUCTION

Indoor playrooms offer multiple benefits for children, including physical, social, and cognitive development [1]. Physical benefits include increased exercise and movement, improved motor skills, and improved coordination. Social benefits include the opportunity to interact with other children and develop social skills such as sharing, taking turns and conflict resolution. Cognitive benefits include the opportunity to engage in educational activities and develop problem-solving skills, such as jigsaw puzzles, educational games, and reading materials [2]. Indoor playrooms have become an increasingly popular trend in urban areas, especially high-density cities, to provide a safe and engaging environment for children to play and socialize. With more families living in high-rise buildings or without access to outdoor spaces, indoor playrooms provide a space where children can release their energy and develop their social skills. In this article we will examine the importance of indoor playrooms in relation to population density analysis and why it is important to research this growing trend.

To better understand the spatial distribution of business objects, particularly those related to population, as well as their density, it is crucial to understand their geographic location and model them appropriately. An effective approach that is widely used in various research disciplines for optimal site planning and spatial analysis of the customer base is the use of Geographic Information Systems (GIS). GIS refers to a computerized system designed to collect, store, manipulate, analyze, and manage spatial or geographic data. This system provides a powerful platform for integrating diverse sources of information, including maps, satellite imagery, and other location-based data, enabling researchers to create a visual representation of the data and uncover patterns and relationships that may not be readily apparent through traditional analysis methods. GIS has gained attention in recent years as an important research area in urban planning [3]. GIS is a powerful spatial analysis and visualization tool that enables researchers to study and understand the relationships between the built environment, population density, and access to amenities such as indoor playrooms.

In line with this notion, a recent study by the authors in [4] delved into the use of GPS, accelerometer, and GIS data to scrutinize the extent of physical activity in school playgrounds among teenagers. The results suggest that certain types of play equipment, such as soccer fields and climbing equipment, can encourage physical activity among both boys and girls, while the location of play equipment also plays a crucial role. This study can have practical implications for the design and development of playrooms in schools and other similar environments. The researchers also identified areas where physical activity was comparatively low, such as playgrounds without any play equipment and playgrounds designed for younger children. In recent years, GIS has been used extensively to identify the most favorable locations for the development of new facilities or attractions, such as: wind farms [5], rainwater harvesting systems [6], fire stations [7], photographic studios [8] or book-related services in relationship with schools location and population [9] and many more. Several of these studies have considered the geographic location of stores in relation to the population based on factors such as age, socioeconomic status, or ethnic groups.

The goal of this study is to show users and allow them to explore critical features related to indoor playgrounds. The geodatabase contains essential attributes such as company name, location, and geographic location coordinates. Furthermore, additional attributes have been assigned to each indoor playroom to give users a better understanding of the services they offer. The rating of customer user satisfaction and the number of reviewers per facility give a clear indication of the quality of the services offered. In addition, the project will conduct a

density analysis of the population within the city limits of Skopje, focusing on their proximity to indoor playrooms. This analysis provides valuable visual insight into the accessibility of these recreation centers to residents of the area of interest. This will enable us to improve customer service and identify potential new locations for the construction of additional indoor playgrounds. However, it must be emphasized that this is an approach based on current available data, as the selection of a suitable location for starting a business depends on many other factors in addition to population size.

The remainder of this paper is structured as follows: Section II provides an overview of the data collection process and methodology used for mapping indoor playrooms and population densities in a given area. In Section III, we analyze the relationship between the locations of these facilities and the distribution of population clusters and discuss the implications of our results for identifying new business opportunities. Finally, we summarize our main observations and conclusions in Section IV.

II. Materials and Methods

To map and analyze the necessary data for indoor playrooms and population density, we used ArcGIS software [10]. It is worth noting that the data and analysis presented in this study are limited to the selected geographical location around in the city of Skopje.

A. Data collection and geodatabase

In this section we provide an overview of the data collection process and geodatabase creation for the indoor playrooms and population density analysis in selected area in the city of Skopje. The data collection process involved gathering information about the location of indoor playrooms, the names of their respective companies, and the XY coordinates of each object. Additionally, we collected data on customer satisfaction ratings and the number of reviewers for each facility. This is done by web extraction from several different web sources, such as the Google Search, "Zlatna Kniga" [11] and "Reklama 5" [12]. In total, we collected information on 41 indoor playroom objects.

The data collected for the indoor playrooms and population density analysis is carefully sorted and organized into relevant attributes to describe each object of interest. To effectively manage and analyze this data, an Excel spreadsheet was created. Each row in the table is identified by a unique object ID, and each column corresponds to a specific attribute of interest, given further:

- Name – name of the indoor playroom
- Address – location of the indoor playrooms
- Latitude and longitude – Latitude and longitude expressed as decimal degrees
- Rating – user satisfaction rating according to Google.
- Customers – number of users.

The first step in the conversion process is to import the Excel spreadsheet into ArcGIS software, which then we convert into a geodatabase. The procedure can be performed in ArcMap by selecting and then adding data. This is done using "Add XY data" option and finally connecting to the desired database/document. This process automatically analyzes the latitude and longitude coordinates and we set the coordinate system to the WS 1984 geographic coordinate system. Once all the required parameters are set, the geographic coordinates are displayed as points on the map. Thanks to the use of an accurate coordinate system, most of the displayed coordinates are quite accurate. To label the features of the newly formed layer we use the Label Features tool in ArcMap. To simplify data manipulation and facilitate communication, the necessary map and shape layers we downloaded from a freely accessible website [13]. By organizing the data in this way, we can more easily identify trends and patterns in the distribution of the indoor playroom facilities and population clusters in the desired region.

B. Data on a map

A crucial next step entails the improvement of the visual representation of objects on the map, as they are currently presented solely as dots. Upon creating each object, we use the Symbology tab to modify the associated symbol to more accurately resemble the object that is of interest. Similarly, the Labels tab enables the user to adjust the font size, thereby enhancing clarity and legibility. Notably, the map depicting Skopje alongside the municipalities of Dracevo and Novo Lisice has been obtained from the [13] webpage and named Skopje_Boundary.shp. The map was created via the Export Data function of the gis.osm shapefile and SQL expression, after eliminating all markers except the relevant city borders. By using the Merge tool we merge the above-mentioned polygons, leaving one large shape that is within the city limits of Skopje. This remaining shape is critical for interpolation in the final stages of the project. Geographical objects have been appropriately symbolized, as shown in Fig. 1. In particular, the red icon representing a children's playroom serves a purely aesthetic function and does not affect the results.

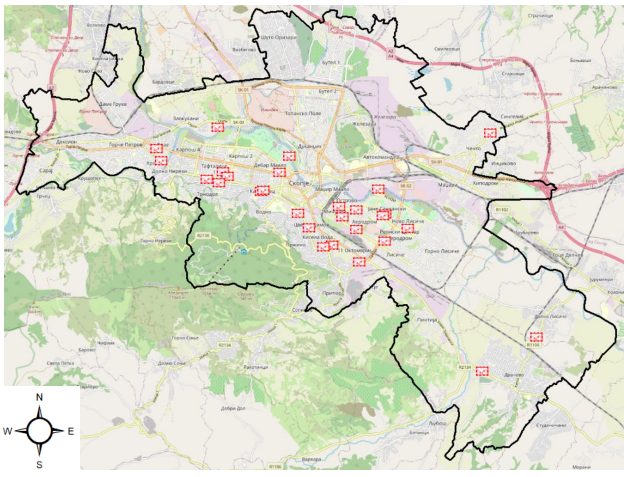


Figure 1. Mapped locations of indoor playrooms objects on the map.

The next step in our analysis involves the import of population density data, which is essential for assessing the child population in Skopje by considering women population density under 50 years [14]. We assume that at least one child accompanies this population. Population density data can be added and viewed in a similar manner to the previous indoor playgrounds data using the ArcMap software with XY Data tool. The program automatically analyzes the longitude and latitude coordinates, which are then displayed as points on the map. It is important to note that while we have population density data for the entire territory of the Republic of North Macedonia, our focus is solely on the city of Skopje and its borders.

III. Results and Disscussion

Spatial interpolation uses points with known values to estimate values at other unknown points and this is a very powerful tool. Spatial interpolation can be used to provide contours for graphical display of data, to calculate some properties of the surface at a given point, or to change the unit of comparison when using different data structures in different layers [18].

A. Interpolation Map

Using the IDW interpolation method [15], we generate an interpolation map from a table containing the locations of indoor playroom objects based on their user satisfaction ratings recorded as rating attributes in the geodatabase. This process enables the identification of zones with better or worse services rated by the customer. This creates of a layer with a heat map showing all the indoor playroom objects, visualized and presented according to the user rating provided by the customers themselves, in this case the children's parents. In order to achieve this, we first need to constrain the space inside the shape itself. This is done from the Geoprocessing - Environments menu. The Processing extent option is used to constrain the workspace by using the Skopje_Boundaries layer created earlier. The Mask option in the Raster analysis is also used to create a mask around this shape layer, giving the heatmap the shape of the layer and a more aesthetically pleasing appearance.

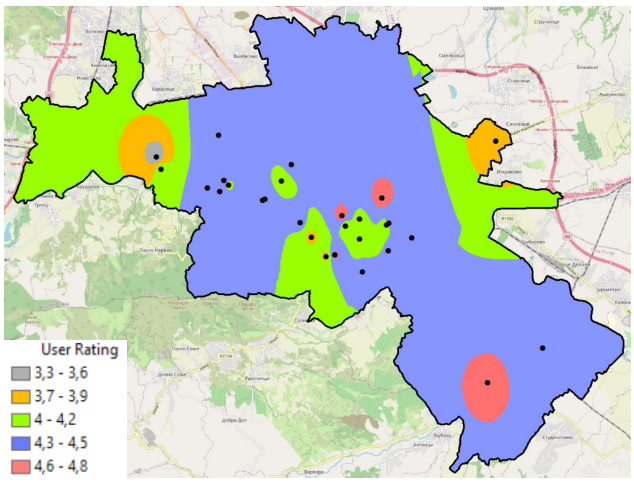


Figure 2. IDW interpolation map for the rating of indoor playrooms objects on the map.

As shown in Fig. 2, the heat map represents all indoor playroom objects where people can bring their children to play and reflects the user satisfaction rating given by the customers themselves. By examining the map results, we found that most of the indoor playrooms have an average rating of 4 stars and some of them outside of the city center have ratings above and below 4. It is important to note that the user satisfactory rating considered not only the feedback from the children, since most of them may not know how to leave ratings, but rather the opinions of the parents about the indoor playroom service. Our next step in the analysis is to relate these results to population densities and their geographical distribution.

B. Hot spot analysis with multiple ring buffers

Population density is a crucial factor that must be taken into account when determining the optimal location for a business object, including indoor playrooms and their services. The accessibility of the location for the target group of customers, in this case children and parents, gives the object a competitive advantage. Our analysis focuses on examining the placement of indoor playroom objects in relation to clusters of targeted population data and the distances between their locations.

In the initial phase of our analysis, we will generate ring buffers to study the optimal location for indoor playroom objects in terms of clusters of female population data and location distance. To achieve this, we will use the Multiple Ring Buffer tool available in ArcGIS software. The Multiple Ring Buffer tool is available in the Proximity toolbox and once the distances are entered the rest of the parameters are set to the default values. Since we have already mapped the indoor playroom objects, we will perform this analysis in addition to these existing data points. Multiple Ring Buffer Analysis is responsible for creating multiple buffers at specific distances around each indoor playroom. The buffers can be displayed without overlapping to avoid discrepancies. For each indoor playroom in selected area, we will create three distance buffers with distances of 100, 300 and 500 meters.

Analyzing population data is a crucial step in understanding the distribution and patterns of people in different regions. Spatial analysis is an effective tool for identifying high and low value clusters within a population dataset. In this regard, the ArcGIS Optimized Hot Spot Analysis (OHSA) tool was instrumental in identifying and visualizing these clusters using different layers of data. Here we discuss the clustering part of the analysis using the ArcGIS OHSA tool over high-resolution population density data [16]. The ArcGIS OHSA tool is a spatial analysis tool that can be used to identify spatial clusters of high and low values. The tool is used to analyze many types of datasets where the spatial distribution of values is important. ArcGIS OHSA tool uses a statistical approach to identify areas of high and low values, taking into account surrounding values and the overall distribution of values within the dataset. We can represent the hot spots, cold spots and middle value areas with different prominent colors. The mean values can be visualized with a color gradient from light to dark to show the deviation from the overall mean value. This allows for a better understanding of the distribution of data on the

population density of target population and areas that require attention. The results of the OHSA are presented in Fig. 3.

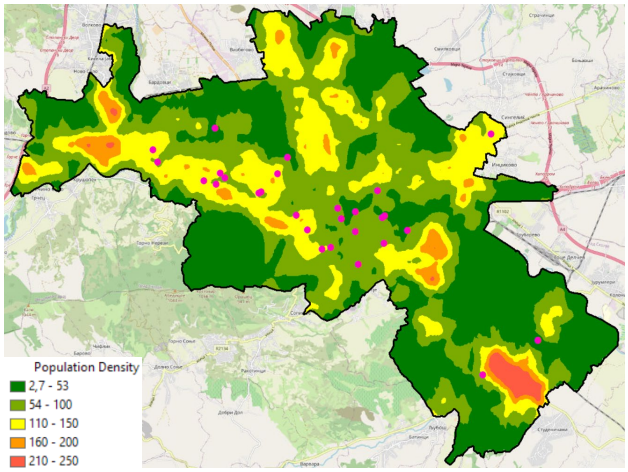


Figure 3. Population clustering as a result from OHSA. Red color depicts a densely populated area, while green covers lightly populated area.

After performing the optimized hot spot analysis with the ArcGIS tool, the output obtained is a point layer that can be difficult to interpret due to overlapping features. To overcome this limitation, we need to convert the point layer to a raster format that can provide a clearer picture of the spatial distribution of the target population. In addition to the optimized hot spot analysis, we also need to integrate the results of the Multiple Ring Buffer tool to get a comprehensive analysis of the distribution of the targeted population. The Multiple Ring Buffer tool is used to create concentric buffers around a point layer so that each ring represents a specific distance from the center of the object. In our case, the center would be the location of the indoor playrooms. By combining the results of the optimized hot spot analysis and the Multiple Ring Buffer tool, we can create a raster layer that shows population densities in various concentric zones around the indoor playrooms objects.

To represent the distribution of the clusters obtained from the earlier analysis, we used an interpolation, specifically using the Kriging tool [17]. The kriging interpolation technique, similar to inverse distance weighting interpolation, is used to generate a surface by analyzing a set of sample points. However, kriging interpolation excels as a geostatistical approach, in contrast to the deterministic nature of IDW interpolation. Although both methods can accept feature classes as input, they rely on the proximity of neighboring sample points to construct the surface. We ran Kriging interpolation on the point features extracted from the optimized hot spot analysis, specifying the NNeighbors attribute as the z-value field. As with the previous IDW task, we configured the environment settings to handle the expansion and masking parameters for the administrative border region of the Skopje_Boundary.shp file.

The resulting interpolation results are shown in Fig. 4 along with the location of indoor playgrounds and multiple ring buffers. Analysis of the results shows that most of the indoor playrooms are in and around the city center where the population density is high. However, we observe that the northern, eastern, and southern regions of Skopje have a lower density of indoor playroom facilities.

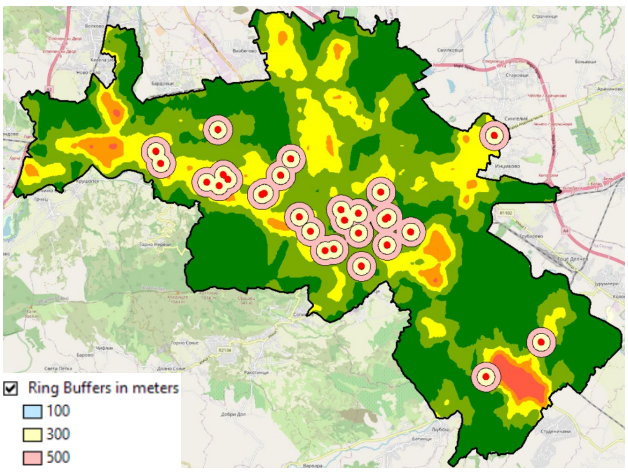


Figure 4. Population clustering map resulted from the OHSA combined with the Multi Ring Buffer tool markers.

Based on these insights, we recommend opening a new indoor playroom with a possible delivery service to attract new customers, improve service quality and increase customer satisfaction. The important thing to note is that the densely populated indoor game room establishments should try to improve their customer service, especially those with ratings of 4 or below. Despite the southeastern region's high population density, there is a shortage of indoor playrooms available to the public. This result suggests potential locations for new business opportunities, as the ability to travel from communities to downtown may not be easily accessible or affordable for some individuals.

IV. Conclusion

This research work focuses on mapping and visualizing the population density related to indoor playroom facilities and services within a given area in the city of Skopje. Our study provides an overview of the geographical distribution and valuation of the indoor playroom markets. Additionally, we identify areas with restricted accessibility by overlaying the locations of indoor play areas with the target audience. Our analysis confirms that the indoor playroom stores are poorly located and present significant access difficulties due to their central location, especially in the North, East and Southeast regions of the study area. To address this issue, we recommend considering opening new facilities in border areas. In summary, this study emphasizes the importance of addressing disparities in accessibility in urban areas in order to better serve the needs of the community and improve their quality of life.

Future research could expand our findings by including an analysis of the average travel time and costs incurred by people using these services. Additionally, including a rating or satisfaction factor from the target audience could provide further insights to balance store placements and move to a more distributed location model.

ACKNOWLEDGMENT

This work was partially financed by the Faculty of Computer Science and Engineering at the Ss. Cyril and Methodius University in Skopje.

REFERENCES

[1] G. W. Evans, “Child development and the physical environment,” Annu. Rev. Psychol. vol. 57, 2006, pp. 423-451.

[2] K. T. Ferguson, R. C. Cassells, J. W. MacAllister, and G. W. Evans, “The physical environment and child development: An international review,” International Journal of Psychology, vol. 48, no. 4, 2013, pp.437-468.

[3] Murray, A.T., Xu, J., Wang, Z. and Church, R.L., 2019. Commercial GIS location analytics: capabilities and performance. International Journal of Geographical Information Science, 33(5), pp.1106-1130.

[4] T. Amholt, and J. Schipperijn, “Where are tweens active in school playgrounds? A hot-spot analysis using GPS, accelerometer, and GIS data,” Mini-Oral Presentation B1. 4. The Health & Fitness Journal of Canada, vol. 14, no. 3, 2021.

[5] M. Szurek, J. Blachowski, and A. Nowacka, “GIS-based method for wind farm location multi-criteria analysis,” Mining science, vol. 21, 2014, pp. 65-81.

[6] K. N. Sayl, A. S. Mohammed, and A.D. Ahmed, “GIS-based approach for rainwater harvesting site selection,” In IOP Conference series: Materials science and engineering, vol. 737, No. 1, p. 012246, 2020.

[7] P. H. Nyimbili, and T. Erden, “GIS-based fuzzy multi-criteria approach for optimal site selection of fire stations in Istanbul, Turkey,” Socio-Economic Planning Sciences, vol. 71, p.100860, 2020.

[8] F. Gjorgjioski and A. Naumoski, "Spatial Analysis of Elderly Access to Photographic Services and Their Rating in Skopje," 2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, 2022, pp. 230-233, doi: 10.1109/ZINC55034.2022.9840605.

[9] N. Netkovska, G. Mirceva and A. Naumoski, "Spatial Relationship between Schools Population and Book Related Services in Skopje," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 997-1001, doi: 10.23919/MIPRO55190.2022.9803578.

[10] ArcGIS Desktop 10.8, Environmental Systems Research Institute.

[11] Zlatna Knig, website. https://www.zk.mk Last accessed: 14 January 2023.

[12] Reklama5, website. https://www.reklama5.mk Last accessed: 12 Octomber 2022.

[13] Geofabrik official webiste. https://www.geofabrik.de/. Last accessed: 12 June 2022.

[14] Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN – Columbia University, “High resolution settlement layer (hrsl).” [Online]. Available: https://data.humdata.org/dataset/north-macedonia-high-resolution-population-density-maps-demographic-estimates. Accessed: 9 October 2022.

[15] D. F. Watson, and G.M. Philip, “Refinement of Inverse Distance Weighted Interpolation,” Geoprocessing 2, 1985, pp.315–327.

[16] A. Getis, and J. K. Ord, “The Analysis of Spatial Association by Use of Distance Statistics,” Geographical Analysis, vol. 24, no. 3, 1992, pp. 189-206.

[17] M. A. Oliver, "Kriging: A Method of Interpolation for Geographical Information Systems," International Journal of Geographic Information Systems, vol. 4, 1990, pp. 313–332.

[18] Spatial Interpolation I, website. https://ibis.geog.ubc.ca/courses/klink/gis.notes/ncgia/u40.html, Last accessed: 18 March 2023.

Predicting Customer Behavior in Support Channels Using Machine Learning

Muhamed Begović, Elma Avdagić-Golub, Belma Memić and Amel Kosovac
University of Sarajevo/ Faculty of Traffic and Communication, Sarajevo, Bosnia and Herzegovina
muhamed.begovic@fsk.unsa.ba, elma.avdagic@fsk.unsa.ba, belma.memic@fsk.unsa.ba, amel.kosovac@fsk.unsa.ba

***Abstract* - Support channels represent a unique opportunity to improve customer satisfaction by offering a consistent experience in resolving customer issues. Several surveys show that customers have raised their standards of customer support services. While only a few years ago customers willingly waited a long time to speak with one of the service agents and were patient for their problem to be resolved, today's customers have very limited patience and want a solution to the problem immediately. Customers don't want to settle for a mediocre support channel experience. Support channels must provide superior service capacities so that customers see that the company values their choice and time. Efficient management of support centers implies accurate modeling of customer behavior on hold. The subject of our research is the application of data research techniques for predicting customer behavior in support channels. In this paper, we apply machine learning methods to predict customer behavior. Based on historical data in the service system, we use classification algorithms to predict customer patience in service channels.**

Keywords - support channels, machine learning, customer patience

I. INTRODUCTION

Service companies face huge competition in the global market. Strategic planning of long-term relations with customers is the key to business success. Contact centers are complex systems where customer requests arrive through different media, which we call support channels. Support channels are calls, e-mails, chatbots, live chats, etc. They have a direct impact on customer satisfaction with the company's service. According to the type of communication, we distinguish two types of support call channels: incoming and outgoing calls[1]. Incoming call channels are most often used to provide information to customers, aiming to solve their problems, while outgoing calls have the purpose of proactive communication with existing or potential customers[2]. Support channels are service systems. The support call channel service process consists of three basic components: the process of the customer's call arriving in the system, the service process, and the customer's waiting time[3]. The standard assumption is that call arrivals follow a Poisson process with a (potential) random arrival rate, assumed to be constant over a given time [4]. A service process is the duration of a conversation between a customer and a customer service representative (CSR). Waiting time is the time required for the customer to receive the necessary information. This time is called the time of patience. Patience time is defined as the time the customer is willing to wait for the service [3]. It depends on the relationship between demand and supply, that is, the number of incoming calls in relation to the service capacity of CSR. When calling, the customer first encounters the Interactive Voice Response (IVR) and triages for the requested services, providing input information. IVR systems for individual customers can be complex and a labyrinth from which the customer leaves the service system. Therefore, efficient management of the service system implies accurate modeling of customer behavior in the service system

Data science is one of the main drivers of changes in business processes, from the way we treat customers to modeling their behavior. Accurate predictive models for customer behavior are key to the design and optimization of business processes[5]. Predicting the customer's patience, i.e. predicting the waiting time of the upcoming call, incoming calls can be treated and routed automatically and intelligently. The biggest challenge of contact center management is to respond to each incoming customer call uniquely in real-time, taking care that all customers do not react in the same way while waiting in the service queue.

In this paper, we observe a telephone support channel with incoming calls. Each call consists of a series of attributes that describe the customer's interaction with the support channel, such as the date, time of the call, reason for the call, and others. We use the data of the user's interaction with the service system, such as the date and time of the call, and use them to describe previous contacts. This creates the possibility for profiling and segmenting customers according to this attribute in the support channel. The main goal of our work is to identify customer behavior patterns. Based on these patterns of behavior, we conclude the premature abandonment of the customer from the service system by applying machine learning. We predict customer behavior at the individual level, based on attributes that describe the customer's recent contacts, as well as patience to speak with CSR. This kind of work can provide instructions for practical application in all companies, especially service companies, through any support channel.

The rest of the paper is organized as follows: Section II presents existing research efforts in this study field, Section III presents the simulation model, and Section IV presents the results of the prediction model and the discussion. Finally, conclusions and guidelines for future research are presented.

II. LITERATURE REVIEW

Quantitative models for designing and managing service systems can be observed through three common analytical techniques: queuing theories, text mining models, and data mining[6]. Problems investigated in the mathematical modeling of call centers can be found in papers [7], [8], [9], and [10]. In particular, papers related to the application of the impatient waiting model in the management of service systems are [11], [12], and [13]. Text mining models refer to obtaining information from customers' text messages, such as detecting emotions, monitoring conversations between customers and CSRs to reject problematic calls, etc.[14], [15]. Data mining models are aimed to predict the intensity of call arrivals [16], [17], predicting the reason for the call based on previous interactions [18] and routing calls to the most appropriate CSR [2].

Customer patience is an attribute that describes behavior in a service system. Patience time is defined as the time the customer is willing to wait, and waiting time is defined as the time the customer needs to wait to receive the service [3]. Patience time is mostly interrupted, as the vast majority of customers get service before their patience time runs out. Among the factors that influence patience is the time already spent on the call, i.e. the total time spent in the IVR and the waiting time in the queue [19]. Positive waiting time is a manifestation of the mismatch between demand (customer service requests) and supply (availability of CSR). In the last case, such a wait results in system failure, i.e. customer abandonment [4].

Our model differs from the mentioned studies. First, the model focuses on predicting the behavior of customers in the queue. Second, the model includes attributes that characterize the customer's previous contacts through the telephone channel. Customers behave differently in the queue, and not all of them are capable of navigating the IVR system. Individual customers may have a greater tendency to opt out of the queue. Based on previous interactions, it can be assessed whether the customer tends to abandon the service system. To the best of our knowledge, our model is the first attempt to incorporate opt-out information from a customer's previous contacts, to predict their behavior in a future telephone channel approach.

III. METHODS

A. Dataset

For this research, we used the operational data of the bank's contact center. The data covers twelve months of one year for each phone call. The database consists of 444,448 phone calls, with a complete history of call movement through the contact center. In the dataset, we distinguish attributes that describe the current call (transaction), such as the date and time of the call, the reason for the call, and the time spent in the queue. These attributes are very important because they affect the customer's patience. As we can see in Fig. 1, the intensity of calls is higher at the beginning of the month, especially in the first week of the month.

In addition, it can be concluded that in the first and last week of the month, there is an increase in unanswered calls compared to answered ones. This means that maybe customers are more impatient or the CSR shift schedule is bad. Also, the system is the busiest around 9 a.m., then the number of incoming calls gradually decreases until around noon and increases again around 4 p.m., as shown in Fig. 2.

The reason for the customer's call is one of the crucial attributes for analyzing the customer's patience. We will consider the reason for the call as the type of problem reported by the customer. Each type has a different intensity of arrival and serving time, and thus the waiting time for a given service [20] [8], [21], [22], as shown in Fig. 3.

We used only those calls that are recognized in the system and have assigned Customer ID attributes since the goal of modeling customer behavior is to predict the class of patience based on customer characteristics.

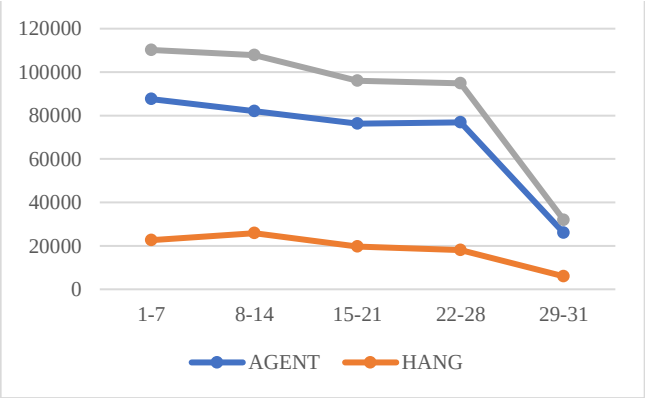


Figure 1. Call volume for one month

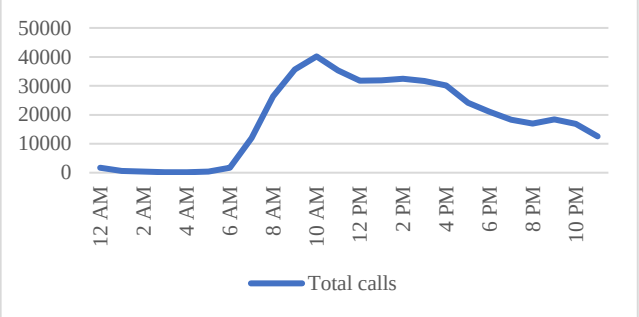


Figure 2. Call volume per hour

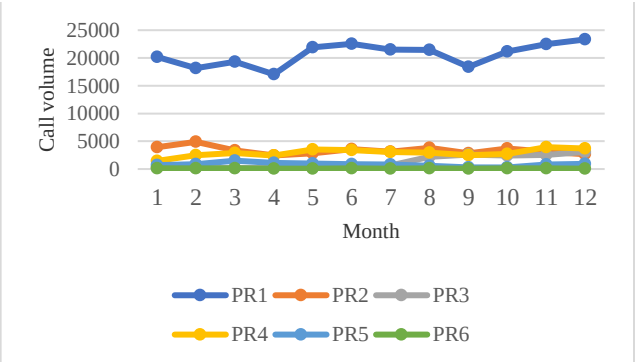


Figure 3. Call volume per type of problem

Our dataset contains 15 columns, and 192,889 individual records (customer calls). 154,441 calls have an "agent" call outcome, and 38,447 calls have a "hang" outcome. The first 6 columns represent categorical variables and need to be encoded. The categorical variables were encoded to convert them into a numerical format so that they can be used as input features in the logistic regression model.

We used One-Hot Encoding, where each unique value in a categorical column is represented as a new column and is assigned a binary value of either 1 or 0, depending on its presence in the row. Each row would have values of either 1 or 0 for these columns, indicating the presence or absence of the respective value in the row. The remaining columns are numerical. The final column, is the target variable that needs to be predicted and it has two possible values: "agent" and "hang". The representation of individual classes in the target variable was uneven in the ratio of 8:1 in favor of the "agent" class. Therefore, we equalized the number of samples for classes and the resulting dataset had 76,894 records.

The data was split into a training set and a testing set, with 80% of the data being used for training and the remaining 20% for testing. The model was fine-tuned using grid search to determine the best hyperparameters for the solver, penalty, and regularization strength.

B. *Prediction of the customer's patience using classification algorithms*

The paper aims to develop a predictive model based on customer characteristics to predict the customer's patience in the service queue during the next call to the contact center, including previous interactions. The target variable for this algorithm is the customer's patience class. So, our problem is to predict whether the call belongs to call class 1 (if the customer will talk to CSR) or 0 (if the customer will leave the service system - hang up).

Let's denote the total number of calls with *I*, for each call i∈{1,2,...*I*} we define the variable y_i as:

$$y_i|x_i = \begin{cases} 1 & \text{if a customer was served by the agent} \\ 0 & \text{if the customer abandoned the system} \end{cases}, \quad (1)$$

where x_i denotes a vector of customer attributes. Therefore, we divide customers into patient and impatient ones. To build a predictive model for the target variable, we used logistic regression, a supervised machine learning algorithm commonly used for binary classification problems.

Referring to linear regression and model training, class probability y_i for characteristic x_i is:

$$p(y_i|x_i) = w_0x_0 + w_1x_1 + \cdots w_Ix_I\,, \quad (2)$$

where *w* is the weight of each feature per class. The vector of characteristics represents the profile of the customer. Customer patience is the aggregate result of assessments based on fourteen characteristics. The attributes used in our analysis are shown in Table I. Customer profile attributes are divided into categorical characteristics and numerical values. Each categorical attribute in the data is modeled by a binary characteristic.

Some attributes require special modeling, which was done in the paper [23]. The attributes of the customer's job, marital status, education, and contact are static and represent the customer's personal characteristics. The attribute of customer priority indicates the importance of the customer to the company.

Customers have different preferences, expectations, and needs from the company, but they also have different incomes and spending profiles, so they must be managed in different ways [24]. This is especially evident in situations when the company changes its business strategies or solves management problems. Not all customers will react the same. This attribute has three categories, normal, priority, and higher priority.

We include four attributes to represent the date and time of past contacts, namely, the month, the regular week of the call, the day period of the call, and the elapsed time since the last call measured in days. The ordinal number of the week shows in which working week in a given month the call is made. Following the example of [25] the day is divided into relatively short time blocks, short enough that the call arrival rate does not change significantly in the block. The day is divided into five segments (time blocks): 7:00 - 11:00, 11:00 - 14:00, 14:00-17:00, 17:00-20:00, and 20:00-24:00.

We also distinguish attributes that show the frequency of recent contacts with the company, i.e. the frequency of calls within a week and the frequency of calls within the same day. The customer's patience can be affected by the period during which the problem is unresolved and how often the customer has to call to get it resolved.

The type of Call is the reason why the customer calls customer support. This is the type of service that the customer requests from the company. The reason for the call can be a decisive factor in the customer's patience. If the problem is urgent to solve, the customer will be ready to wait even longer [19].

The last attribute in the observation is the waiting time. It is important to distinguish between the time the customer has to wait before contacting the CSR and the time the customer is willing to wait before leaving the system. The first time is set by the system according to the number of CSR that are available. The time the customer is willing to wait is a measure of the customer's patience. Clearly, the time before being served is not known in advance. Therefore, in order to be able to use this time as an input variable, it is necessary to make an estimate earlier. The assessment of that time was not the subject of our work. However, this can be made based on parameters such as the intensity of incoming calls, the number of available agents, the period of the day, etc. In addition, it is possible to use advanced methods and machine learning such as those in the papers [26], [27].

We especially emphasize that the target variable (1) is defined individually for each call. Based on previous calls, it is necessary to conclude the characteristics of the customer's patience for the current call. The steps of the prediction algorithm implementation are presented in Fig. 4.

For the proposed model, the algorithm code was written with the help of the Google Collaboratory platform as part of the Google Research project. It is an online platform that enables the implementation of Python code. The Google Collaboratory platform enables the implementation of machine learning algorithms with the use of the necessary processing and memory capacities, which, in the case of large databases, are very difficult to provide on standard computers.

TABLE I. CUSTOMER PROFILE ATTRIBUTES

Customer profile		
Attribute	Data type	Sample
Customer ID	Integer	25365864
Customer job	String /Category	management
Customer marital status	String /Category	single
Customer education	String/Category	Professional, course
Customer contact	String/Category	mobile, telephone
Customer age	Integer	32,54,72
Customer priority	Integer/Category	1,2 or 3
Type of problem	String/Category	PR1,PR2,PR3
Date of call (month)	Integer	1,5,12
Date of call (week)	String/Category	I, II, III, IV, or V
Date of call (time of call)	String/Category	time1, time2
Time from the last call (days)	Integer	2,4,50,60
Call frequency per week	Integer/Category	1,2,3
Call frequency per day	Integer/Category	1,2,3
Queue waiting time (sec)	Integer	15, 25, 54

Start

Step 1. *Importing a dataset D*

Step 2. *Splitting dataset D into training (80%) – D* and testing (20%) – D^{test} dataset*

Step 3. *Training the model on the training dataset D**

Step 4. *Call class prediction on test case*

Step 5. *Assessment of classification efficiency - comparison with the actual call class*

Step 6. *Storing the classification model*

End

Figure 4. Pseudo code of the prediction algorithm

IV. RESULTS AND DISCUSSION

The process used to define hyperparameters in our code involved Grid Search, which is a commonly used technique in machine learning to find the optimal set of hyperparameters for a model. We used Grid Search to find the optimal values for the hyperparameters "penalty", "solver", and "C" of a logistic regression model. The hyperparameters "penalty" had the possible values of "l1" and "l2", "solver" had the possible values of "liblinear", "lbfgs", "sag", "saga", while "C" had a range of values from 0.01 to 1. After defining the grid of hyperparameter values logistic regression model was trained and evaluated for each combination of hyperparameters. The GridSearchCV function from scikit-learn was used, which performs cross-validation for each combination of hyperparameters and returns the set of hyperparameters that result in the best performance. In this way, Grid Search found the best hyperparameters for the model: "C": 0.01, "penalty": "l2", "solver": "lbfgs". These optimal hyperparameters led to improved performance and accuracy for the model.

To evaluate the success of the algorithm, we used the following measures: *Accuracy*, *Precision*, *Recall*, and *F1 scores* [28]. The achieved results show that the model has a high accuracy of 95.46%. This means that out of all the predictions made by the model, 95.46% of them were correct. The precision of 92.05% shows that out of all the positive predictions made by the model, 92.05% of them were actually positive. The recall of 99.39% means that the model was able to correctly identify 99.39% of all the actual positive cases. The F1 Score of 0.96 is a metric that combines both precision and recall and gives a single value as a measure of the model's performance.

The Confusion Matrix provides a more detailed view of the model's performance by showing the number of true positive, false positive, true negative, and false negative predictions. In this case, the matrix shows 7132 true positive predictions, 652 false positive predictions, 46 false negative predictions, and 7549 true negative predictions (Table III).

TABLE II. MEASURE OF SUCCESS OF ALGORITHMS FOR CALL CLASSIFICATION

	Accuracy	***Precision***	***Recall***	***F1 Score***
%	95.46	92.05	99.39	0.96

TABLE III. CONFUSION MATRIX

		ACTUAL	
		positive	**negative**
PREDICTED	**positive**	7132	652
	negative	46	7549

V. CONCLUSION

One of the biggest challenges of managing support call channels is to respond to each incoming customer call uniquely in real time. The job is additionally complicated if we take care that all customers do not react in the same way while waiting in the service queue. In this paper, we have shown that prediction models based on previous customer interactions can predict the customer's propensity to the risk of leaving the queue. To predict the patience of the customer, we ask a key question: "Did the customer wait to be served by the CSR or did he give up?“. This analysis is based on a large data set that includes 76,894 customer calls from a single customer support telephone channel. An important advantage of this research is that we focus on the prediction of call arrival at the individual customer level. This paper demonstrates that by utilizing a specific set of customer attributes as algorithm input, it is feasible to

construct a customer patience prediction model with a satisfactory degree of precision.

The classification model has a satisfactory accuracy and opens more questions for further research. First of all, the dynamic attributes of customer behavior created in this way, which are evaluated during each call, are suitable for classifying calls into different classes of call urgency, which could be a topic for future research. If the customer is characterized as impatient for an immediate call, he can have priority when routing the call to a free CSR. Certainly, respecting possible ethical issues, patience would not be the only factor when deciding the priority of answering calls in the practical implementation. But by combining it with other customer profile attributes (especially the priority and the reason for the call) it is possible to avoid "punishing" patient and "rewarding" impatient customer. Second, it would be interesting to apply other classification algorithms and compare their accuracy. The waiting time is the lower limit of the time customer are willing to wait. It would be interesting to use other authors' patience evaluation models, then apply classification algorithms to them, and compare them with our results. Based on the customer's previous interactions, a certain time can be estimated for which the customer must be approached in order not to leave the queue, which will be the subject of future research.

REFERENCES

[1] A. R. Ali, "Intelligent call routing: Optimizing contact center throughput," *Proc. 11th Int. Work. Multimed. Data Mining, MDMKDD'11 - Held Conjunction with SIGKDD'11*, pp. 34–43, 2011, doi: 10.1145/2237827.2237832.

[2] E. Avdagić-Golub, M. Begović, and A. Kosovac, "Optimization of agent-user matching process using a machine learning algorithms," *TEM J.*, vol. 9, no. 1, 2020, doi: 10.18421/TEM91-22.

[3] G. Li, J. Z. Huang, and H. Shen, "To Wait or Not to Wait: Two-Way Functional Hazards Model for Understanding Waiting in Call Centers," *J. Am. Stat. Assoc.*, vol. 113, no. 524, pp. 1503–1514, 2018, doi: 10.1080/01621459.2018.1423985.

[4] R. Ibrahim, H. Ye, P. L'Ecuyer, and H. Shen, "Modeling and forecasting call center arrivals: A literature survey and a case study," *Int. J. Forecast.*, vol. 32, no. 3, pp. 865–874, 2016, doi: 10.1016/j.ijforecast.2015.11.012.

[5] A. D. Kelleher, J Namee B, *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. The MIT Press, 2015. [Online]. Available: https://www.ptonline.com/articles/how-to-get-better-mfi-results

[6] M. Saberi, O. Khadeer Hussain, and E. Chang, "Past, present and future of contact centers: a literature review," *Bus. Process Manag. J.*, vol. 23, no. 3, pp. 574–597, 2017, doi: 10.1108/BPMJ-02-2015-0018.

[7] V. M. Zeynep Aksin, Mor Armony, "The Modern Call Center: A MultiDisciplinary Perspective on Operations Management Research," *Prod. Oper. Manag. Soc.*, vol. 16, no. 6, pp. 657–664, 2007, doi: 10.3401/poms.

[8] B. Legros, O. Jouini, and Y. Dallery, "A flexible architecture for call centers with skill-based routing," *Int. J. Prod. Econ.*, vol. 159, pp. 192–207, 2015, doi: 10.1016/j.ijpe.2014.09.025.

[9] A. Senderovich, M. Weidlich, A. Gal, and A. Mandelbaum, "Queue mining for delay prediction in multi-class service processes," *Inf. Syst.*, vol. 53, pp. 278–295, 2015, doi: 10.1016/j.is.2015.03.010.

[10] F. Iravani and B. Balcioğlu, "On priority queues with impatient customers," *Queueing Syst.*, vol. 58, no. 4, pp. 239–260, 2008, doi: 10.1007/s11134-008-9069-6.

[11] A. Dudin, C. Kim, O. Dudina, and S. Dudin, "Multi-server queueing system with a generalized phase-type service time distribution as a model of call center with a call-back option," *Ann. Oper. Res.*, vol. 239, no. 2, pp. 401–428, 2016, doi: 10.1007/s10479-014-1626-2.

[12] O. Dudina and S. Dudin, "Queueing system MAP/M/N/N + K operating in random environment as a model of call center," *Commun. Comput. Inf. Sci.*, vol. 356 CCIS, pp. 83–92, 2013, doi: 10.1007/978-3-642-35980-4_10.

[13] M. G. Nosova and M. V. Degtyareva, "Building and Analysis of the Contact Center Model As a Queuing System With Multi-Level Ivr and Impatient Applications," *Вестник Алтайской Академии Экономики И Права*, no. №9 2019, pp. 101–106, 2019, doi: 10.17513/vaael.705.

[14] M. Abhishek Pandharipande and S. K. Kopparapu, "A novel approach to identify problematic call center conversations," *JCSSE 2012 - 9th Int. Jt. Conf. Comput. Sci. Softw. Eng.*, no. April, pp. 1–5, 2012, doi: 10.1109/JCSSE.2012.6261805.

[15] J. Chatterjee, A. Saxena, and G. Vyas, "An automatic and robust system for identification of problematic call centre conversations," *Int. Conf. Micro-Electronics Telecommun. Eng. ICMETE*, pp. 325–330, 2016, doi: 10.1109/ICMETE.2016.48.

[16] V. U. of W. Rafiq A. Mohammed1, "Using Personalized Model to Predict Traffic Jam in Inbound Call Center," *EAI Endorsed Trans. Scalable Inf. Syst.*, 2017, doi: http://dx.doi.org/10.4108/eai.18-1-2017.152101.

[17] R. A. Mohammed and P. Pang, "Agent personalized call center traffic prediction and call distribution," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7063 LNCS, no. PART 2, pp. 1–10, 2011, doi: 10.1007/978-3-642-24958-7_1.

[18] S. Moazeni and R. Andrade, "A Data-Driven Approach to Predict an Individual Customer's Call Arrival in Multichannel Customer Support Centers," *Proc. - 2018 IEEE Int. Congr. Big Data, BigData Congr. 2018 - Part 2018 IEEE World Congr. Serv.*, pp. 66–73, 2018, doi: 10.1109/BigDataCongress.2018.00016.

[19] P. D. Feigin, "Analysis of Customer Patience in a Bank Call Center," 2006.

[20] G. Koole and A. Pot, "An overview of routing and staffing algorithms in multi-skill customer contact centers," *Submitt. Publ.*, pp. 1–42, 2005, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.1714&rep=rep1&type=pdf

[21] J. Chen, J. Dong, and P. Shi, "A survey on skill-based routing with applications to service operations management," *Queueing Syst.*, vol. 96, no. 1–2, pp. 53–82, 2020, doi: 10.1007/s11134-020-09669-5.

[22] X. Sun, "Multi-attribute intelligent queueing method for onboard call centers," *Cluster Comput.*, vol. 22, pp. 5207–

5215, 2019, doi: 10.1007/s10586-017-1173-0.

[23] E. Avdagic-Golub, M. Begovic, S. Causevic, and A. Kosovac, "Profiling Contact Center Customers for Optimization of Call Routing Using Data Mining Techniques," *2021 20th Int. Symp. INFOTEH-JAHORINA, INFOTEH 2021 - Proc.*, no. March, pp. 17–19, 2021, doi: 10.1109/INFOTEH51037.2021.9400671.

[24] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusof, Fachrudin, and T. M. A. Aziz, "Segmentation model of customer lifetime value in Small and Medium Enterprise (SMEs) using K-Means Clustering and LRFM model," *Int. J. Integr. Eng.*, vol. 11, no. 3, pp. 169–180, 2019, doi: 10.30880/ijie.2019.11.03.018.

[25] L. Brown *et al.*, "Statistical analysis of a telephone call center: A queueing-science perspective," *J. Am. Stat. Assoc.*, vol. 100, no. 469, pp. 36–50, 2005, doi: 10.1198/016214504000001808.

[26] M. Thiongane, W. Chan, P. Ecuyer, M. Thiongane, W. Chan, and E. Waiting, "Waiting time predictors for multi-skill call centers," in *2015 Winter Simulation Conference, Huntington Beach, United States*, 2015.

[27] A. I. Kyritsis and M. Deriaz, "A machine learning approach to waiting time prediction in queueing scenarios," *Proc. - 2019 2nd Int. Conf. Artif. Intell. Ind. AI4I 2019*, pp. 17–21, 2019, doi: 10.1109/AI4I46381.2019.00013.

[28] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions," *IEEE Access*, vol. 7, pp. 144907–144924, 2019, doi: 10.1109/ACCESS.2019.2944243.

Modeling Tennis Matches Using Monte Carlo Simulations Incorporating Dynamic Parameters

J. Krčadinac, E.M. Maruševec, L. Jerković, I. Kovač, J. Zloić, A. Šarčević, M. Vranić
University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
jakov.krcadinac@fer.hr

***Abstract* - Although it may seem to be one of the more unpredictable sports, tennis can be rather accurately modelled using the Monte Carlo method. This study aims to evaluate the accuracy of a Monte Carlo simulation that integrates dynamic tennis parameters in forecasting the outcome of a specific match. To predict the outcome of a tennis match, a conventional Monte Carlo simulation based on the identical and independent point distribution assumption requires only two parameters: the probabilities of both players winning a point on their own serve. A more sophisticated method proposed in this paper considers how fatigue affects a player's performance and it analyses and implements the change in the probability of winning a service point after "breaking" an opponent's service game. Calculating the relevant statistics required for player profiling was a critical step in this study. Following that, both previously mentioned variations of the Monte Carlo simulation were implemented to compare their performance. Finally, the method was tested on real-world tennis data.**

Keywords – ***tennis; Monte Carlo simulation; sport match dynamic parameters; prediction; comparing different Monte Carlo simulation versions***

I. INTRODUCTION

Sports analytics has been practiced for decades, but the breakthroughs in data collection have contributed to the popularization of that field in recent years. Sports analytics refers to managing structured data and applying predictive analytics models to analyse diverse aspects of the sports industry. The results of sports analysis are used to assist decision-makers in gaining a competitive advantage in the field of play [1]. Additionally, extensive analysis of historical data can even help in optimizing players' training programs to achieve the best results [2]. In most popular sports, data analysis and the usage of statistics have become commonplace.

Enormous investments in the sports industry are contributing to the immense growth of the sports analytics market in recent years. Consequently, this led to the rapid development of the online gambling and betting industry. Online gambling amounted to 61.5 billion U.S. dollars in 2021 and is expected to rise to 144.4 billion U.S. dollars by 2028 [3].

One of the most popular betting sports is tennis. Tennis is considered one of the hardest sports to play professionally because of its intense physical nature, the need for near-perfect hand-eye coordination during the whole match, and very little financial support for up-and-coming players. The players start their careers by playing in local and regional tournaments, and if they rank highly, they'll get a chance to play in ITF tournaments. If the players are also successful in ITF tournaments, they might get a chance to compete in WTA and ATP tournaments, which are the most important and prestigious tournaments in tennis [4]. Both ATP and WTA tournaments have a sophisticated method of ranking players based on their success in previous tournaments at that level, but most tennis fans are quite aware that those rankings are a very poor indicator when it comes to predicting the winner because lower-ranked players often beat higher ranked players [5].

Despite its seeming unpredictability, tennis is one of the few popular sports that can be rather accurately modelled using methods such as the Monte Carlo simulation, which is the subject of this research. Tennis is very suitable for modelling because a tennis match has no time limit, it is played until one of the players wins and there is no possibility of a draw. It is also an individual sport (in this paper we are not considering the outcome of doubles matches), so there are no complicated team relationships. All these factors contribute to tennis being one of the most popular betting sports [6].

The Monte Carlo method is a mathematical technique used for modelling and simulating complex systems by generating random samples. It generates a set of random numbers according to the data distribution and parameters for each variable to solve mathematical problems and make predictions [7]. This technique is widely used in various fields such as finance, physics, engineering, and computer science. It is particularly useful in problems where an analytical solution is not possible, and instead, the solution is estimated by repeating random trials [8]. Only one parameter for each player is needed to simulate a whole tennis match using the Monte Carlo method, and that is the **probability of a player winning the point when they are serving** [9]. The probability of winning a set, and consequently a match, is independent of which player serves first [10]. By alternating the serving player between games just like in a real tennis match and using

their probability of winning the point whilst simultaneously tracking the result, we can model the whole match. By simulating the match many times and counting how many times each player won, it's possible to estimate which of the two players is more likely to win the match.

The main idea of this paper is to test whether it is possible to improve this simple, already-known approach to predicting the winner of a match by also considering **dynamic parameters** of tennis, such as player **fatigue** and the effect of **one player "breaking" his opponent's service game**.

The research consisted of 3 main stages:

- Data cleaning and extraction of relevant player statistics from the dataset,
- Implementing the calculated statistics in custom versions of the Monte Carlo simulation to predict the winner of tennis matches that already took place in real life,
- Analyzing the results.

The upcoming sections of this paper describe the aforementioned stages of the research and provide further theoretical background on the Monte Carlo simulation. To conclude the research, we compared the winner prediction success rate of different Monte Carlo simulation versions.

II. DATA PREPARATION AND EXTRACTION OF RELEVANT PLAYER STATISTICS

The standard train-test machine learning method was used to test Monte Carlo simulation accuracy. To extract all the needed player statistics and later test the methods' accuracy, a large dataset was used that contains almost all tennis matches that took place in Challenger, ITF, WTA, and ATP tournaments during 2015 and 2016, along with every change in the current score during the match. The dataset was split into two parts – 70% of the data was used for training and 30% of the data was used for testing. The training dataset was also used to calculate the relevant statistics required for player profiling.

In this analysis, a real-life dataset consisting of the results of various tennis matches that were played in 2015 and 2016 was used. The dataset used is the so-called *in-play* dataset, which means that it is supposed to contain all the in-play changes in the score and the corresponding betting odds made by the betting house itself [11]. Apart from the in-play datasets, there are also so-called *play-by-play* datasets. Play-by-play datasets are the most insightful types of datasets in the world of sports. They contain much more information about the within-match events than in-play datasets, but appropriate tools for modelling and simulating that use play-by-play data have not yet been developed, as opposed to the ones that use in-play datasets [12]. Play-by-play is also better for analyzing sport events with more interaction between various of its components, e.g., football matches, where a play-by-play analysis is needed to reflect actual interplay and to qualify different interaction phases [13]. On the other hand, the in-play method is suitable for sport events with more discrete nature of the game, such as tennis, or cricket [14].

Before extracting the relevant statistics from the real-life dataset, the dataset must be thoroughly examined. Any matches that don't have the final result should be discarded, along with doubles matches because only singles matches are relevant for this research. Rows in which the current result is the same as in the previous row were also removed because they are redundant. Since there is much more available data on well-established players who regularly take part in WTA and ATP tournaments, the decision was made to also remove all matches from tournaments that are not of WTA or ATP level. The distribution of the number of matches played for all players in the dataset is shown on Figure 1. To be able to model the performance of individual players well, players that have fewer than 10 matches in ATP or WTA tournaments were removed from the dataset along with all the matches they participated in. Only matches where both players have played more than 10 matches in ATP or WTA tournaments were modelled using the Monte Carlo methods. After this step, 55% of the matches and 28% of the players remained in the training part of the dataset with an average of 11.4 matches per player.

One other problem arose when we tried to analyze the data, and that are the names of players containing non-ASCII characters. There were a few different versions of those players' names in the dataset. For example, there were 4 versions of the name "Marin Čilić", each containing different non-ASCII characters. After identifying all players with non-ASCII characters in their names, those names were lemmatized and standardized (all variations of "Marin Čilić" were changed to "Marin Cilic").

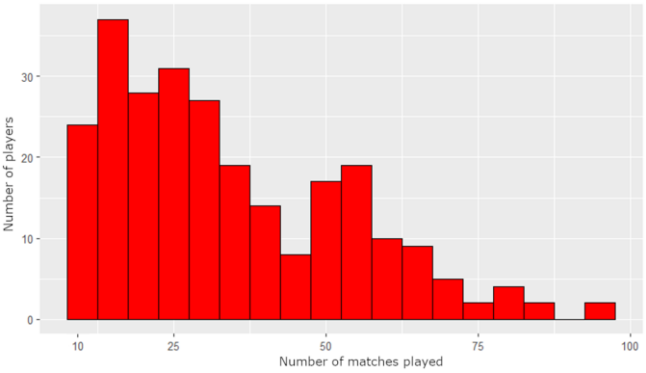


Figure 1. Distribution of the number of matches played for all players in the dataset

After cleaning the data, it was possible to calculate the probability of winning a point when serving for each player in the dataset. That statistic is later used in the basic Monte Carlo prediction model that doesn't take the dynamic parameters of tennis into consideration. One of the dynamic parameters is fatigue, and to test how players perform under fatigue it was decided to differentiate between the probability of winning a point when serving during the first 2 sets and if the game goes past the second set, when almost all players will start to feel fatigued. After calculating the probability of winning a point when serving in the first two sets and the probability of winning a point after the second set, it was clear that the top players on the tour like Djokovic, Nadal, and Federer

manage to either keep or even slightly improve their probability of winning the point when serving during the whole match. When taking into consideration the effect of one player "breaking" his opponent's service game, which is the second dynamic parameter that we propose to be used in the Monte Carlo method, the probability of a player winning the point after breaking their opponent was calculated, along with the probability of winning when the opponent breaks them.

III. MODELLING THE MATCHES USING MONTE CARLO METHODS

A. Monte Carlo method

The term *Monte Carlo methods* refers to a class of computational algorithms that use random sampling from a probability distribution to solve a problem that involves a large number of degrees of freedom and is usually of mathematical or statistical nature. This research uses the Monte Carlo method to simulate real-life tennis matches many times and determine the more likely winner. Monte Carlo methods have a wide range of applications in many different fields of science such as radiological sciences [15], as well as in engineering [16] and finance [17].

One of the earliest applications of Monte Carlo methods was in late 1940s, when physicists Nicholas Metropolis and Stanislaw Ulam developed a method to solve complex integrals that were not analytically computable [18]. Monte Carlo methods have become widespread ever since, finding applications in numerous fields of scientific computing [19].

A variety of problems, from simple integrals to complex simulations of advanced multivariable systems, can be solved by using Monte Carlo methods. Handling systems with uncertain or random inputs is one of the crucial strengths of these methods, which allows researchers to precisely estimate the probability of rare events, as well as to measure the quantity of their results [20].

There are many variations on the basic Monte Carlo method, including importance sampling [21], Markov chain Monte Carlo [22], and quasi-Monte Carlo methods [23]. Importance sampling can improve the efficiency of Monte Carlo simulations by focusing on the most important parts of the probability distribution. Markov chain Monte Carlo methods use a series of dependent random samples to explore complex distributions, while quasi-Monte Carlo methods use low-discrepancy sequences of points to generate more representative samples [24].

One of most widely known applications of Monte Carlo methods is in the financial domain. Experts use them to simulate the behaviour of financial markets [25] and to find the oprimal way of trading financial instruments such as stocks or ETFs. They can be used to assess the probability of a stock portfolio losing certain amount of value over a specific interval, or to appraise an option that depends on the value of an underlying asset [26].

Along with numerous advantages, Monte Carlo methods also have some disadvantages. They can require extremely high computational resources, especially for problems with multidimensional data space or multivariate dependencies, which can be difficult to implement [27]. However, as the system size grows, the problem converges to a simpler, deterministic limit that is cheap to solve [28]. To ensure the accuracy and reliability of Monte Carlo simulations, researchers must pay close attention to the quality and representativeness of the random samples used in the simulation.

In general, Monte Carlo methods are an efficient, multifunctional and very flexible tool used for predicting the most likely outcome and simulating complex systems. That is why researchers in a wide range of scientific fields tend to use them for gaining insight into a variety of problems.

B. Standard Monte Carlo approach

The basic Monte Carlo model used in this research is a point-based model because it only needs the probability of Player A winning any given point versus Player B for predicting the winner of the match [29]. Monte Carlo tennis simulation starts from a random number generator that generates values for the success or failure of both players p_a and p_b sampled from a uniform distribution on the interval [0, 1]. When player A is serving, a value on the unit interval is sampled and if it lies in the range [0, p_a], player A wins the point. If not, player B wins that point. When player B is serving, if the value on the unit interval is in the range [0, p_b], player B gets the point. If not, player A gets that point. [30]. By keeping score after simulating each point and alternating the serve between players just like in a real tennis match, the model predicts the possible result and match winner. Figure 2 shows this scoring system with probabilities used when the player A is serving.

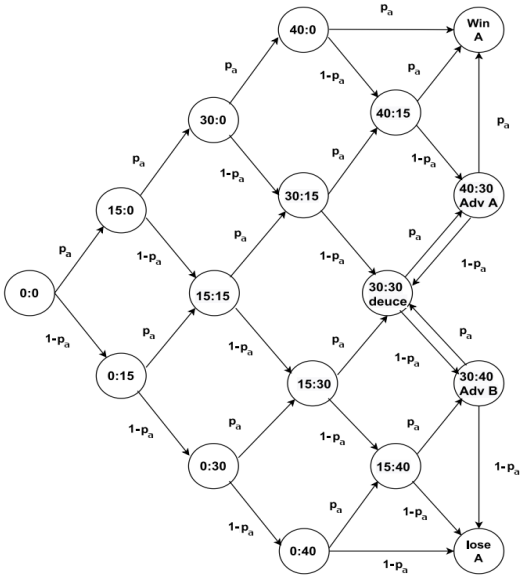


Figure 2. Markov chain for a game in which Player A is serving [37]

Of course, the basis of a Monte Carlo method is not to simulate the whole match just once, but many times and reach the conclusion of who is the more likely winner

[31]. Each match in our test dataset was simulated 400 times. Two other versions of the Monte Carlo model that incorporate dynamic parameters of tennis were also used to predict the outcome of tennis matches, and the accuracy of all three Monte Carlo simulations was compared.

C. *Monte Carlo simulation incorporating dynamic parameters*

As in any sport, fatigue is one of the biggest factors when it comes to an athlete's performance. In tennis, the players need to be very explosive and precise when returning or serving the ball whilst being able to keep that level of explosive strength during the entire match. Because of that, fatigue seems to be one of the most important dynamic parameters to consider when predicting tennis matches. Of course, almost all of the players in ATP and WTA tournaments, especially those at the highest level, try to adjust their style of play during the match to better deal with fatigue [32]. Despite that, in this paper it was assumed that almost all players will start to feel significantly fatigued after 2 sets of play, which equates to approximately an hour and 20 minutes, not taking rest breaks into consideration [33]. That's why it was chosen to differentiate between the percentage of points won when serving during the first two sets, and the percentage of points won after the game goes past the second set. Both percentages were calculated by looking at the in-play data from all the matches in the training part of the dataset. The results from the first two sets were then used when calculating the percentage of points won during them, and when calculating the percentage of points won after the game goes past the second set only results from the third, fourth, or fifth set were considered. Both percentages were used accordingly in the Monte Carlo simulation that incorporates dynamic parameters. This is shown in Figure 3, where it is visible that after the game goes past the second set, the percentage of points won when tired is used instead of the percentage of points won during the first two sets.

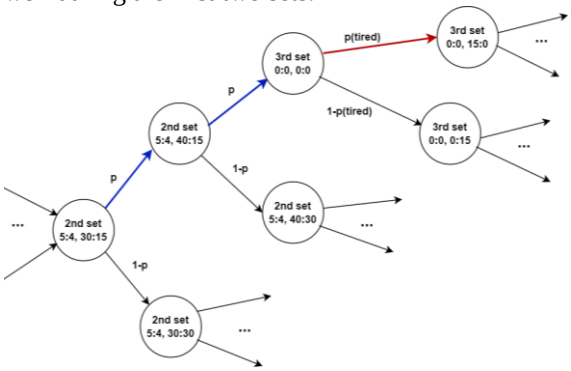


*Figure 3. When the simulated match enters the third set, the Monte Carlo simualtion starts using the calculated percentage of points won after the game goes past the second set (**p(tired)**) instead of the the percentage of points won during the first two sets (**p**)*

In the world of sports, another widely researched parameter of sport dynamic is momentum or the so-called *"hot hand"*. "Hot hand" is a situation where one player exceeds their normal expectations because of a previous victory [34]. This effect is considered very important in table tennis [35] as well as in regular tennis matches [36]. In this paper, we considered the momentum in break situations. In tennis, a break is when a player wins a game as the receiving player, thereby they break the opponent's service game. To account for break situations in matches, we calculated the percentage of points won only in the moments when one player broke more of their opponent's serves in a set than the opponent because in those situations one player has the advantage over the other. Those situations were registered by observing results in breaks whilst considering is serving. If one player is serving and he has more break points than the opponent, that indicates he/she broke at least one more of opponent's serves than his opponent did, and from that point on his/her percentage of points won was recorded. This observed situation lasted until the set was over or when the number of broken opponent's serves from both players become equal. That situation was registered by checking when the player who is serving will have equal or just one break point less than his opponent. Percentages calculated by this method were used in the Monte Carlo simulation that incorporates dynamic parameters in the previously described moments. If there was both a situation of advantages in broken serves and long match duration at the same time, the mean of both percentages was used.

Three versions of the Monte Carlo simulation were implemented in this research. First was standard Monte Carlo. Second was Monte Carlo simulation which only considers fatigue and the third that considers both fatigue and the effect of one player breaking the other.

IV. RESULTS

The accuracy of all the aforementioned Monte Carlo simulations was tested on a testing part of the dataset which consisted of 1232 matches that took place in 2016 on ATP or WTA tournaments which were not used to calculate relevant player statistics. Each match was simulated 400 times to determine the more likely winner, after which the simulation winner predictions were compared to the real winner of each match. For example, in the match between Novak Djokovic and Kei Nishikori that took place at the Australian Open on 26 January 2016, which Djokovic won, the basic version of the Monte Carlo simulation calculated that the chance of Djokovic winning is 58, 50%, the version of Monte Carlo that considers only fatigue calculated the chance of Djokovic winning to be 61,25% and the third version the Monte Carlo method that considers both fatigue and the effect of one player breaking the other calculated the chance of Djokovic winning to be 57,25%. The likelihood of a player winning the match is a percentage calculated by dividing the number of simulated matches in which they won and the number of simulated matches.

Table I presents the number and the percentage of correct winner predictions using each Monte Carlo simulation variations. The accuracy of all methods ended up being almost identical, with all having around 60% correct guesses.

TABLE I
PREDICTION ACCURACY OF DIFFERENT MONTE CARLO METHODS

Monte Carlo simulation	Number of correct winner predictions (out of 1232)	Percentage
Basic	756	61,36%
Including player fatigue	731	59,33%
Including both player fatigue and the probability of winning after a break	725	58,85%

If the Monte Carlo methods calculate that the chance of one player winning the match is just slightly higher than 50%, that means the players are very evenly matched in terms of their skill, and the prediction is highly uncertain. By considering only the matches where the Monte Carlo method is significantly more than 50% certain of the win of one player, each version becomes more accurate in predicting the winner. Table II shows the number and the percentage of correct predictions when one player is more than 70% likely to win. In these circumstances, the percentage of correct guesses increases to 67,21%. Figure 4. shows how accuracy of the Monte Carlo simulation improves when one player has a 70% chance of winning rather than both players having almost equals chance of winning.

TABLE II
PREDICTION ACCURACY OF DIFFERENT MONTE CARLO METHODS WHEN ONE PLAYER IS **MORE THAN 70% LIKELY TO WIN**

Monte Carlo method	Correct winner predictions	Percentage
Basic	378 out of 572	66,08%
Including player fatigue	295 out of 439	67,20%
Including both player fatigue and the probability of winning after a break	291 out of 433	67,21%

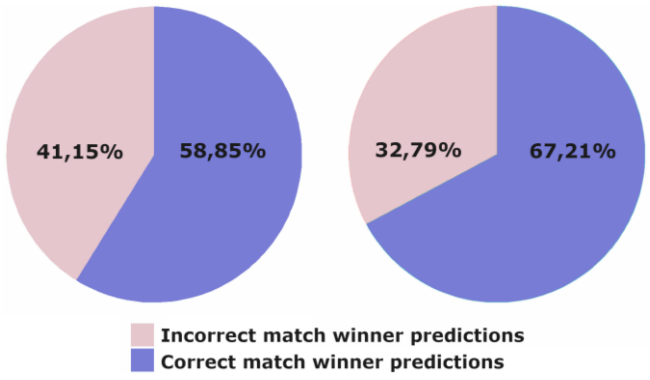


*Figure 4. Accuracy of the Monte Carlo simulation that consideres both player fatigue and the probability of winning after a break when the chance of one player winning is **50%** or more (**left pie chart**) and when the chance of one player winning is **70%** or more (**right pie chart**)*

V. CONCLUSION

The goal of this study was to compare the accuracy of predicting the match winner between different versions of the Monte Carlo simulation. One version uses only the percentage of points won, another version considers both the percentage of points won and player fatigue during the match and the last version uses all of the previously mentioned factors as well as the effect of one player breaking his opponent's service game. The dataset used was a real-life in-play dataset. It contained a lot of irrelevant data that needed to be removed, such as matches without the final result, doubles matches etc. Afterwards, the impact of fatigue on players' performance and the change in the probability of winning a service point after "breaking" an opponent's serve could be calculated. Lastly, all of the previously mentioned Monte Carlo simulation versions were implemented using the calculated player statistics and tested on a different dataset containing matches from 2016.

Even when all the tested versions of Monte Carlo simulations are more than 70% sure of the win of one player, the accuracy of their predictions is about 67%. By using different, perhaps more theoretical models instead of the Monte Carlo simulation such as point-based models, paired comparison models, and machine learning models, the prediction accuracy still doesn't surpass 70% [37]. Wilkens [38] also concludes that the average prediction accuracy cannot be increased to more than about 70% regardless of the model used. This further proves that tennis is a highly unpredictable and dynamic sport, even though it can be easily and accurately modelled. It is impossible to quantify dozens of factors such as emotional state, injuries, and fan support that can influence the outcome of tennis matches [39]. Incorporating dynamic parameters of tennis and building a more sophisticated prediction model based on the Monte Carlo method does not necessarily result in a significant increase in correctly predicted match winners. The results could perhaps be improved by testing different ways of using the calculated player statistics in the simulation that considers dynamic parameters, but there is little hope that the increase in accuracy would be significant.

ACKNOWLEDGMENT

This research has been supported by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS).

REFERENCES

[1] E. Morgulev, O. H. Azar, and R. Lidor, "Sports analytics and the big-data era", International Journal of Data Science and Analytics, 2018, pp. 213-222.

[2] A. Knobbe, J. Orie, N. Hofman, B. Burgh, and R. Cachuco, "Sports analytics for professional speed skating", Data Mining and Knowledge Discovery, 31, 2017, pp. 1872-1902.

[3] Statista, Market size of the online gambling and betting industry worldwide in 2021, with a forecast for 2028, viewed 16 January 2022, <https://www.statista.com/statistics/270728/market-volume-of-online-gaming-worldwide/>.

[4] TennisGate LTD, You do not choose to be a professional tennis player – it chooses you, viewed 6 February 2023, <https://www.tennisgate.com/you-do-not-choose-to-be-a-professional-tennis-player-it-chooses-you/>.

[5] Tennis rankings: How they work and difference between ATP and WTA systems (ATP Rankings – men's singles and doubles tennis rankings), viewed 6 February 2023 <https://olympics.com/en/news/tennis-rankings-atp-wta-men-women-doubles-singles-system-grand-slam-olympics>.

[6] Top 7 biggest betting sports in the world, viewed 19 February 2023 <https://mwwire.com/2022/11/02/top-7-biggest-betting-sports-in-the-world/>.

[7] K. R. Koch, "Monte Carlo methods", GEM - International Journal on Geomathematics, 9, pp. 117-143, 2018

[8] D. Luengo, L. Martino, M. Bugallo, V. Elvira, and S. Simo Särkkä, "A survey of Monte Carlo methods for parameter estimation", EURASIP Journal on Advances in Signal Processing, 25, 2020

[9] P. K. Newton, and K. Aslam, "Monte Carlo tennis", SIAM Review, 48(4), 2006, pp. 722-742.

[10] P. K. Newton, and J. B. Keller, "Probability of Winning at Tennis I. Theory and Data," Studies in applied Mathematics, 114(3), 2005, pp.241-269.

[11] A. Šarčević, D. Pintar, M. Vranić, and A. Gojsalić, "Modeling in-match sports dynamics using the evolving probability method", Applied Sciences-Basel, 11(10), 2021

[12] P. Vračar, E. Štrumbelj, and I. Kononenko, „Modeling basketball play-by-play data", Expert Systems With Applications, 44, 2016, pp. 58-66

[13] F. Korte, D. Link, J. Groll, and M. Lames, 2019., „Play-by-Play network analysis in football", Frontiers in Psychology, 10, 2019.

[14] M. Asif, and I.. G. McHale, „In-play forecasting of win probability in One-Day International cricket: A dynamic logistic regression model", International Journal of Forecasting, 32(1), 2016, pp. 34-43

[15] R. L. Morin, Monte Carlo Simulation in the Radiological Sciences, CRC Press, Boca Raton, 2019

[16] X. Zhang, Z. Lu, K. Cheng, and Y. Wang, "A novel reliability sensitivity analysis method based on directional sampling and Monte Carlo simulation", Proceedings of the Institution of Mechanical Engineers Part O Journal of Risk and Reliability, 234(19), pp. 622-635, 2020

[17] M. S. Shahriar, K. B. M. R. Hasan, T. Hossain, T. H. Beg, K. M. A. Islam, and N. M. Zayed, "Financial decision making and forecasting techniques on project evaluation: A planning, development and entrepreneurial perspective", Academy of Entrepreneurship Journal, 27(4), pp. 1-7, 2021

[18] M. Betancourt, "The Convergence of Markov Chain Monte Carlo Methods: From the Metropolis Method to Hamiltonian Monte Carlo", Annalen der Physik, 531(3), 2017

[19] A. Mücke, R. Engel, J. P. Rachen, R. J. Protheroe, and T. Stanev, "Monte Carlo simulations of photohadronic processes in astrophysics", Computer Physics Communications, 124(2-3), 2000, pp. 290-314

[20] D. Xiu, and J. S. Hesthaven, "High-order collocation methods for differential equations with random inputs", SIAM Journal of Scientific Computing, 27(3), 2005, pp. 1118-1139

[21] T. Müller, B. Mcwilliams, F. Rousselle, M. Gross, and J. Novák, "Neural Importance Sampling", ACM Transactions on Graphics, 38(5), pp. 1-19, 2019

[22] C. Nemeth, and P. Fearnhead, "Stochastic Gradient Markov Chain Monte Carlo", Journal of the American Statistical Association, 116(533), pp. 433-450, 2021

[23] G. Ökten, and Y. Liu, "Randomized quasi-Monte Carlo methods in global sensitivity analysis", Reliability Engineering & System Safety, 210, 2021

[24] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Markov chain Monte Carlo in practice, Chapman & Hall/CRC, Boca Raton, 1996

[25] L. Gan, H. Wang, and Z. Yang, "Machine learning solutions to challenges in finance: An application to the pricing of financial products", Technological Forecasting and Social Change, 153, 2020

[26] D. W. H. Alrabadi, N. I. A. Aljarayesh, "Forecasting stock market returns via Monte Carlo simulation: The case of Amman stock exchange", Jordan Journal of Business Administration, 11(3), 2015.

[27] R. Bardenet, A. Doucet, and C. Holmes, "On Markov chain Monte Carlo methods for tall data", Journal of Machine Learning Research, 18, pp. 1-43, 2017

[28] D. F. Anderson, D. J. Higham, and Y. Sun, "Computational complexity analysis for Monte Carlo approximations of classically scaled population processes", Society for Industrial and Applied Mathematics, 16(3), pp. 1206-1226, 2018

[29] Chew the Stat, MATHS POINT: Finding Tennis' Winning Formula, viewed 12 February 2023, <https://chewthestat.com/finding-the-winning-formula/#:~:text=Points%2Dbased%20modelling%20operates%20on,given%20point%20versus%20Player%20B.>

[30] K. Aslam, "A stochastic Markov chain approach for tennis: Monte Carlo simulation and modeling", Thesis(Ph. D.), University of Southern California, 2012.

[31] E. Koehler, E. Brown, and S. J.-P. A. Haneuse, "On the assessment of Monte Carlo error in simulation-based statistical analyses", The American Statistician, 63(2), 2009, pp. 155-162.

[32] M. Reid, and R. Duffield, "The development of fatigue during match-play tennis", Br J Sports Med., 48, 2014.

[33] M. Ferrante, G. Fonseca, S. Pontarollo, 2017, "How long does a tennis game last?", in Proceedings of MathSport International 2017 Conference, 2017.

[34] M. Bar-Eli, S. Avugos, and M. Raab, "Twenty years of "hot hand" research: Review and critique", Psychology of Sport and Exercise, 7(6), 2006, pp. 525-553

[35] R. J. R. Den Hartigh, and C. Gernigon, "Time-out! How psychological momentum builds up and breaks down in table tennis", Journal of Sports Sciences, 36(23), 2018, pp. 2732-2737

[36] H. Dietl, and C. Nesseler, "Momentum in tennis: controlling the match", UZH Business Working Paper Series 365, University of Zurich, 2017.

[37] Ana Šarčević, Mihaela Vranić, Damir Pintar, Agneza Krajna, 2022., "Predictive modeling of tennis matches: a review", Point based models

[38] S. Wilkens, "Sports prediction and betting models in the machine learning age: The case of tennis", Journal of Sports Analytics, vol. 7, no. 2, pp. 99-117, 2021.

[39] M. de Araujo Fernandes, "Using Soft Computing Techniques for Prediction of Winners in Tennis Matches", Machine Learning Research, 2(3), 2017, 86–98.

Data Warehouse-Based Analytical System in Private Higher Education Institution

Mario Fabijanić*, Domagoj Ružak*, Ana Novosel*,** and Tomislav Hlupić*,**
* Algebra University College, Zagreb, Croatia
** Poslovna Inteligencija, Zagreb, Croatia
tomislav.hlupic@racunarstvo.hr

***Abstract* – Data warehouses have been present in the business environment for several decades, providing businesses with deeper insight into their business data through which they gain competitive advantage and decision-making support. However, the integrated characteristics of data warehouses can be a key factor in improving process automatization through the integration of various heterogeneous sources, speeding up certain processes, and providing process integration in the organization without impacting the existing infrastructure. In this paper, an analytical system based on a data warehouse is presented which integrates data sources in a high education institution, such as internal operational applications and structured and semi-structured files, and incorporates new data sources such as websites and their interfaces. By introducing the integration of the data sources, processes that share the same data scattered across multiple files or need multiple manual entries can use a single source of truth, while also having the possibility of automatic integration of open-access data. In the paper, the necessary processes are identified, and their respective data marts are proposed, while also providing proof-of-concept for several chosen processes.**

Keywords – Data warehouse; data mart; education institution; teaching staff; evaluation system; evaluation

I. INTRODUCTION

Taking care of the quality of the teaching staff and the consequent quality of the education they provide is one of the most important tasks of every higher education institution. This is especially pronounced in private institutions, which must continuously confirm and improve their representation in the market and society. To make this possible, higher education institutions collect relevant data, process them, and categorize teaching staff, on the one hand, to reward the most successful, on the other hand, to empower the less successful.

Data warehouses are a well-known part of higher education information systems. They are introduced at some point during the development of those information systems to provide an efficient analysis of the provided data through an intuitive service [1]. Throughout time, data warehouses grow through time, face typical data warehousing issues (like data quality), and require adjusting the ETL loads based on the analysis of usage statistics [2] [3]. Also, similar systems were introduced on other universities [4], where some have been focused on calculating performance through multiple academic years [5].

At Algebra University College, an evaluation system has been established. It implements the evaluation criteria for the lecturers and their assistants, which are transparently scored and displayed in real-time, at each semester and academic year, for each member of the teaching staff. The evaluation system generates four ranking lists of courses: a list of lecturers, a list of assistants in undergraduate study courses, a list of lecturers, and a list of assistants in graduate study courses. The number of points for the lecturer on a certain course and the number of points for the assistant on a certain course will be within each list. For example, a lecturer who teaches three undergraduate courses will be shown on three lists and will potentially have a different number of points earned for each course.

All lists are divided into five categories. The number of points is monitored continuously, and the final calculation will be made at the end of the academic year for each course since certain elements of scoring depend on the annual work, and some of them depend on the work on the course itself within a particular semester.

The evaluation criteria are divided into categories:

- The result of the first survey in the semester
- The result of the survey at the end of the semester
- Exam quality
- Innovative approaches and good practices in teaching
- Absence of lateness to classes
- Absence of delay in entering grades
- Absence of points input delay
- Absence of delay in setting up courses at the beginning of the semester
- Participation in the preparation of students for competitions / extracurricular activities, or mentoring of final and graduate theses and practice
- Attending training for lecturers and their assistants (participation + results)
- Participation in international exchange, teaching in English, or professional/scientific production in English, related to our institution
- Publication of papers and their registration in CROSBI (*Croatian Scientific Bibliography* [7]).

One aspect of the above categories is to cover the lecturer and lecturer assistant's everyday work quality

during the semester. At Algebra University College, an internal operating system is used to handle the whole teaching process, e.g., the timetable for teaching, course materials, student, and lecturer's (assistant's) attendance lists, exam terms, exam applications, and student survey results. Besides the information needed to support those activities, there are also time stamps for each activity which help to evaluate some of the mentioned criteria.

Another aspect is related to the thrive to strengthen students and teaching staff, e.g., helping students to participate in extra curriculum activities and competitions, or lecturers attending additional training targeted to their needs. The publication of scientific papers also belongs to this aspect, and all those activities must participate in the evaluation process.

The exact formulations for each of the above evaluation criteria nor their weight coefficients are not the subject of this paper, and therefore they will not be explained further. Rather, this paper is dedicated to the warehouse architecture system which may support the whole process.

From the above, the teaching staff evaluation system is complex to implement because the data required for the evaluation itself are found in different IT systems, databases, websites, and individual documents. Manual data collection is not time-optimal, does not always provide information, and is subject to human error. Such mistakes may significantly reduce confidence in the entire evaluation system and can increase dissatisfaction among all stakeholders.

This paper proposes a solution to the described problem in the form of a data warehouse that collects data from different sources, continuously, and transparently, to make the whole process consistent and fair to evaluated lecturers and their assistants. The proposed solution consists of an architecture that is not only limited to the currently valid evaluation criteria but can subsequently be supplemented with new data marts to support some new criteria that may be added in the future. As an example of the usefulness of the proposed system, this paper presents the data mart developed to display data about the scientific production of an individual member of the teaching staff. In addition to the evaluation process, the proposed data mart is an example of usefulness in other needs also, such as discovering the same or similar scientists' research areas, who may not be aware of it.

II. METHODOLOGY

The proposed data warehouse serves as a central point for gathering various data from multiple sources. Examples of these sources are internal operational software with transactional databases, structured and semi-structured data scattered in multiple files that typically have the same data stored for different purposes, and outside sources such as websites with their APIs. Through introducing a data warehouse, these data are cleansed, integrated, and stored in a way that previously independently stored data can potentially provide more insight into existing data.

A. High-level architecture

In Figure 1, the high-level analytical system architecture is shown which will is described in the next paragraphs. The architecture can be divided into three main layers: the staging area, with a staging database and landing zone, the data warehouse, and the business intelligence layer. Data, arriving from the data sources described later in this chapter, are either loaded directly into the staging database or they are stored in the landing zone from where the integration processes fetch the files and load them to appropriate staging database areas. All entities in the staging database reflect the structure of loaded tables (when the sources are other databases) and files that have fixed, known structures. External semi-structured data sources require either flattening the data if applicable (such as JSON and XML formatted data) or preprocessing the data when the flattening process would be intensive performance-wise and would lead to a lengthening of the whole integration process.

The data warehouse is modeled using Kimball's star-schema, bottom-up approach by modeling one data mart at a time. Through modeling, conformed dimensions are easily identified, although several of them fall into this category, such as teacher dimension, student dimension, course dimension, etc. Each data mart also introduces new attributes for existing dimensions which are determined by their final use by a certain business process. Dimensional tables, when describing hierarchical data, are flattened to the highest granularity level to represent the hierarchy in the dimension, rather than in the fact table. Initially, the dimensions are of SCD1 and SCD2 type (when applicable), while they might later be redesigned to match the hybrid slowly changing dimension structure if the business processes will need such an approach.

The business intelligence layer is intended for adjusting the data mostly as a support for building models in various business intelligence applications but also can provide data for various ad-hoc analyses or even advanced analytics usage, such as data mining and machine learning. This architecture area is easily adjustable, representing the data in a user-friendly and application-ready structure, and should handle all necessary adjustments so the data models in different applications would be aligned.

B. Categorization of data sources and integration specifics

Most of the operational data come from a single source, the internal operational software, holding different data about student performance and assignments, logs containing the data of teachers' and assistants' actions, and many other data that can, for example, affect the final teacher's evaluation score. These data are stored in a relational database that can be accessed through direct queries or views adjusted for specific purposes.

Other operational data are typically stored in Excel or CSV files created from manual entries. Some of these data are a result of entries in the online forms, while the others are wrangled, adjusted data that is exported from the previously mentioned operational software.

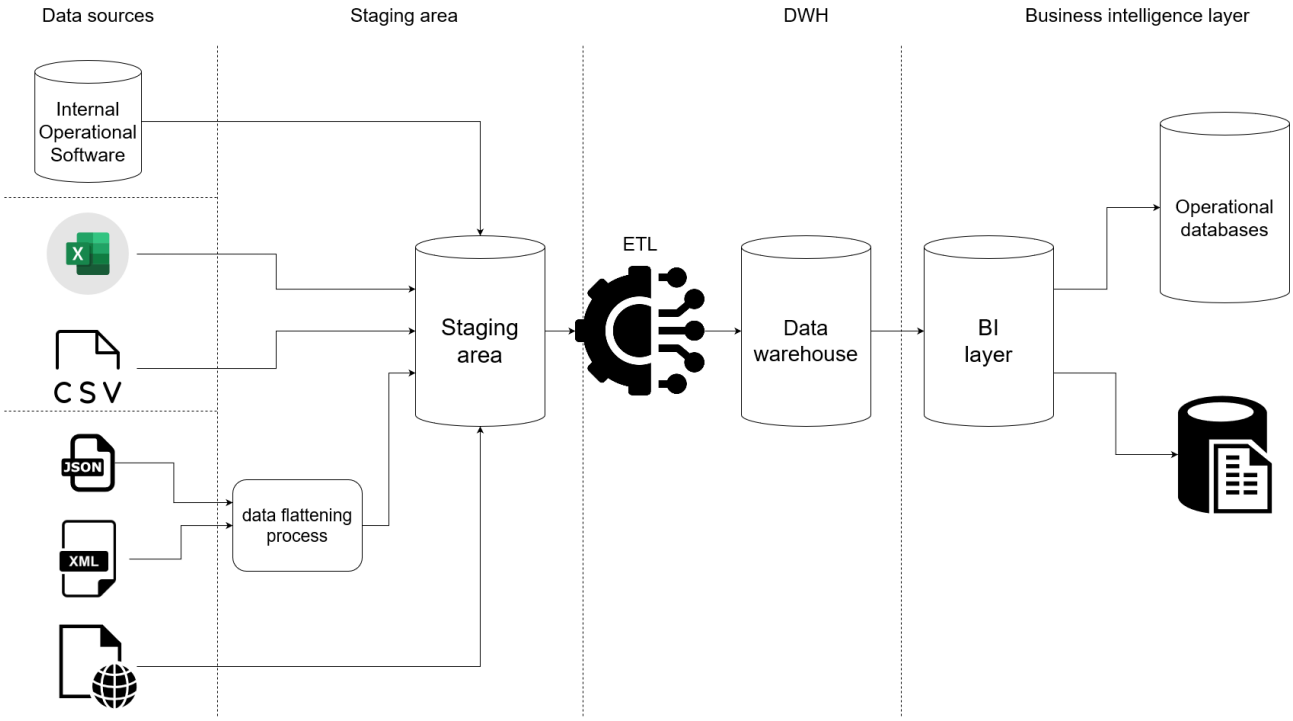


Figure 1. High-level analytical system architecture

Finally, data that is currently largely unused, but can supplement existing data or even become newly used data comes from outside the organization. These data are provided by external websites, each of them being a source for a certain process. An example of the external data from the website and its integration is described in the next chapter. All three data source categories are described in the next paragraphs.

The first data source category includes the data from operational software. These data are integrated on a daily level, which makes the data warehouse a copy of the transactional data, following the definition given by Kimball [6]. That way the basic performance of teachers, such as entering students' scores and grades or course preparations are available for analysis on the next batch load, and the provided data can be shown on dashboards if needed. Moreover, the stored data can be used for further calculations or ad-hoc analysis without impacting the data in the operational software.

The second data source category is various Excel and CSV files whose structure is previously defined. Data stored in the files belonging to this category serve many different purposes. As internal operational software adjustment to new features is lengthy, manual entries serve as data sources for various analyses and reports. Some of these files hold the data originating from the operational software itself but are adjusted or manually wrangled to conform to structures used later for operational and/or analytical purposes. Typically, adjusting the data requires manual labor which generates additional overhead in the processes, often becoming a bottleneck.

Moreover, all the data originating from online forms also fall in this category, as the forms, once created, have a known structure. Even though these data, strictly speaking, are indeed created outside the organization, they belong to the operational data since their structure is defined by the employees inside the organization.

The files containing these data can be integrated both manually and on schedule – both processing approaches include placing files in the landing zone, their ingestion into the staging database, running a defined data quality check, and storing them in the data warehouse.

Fully external data are the third data source category. This category includes all relevant web pages, APIs with semi-structured data, files created outside of the organization (whose structure and content are governed by external users), and generally all files containing data that can augment or supplement existing data from within the organization.

Data sources in this category are usually integrated automatically as the emergence of new external data is unpredictable. The velocity of integration depends on the importance and volume of the data belonging to a certain data source. For example, some data sources create new data daily, but those data are needed periodically (monthly, quarterly…) or only after a certain date, so the integration process schedule is adjusted to the final usage, creating the on-time data in the data warehouse. All automated integrations must also support ad-hoc loads as reloads may be required in certain cases where refreshing the data could be an additional burden.

C. Proposed data mart list

Each data mart in the proposed data warehouse model typically holds the data for a separate business process. Most of the data marts have one or more conformed dimensions, as the dimensionality of this data warehouse

lies in the several inter-joined academic entities such as students, teachers, and courses.

In Table I, a list of proposed data marts with a short description and examples of dimensions are provided.

D. Scientific production data mart

The scientific production data mart is chosen as an example since this data mart serves as a source of data for multiple teaching staff evaluation processes and reports. Currently, three different business processes consume data from this database: the annual scientific publication of an individual lecturer or lecturer assistant, the annual assessment of work, and promotion to a higher teaching or scientific position. In the model, shown in Figure 2, this fact table is named *f_publications*. The main source for the scientific publications data is CROSBI [7], which allows the automatic fetching of all publications belonging to certain authors.

The main prerequisite for the integration is storing the CROSBI identifier in the teacher dimension (marked as *d_teacher* in the data warehouse model), from which a file in BibTeX format with publications can be exported. Other prerequisites include the publication's identifier, depending on its category (DOI for the published papers, ISBN for the published books, etc.) that serves as a unique identifier in the fact table. In this case, a single published publication is seen as a business event with its measures.

The second part is building a bridged table that holds the connection between the publication and its authors. This certain data warehouse design pattern allows a fact table to be related to a dimension in a many-to-many relationship without a time variance since the time variance of the publication is already held in its fact table. Moreover, this supports the occurrences when the author of the existing publication that was published before his tenure at the university entered the dimension – in the next data load, the connection between the publication and the author will be established immediately.

Besides with identifier (DOI/ISBN), each publication is represented by measures, such as the number of authors (calculated during integration), the journal's quartile (derived from the dimension, when applicable), different flags on whether the publication is associated with Algebra University College and whether it was written by an employee or external associate, and dimensional data such as date of publication, the academic year of publication (calculated based on dimensional data), recension type, publication type, etc.

The load of the data in the data mart is done in several steps, as shown in Figure 3. Loading of the data starts after the data fetching from the CROSBI portal when the data about publications from all lecturers and lecturer assistants are gathered using their CROSBI identifiers.

During the preprocessing stage, the DOI and ISBN identifiers are extracted from the data because they can be used to enrich the existing publication description with the data from other sources, such as Scopus journal data, or publication type. Initially, all the data are retrieved, and it is expected that a system that identifies only new entries will be introduced during the next phases of development.

TABLE I. LIST OF PROPOSED DATA MARTS

Data mart	Description	Dimensions
Operational finances data mart	Holds the data for contract payments based on the delivered workload. Data are derived from other data marts	Teacher (d_teacher)
Teacher workload data mart	Holds the data on the delivered workload (lectures, labs, thesis committee…)	Teacher (d_teacher) Course (d_course)
Exams data mart	Holds the data on conducted exams	Teacher (d_teacher) Student (d_student) Course (d_course)
Exam quality data mart	Holds the data on the quality of conducted exam	Teacher (d_teacher) Course (d_course)
Course operational data mart	Holds the course data operational metrics (such as timestamps of the course preparations)	Teacher (d_teacher) Course (d_course)
Thesis data mart	Holds the data of thesis proposal, defense, and committee workload	Teacher (d_teacher) Student (d_student)
Survey data mart	Holds the data of surveys conducted for each course	Teacher (d_teacher) Course (d_course)
Evaluation system reporting data mart	Holds the calculated measures based on the data from other data marts	Teacher (d_teacher) Course (d_course)

The first step in loading the fact table includes fetching the basic publication data from the staging database. The basic data are retrieved from CROSBI data staged in the database, in which BibTeX entries are modified and stored in a single database record.

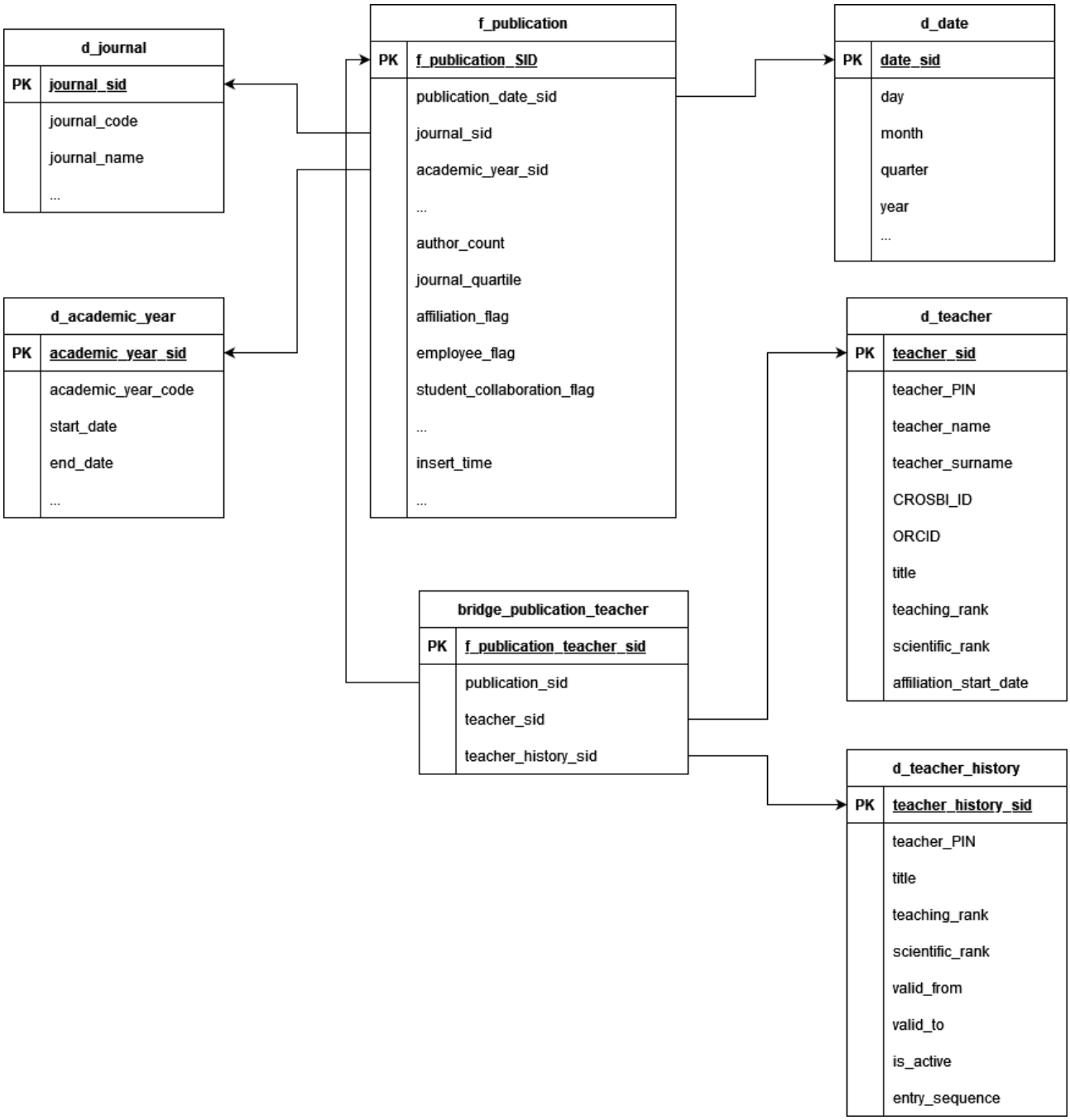


Figure 2. Star schema of the proposed data mart

Further on, those data are enriched with additional data from other data sources, to have the most possible detailed description of a certain entry. Initially, only several data sources are planned besides CROSBI [8] [9], although the extension is easily possible by adding new attributes and updating the existing entries by using their identifiers. After creating the full scope of data for a certain entry, the next step is to integrate them and adjust them to the fact table model through integration packages and lookups to dimensional tables inside them. It should be emphasized that the time-variance of the entry depends on the available information and the publication date, if not available through the data, should be derived based on the set of rules (e.g., if the day of the publication is not available, but month and year are, the date is set to the last day of the given month). Some of the mappings between the fact table and dimension tables are deliberately redundant to enable easier ad-hoc analysis and report creation. The last step before loading the data into the fact table is creating derived measures and flags. These measures and attributes are created based on the set of previously defined rules, such as calculating the number of authors, marking the affiliation to the parent institution, defining whether the publication is a collaboration between a teacher and a student, etc. Those measures are later used in various reports, impacting the performance of the reports (since the reporting model is simpler and the calculations do not have to be done during the report creation), and can be used in multiple different reports. Finally, after the fact table is loaded, the bridge table that connects publications with authors is filled.

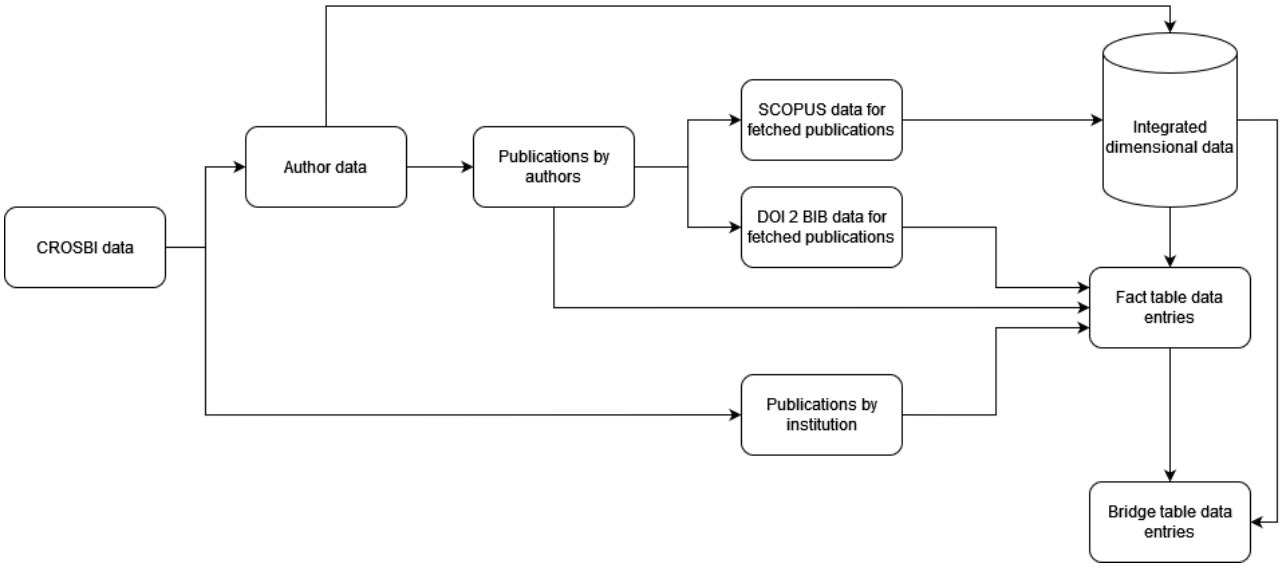


Figure 3. Publication data mart load data flow

The load of the bridge table comes in two phases: during the first phase, the data about the publications currently located in the staging database are loaded, while the second phase identifies the new teachers loaded in the teacher's dimension and loads the history of their publications. The identification of new entries is done through a timestamp stored as a technical attribute in the dimension. As the publications belonging to a certain person are retrieved during the staging phase, this allows adding entries for both new and existing publications.

III. CONCLUSION

In this paper, an analytical system built on a data warehouse is introduced, to implement an architecture to support teaching staff evaluation within private higher education institution. It merges existing data sources, like internal operational applications, structured and semi-structured files, and websites and their interfaces, as well as adds new sources of data. By providing the integration of the data sources, operations that require several human entries or share the same data dispersed across multiple files may now use a single source of truth and automatically integrate open-access data. The paper identifies the required processes and proposes the appropriate data mart as the proof-of-concept for a few selected operations. Future work will cover the implementation of new data marts needed to make the proposed analytical system more comprehensive, covering the majority or even all the needs that higher education institutions have regarding their teaching staff evaluation.

REFERENCES

[1] Baranović, M., Madunić, M., & Mekterović, I. (2003). Data warehouse as a part of the higher education information system in Croatia. Proceedings of the International Conference on Information Technology Interfaces, ITI, 121–126.

[2] Mekterovic, I., Brkic, L., & Baranovic, M. (2009). Improving the ETL process of higher education information system data warehouse. Proceedings of the 9th WSEAS, 8(10), 265–270.

[3] Mekterović, I., Brkić, L., & Baranović, M. (2009). Improving the ETL process and maintenance of higher education information system data warehouse. WSEAS Transactions on Computers, 8(10), 1681–1690.

[4] Dell'Aquila, C., & Tria, F. Di. (2007). An academic data warehouse. AIC'07: Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Informatics and Communications, 229–235.

[5] Mihai, P., Marian, L. M., & Alexandru, L. M. (2010). Measuring The Performance Of Educational Entities With A Data Warehouse. Annales Universitatis Apulensis Series Oeconomica, 1(12), 176–184.

[6] Kimball, R., & Ross, M. (2002). The data warehouse toolkit: the complete guide to dimensional modelling. In Nachdr.]. New York [ua]: Wiley.

[7] CROSBI - Croatian Science Bibliography. https://www.bib.irb.hr/ (Date of access 6. 2. 2023).

[8] Rose, M. E., & Kitchin, J. R. (2019). pybliometrics: Scriptable bibliometrics using a Python interface to Scopus. SoftwareX, 10, 100263.

[9] DOI2BIB. (n.d.). https://www.doi2bib.org/bib/ (Date of access 6. 2. 2023.)

Opinion Mining on E-commerce Live Broadcast of Agricultural Products Based on WLDA-Apriori Model

Binghao Wang*, Yifan Wang†, Zeqi Wang†
* Chengdu City Construction Investment and Management Group Co., Ltd, Chengdu, China
† Wuhan University, Wuhan, China
yifwang@whu.edu.cn

***Abstract*—With the popularization of information technology in rural society, the e-commerce live broadcast of agricultural products has provided an impetus for agricultural marketing from rural to urban areas. But as an emerging industry, it has problems in live broadcast content, marketing strategies, market supervision, industry norms and agricultural aid effects. In order to fully understand the public opinion and promote the sustainable development of this industry, this paper proposes an opinion mining method based on WLDA-Apriori model. The method combines topic model with association rule algorithm, which can further analyze the implicit semantic relations among topic words. This paper also conducts the empirical study with social media data from Weibo, and obtains eight topics the Chinese public concerned. Based on topic words and their relations, the public opinion are identified and discussed from aspects of Industry Development, Market Effect, Live Broadcast Subject and Negative Effect. By encompassing as many netizens as possible, the research scope of this paper surpasses the constraints of conventional questionnaire surveys and case studies, enhancing the generalizability of the research findings and recommendations. But this paper does not integrate and analyze social media data from multiple platforms, nor does it analyze public opinion from the sentiment level.**

***Keywords*—public opinion, e-commerce live broadcast, agricultural products, topic model, association rule**

I. Introduction

In the era of mobile Internet, e-commerce live broadcast has become a new online marketing method in which merchants interact with consumers instantly in order to achieve the goals of gathering popularity, increasing sales and building brands. With the continuous expansion of information technology application in rural, e-commerce live broadcast has gradually become one of the main ways to sell agricultural products. Compared with the image display provided by traditional e-commerce, the e-commerce live broadcast of agricultural products enable users to perceive information from more aspects, which has a positive impact on users' purchasing behavior. Especially since the outbreak of the COVID-19 pandemic in early 2020, the traditional circulation channels of agricultural products have been blocked, and the sales of a large number of fresh agricultural products have stagnated. With its innovative advantages in the transaction mode, e-commerce live broadcast of agricultural products has realized the efficient match between the agricultural supply and demand information, and become an important mode of helping farmers and poverty alleviation. In Q1 2020, online sales of agricultural products in China surpassed 90 billion yuan (around 13.9 billion US dollars), while the number of live broadcasts for agricultural e-commerce exceeded 4 million. [1].

However, as an emerging industry, the e-commerce live broadcast of agricultural products needs to be improved in terms of live broadcast content, marketing strategies, market supervision, industry norms and agricultural aid effects. How to give full play to its role and realize the sustainable development of this industry is a common concern of the government, enterprises and academia. Previous research has predominantly utilized questionnaires [2], mathematical models [3], and case analyses [4] to investigate the factors impacting the development of agricultural products e-commerce live broadcasting and suggest corresponding recommendations. Due to the limited scope of the questionnaire audience and cases, these studies could only meet the needs of specific populations, and the general applicability is relatively poor. At the same time, social media provides a rich source of information, including the basic opinions of the public on some issues. Making good use of this resource can help managers fully understand the public's expectations for the e-commerce live broadcast of agricultural products, and clarify the priorities of various matters to serve as an important basis for future work.

From the perspective of public opinion mining, this paper adopts topic model and association rule algorithm to carry out topic extraction and topic word association on the social media data of agricultural products e-commerce live broadcast. On this basis, this paper identifies the public opinion on the e-commerce live broadcast of agricultural products, and discusses the meanings or problems behind them. The research questions in this paper are as follows:

- RQ1: How to combine topic model and association rule algorithm to mine public opinion?
- RQ2: What are the topics concerned by the Chinese public about the e-commerce live broadcast of agri-

This paper is supported by the Chinese NSFC Major Program (Grant No. 72074172). Yifan Wang is the corresponding author.

cultural products?
- RQ3: What are the Chinese public opinion on the e-commerce live broadcast of agricultural products?

II. Related Work

A. E-commerce Live Broadcast of Agriculture Products

With the emergence of e-commerce live broadcast of agricultural products, researchers have studied its significance and mechanism. Liu et al. [2] took the rural e-commerce in Xinhua County as the research object and conducted a questionnaire survey. They found that under the influence of COVID-19, the e-commerce live broadcast had become the main way to promote poverty alleviation in counties. Lu [5] conducted empirical analysis on the questionnaire data issued by sellers selling navel oranges through e-commerce live broadcasts. It turned out that e-commerce live broadcast can effectively drive the development of fresh agricultural products e-commerce. Zeng et al. [6] demonstrated the internal mechanism of agricultural marketing from rural to urban areas shaped by e-commerce live broadcast. It is specifically reflected that e-commerce live broadcast transforms the agricultural marketing into short-chain value transmission, multi-value output and value co-creation.

In addition, researchers have also studied the influence factors and development countermeasures of e-commerce live broadcast effect of agricultural products. Some studies explored the influencing factors of consumers' purchase intention in e-commerce live broadcast, and made suggestions on e-commerce live broadcast of agricultural products. Shi et al. [3] presented a model for precise marketing classification of agricultural products using clustering, and used customer development space to achieve precise marketing for the e-commerce live broadcast platform. Combined with Stimulus Organism Response and Perceived Value Theory, Yang et al. [7] constructed the model of consumers' purchasing intention of fresh agricultural products under the live broadcast situation, and used SPSS and AMOS to verify the reliability and validity of the questionnaire data. Other studies innovatively reveal the influencing factors of whether the live broadcast of agricultural products can be adopted from the perspective of operators. For instance, Gao and Sheng [4] carried out an empirical study on the influencing factors and mechanism of e-commerce live broadcast adoption intention of new agricultural operation entities. The result showed that the acknowledge factors, subject factors and operation factors of the operators are the key influencing factors.

Existing studies have demonstrated the value and mechanism of agricultural products e-commerce live broadcast from the aspects of poverty alleviation, rural revitalization and e-commerce development. Besides, from the perspective of consumers and operators, they have respectively explored the factors affecting the development of agricultural products e-commerce live broadcast, and put forward suggestions for the government, industry and various participants. However, due to the limitations of data and experimental methods, the existing research fails to view from the perspective of public opinion, and lacks of general applicability, which needs to be further supplemented.

B. Opinion Mining in Social Media

Text mining technology has been extensively applied and validated in various domains such as big data analysis in finance [8], news text classification [9], sentiment analysis on online platforms [10], and many others. Public opinion mining is another significant application of text mining technology, which involves analyzing text to identify and extract opinions expressed by authors or groups regarding a particular object or subject from a vast amount of information. In the context of social media, opinion mining refers to the process of extracting sensitive or highly relevant information from social media data, analyzing public sentiment and trends, and providing guidance for decision-making. [11]. Because public perspective has long been ignored in ex-post evaluation of large cross-regional projects, Wan et al. [12] used lexicon-based sentiment analysis and LDA topic modeling technologies to identify the patterns of public sentiments and topics from spatial and temporal perspectives. Fu et al. [13] used the GDELT and TWITTER platforms' data to mine the public opinion on online learning during the COVID-19 pandemic and found that there was a difference between the topics that the media focused on and the content that the audience talked about.

Researchers have mainly mined public opinion from the topic or sentiment level so far. Among them, opinion identification at the topic level is only based on separate topic words, and the implicit semantic associations among topic words have not been found. Besides, The method of combining topic model with association rule algorithm has been applied to other research problems to improve the interpretability of topics. Wang [14] constructed a mining method for ideological and political knowledge elements based on LDA model and Apriori algorithm. Ruan et al. [15] used LDA model to extract topics from text data, and then used association rules to enhance the semantics of words describing topics, so as to mine the semantic associations of text topics. Wu et al. [16] combined BTM (Biterm Topic Model) with association rules to conduct topic discovery on COVID-19 vaccine.

III. Methodology

This paper proposes an opinion mining method based on WLDA-Apriori model. This method can be divided into the following steps. First, construct the WLDA topic model weighted by TF-IDF algorithm, and extract topic information from social media data. In this step, each topic and its separate topic words will be obtained. Next, based on the Apriori association rule algorithm, the association rules among separate topic words are mined to obtain the implicit semantic relations. Finally, the public opinion of each topic will be identified by analyzing the topic words and the semantic relations among them.

A. Weighted LDA Topic Model

LDA (Latent Dirichlet Allocation) [17] is a three-layer Bayesian probability model that includes words, topics and documents. It obtains topic features through the probability distribution of topics, which can be used to identify topic information hidden in large-scale corpora. The LDA model uses Gibbs sampling method to estimate the posterior distribution of topics, but this method causes the distribution of topics and words to incline to high-frequency words rather than words that can represent documents, thus reducing the model's ability to describe text information. Besides, deleting stop words can not completely filter out these high-frequency words with poor representation.

To solve this problem, Wilson et al. [18] have proved that reasonable weighting strategies based on LDA model will enable the natural allocation of high-frequency words. The current weighting strategies mainly use the text style structure or the distribution characteristics of items. The latter introduces effective mathematical functions into the representation model by considering the distribution characteristics of items, so that the term set is not submerged by a large number of high-frequency words. TF-IDF (Term Frequency - Inverse Document Frequency) is a widely used weight calculation formula, and previous studies [19], [20] have combined LDA topic model with TF-IDF to improve the accuracy of text representation. TF-IDF considers the frequency of terms on the one hand, and the inverse document frequency on the other. The specific formula is as follows:

$$\mathrm{WT}\left(w_{ij}\right)=TF\left(w_i\right)*\mathrm{IDF}\left(w_i\right)=tf_{ij}*\log\left(\frac{n}{n_i+\mu}\right) \tag{1}$$

Where $TF\left(w_i\right)$ calculated by tf_{ij} represents the number of occurrences of term w_i in document j, $\mathrm{IDF}\left(w_i\right)$ calculated by $\log\left(\frac{n}{n_i+\mu}\right)$ represents the inverse document frequency of term w_i, n is the total number of documents, n_i is the number of documents containing term w_i, and μ is the adjustment factor.

In this paper, the weight value $WT\left(w_{ij}\right)$ based on TF-IDF is assigned to each term and extended to the WLDA model. A new Gibbs sampling formula is formed as follows:

$$\begin{aligned} &p\left(z_{ij}=k\mid z^{-ij},\alpha,\beta\right)\\ &\propto\frac{\sum_{j=1}^{M}WT(w_{ij})n_{ijk}^{-ij}+\beta}{\sum_{i=1}^{V}\sum_{j=1}^{M}WT(w_{ij})n_{ijk}^{-ij}+V\beta}\times\frac{\sum_{i=1}^{V}WT(w_{ij})n_{ijk}^{-ij}+\alpha}{\sum_{k=1}^{K}\sum_{i=1}^{V}WT(w_{ij})n_{ijk}^{-ij}+K\alpha} \end{aligned} \tag{2}$$

Where z_{ij} represents the topic variable of the term w_i in the document d_j, $-ij$ represents the term w_i in the excluded document d_j, n_{ijk} represents the number of times the term w_i in the document d_j is assigned to the topic k, α represents the Dirichlet prior distribution of topics, β represents the Dirichlet prior distribution of words, K represents the number of topics, and V represents the total number of terms in the set.

B. Apriori Association Rule Algorithm

Association rule is a common data mining method, which is originally designed for Market Basket Analysis, and then used to discover potential rules in a large amount of data, so as to explore the correlation among data. Apriori is a classical algorithm for association rule analysis. It adopts a layer-by-layer search method, and is easy to implement. This paper uses Apriori association rule algorithm to mine association rules among topic words.

To illustrate the process of mining association rules using Apriori, the following assumptions are made: the word set is I, and any subset of I is called the itemset, which is represented by X; the number of words contained in each itemset is called the length of the itemset, and an itemset whose length is k is called the k-itemset; the whole corpus is D, and $|D|$ indicates the number of documents contained in corpus D; the number of occurrences of an itemset in corpus D is called frequency, which is expressed as $count(X)$.

First, generate frequent itemsets based on support. Support indicates the probability of occurrence of itemset X in all itemsets. Given a minimum support, if the support of X is bigger than the minimum support, it is called a frequent itemset. The specific formula is as follows:

$$\mathrm{support}(X)=\mathrm{count}(X)/|D| \tag{3}$$

Second, generate strong association rules based on confidence. Confidence indicates the possibility of containing the feature word Y in the item set containing the feature word X. Given a minimum confidence, when the confidence is bigger than the minimum confidence, the corresponding rule is called a strong association rule. The specific formula is as follows:

$$\text{confidence }(X\Rightarrow Y)=\mathrm{support}(XY)/\text{ support }(X) \tag{4}$$

Third, filter effective strong association rules based on lift. Lift indicates the ratio of the probability of containing the term Y at the same time with the term X to the probability of the occurrence of the term Y. Association rules that meet the minimum support and minimum confidence need to calculate corresponding lift. Generally, when the lift is bigger than 3, the association rule can be considered as a effective strong association rule. The specific formula is as follows:

$$\text{lift }(X\Rightarrow Y)=\mathrm{support}(XY)/\text{ support }(X)\text{ support }(Y) \tag{5}$$

IV. Result & Discussion

A. Data Preparation

With an increasing number of users turning to it as their primary channel for event exposure and emotional expression, Weibo has become a platform in China that has a significant impact on shaping public opinion. Therefore, this paper selected Weibo as the data source, and used "e-commerce live broadcast" combined with "agricultural

products", "help farmers" and other related words as keywords to obtain social media data. At the same time, in order to better reflect the users' opinion, only original microblogs were collected. In August 2022, 8488 microblogs related to agriculture products e-commerce live broadcasts were collected, which included information such as the texts, release time, and publishers. These microblogs spanned from September 2017 to August 2022. The collected data underwent preprocessing, which involved segmenting the microblog texts using the Jieba toolkit in Python, with continuous improvement through the addition of a custom dictionary. A stopword dictionary was also created to remove irrelevant words. As a result of these steps, a social media dataset about agriculture products e-commerce live broadcasts was created.

B. Topic Extraction

Based on the Gensim toolkit in Python, this paper introduced TF-IDF algorithm and constructed a weighted LDA model to achieve topic extraction of the data set. After the calculation of perplexity, it is found that when the number of topics is 8, the perplexity is lower. Therefore, this paper set the optimal number of topics K to 8. Other experimental parameters were set as follows: α is $50/K$, β is 0.01, and the number of iterations is 100.

After using the weighted LDA model to extract topics from the data set, eight topics and the term distribution of each topic were obtained. Under each topic, ten words with high probability were selected as the topic words, and the topic name was inferred according to the topic words. The results are shown in TABLE I.

TABLE I: Concerned Topics Related to E-commerce Live Broadcast of Agricultural Products

ID	Topic Name	Topic Words
1	Industry Convergence	development, industry, construction, agriculture, the masses, promotion, rural, tourism, base, service
2	Rise of E-commerce Live Broadcast Industry	rural vitalization, activity, poverty alleviation, characteristic, culture, practice, sales, in field, youth, increase revenue
3	Participation of Public Welfare Forces	corn, Yugang Li, homeland, hairy crab, support for farmers, price, public welfare, deputy to the People's Congress, Back to Field, CCTV
4	Employment and Entrepreneurship	enterprise, employment, training, innovative undertaking, reward, support, policy, migrant worker, innovation, talent
5	Market and Consumption	consumption, digitization, epidemic situation, growth, prevention and control, market, economy, data, accelerate, technology
6	Participation of Well-known Enterprises or Stars	Oriental Selection, New Oriental, Minhong Yu, background, fake, Yuhui Dong, garlic, Zhan Xiao, challenge, fans
7	Violation Behavior	foodstuff, food safety, problem, renovate, production, illegal, fake, quality, propagate, joint action
8	Damage to Consumers' Rights and Interests	weather, mildew and rot, New Oriental, grape, customer, peach, fresh food, unsalable, complaint, refund

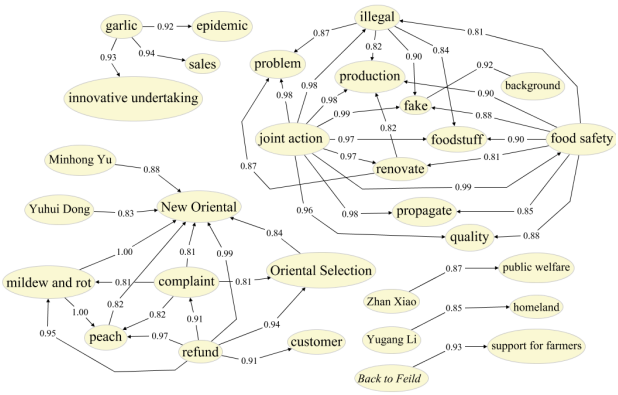


Fig. 1: Network of binomial association rules for topic words.

C. Topic Word Association

This paper extracts association rules between subject words using the Apriori algorithm with a minimum support of 0.01, minimum confidence of 0.8, and minimum lift of 3. We obtained 1009 strong association rules (48 binomial, 961 trinomial), selected representative rules, and visualized the association rules network using CmapTools. Fig. 1 shows a network of binomial association rules for topic words. Each node represents a topic word, each directed edge represents the direction of an association rule, and the number on each edge represents the confidence of the association rule. We can see that there are lots of separate binomial association rules. For example, under the condition that the topic word "Back to Field" (a Chinese variety show) appears, there is 0.93 probability that "support for farmers" will appear. There are even several networks among the topic words. For example, the network centered at "New Oriental", "complaint" and "refund" reflects that consumer complaints and refunds caused by the rotten peaches sold by New Oriental; the network centered at "illegal", "food safety", "joint action", "fake" and "renovate" reflects that a series of joint regulatory actions have been taken on account of the food safety and false publicity.

D. Public Opinion Identification

In this paper, the eight topics obtained from the experiment were summarized into four aspects: Industry Development, Market Effect, Live Broadcast Subject and Negative Effect. The public opinion in each topic were identified according to the topic words and the association rules among them. The results are shown in TABLE . According to the above topic words and their associations, this paper identifies the public opinion under each topic, as shown in Table II. At the same time, this paper divided the eight topics into four aspects: Industry Development, Market Effect, Live Broadcast Subject and Negative Effect. And the specific meaning of public opinion were explained and discussed from these aspects.

TABLE II: Public Opinion on the E-commerce Live Broadcast of Agricultural Products

Aspect	Topic	Public Opinion
Industry Development	Rise of E-commerce Live Broadcast Industry	to increase farmers' incomes, Conducive to poverty alleviation and rural revitalization, to propagate local characteristics and culture, to attract young people to the countryside
	Industry Convergence	to promote the integration and upgrading of traditional industries, to do practical work for the people, to improve rural service level
Market Effect	Market and Consumption	to implement poverty alleviation through consumption, to revitalize the agricultural product market, to take advantage of the potential of the sinking market, to integrate technology and other resources, to achieve economic growth
	Employment and Entrepreneurship	to support farmers as streamers, to increase employment opportunities, to carry out entrepreneurship, to issue incentive policies, to provide training
Live Broadcast Subject	Participation of Public Welfare Forces	to support farmers, to speak for the agricultural products of hometown, to lower the price
	Participation of Well-known Enterprises or stars	to launch the business of agricultural products e-commerce live broadcast, to attract followers, to raise doubts
Negative Effect	Violation Behavior	false propaganda, cannot guarantee the quality, food safety problems, to conduct joint action
	Damage to Consumers' Rights and Interests	mildew and rot, complaints, to request a refund, loss of trust

Industry Development mainly discusses the influence brought by *Rise of E-commerce Live Broadcast Industry* and *Industry Convergence*, such as rural revitalization, cultural export, talent introduction and upgrading of traditional industries. The marketing from rural to urban areas of agricultural products driven by the e-commerce live broadcast has increased farmers' income and promoted the upgrading of local industries, which is of great significance for the implementation of strategies such as poverty alleviation and rural revitalization. The unique live scenes of agricultural products in the field can better present agricultural products and bring more intuitive experience to users [21], which increased the product reliability and users' purchase intention. Besides, this way also plays a role in the promotion of local culture and scenery. However, the current live broadcasts are single in form and repetitive in content, which is difficult to highlight the local characteristics and cultural depth.

Market Effect includes *Market and Consumption* and *Employment and Entrepreneurship*, focusing on the phenomenon of market sinking and consumption upgrading brought by the e-commerce live broadcasts of agricultural products, as well as their role in market activation, economic growth and technological development. The e-commerce live broadcast of agricultural products enabled many links of agricultural products to be integrated into the Internet, and a large number of advantageous resources (such as talents, capital and technology) poured in [22]. This has led to the revitalization of the rural market. At the same time, e-commerce platforms have begun to use the huge potential of sinking markets, expanded the market of agricultural products. Due to the expansion of agricultural products market and low threshold of e-commerce live broadcast, farmer streamers came into being. There are also farmers, migrant workers or the young generation [23] who start their own businesses through e-commerce live broadcast of agricultural products. In addition, local governments and live broadcast platforms also encourage farmer streamers through a series of measures, and provide public welfare training opportunities, which has brought a virtuous circle to the development of this industry [24]. On the other hand, it is not uncommon to fail to bring about an increase in real income to operators, or even cause losses.

Live Broadcast Subject mainly discusses the advantages and doubts caused by the participation of multiple live broadcast subjects such as *Participation of Public Welfare Forces* and *Participation of Well-known Enterprises or stars*. Public welfare forces mainly use low prices and subsidies to promote unsalable agricultural products or advocate for agricultural products in their own hometown, of which officials' live broadcasts is a typical representative. Officials' live broadcasts means that local officials act as the streamers to endorse local products, so as to facilitate the export of local agricultural products or the promotion of cultural and tourism services. In addition, the participation of well-known enterprises and stars has brought huge flow and purchasing power to the e-commerce live broadcast of agricultural products [25]. But it is difficult for these streamers to promote specific agricultural products for a long time. The participation of these subjects has also raised doubts. For instance, the demand for agricultural products has surged in a short period, resulting in temporary production and hasty packaging, which could affect consumers' purchasing experience [26].

Negative Effect includes *Violation Behavior* and *Damage to Consumers' Rights and Interests*, and discusses the problems arising from obstructed agricultural sales processes and inadequate market regulation. On the one hand, Due to the lack of a mature market supervision mechanism, a series of violations (such as false publicity, incomplete product information and food safety problems) occur in the e-commerce live broadcast of agricultural products, which may affect consumers' health in serious cases. Besides, A series of joint regulatory actions have been carried out to monitor these issues. On the other hand, whether because of substandard quality at the source or problems during transportation, agricultural products purchased are easy to become stale. This may lead to a series of consumer rights protection behaviors such as complaints and refunds, and even reduce consumers' trust in e-commerce live broadcast of agricultural products [25].

V. Conclusions

To gain a comprehensive understanding of public opinion regarding e-commerce live broadcasting of agricultural products, this paper has undertaken two key initiatives. Firstly, a public opinion mining approach based on the WLDA-Apriori model is presented. This approach employs a combination of topic modeling and association rule algorithms, enabling the analysis of implicit semantic relationships between subject words and increasing interpretability. Secondly, an empirical study is conducted using social media data from Weibo. The results show that the topics of Chinese public concern are *Industry Convergence*, *Rise of Live Broadcast Industry*, *Participation of Public Welfare Forces*, *Employment and Entrepreneurship*, *Market and Consumption*, *Participation of Well-known Enterprises or stars*, *Violation Behavior* and *Damage to Consumers' Rights and Interests*.

Taking public opinion into account, this paper provides an analysis of the current state and prevalent issues surrounding e-commerce live broadcast in the agricultural industry.Compared with traditional questionnaires and case studies, the results of this paper are more universal, and the suggestions based on these will be more consistent with the needs of the public. However, this article is subject to a limitation whereby the data is derived from a sole platform. Unlike social media platforms, users on e-commerce platforms such as Douyin and Taobao exhibit a greater proclivity towards live-streamed content, live-streamed formats, and product quality, which results in data that encompasses semi-structured and unstructured formats. Thus, it is imperative to incorporate these additional data sources to facilitate a more comprehensive interpretation of user demand. Furthermore, incorporating emotional tendencies in opinion mining can provide a more comprehensive representation of public sentiment. Future research should strive to construct a multidimensional opinion mining model that accounts for topics, emotions, and other relevant factors.

References

[1] W. Shang, "More than 4 million live e-commerce broadcasts in the first quarter," *The People's Daily*, no. 19, 2020.

[2] K. Liu, B. Liu, H. Xu, Y. He, and Y. Cao, "Research on e-commerce live broadcasts helping poverty alleviation under the influence of the covid-19:———take xinhua county, hunan province as an example," in *2020 2nd International Conference on Economic Management and Model Engineering (ICEMME)*. IEEE, 2020, pp. 858–865.

[3] Y. Shi, X. Li, S. ZhouGong, X. Li, and H. Wang, "Precise marketing classification of agricultural products for e-commerce live broadcast platform using clustering," *Mobile Information Systems*, vol. 2022, 2022.

[4] K. Gao and Y. Sheng, "Research on the influencing factors and mechanism of livestreaming commerce adoption intention of new agricultural operation entities," *China Business and Market*, vol. 35, no. 10, p. 9, 2021.

[5] Z. Lu, "Webcasting and the development of online retailers of fresh agricultural products: Driving mechanism and empirical test," *China Soft Science*, 2021.

[6] Y. Zeng, C. Ma, L. Li, and H. Guo, "Live streaming e-commerce and the value recreation of agricultural marketing from rural to urban areas: Mechanism and implementation paths," *Issues in Agricultural Economy*, no. 2, p. 10, 2022.

[7] W. Yang, X. Yang, Y. Bai, H. Yi, M. He, and X. Li, "Research on purchase intention of fresh agricultural products based on sor theory under live broadcast situation," in *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)*. IEEE, 2021, pp. 416–424.

[8] S. Pejić Bach, Krstić and L. Turulja, "Text mining for big data analysis in financial sector: A literature review," *Sustainability*, no. 11, p. 5, 2019.

[9] A. Kadriu and L. Abazi, "A comparison of algorithms for text classification of albanian news articles," *Proceedings of the ENTRENOVA - ENTerprise REsearch InNOVAtion Conference*, no. 3, pp. 1–7, 2017.

[10] B. Nguyen, V.-H. Nguyen, and T. Ho, "Sentiment analysis of customer feedback in online food ordering services," *Business Systems Research Journal*, no. 12, pp. 46–59, 2021.

[11] K. Guo, L. Shi, W. Ye, and X. Li, "A survey of internet public opinion mining," in *2014 IEEE international conference on progress in informatics and computing*. IEEE, 2014, pp. 173–179.

[12] X. Wan, R. Wang, M. Wang, J. Deng, Z. Zhou, X. Yi, J. Pan, and Y. Du, "Online public opinion mining for large cross-regional projects: Case study of the south-to-north water diversion project in china," *Journal of Management in Engineering*, vol. 38, no. 1, p. 05021011, 2022.

[13] Z. Fu, M. Yan, C. Meng, W. Wang, X. Hu, Y. Li, J. Wang, Z. He, and Z. Wang, "Opinion mining about online education basing on gdelt and twitter data," in *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. IEEE, 2020, pp. 741–745.

[14] L. Wang, "Research on the mining of ideological and political knowledge elements in college courses based on the combination of lda model and apriori algorithm," *Applied Mathematics and Nonlinear Sciences*.

[15] G. Ruan and L. Xia, "Mining document topics based on association rules," *Data Analysis and Knowledge Discovery*, no. 12, pp. 50–56, 2016.

[16] P. Wu, T. Shi, and C. Ling, "Topic discovery of covid-19 vaccines based on weibo platform," *Information Science*, vol. 40, no. 7, p. 8, 2022.

[17] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[18] A. Wilson and P. A. Chew, "Term weighting schemes for latent dirichlet allocation," in *human language technologies: The 2010 annual conference of the North American Chapter of the Association for Computational Linguistics*, 2010, pp. 465–473.

[19] N. Lee, E. Kim, and O. Kwon, "A real-time combination of text-mining methods for flexible movie recommendation in human robot interaction," in *2016 International Conference on Platform Technology and Service (PlatCon)*. IEEE, 2016, pp. 1–6.

[20] L. Posch, M. Panahiazar, M. Dumontier, and O. Gevaert, "Predicting structured metadata from unstructured metadata," *Database*, vol. 2016, 2016.

[21] M. Zan and Z. Wang, "E-commerce live streaming of agricultural products: a new mode of poverty alleviation," *Issues in Agricultural Economy*, no. 11, p. 10, 2020.

[22] L. Song and X. Ni, "E-commerce consumption for poverty alleviation helps agricultural products out of difficulties," *People's Tribune*, no. 29, p. 2, 2020.

[23] Z. Wang, "How "live broadcast + e-commerce" helps rural revitalization," *People's Tribune*, no. 15, p. 2, 2020.

[24] J. Dai and W. Luo, "Research on the key path to improve user stickiness under the scenario of agricultural products live streaming," *China Business and Market*, vol. 36, no. 5, p. 12, 2022.

[25] Y. Yuan and W. Zhang, "The dilemma and its solution of rural livestreaming marketing from the perspective of social capital," *China Business and Market*, vol. 35, no. 10, p. 8, 2021.

[26] H. Guo and J. Qv, "Study on the sustainable development of e-commerce live broadcast assisting agriculture," *People's Tribune*, no. 20, p. 3, 2020.

CroStats — visualization of population in Croatia

V. Perković, F. Belac, K. Boroš, M. Kos, L. Humski, D. Pintar
University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia

***Abstract*—Population census is often a popular topic in everyday conversations. After extensive research, a conclusion was made that currently no services exist that offer a convenient, easily interpretable information about the census. Our solution for this problem is CroStats – a simple web interface for population census data visualization intended for general public. The central feature of the CroStats is its intuitive and simple graphical display. It analyzes the most important categories of the population census by county such as population, age, birth rate and mortality. All this information is displayed on an interactive map of the Republic of Croatia. Extra features offered are graphs of changes by year and interesting historical and demographic facts about counties that the broader audience may have interest in.**

Keywords*—*Visualization, population census, graphs, web interface, geography, demography.

I. INTRODUCTION

Population census data is a crucial aspect of understanding the demographic trends and patterns of a country. When it comes to understanding large amounts of data, people tend to rely on simple and clear visualizations. While various static visualizations may be informative and intuitive, they do not offer a way to easily adapt or expand on the depicted information.

In light of this issue, CroStats, a simple web interface for population census data visualization was developed.

The main focus of the application was to provide a user-friendly and intuitive graphical display of the most important categories of population census data, including population, age, birth rate, and mortality, presented on an interactive map of the Republic of Croatia. The web interface includes not only this graphical display but also graphs of changes by year and informative historical and demographic facts about counties, making the data more accessible and comprehensible to the general public.

Through the following sections, the paper will go over a brief history on population census, related work to CroStats, overview of the used technologies, a thorough overview of the application including examples of use and the application architecture, future work and a conclusion.

II. POPULATION CENSUS

The population census is designed to provide the government with insight into the number of residents and general demographic characteristics. It dates back far into history, with the first census being conducted approximately 2000 years ago by the Han dynasty in China. At that time, China had 57 million inhabitants, and today the number has grown significantly [1]. Much has changed since then, but population census is still used in almost all countries in the world. With the information collected through the population census, it is possible to track demographics and notice new demographic trends. In Croatia, the first population census was conducted by Emperor Joseph II in 1857 [2], and during the existence of the Republic of Croatia, three censuses have been conducted in 2001, 2011, and 2021, which is also the basis of CroStats. The data was collected from the State Bureau of Statistics [3] where it is available to everyone. Unfortunately, the data format is poorly designed, organized, and displayed. The data is altogether hard to read and makes it difficult for everyday people to navigate a large amount of numbers. Sorting the data logically was a challenge due to these shortcomings, but in the end it was successfully resolved.

III. RELATED WORK

Eurostat Regional Yearbook [4] is an example of a web interface that is available to everyone, displaying a graphical representation of Europe with data such as the median age by county and predictions for the year 2050. CroStats differs in the fact that it currently focuses on Croatia. It will also have various categories of information to display. The only widely used such program that we found is Eurostat.

IV. TECHNOLOGY

While developing CroStats, a variety of technologies were used to create an interactive and informative experience for users.

First and foremost, R Shiny [5] was used as the framework for the application. Shiny is a web application framework for the R programming language. It allows users to build interactive web applications and dashboards that can be easily shared online.

The package is based on a reactive programming model, which means that the user interface (UI) components automatically update when the underlying data changes.

Shiny applications consist of two parts: the user interface (UI) and the server logic. The UI defines the layout and components of the web application, such as buttons, sliders, and text boxes. The server logic defines the behaviour of the application, such as data processing,

analysis, and visualization. The Shiny package provides a variety of UI components and server functions that can be used to build complex applications. One of the main advantages of Shiny is its ability to handle large datasets and complex calculations in real time. The reactive programming model allows the application to respond quickly to user inputs, making it ideal for data exploration and visualization. Shiny also supports a wide range of data visualization libraries, such as ggplot2, plotly, and tmap.

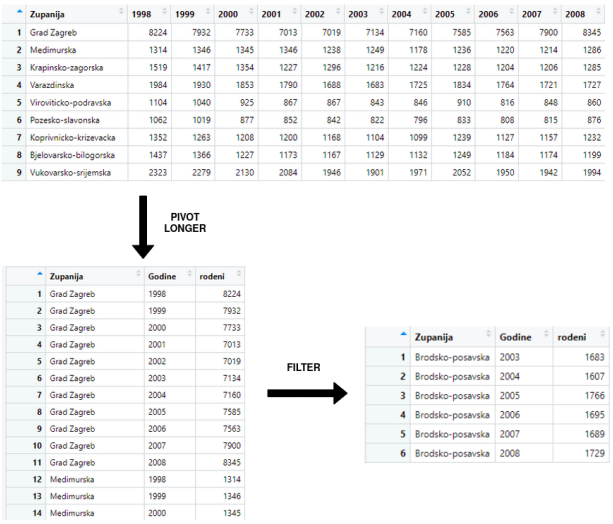
To display the data, several popular data visualization libraries in R were used. These included tmap [6], plotly [7], and ggplot2 [8].

Tmap is a package for R that provides a set of tools for creating thematic maps. Thematic maps are a type of data visualization that represents spatial data as colours, patterns, or symbols on a map. Tmap provides a simple syntax for creating thematic maps that can be customized with a variety of colour palettes, legends, and other design elements. One of the main advantages of tmap is its support for a wide range of map projections, including the popular Mercator and Robinson projections. Tmap also supports interactive map features such as zooming, panning, and tooltip displays, making it ideal for creating interactive data visualizations.

Plotly is a package for R that provides a set of tools for creating interactive data visualizations. Plotly allows users to create a wide range of charts and graphs, including scatter plots, bar charts, and line charts. The package also supports 3D plotting and animations. One of the main advantages of plotly is its ability to create interactive plots that can be easily shared online. Users can zoom, pan, and hover over data points to view more detailed information. Plotly also provides a wide range of customization options, including colour palettes, legends, and annotations.

Ggplot2 is a package for R that provides a set of tools for creating data visualizations using the grammar of graphics. The grammar of graphics is a theoretical framework for creating visualizations that consist of a set of rules for mapping data to visual elements, such as points, lines, and bars. One of the main advantages of ggplot2 is its ability to create complex visualizations with a few lines of code. The package provides a wide range of customization options, including colour palettes, legends, and themes. Ggplot2 also supports a wide range of data types, including categorical, continuous, and time-series data.

To prepare data for visualization, both the dplyr [9] and tidyr [10] libraries were used. The dplyr package provides a set of functions for manipulating and summarizing data, while the tidyr package is designed to reshape data into a tidy format for easier analysis and visualization. One of the functions from tidyr used in this analysis was "pivot_longer" (see figure 1), which is used to transform data from wide to long format, making it easier to work with and visualize. One of the main advantages of dplyr is its efficiency when handling large datasets. The package uses a set of optimized functions that can perform



	Zupanija	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1	Grad Zagreb	8224	7932	7733	7013	7019	7134	7160	7585	7563	7900	8345
2	Medimurska	1314	1346	1345	1346	1238	1249	1178	1236	1220	1214	1286
3	Krapinsko-zagorska	1519	1417	1354	1227	1296	1216	1224	1228	1204	1206	1285
4	Varazdinska	1984	1930	1853	1790	1688	1683	1725	1834	1764	1721	1727
5	Viroviticko-podravska	1104	1040	925	867	867	843	846	910	816	848	860
6	Pozesko-slavonska	1062	1019	877	852	842	822	796	833	808	815	876
7	Koprivnicko-krizevacka	1352	1263	1208	1200	1168	1104	1099	1239	1127	1157	1232
8	Bjelovarsko-bilogorska	1437	1366	1227	1173	1167	1129	1132	1249	1184	1174	1199
9	Vukovarsko-srijemska	2323	2279	2130	2084	1946	1901	1971	2052	1950	1942	1994

	Zupanija	Godine	rodeni
1	Grad Zagreb	1998	8224
2	Grad Zagreb	1999	7932
3	Grad Zagreb	2000	7733
4	Grad Zagreb	2001	7013
5	Grad Zagreb	2002	7019
6	Grad Zagreb	2003	7134
7	Grad Zagreb	2004	7160
8	Grad Zagreb	2005	7585
9	Grad Zagreb	2006	7563
10	Grad Zagreb	2007	7900
11	Grad Zagreb	2008	8345
12	Medimurska	1998	1314
13	Medimurska	1999	1346
14	Medimurska	2000	1345

	Zupanija	Godine	rodeni
1	Brodsko-posavska	2003	1683
2	Brodsko-posavska	2004	1607
3	Brodsko-posavska	2005	1766
4	Brodsko-posavska	2006	1695
5	Brodsko-posavska	2007	1689
6	Brodsko-posavska	2008	1729

Fig. 1: Visualization of data wrangling

data manipulation operations much faster than the base R functions. Using the "filter" function from the dplyr package is more optimized for handling large datasets compared to the base R function "subset". This is because the "filter" function uses a lazy evaluation approach that only evaluates the necessary conditions to filter the data. In contrast, the "subset" function evaluates all conditions, which can be time-consuming with large datasets.

This makes it ideal for data analysis and visualization tasks that involve large datasets. Dplyr provides a set of functions that can be used for data manipulation and summarization tasks, such as filtering, selecting, grouping, and aggregating data.

In order to load and handle the data related to drawing the map, the sf library [11] was used. The sf package in R is a popular library for working with spatial data. It provides a set of functions and classes for manipulating, visualizing, and analyzing geospatial data in R. One of the main advantages of sf is its ability to handle a variety of spatial data formats, including shapefiles, GeoJSON files, and other common GIS data formats such as KML, TIGER and GML. It provides a set of functions for reading and writing spatial data from these formats, as well as for performing common spatial operations such as spatial joins and intersections.

Besides the technologies above, ShinyThemes [12] is a package that provides a variety of premade themes that can be easily applied to a Shiny app. It was used to add visual flair to the app. Lastly, Git [13] and Gitlab [14] were used for version control throughout the development process. Version control is an essential tool for any software development project, and it allowed for keeping track of changes made to the codebase, collaboration with other team members, and easily roll back changes if needed.

There are other data dashboarding frameworks available besides Shiny, such as Dash [15], Bokeh [16], and Streamlit [17]. However, Shiny was superior to the other

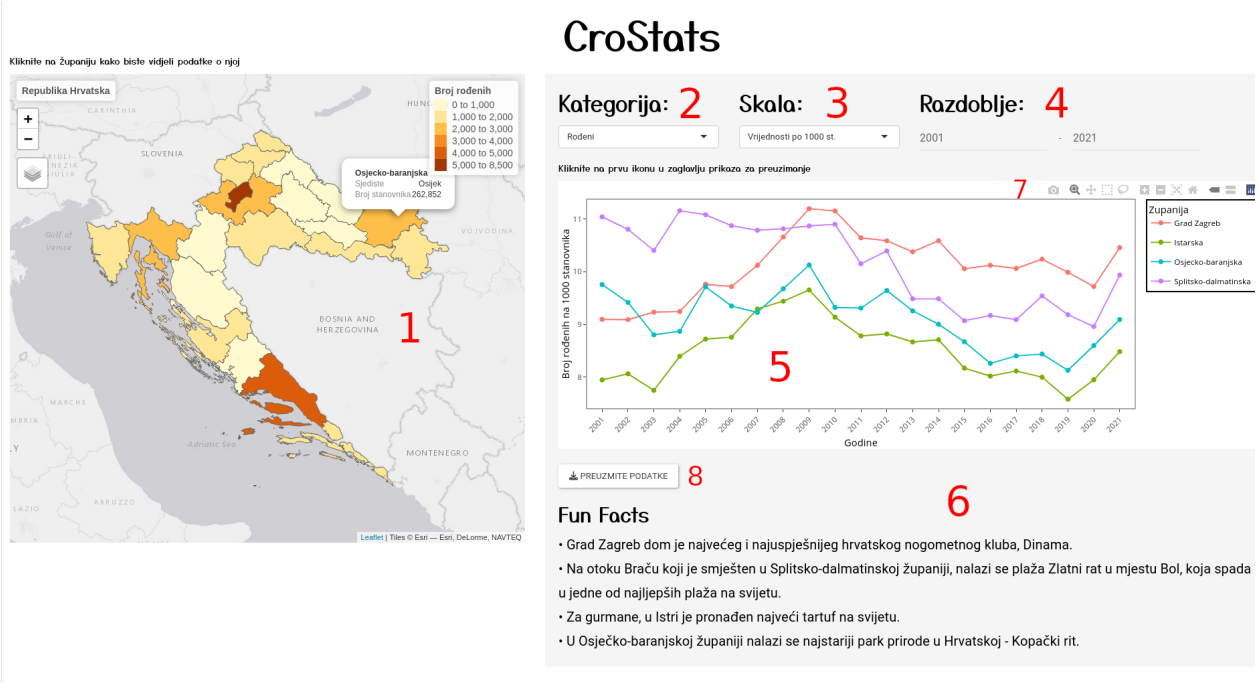


Fig. 2: CroStats app interface

options because it is the most seamlessly integrated with R ecosystem. Streamlit, Dash, and Bokeh are Python-based dashboarding solutions, running on the Tornado [18] and Flask [19] web frameworks.

Integration of R Shiny, tmap, plotly, and ggplot2 libraries with the dplyr package enabled the development of a functional and visually appealing application for data visualization and analysis. These technologies provided a range of features and customization options that allowed for the manipulation and summarization of data, filtration and selection of relevant data points, and creation of informative and visually appealing maps and graphics. The result was an application that provided an intuitive user experience and showcased the data in a visually appealing and informative way.

V. APPLICATION OVERVIEW

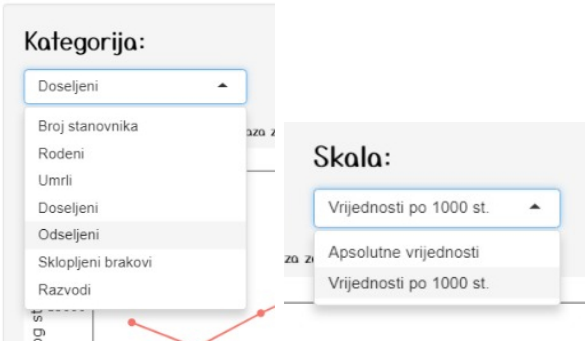
The CroStats application was realized as a web interface.

The application's interface features a map of the Republic of Croatia on the left side (see figure 2, mark 1), while on the right side, there are two drop-down menus for category and scale selection and a range selection for a specific period (see figure 2, marks 2, 3, 4 and figure 3). Scatter plots of the selected county can be viewed below the drop-down menus and range selection (see figure 2, mark 5), along with historical and demographic facts about the selected county (see figure 2, mark 6).

The application allows for the selection of any county on the map by clicking with a mouse. The selected category, scale and period for the selected county will then be displayed on a scatter plot. Up to four counties can be selected for comparison, and four scatter plots with accompanying facts will be displayed. The map of the Republic of Croatia will be re-rendered and coloured based on the selected category each time a different category is chosen (Number of inhabitants, Inhabited population, Migrated population). If a fifth county is selected, the earliest county selected will be replaced, so that there is always a display of up to four county trends on the graph with accompanying facts.

The application offers the functionality to save the graph of user's interest (see figure 2, mark 7 and figure 4) and data (see figure 2, mark 8) to their device storage so that the users can download them locally as a CSV file containing the current category.

Downloading graphs and data gives the users the flexibility to customize the information in their reports, presentations, and projects.



(a) Category selection (b) Scale selection

Fig. 3: Selection

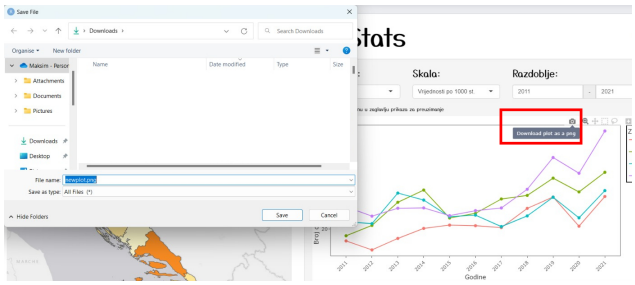


Fig. 4: Graph download

The application provides a platform for examining population data and gaining new insights for anyone with a curious interest in Croatia's population patterns.

A. Architecture

The architecture of the system is designed to support the creation of an interactive web application using the Shiny framework (see figure 5). The Shiny application consists of two main components: the Shiny server and the Shiny UI. The Shiny server is responsible for processing user requests, executing R code, and returning the results to the user's web browser. It handles all the logic and calculations required for data analysis and visualization. The Shiny UI is responsible for defining the layout and visual elements of the web application. It allows users to interact with the data and provides a graphical interface for displaying the results of the analysis. The UI is defined using R code, which creates HTML, CSS, and JavaScript code that is sent to the user's web browser. Data for processing and visualization are sourced from an Excel spreadsheet.

The spreadsheet contains multiple sheets, each of which represents a different data source. The application uses R packages such as readxl [20] to import the data into the application. The Shiny framework leverages a variety of packages and tools within the R ecosystem, including ggplot2 and plotly, to create interactive and responsive web applications. Overall, a scalable and adaptable solution for data processing and visualization is intended to be provided by the system architecture. By leveraging the power of the Shiny framework and the R programming language, the application can easily adapt to changing data and analytical requirements. A simple and effective method of organizing data for analysis is the ability to source data from various sheets within a single Excel spreadsheet.

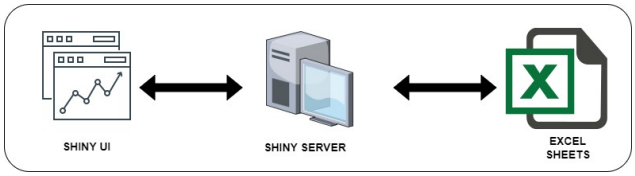


Fig. 5: Application architecture

B. Examples of use

Due to the lightweight interface that the application CroStats provides, the users should feel fairly comfortable when using it for the first time.

There are several ways to use the application. The main ones are finding the exact values of the categories for a certain year, comparing the absolute values of several counties and comparing the counties based on the values per thousand inhabitants. The user can compare up to four counties so that the view doesn't get crowded.

Firstly, an example where the user just wants to find out the population of a certain county in a certain year. Since that is the default layout, all the user has to do is hover over the point that coresponds to the wanted year (see figure 6).

Secondly, if the user wanted to check out more than one county, the user could just choose more counties by clicking on the map as described in the applications hint placed nearby. Deselection is made by clicking on the selected county. All of the above works the same for all the other categories prompted in the drop down menu under "Kategorija:". Figures 6 and 7 showcase the implementation of these examples.

In order to compare counties it should be noted that it's not a good idea to base the comparison on the absolute values. The number of inhabitants plays a big role if the counties vastly differ from one another. Because of that, scale is incorporated into the application. Aside from absolute values, the application provides values per thousand inhabitants. This way a user can confidently compare counties without having to worry about the size of each of the counties (see figure 2).

And while the user plays with the settings of the interface, historical and demographic facts are generated at the bottom to give a few insights about the selected counties.

VI. FUTURE WORK

For future work, even though Shiny made the development of the interface possible and intuitive, the lack of versatility in hosting options and overall design options resulted in considering scaling the application using a different technology.

One technology that is particularly interesting is D3.js [21]. D3.js, or Data-Driven Documents, is a JavaScript library that allows for the creation of interactive data visualizations on the web. It is a powerful tool that can be used to create a wide range of visualizations, from simple charts and graphs to more complex and interactive maps and animations.

There are several reasons why using D3.js for future work is considered. First, D3.js is a widely used and well-established technology that has a large and active community of developers. This means that there are many resources and tutorials available for learning and working with the library. Additionally, D3.js has a wide range of

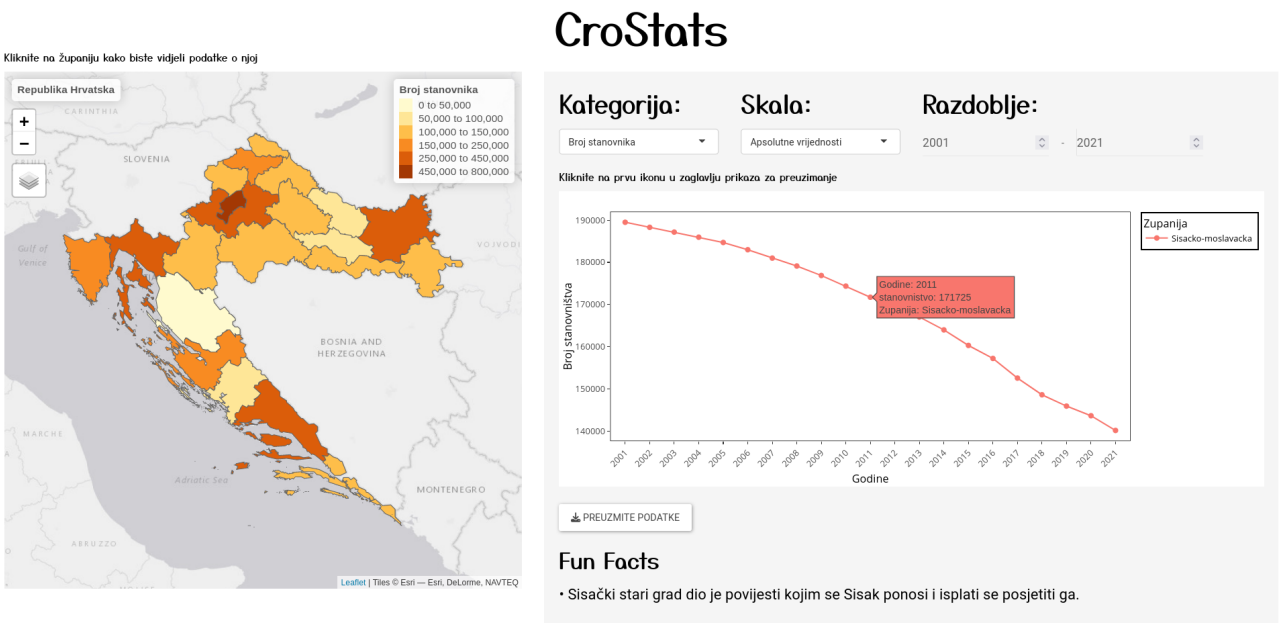


Fig. 6: CroStats example 1

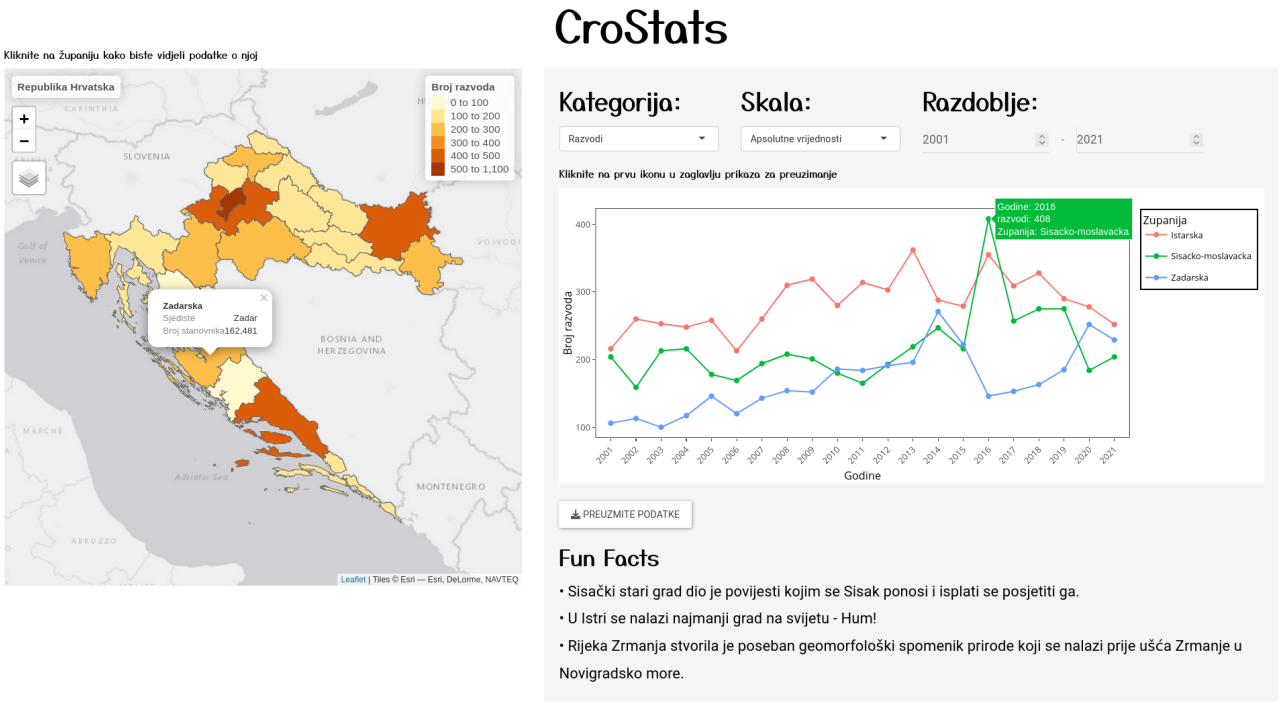


Fig. 7: CroStats example 2

built-in functions and tools that can be used to create advanced visualizations and animations that are not possible with R Shiny. There is also a significant advantage when interactivity is taken into account. Animations can be added to visualizations, which can help the data become more engaging and easier to understand. R Shiny has a more streamlined way of displaying visualizations but the big drawback is loss of flexiblity in shaping interactivity. Another advantage of D3.js is that it is built on JavaScript, which is a widely-used programming language that is supported by most web browsers.

The structure of the data could also be implemented in a better way. Instead of Excel spreadsheet, the data could be imported into a database that communicates with the app. It would then be extracted and used for the visualization. That could simplify the process and make the app faster which would definitely be an improvement.

As far as features go, the number of different graphs that are displayed on the app would be expanded. Further data analysis combined with new insights for what is possible with other frameworks revealed that data can be visualized in many ways. Some of them are not traditional but they can easily provide the same, if not more intuition on what the data is trying to tell.

There would be rules set in place for which column means what, and if the users were motivated enough to clean their data, the users could use the tool for their own (real or imaginary) demographic. Of course, a user might want to store the data for later analysis, and adding a data persistence layer should be pretty simple because of available infrastructure as a service (IaaS) solutions like Microsoft Azure [22]. The app would then connect to a web service that would handle user authentication and authorization, and then serve the saved data from a database. Both the web service and the database would be handled by the cloud provider, which would enable the developers to quickly build the business logic.

VII. Conclusion

CroStats is a great tool for visualizing data from population census. It is based on a user-friendly web interface. The main idea was to build a robust tool that helps users gain insights into population trends and patterns. The use of interactive graphs and maps in the web interface improves the user experience as it provides a dynamic way to explore data. Using that kind of data presentation enables users to better understand the demographic data of Croatia. Furthermore, the simplicity of the web interface makes it accessible to a broader audience, regardless of the technical background.

Acknowledgment

This research has been supported by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS).

References

[1] "history," Mar 2023. [Online]. Available: https://www.abs.gov.au/websitedbs/d3310114.nsf/51c9a3d36edfd0dfca256acb00118404/eadaffffb171cab6ca257161000a78d7!OpenDocument
[2] "Cro history," Mar 2023. [Online]. Available: https://www.croatian-genealogy.com/educational-articles/19th-century-census-source-croatian-genealogical-data/
[3] "The state bureau of statistics," Feb. 2023. [Online]. Available: https://dzs.gov.hr/
[4] "Eurostat regional yearbook," Feb 2023. [Online]. Available: https://ec.europa.eu/statistical-atlas/viewer/
[5] "R shiny," Feb 2023. [Online]. Available: https://shiny.rstudio.com/
[6] "tmap," Feb. 2023. [Online]. Available: https://cran.r-project.org/web/packages/tmap/vignettes/tmap-getstarted.html
[7] "plotly," Feb. 2023. [Online]. Available: https://plotly.com/
[8] "ggplot2," Feb. 2023. [Online]. Available: https://ggplot2.tidyverse.org/
[9] "dplyr," Feb. 2023. [Online]. Available: https://dplyr.tidyverse.org/
[10] "tidyr," Mar 2023. [Online]. Available: https://cran.r-project.org/web/packages/tidyr/index.html
[11] "sf," Feb. 2023. [Online]. Available: https://cran.r-project.org/web/packages/sf/index.html
[12] "Shinythemes," Feb. 2023. [Online]. Available: https://rstudio.github.io/shinythemes/
[13] "Git," Feb. 2023. [Online]. Available: https://git-scm.com/
[14] "Gitlab," Feb. 2023. [Online]. Available: https://gitlab.com/
[15] "dash," Mar 2023. [Online]. Available: https://plotly.com/dash/
[16] "bokeh," Mar 2023. [Online]. Available: https://bokeh.org/
[17] "streamlit," Mar 2023. [Online]. Available: https://streamlit.io/
[18] "tornado," Mar 2023. [Online]. Available: https://www.tornadoweb.org/en/stable/
[19] "flask," Mar 2023. [Online]. Available: https://flask.palletsprojects.com/en/2.2.x/
[20] "readxl," Mar. 2023. [Online]. Available: https://readxl.tidyverse.org/
[21] "Data-driven documents," Feb. 2023. [Online]. Available: https://d3js.org/
[22] "azure," Mar 2023. [Online]. Available: https://azure.microsoft.com/en-us

A New Approach to Semantic Parsing of Metonymic Phrases by a Business Intelligence System

Vladimir A. Fomichov
Moscow Aviation Institute (National Research University), Department of Intelligent Monitoring Systems, Moscow, Russia
vfomichov@gmail.com

***Abstract* - The functioning of numerous business intelligence systems (BIS) substantially depends of the content of databases. In many thematic domains, this content is highly dynamic, e.g., in case of the prices for oil, aluminum, wheat, etc. It is the reason for the birth of a BIS project aimed at updating the content of relational databases by means of extracting actual values of data from the texts in natural language (NL). The principal subject of the paper is theoretical foundations of building semantic representations (SRs) of the metonymic phrases. In linguistics, a metonymic construction is a linguistic construction in which an entity is referred to by the name of something closely associated with that thing or situation. The main scientific results set forth in the paper are as follows: (a) a logical structure of a minimalist terminological knowledge base helping to fulfil semantic parsing of metonymical phrases, it is motivated by the theory of K-representations; (b) an original algorithm of finding a semantic interpretation of metonymic phrases of some kinds.**

Keywords - metonymy; logical structure of terminological knowledge base; natural language processing; semantic parsing of metonymical phrases; theory of K-representations; SK-languages

I. INTRODUCTION

The history of computer intelligent systems called natural language processing systems (NLPSs) is over five decades long. During several initial decades, the main attention of the researchers was attracted to developing computational methods of processing regular phrases and short discourses. It means that the meaning of such texts is determined by their syntactic structure and the correspondence between their lexical units and their meanings, taking into account that one lexical unit may be associated with several meanings, as, e.g., in case of the words “mouse”, “cloud”, “battery”, “green”.

However, the specialists in numerous fields very often encounter the so called irregular texts, first of all, the texts with metaphors, metonymic constructions, and with idiomatic constructions. The focus of the present paper is the processing of metonymical phrases. One says in theoretical and computational linguistics about a metonymical sentence in case a sentence *Expr* mentions an object *X* but uses with this aim a designation of an object *Y* associated in any way (through a knowledge base) with the object *X*. For instance, the phrase S1 = “The winners in this hockey match were the maple leaves” means that the winner in this match was the combined team of Canada.

The problem of developing computational algorithms for analyzing metonymical texts has attracted the attention of many researchers since the 1990s, see, in particular, [1-4]. For the authors of [1], semantic parsing of metonymic texts is a particular kind of abduction, i.e., an inference to the best explanation. The interpretation of a text is considered as the minimal explanation of why the text would be true. In [1], first order logic is used as a formal tool for theoretical considerations. No algorithm of metonymic texts’ semantic parsing is suggested.

Markert and Hahn set forth in [2] deep theoretical foundations of constructing semantic representations of sentences and discourses. With this aim, they use a version of KL-ONE terminological knowledge representation language. The paper [2] provides a methodological basis for developing computer systems being able to analyze metonymic texts. However, for designing a system of the kind, a group of programmers should have a rather high background in descriptive logics’ formalisms. The paper [2] contains no algorithm of metonymic texts’ parsing being clear for the programmers of average qualification.

The paper [3] suggests to use associative relations between the words for discovering and interpreting metonymic phrases. E.g., the stated method uses the information about the triples of the form (A, B, C), where A is a stimulus word (a verb, e.g., “to convey”), B is a deep case (in other terms, a semantic role, a conceptual case), and C is the information about associated nouns and pronouns connected with certain distances.

The subject of the paper [4] is detecting implicit metonymic relations regarding the organizations and locations presented in WordNet.

The common feature of the listed papers is that the stated approaches don’t suggest a domain-independent logical structure of a terminological knowledge base and of objects database helping to find a semantic interpretation of metonymic phrase.

The analysis shows that today there is a need of formal models and algorithms being quite clear for the programmers and being able to become a starting point

for the creation of effective computational systems "understanding" the texts with metonymy.

The main scientific results set forth in the paper are as follows: (a) a logical structure of a minimalist terminological knowledge base helping to fulfil semantic parsing of metonymical phrases, it is motivated by the theory of K-representations; (knowledge representations); (b) an original algorithm of finding a semantic interpretation of metonymic phrases of some kinds.

II. THE CENTRAL IDEA OF THE APPROACH

Example. Let T1 = "On April 12, a match between the football teams "Zenith" and "Spartak Moscow" took place in Saint-Petersburg". It was the victory of the red-whites with the score 2:1". Let's consider a scheme of understanding by a fan what team has achieved the victory. The noun-like construction "the red-whites" denotes an unknown physical object associated with the semantic unit "a value of colour". The text T1 mentions the entities associated with the notions "a match", "a football team", "a city", "a victory". Let Set1 be the set of all characteristics of the listed concepts. Then Set1 includes the characteristic Symbolics associated with the concept "a football team", and its semantic restriction is the concept "emblem". Suppose that this concept has the characteristic Colour with the semantic restriction "a value of colour", and the characteristic Drawn-object with the semantic restriction "a physical object".

Then the information about the emblems of the teams participated in the match is looked for in the objects database. The result of this search is that the emblem of the team "Zenith" combines the colours white and blue, and the emblem of the team "Spartak Moscow" combines the colours red and white. Then the conclusion is drawn that the team "Spartak" (Moscow) is the winner in the match.

Below an algorithm of processing metonymical phrases is introduced. The essence of this algorithm is processing of two relations called terminological knowledge base (TKB) and objects database (ODB). The originality of TKB is, in particular, the lack of fixed semantic interpretation of the most part of this relation's attributes. It is because the considered structure of TKB aims at reflecting the structure of a K-string representing a piece of knowledge (a semantic frame) associated with a concept.

III. SHORTLY ABOUT THE THEORY OF K-REPRESENTATIONS

The methodological basis of the present paper is the theory of K-representations (knowledge representations). It is an original theory of designing semantic-syntactic parsers of natural language (NL) with a broad usage of formal means for representing input, intermediary, and output data. The theory of K-representations (TKR) is presented in numerous papers both in English and Russian written by V. A. Fomichov and his colleagues [5-15].

The *first constituent* of TKR is the theory of SK-languages (standard knowledge languages), stated, in particular, in [10, 14]. The kernel of the theory of SK-languages is a mathematical model describing a system of such 10 partial operations on structured meanings (SMs) of natural language texts (NL-texts) that, using primitive conceptual units as "blocks", we are able to build formal representations of SMs of arbitrary NL-texts (including articles, textbooks, etc.) and of arbitrary pieces of knowledge about the world [10].

The initial version of the theory of SK-languages is called the theory of K-calculuses and K-languages [5-8].

The *second constituent* of TKR is a broadly applicable mathematical model of a linguistic database, it is described in Chapter 7 of [10]. The model describes the frames expressing the necessary conditions of the existence of semantic relations, in particular, in the word combinations of the following kinds: "Verbal form (verb, participle, gerund) + Preposition + Noun", "Verbal form + Noun", "Noun1 + Preposition + Noun2", "Noun1+ Noun2", "Number designation + Noun", "Attribute + Noun", "Interrogative word + Verb".

The *third basic constituent* of TKR is formed by a family of complex, strongly structured algorithms carrying out semantic-syntactic analysis of texts from some practically interesting sublanguages of NL. These algorithms transform NL-texts into their semantic representations being K-representations, i.e., the expressions of some SK-languages. The algorithm *SemSynt1* is presented in the second part of the monograph [10]. The input texts can be from the English, German, and Russian languages. That is why the algorithm *SemSynt1* is multilingual.

An important feature of the algorithm *SemSynt1* is that it doesn't construct any syntactic representation of the inputted NL-text but directly finds semantic relations between text units. The other distinguished feature is that a complex algorithm is completely described with the help of formal means, that is why it is problem independent and doesn't depend on a programming system.

The *fourth constituent* of TKR is a collection of scientific results expanding theoretical foundations of advanced ontologies, cross-lingual conceptual information access, multilingual semantic web, the design of agent communication languages in multi-agent systems] and recording the content of e-negotiations [6-15].

The main output of the algorithm *SemSynt1* is a semantic representation (SR) of the input text being its K-representation. The expressions of SK-languages are constructed from the basic conceptual units and several service symbols by means of certain partial operations Op[1] - Op[10]. Let's see (without mathematical details) how these partial operations do work.

The operation Op[1] allows us to join intensional quantifiers and designations (simple or compound) of notions, in particular, for constructing the formulas *certain airplane1, certain airplane1 * (Manufacturer, Boeing), all airplane1 * (Manufacturer, Boeing).*

The operation Op[2] is used for constructing the expressions of the form $f(t_1, ..., t_n)$, and Op[3] enables us to build the expressions of the form $(c \equiv d)$. Example: *(Age(S.Nosov) ≡ 27/year).*

One uses the operation Op[4] for building the expressions of the form $rel(t_{1,\ldots,}\ t_n)$, where *rel* is the name of a relation with *n* attributes (example: *Earlier (Creation-date(certain file1), #yesterday)*). The operation Op[5] provides the possibility to mark a formula or its part by means of a variable. Example: *all car1 * (Manufacturer, BMW) : S1*.

The operation Op[6] allows us to join the negation connective ¬ to a formula (example: ¬*airplane1*). The operation Op[7] governs the use of the logical connectives ∧ (and) and ∨ (or). Example: *car1 * (Manufacturer, (BMW ∨ Opel))*.

Using the operation Op[8] at the last step of an inference, it is possible to construct compound designations of notions. Example: *file1 * (Extension1, ("doc" ∨ "docx"))(Location, certain desktop)*.

The partial operation Op[9] enables us to use universal quantifier and existential quantifier for building the formulas. The operation Op[10] helps to construct the designations of the *n*-tuples of the form $<c_1, \ldots, c_n>$, where $n \geq 1$.

IV. DEVELOPING LOGICAL STRUCTURE OF TERMINOLOGICAL KNOWLEDGE BASE AND OBJECTS DATABASE

In the middle of the 1970s, professor M. Minsky from MIT suggested the notion of frame as a structure for representing stereotype knowledge [15]. This notion acquired a broad popularity in the 1970s – 1980s. In particular, a new generation of knowledge representation languages was born, they were called the frame-like knowledge representation languages. A frame of a concept distinguishes the main components of the objects qualified by this concept and indicates the semantic restrictions for each component.

Using broad expressive possibilities of SK-languages introduced in the theory of K-representations, let's consider a formal model of a terminological knowledge base in the form of a finite set *Set-Concepts-Frames*, consisting of the expressions of the form

$<concept1, concept2 * (rel_1, value_1)\ldots(rel_n, value_n)>$, (1)

where *concept1* is the focus concept (i.e., the concept to be explained), *concept2* – the concept-stereotype, that is, the known concept used for explaining the meaning of *concept1*; $n \geq 1$, for each $k = 1,\ldots, n$, rel_k is the name of a binary relation or the name of a function with one argument; $value_k$ is ether the second attribute of the relation or the value of the function depending on the kind of rel_k.

Let's demand that the expression of the form (1) be a K-string, that is, be an expression of the SK-language *Ls(B)* in certain conceptual basis *B*. Then it is seen from the structure of the form (1) that the partial operation Op[10] was used during last step of its inference, and the operation Op[8] was applied during the preceding step of inference.

Example. The set *Set-Concepts-Frames* may contain the expression

*<football-team, sport-team * (Called, certain string)(Sport-kind, football)(Symbolics, certain emblem)(Country-relation, certain country1)(City-relation, certain city1)(Senior-coach-relation, certain person)(Captain-relation, certain person)>*.

It is suggested to implement the finite set *Set-Concepts-Frames* as a relation *ConceptsFrames* of a relational database. The structure of this relation is original, in particular, due to the lack of fixed semantic interpretation of the most part of this relation's attributes. It is because the considered structure of TKB aims at reflecting the structure of a K-string representing a piece of knowledge (a semantic frame) associated with a concept.

Only the first – fourth attributes of the relation *ConceptsFrames* have a fixed semantic interpretation. The attribute No. 1 with the index Row contains the ordered numbers of the relation's rows. The attribute No. 2 with the index Focus contains the concepts characterized by the values of other attributes of the same row. For instance, in case of the considered text T1, the column 2 of the relation *ConceptsFrames* may contain the elements *match1, football-team, city1, emblem1*. For arbitrary row p, the element on the intersection of this row and the column 3 is the basic concept for the focus concept on the intersection of the same line and column No. 2. E.g., it may be that ConceptsFrames[p, 2] = football-team, and ConceptsFrames[p, 3] = sport-team.

The column No. 4 has the index Numb. For arbitrary row p, the element Concepts-frames[p, 4] is a positive integer interpreted as the number of the so called *search attributes.* Let numb-search-attr1 = ConceptsFrames[p, 4]. The elements of the columns with ordered numbers 4+1, 4+3, …, 4+2numb-search-attr1 – 1 will be called *search attributes.* The attributes and their values will be (with some exclusions) used for semantic processing of metonymic phrases.

The elements of the columns with ordered numbers 4+2, 4+4, …, 4+2*numb-search-attr1 will be called *semantic restrictions of the search attributes.*

Due to technical reasons, the relations ConceptsFrames and ObjectsBase (the latter is described below and contains the data about various entities from the considered application domain) are represented below not by the tables (as usual) but as the sequences of linear records of their rows. For instance, the table 1

A	B	C
U_1	U_2	U_3
W_1	W_2	W_3

Table 1. A table for constructing a linear record of the matrix. Here A. B. C are the indices of the columns 1. 2, 3, the values U_1 U_2, U_3 are the elements of the first row, and W_1 W_2, W_3 are the elements of the second row.

will be represented by the expressions

A/ U_1, B/ U_2, C/ U_3,

A/W_1, B/W_2, C/W_3.

Let the columns of the relation ConceptsFrames with the numbers 4+1, 4+3, 4+5 have the indices Attr1, Attr2,

Attr3,…, and the columns of the relation ConceptsFrames with the numbers 4+2, 4+4, 4+6 have the indices Restr1, Restr2, Restr3,….

Example 1. Taking into account these agreements, let's consider the following linear record of the relation ConceptsFrames containing semantic frames of the concepts *contry1*, *city1*, *football-team*, and *emblem1*:

Row/1, Focus/country1, Base/space-object, Numb/3, Attr1/Called, Restr1/string, Attr2/World-part, Resrt2/space-object, Attr3/capital, Restr3/city1,
Row/2, Focus/city1, Base/space-object, Numb/3, Attr1/Called, Restr1/string, Attr2/Country-part, Restr2/country1, Attr3/Population, Resrt3/integer*,
Row/3, Focus/football-team, Base/sport-team, Numb/3, Attr1/Called, Restr1/string, Attr2/Sport-kind, Restr2/football, Attr3/Symbolics, Resrt3/emblem1,
Row/4, Focus/emblem1, Base/picture1, Numb/2, Attr1/Colours-combination, Restr1/colour-value*, Attr2/Drawn-object, Restr2/phys-object.

It is important for us that, according to TKR, *the constructs have no semantic frames*. This fact is used below in the description of the suggested algorithm.

In order to "understand" metonymic phrases, an intelligent computer system is to use not only a terminological knowledge base but also a database containing the information about the objects of external world and coordinated with TKB.

Example 2. With respect to this purpose, let's consider (continuing the example 1) the following logical structure of the relation ObjectsBase:

Row/1, Identifier/Voronezh-city, Concept/city1, Numb/3, Attr1/Called, Restr1/ "Voronezh", Attr2/Country-part, Restr2/Russian-Federation, Attr3/Population, Restr3/1.05#mln,
Row/2, Identifier/Moscow-city, Concept/city1, Numb/3, Attr1/Called, Restr1/ "Moscow", Attr2/Country-part, Restr2/Russian-Federation, Attr3/Population, Restr3/13#mln,
Row/3, Identifier/emblem73, Concept/emblem1, Numb/2, Attr1/Colours-combination, Restr1/(*red* $\wedge$ *white*), Attr2/Drawn-object, Restr2/sport-ball,
Row/4, Identifier/Spartak (Moscow), Concept/football-team, Numb/3, Attr1/Called, Restr1/ "Spartak Moscow", Attr2/Sport-kind, Restr2/football, Attr3/Symbolics, Restr3/emblem73.

The meaning of the first column in the relations ConceptsFrames and ObjectsBase is the same: the elements of these columns are the ordered numbers of the rows. The second column of the relation ObjecstBase has the index Identifier, the elements are the unique identifiers of various entities. The column No. 3 has the index Concept, the elements of this column are conceptual characteristics of the entities represented in the corresponding rows. The columns with the numbers ≥ 4 have the meanings similar to the meanings of the columns with the numbers ≥ 5 of the relation ConceptsFrames. The difference is as follows: instead of formulating a semantic restriction for a search attribute, an exact value of this attribute is indicated (for instance, compare Restr1/string and Restr1/"Moscow").

V. THE CENTRAL IDEA OF THE FIRST EXTERNAL CYCLE ACROSS SEARCH ATTRIBUTES OF THE RELATION CONCEPTSFRAMES

Suppose that the algorithm to be introduced below is processing the notion in the position *j* of the array *MentionedConcepts*, where j≥1, and the array contains the concept *football-team* in the position *.j* Let *row1* be the minimal integer ≥1 such that *ConceptsFrames*[*row1*, 2] = *football-team*. This situation means that the input text T1 mentions an entity A, and it is a football team.

The goal of our current step is to be as follows: to verify that the concept *current-sema* qualifies an entity A, and A is associated with such entity B that the value of one of the properties of B is the *sort-value*. In our example, B is an emblem, the sort-value = *colours-combination*, and the focus-value = (*red* $\wedge$ *white*).

We assume that a connection between the entities A and B may be realized only with the help of any of search attributes of the concept *current-sema*. We should consider not all search attributes of *current-sema* but only the search attributes with semantic restriction satisfying two conditions:

(a) It is not a sort, i.e., it doesn't belong to the set of sorts determined by the considered conceptual basis B (see [14]); in particular, it means that semantic restriction is not the concept *string*;

(b) It is not a sort of a construct, i.e., its semantic restriction doesn't have the final symbol * (for example, the semantic restriction of the search attribute *Population* has the semantic restriction *integer**).

The variable numb-sem-attr1 receives the value ConceptsFrames[row1, 4], it denotes the number of search attributes in the line row1. Then we consider the columns with ordered numbers of the form 4+2k-1 (k=1, 2,..), and for each fixed k fulfill the following actions;

Attribute:= ConceptsFrames[row1, 4+2k-1],

Current-value:=ConceptsFrames[row1, 4+2k].

If *current-value* is either a construct or belongs to the set of sorts St(B) (in particular, it is the sort *string*), then we do nothing and fulfil the same step after executing the assignment operator k:=k+1.

Example 3. Let j=2, then MentionedConcepts[j] = football-team, and row1:=3. Firstly, k:=1, then

ConceptsFrames[row1, 4+2-1] :=Called,

Current-value:=ConceptsFrames[row1, 4+2], hence *current-value*:=*string*. Since string is a sort, we go to the next value of k (by means of the assignment k:=k+1).

VI. A DESCRIPTION OF THE ALGORITHM OF PROCESSING METONYMIC PHRASES

An algorithm called AlgMetonymyProcessing1 is introduced below. The condition of calling it by a computer program has some small peculiarities depending on the considered input language. For English language, this condition is the occurrence in the processed text of the artificial noun attribute constructed from an attribute or several attributes. We say here "artificial noun" in the sense that the dictionaries contain no such expression with the mark that it is a noun. The expression "the red-whites" gives us an example of such construction. In case of Russian language, the condition of calling the algorithm described below is that the considered text contains an attribute or a bunch of attributes relating to no noun (an example: "krasno-belye").

In a situation of the kind, the main module of computer program forms a construct as the value of the string variable *focus-value*. For instance, for the text T1 = "On April 12, a match between the football teams "Zenith" and "Spartak Moscow" took place in Saint-Petersburg. It was the victory of the red-whites with the score 2:1", the variable *focus-value* will have the value (*red* $\wedge$ *white*).

The initial version of the algorithm AlgMetonymyProcessing1 is realized in an intelligent computer system updating the content of relational databases by means of semantic processing of the resources from a full-text database [16]. This initial version is developed by the author of the present paper and is not published anywhere. The algorithm's initial version looks rather similar to a method description, it is not a clear pseudocode. The pseudocode given below is original, and it is a contribution to theoretical foundations of algorithms fulfilling semantic processing of metonymic texts.

External specification of the algorithm

Input:

One-dimensional array **MentionedConcepts** (listing the concepts qualifying the entities mentioned in the input text);

things-number – integer – the number of filled in positions of the array MentionedConcepts;

Sorts – a finite set of symbols called sorts and denoting the most general concepts from the considered thematic domains;

The relations (two-dimensional arrays) **ConceptsFrames, ObjectsBase**;

focus-value – a construct being a unique characteristic of an object to be found;

sort-value – the sort of the construct focus-value.

Auxiliary function First-row. Its argument is any concept *semunit*. The value is such minimal $n \geq 1$ that ConceptsFrames[n, 2] = *semunit* or the value is 0.

Example 4. For the considered text T1, MentionedConcepts[2] = football-team, MentionedConcepts[3] = city1; Things-number = 4; the structure of the relations ConceptsFrames and ObjectsBase is illustrated above in the examples 1 and 2 respectively. Focus-value is the compound construct (*red* $\wedge$ *white*); sort-value is the sort *colour-value.*

Output:

posfound – integer – the number of a row of the relation ObjectsBase indicating a unique entity in the world.

Algorithm AlgMetonymyProcessing1

Begin An one-dimensional array MentionedConcepts is constructed, and the value of the integer variable Things-number is calculated.

var-go-out1 := false; var-go-out2 := false; var-go-out3 := false; j:=0

Cycle 1 with conditional exit (until-cycle) on the parameter j

Begin j:= j+1

Current-sema := MentionedConcepts[j]

{Comment: if j = 2, then current-sema := football-team}

row1 := First-row(current-sema, ConceptsFrames)

{Example: if j=2 then row1:=3 for the input text T1, see Example 1 in Section IV}

Numb-search-attr1 := ConceptsFrames[raw1, 4]

K := 0

Cycle 2 with conditional exit (until-cycle) on search attributes of the concept current-sema

{Comment: exit condition is ((var-go-out 1= true) OR (k = Numb-search-attr1))}

Begin k:= k+1

Current-value:= ConceptFrames[row1, 4+2k]

{Comment: It is semantic restriction of the k-th search attribute}

{Example. If j =2, row1= 3, k=1 then current-value:= string}

If ((current-value doesn't belong to the set of sorts) AND (current-value is not a construct))

then begin looked-concept := current-value;

row2 := First-row(looked-concept, ConceptFrames)

numb-search-attr2 := ConceptsFrames[raw2, 4]

if row2 > 0

then begin m := 0

Cycle 3 with conditional exit (until-cycle) on search attributes of the concept *looked-concept*

{Comment: exit condition is ((var-go-out 2= true) OR (m = numb-search-attr2))}

Begin m := m+1

second-search-value := ConceptsFrames[row2, 4+2m]

if second-search-value = focus-value

then var-go-out2 := true

{Comment: end-if}

until ((var-go-out2 = true) OR (m=numb-search-attr2)
{Comment: end-cycle on m}
end
until ((var-go-out1 = true) OR (k=numb-search-attr1))
{Comment: end-cycle on k}
end
until (((var-go-out1 = true) AND (var-go-out2 = true)) OR (j=ThingsNumber(MentionedConcepts)))
{Comment: end-cycle on j}
End
{Comment: Now go to processing the relation ObjectsBase}
var-go-out3= false
thing-row:= 0
Cycle with conditional exit (until-cycle) on the rows of the relation ObjectsBase
{Comment: exit condition is ((var-go-out3= true) OR (thing-row is the last row of the relation ObjectsBase))}
thing-row := thing-row + 1
current-thing:= ObjectsBase[thing-row, 4+2m]
if current-thing = focus-value
then var-go-out 3 ;= true, posfound:=thing-row
until ((var-go-out 3= true) OR (thing-row is the last row of the relation ObjectsBase))
{Comment: end-cycle on thing-row}
Output(k, m, posfound)
End End

Example 5. Taking into account the examples 1 and 2 in section IV, the successful result for the considered text T1 will be achieved for the following values of the variables: k=3 (according to the example 1, k is the ordered number of the search attribute *Symbolics* of the concept *football-team*), m = 1 (according to the example 1, m is the ordered number of the search attribute *Colours-combination* of the concept *emblem1*), posfound = 4 (according to the example 2, it is the number of the row containing information about the football-team "Spartak Moscow').

That is why the winner in the match is the football team "Spartak Moscow". This information will be inscribed into a corresponding sportive relational database.

VII. CONCLUSION

The purpose of the introduced algorithm is to process metonymical phrases. It seems that it would be easy to implement it with the help of arbitrary programming environment. The content of the next stage of the research would be the expansion of the algorithm in order to process several different kinds of metonymical phrases. In particular, to process the phrases of the kind "The winners in this hockey match were the maple leaves" with the meaning "The combined hockey team of Canada became the winner in this match".

REFERENCES

1. J. R. Hobbs, M. Stickel, P. Martin, D. Edwards, "Interpretation as Abduction", Artificial Intelligence, 1993, vol. 63, pp. 69-142.
2. K. Markert and U. Hahn, "Understanding metonymies in discourse", Artificial Intelligence, 2002, vol. 135, pp. 145-198.
3. T. Teraoka, "Metonymy analysis using associative relations between words", in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pp. 4614–4620, Portorož, Slovenia. European Language Resources Association (ELRA).
4. I. Gonzalez-Dios, J. Alvez, and G. Rigau, "Exploiting metonymy from available knowledge resources", in Computational Linguistics and Intelligent Text Processing: 20th International Conference, CICLing 2019, La Rochelle, France, April 7–13, 2019, Revised Selected Papers, Part I, Apr 2019, pp. 34–43.
5. V. A. Fomichov, "A mathematical model for describing structured items of conceptual level", Informatica. An International Journal. of Computing and Informatics (Slovenia), 1996, vol. 20, pp. 5-32.
6. V. A. Fomichov, "Theory of restricted K-calculuses as a comprehensive framework for constructing agent communication languages", in: Fomichov V.A., Zeleznikar A.P. (eds.). Special Issue on NLP and Multi-Agent Systems. Informatica. An International J. of Computing and Informatics (Slovenia), 1998, vol. 22, pp. 451-463.
7. V. A. Fomichov, "An ontological mathematical framework for E eectronic commerce and semantically-structured Web", in: Zhang Y., Fomichov V.A., Zeleznikar A.P. (eds.), Special Issue on Database, Web, and Cooperative Systems. Informatica. An International J. of Computing and Informatics (Slovenia), 2000, vol. 24, pp. 39-49.
8. V. A. Fomichov, "Theory of K-calculuses as a powerful and flexible mathematical framework for building ontologies and designing natural language-processing systems". in: *Andreasen, T., Motro, A., Christiansen, H., Larsen, H.L. (Eds.), Flexible Query Answering Systems, 5th Intern. Conference, FQAS 2002, Proceedings, Lecture Notes in Artificial Intelligence.* 2002. Berlin, Heidelberg, New York: Springer, 2002, vol. 2522, pp. 183-196.
9. V. A. Fomichov, "Comprehensive Mathematical Framework for Bridging a Gap between Two Approaches to Creating a Meaning-Understanding Web". *Intern. Journal of Intelligent Computing and Cybernetics*, 2008, vol. 1, pp. 143-163.
10. V. A. Fomichov, Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms. Springer, New York, Dordrecht, Heidelberg, 2010.
11. V. A. Fomichov, "Theory of K-representations as a Comprehensive Formal Framework for Developing a Multilingual Semantic Web", Informatica. An Intern. Journal of Computing and Informatics, 2010, vol. 34, pp. 387-396 .
12. V. A. Fomichov, "A Broadly Applicable and Flexible Conceptual Metagrammar as a Basic Tool for Developing a Multilingual Semantic Web", in: Metais, E., Meziane, F., Saraee, M., Sugumaran, V., Vadera, S. (eds.) NLDB 2013. LNCS, 2013, vol. 7934, pp. 249-259. Springer, Heidelberg.
13. V. A. Fomichov, "SK-languages as a Powerful and Flexible Semantic Formalism for the Systems of Cross-Lingual Intelligent Information Access", *Informatica. An International Journal of Computing and Informatics* (Slovenia), 2017, vol. 41, pp. 221-232.
14. V. Fomichov, "Intelligent monitoring of news on economics and finance based on formal semantics of the movement verbs", in 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), September 27 – October 1, 2021, Opatija, Croatia. Proceedings. Rijeka: Croatian Society for Information, Communication and Electronic Technology - MIPRO, 2021.
15. M. Minsky, "A Framework for Representing Knowledge". MIT, Cambridge, 1974.
16. D. V. Podrezov, "Updating the content of relational databases by means of semantic processing of the resources from a full-text database". Bachelor Thesis (academic advisor V.A. Fomichov). Moscow, Moscow Aviation Institute (National Research University), 2022. 75 p.

A Common Pentest Output Schema for Business Intelligence System Ingestion

Shivangi Sharma, Justin M. Pelletier, Bill Stackpole
Rochester Institute of Technology
Rochester, NY 14623
Email: jxpics@rit.edu

***Abstract*—Data-driven Business Intelligence process improvements demand proactive digital vulnerability discovery and exploitation enumeration. These risk discovery and analyses practices are commonly known as penetration tests (pentests), and have emerged as requirement for most organizations under many compliance regimes such as the global Payment Card Industry Data Security Standard (PCI-DSS) and the federal Gramm-Leach Bliley Act (GLBA) in the United States. Though a growing body of research addresses the utility of pentests as a valuable Business Intelligence method, pentest outputs are not standardized and require time-consuming generation and ingestion before they can provide inputs to Business Intelligence Systems. In this study, we examined several pentest reports and propose a standardized schema for output generation in a common format that is easy for Business Intelligence Systems to ingest. We propose that this improvement will allow more rapid delivery of accurate risk information to executives, managers, and information technology professionals. The thirteen elements of our standardized schema represent the first attempt at building a pentest dimensional model, which could further ease the translation of cybersecurity risk information into business impact analyses and organizational risk registers.**

***Index Terms*—business intelligence; penetration test; pentest report; data driven BIS**

I. Introduction

Penetration Testing (pentesting) is the authorized intrusion into digital systems. The outputs provide a list of exploitable vulnerabilities that are present in the system(s), a vulnerability criticality assessment, and a prioritized list of recommended remediation actions [1]. Persons who perform these activities are known by a variety of names such as pentesters, ethical hackers, and white hat hackers. Organizations either do the pentesting themselves or they take help from third parties. These activities help prevent the loss of confidentiality, integrity, and/or availability for the organization's digital systems and seek to prevent attacks that could cause financial or reputational damage [2]. Though pentests are increasingly common, the translation of those findings into business intelligence platforms remains a problem. Many pentesters have difficulty conveying their results to the executives. Reporting the results can be a tedious task that requires a considerable amount of time and effort, as there is no standard framework for reporting pentest results. This can lead to omission of important attributes of the test results in the report, or an insufficient comprehensibility of technical details. This calls into question the tester's ability and knowledge and undermines the efficacy of the pentesting process itself.

For these reasons, we provide a standardized format for penetesting reports. This is designed to reduce the time and effort required to produce an intelligible and complete report. By providing a standardized schema for pentesting report outputs, we seek to support automation of these outputs for ingestion by business intelligence systems. To do this, we compare different penetesting reports and findings to extract the necessary ingredients for a standardized reporting schema.

II. Background

There are many tools and technologies–both manual and automatic–that are used in pentesting to perform reconnaissance, scanning, gaining, and maintaining access.

A. Reconnaissance

Reconnaissance provides initial information is collected as part of the observation, which helps to understand the target and define the scope of the engagement. The information includes gathering of IP addresses, mail servers, public domains, higher-level network topology, and so on. This can be deliberately shared during the contracting of the pentest (*white-box testing*), or part of the mandate defined by a statement of work to discover what is available (*black-box testing*). Tools like Maltego provide a relatively automated mechanism for reconnaisance activities.

B. Scanning

Scanning involves sending probes and saving the responses. Network monitoring tools such as Wireshark, network scanning tools such as Nmap, webserver vulnerability scanners like Nikto, web-application vulnerability scanners like OWASP Zap, and network vulnerability scanners like Nessus are popular tools for discovering common vulnerabilities. If testing a specific application, pentesters can also use automated tools to perform static or dynamic code analysis.

Static Analysis The application's code is examined at rest. Static Application Security Testing (SAST) tools use predefined rules to help detect security vulnerabilities. Example SAST tools include Bandit, SonarQube, Veracode, Flawfinder, Cppcheck, and RATS [3]. A relatively recent addition to the field of Static Analysis allows Security Vulnerability Prediction (SVP). SVP requires models to learn the rules and detect patterns within the code that could yeild security vulnerabilities [4], [5]. SVP, and the Vulunerability Prediciton

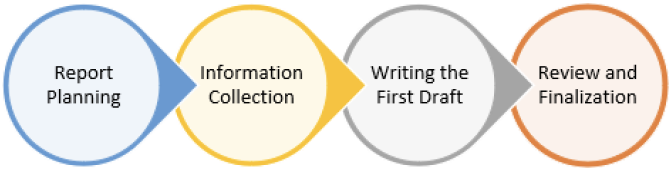


Fig. 1: Phases of Report Writing

Models they are built upon, are an active area of inquiry and have well-documented challenges to integration in automated scanning tools [6].

Dynamic Analysis The application's code is examined while running. Both commercial and open source Dynamic Application Security Testing (DAST) can be useful when the source code is unavailable and, unlike SAST tools, are not designed for any specific programming language. Example DAST tools include Burp Suite, InsightAppSec, and AppScan. Automatable DAST techiques include Web Application Security Testing, Security API Scanning, and Behaviour Driven Security Testing [7].

C. Gaining and Maintaining Access

To gain access, pentesters use information from previous stages to gain access to the application or underlying system. To maintain access, they will deploy various types of tools such as Remote Access Trojans, covert channels, and/or backdoors to escalate privileges and move laterally through a digital network. Though Gaining Access and Maintaining Access are considered two separate steps of a pentest [8], common frameworks such as Metasploit, Powershell Empire, and Cobalt Strike integrate tools that help pentesters both gain and maintain access. Some–such as Caldera–specifically seek to replicate the known tactics, techniques, and procedures of advanced persistent threats.

Generally, these tools find and exploit vulnerabilities. Applying the findings to the organization's business context and providing actionable insights are beyond the scope of currently available tools.

Once the pen testers are through with scrutinizing the weaknesses of the system, all the details have to be combined and presented in a report. Penetration Test Reports are a vital part of the entire penetration testing process since it is the proof of the test results which are presented to the company's Executives and Information Security team for them to take action according to the findings. For example, which system needs to be patched or which system needs to be upgraded, how to take further action in terms of increasing the security, and how to plan the security budget.

As early as 2010, Alharbi formalized through the SANS institute an industry standard for the phases of effective penetration test report writing [9]; this is reflected in Figure 1:

1) Report Planning – This is the first phase of report creation. In this, the testers will focus on the Report Objective and the timeline and will consider the target people for whom the report is being generated.
2) Information Gathering – In this phase the tester uses the collected data that he gathered during each step of penetration testing. The data includes the notes that he took during each phase of testing, the important screenshots which he took while running the tools, the log files of the activities, etc.
3) Writing First Draft – According to Alharbi, this step is highly recommended since it will give the tester a good idea what all things to mention in the report [9]. Also, 60% of the tester's time will be utilized in creating the first draft.
4) Review and Finalization – The first draft should be peer-reviewed before finalizing it. All the team members should participate in this process. Having other opinions could definitely be beneficial for the penetration tester as it will enhance the quality of the report and will cover all the edge cases which the tester might have left. This step is also beneficial for a situation where a team of testers were involved in the same project. Once the reviews have been collected from the peers, the report should be finalized and should be made ready to send to the clients.

III. Related Work

As early as 2010, authors such as Alharbi called for improvement in pentesting report schemas because their disparate report formats led to challenges in interpreting results [9]. More recently – in 2019 – Zakaria et al. compared eight penetration testing reports that they found online [10]. They found some of the reports missed key information, and other reports included much extraneous information. These findings led the authors to call for the community to develop a common schema.

In 2021, Alghamdi considered 20 pentest reports and described techniques of effective report writing, which provides some input to the minimum essential elements for pentest reports [11]. In the same year, Barik et al. analyzed several penetration testing tools, found disparate output formats, and proposed a standardized format for penetration testing reporting [12].

In 2022, Reddy and Pelletier considered five historical breach cases and found that the pentest output data might have helped businesses better manage risk decisions and ultimately avoid losses [13]. Though there has been a relatively robust call for pentest report output standardization, as well as a recent recognition that pentest outputs provide useful business intelligence system inputs, to date no authors have explicitly considered how a schema might allow ingestion to Business Intelligence Systems.

IV. Schema Design

For creating the standardized schema of the penetration testing report, we employed the Waterfall Design Model [14].

The remainder of this section describes our findings and analysis while following this design process, organized accord-

ing to the *a.* Research, *b.* Prototyping, *c.* Coding, *d.* Testing, and *e.* Integration phases.

A. Step 1: Research

In Step 1, we continued research to examine best practices in pentest reporting outputs and considered three tools to generate outputs for our prototype schema.

We gathered six new pentest reports, which had not previously been analyzed for format and output content. Four reports came from the Collegiate Penetration Testing Competition (CPTC) – a global penetration testing competition that simulates real-world pentests [15], [16]. CPTC competitors begin at eight regions across the globe and winners from the regional level ultimately compete in a final championship round. The CPTC organization hosts a GitHub repository containing final participant team reports, which we included in our inquiry as exemplars of world-class pentest reports [17], [18], [19], [19]. Two reports came from professional pentesters. One was from Red Siege Information Security [20], a cyber security consulting firm that provides security solutions to clients. Another came from a professional example and included a video instruction on report formatting [21].

We then used the Penetration Testing Execution Standards (PTES) baseline guidance [22] to inform the necessary components from each of the six reports we selected. According to PTES, the reports should be broken down into two major sections, each meant for a different audience: Executive and Technical.

The **Executive-level** summary should contain high-level information about the test. It includes a *high-level summary*, *findings overview*, *risk severity*, and *strategic recommendations*. The target readers are upper-level management, who are also likely to be engaged in strategic and operational level risk management practices. These data are most likely to inform the Business Impact Analysis and the Organizational Risk Register. This section is written for senior managers, who are the most frequent consumers of Business Intelligence System information [23].

The **Technical-level** detail should contain all the in-depth information of the test in a way that informs the actual risk-management practices at the tactical and operational levels. It includes an *introduction/summary*, the *methodology* for the pentest, *findings*, additional context as *informational findings*, a *conclusion*, and detailed *appendices* containing scan results and additional/non-critical evidence. The findings section is arguably the most important component of the report and it contains six key items:

1) Threat Name/Severity Level: The name of of the vulnerability/exploit and its severity level (i.e. critical, high, medium, or low).
2) Description: Explains what the vulnerability and exploit are.
3) Affected Assets: The machines in the environment affected by vulnerability/exploit.
4) Recommendation/remediation: Recommended technical solution to remediate the risk/threat.
5) References: An external validation of the finding, usually a listing in a public vulnerability or exploit database, and further information on remediation.
6) Business Impact: A broader picture of the effect of these risks should be mentioned in this section.

This detail allows Information Technology professionals to understand the root of the problem and fix it. Though this level of detail is unlikely to make it to board-level products such as the Organizational Risk Register, granular data in this section allow for a more data-driven Business Intelligence System, which is important in increasingly automated sectors such as logistics [24].

TABLE I: Executive-level report components

Executive Summary Sections	PTES [14]	IT Pro TV [18]	Red Siege Information Security [19]	CPTC A [20]	CPTC C [21]	CPTC F [22]	CPTC O [23]
Summary	✓	✓	✓	✓	✓	✓	✓
Finding Overview/ Scope	✓	✓	✓	✓	✓	✓	✓
Risk Severity Representation	✓	✓	✓	✓	✓	✓	✓
Strategic Recommendation	✓	✓	✓	✓	✓	✓	✓
Roadmap	✓					✓	

TABLE II: Technical-level report components

Technical Report Sections		PTES [14]	IT Pro TV [18]	Red Siege Information Security [19]	CPTC A [20]	CPTC C [21]	CPTC F [22]	CPTC O [23]
Introduction/ Summary		✓	✓	✓	✓	✓		
Methodology		✓	✓	✓	✓		✓	✓
Findings								
1.	Threat Name and Severity Level	✓	✓	✓	✓	✓	✓	✓
2.	Description	✓	✓	✓	✓	✓	✓	✓
3.	Affected Assets	✓	✓	✓	✓	✓	✓	✓
4.	Recommendations/ Remediation	✓		✓	✓	✓	✓	✓
5.	References	✓		✓	✓			✓
6.	Business Impact	✓			✓	✓		✓
Informational Findings		✓			✓	✓		✓
Conclusion		✓				✓	✓	✓
Appendices		✓		✓	✓	✓	✓	✓

Tables I and II depict a comprehensive comparison of the papers according to these components. Of note, the orange-colored fields are addressed by this project through a standardized JSON schema, which is described in later steps. These orange-colored fields directly inform what might be generated by automated tools and therefore should be considered in a

prototype automation schema for Business Intelligence System ingestion. In the next step, we describe this data generation and prototyping.

B. Step 2: Prototype

To generate data for our prototype schema, we considered outputs from three of the most popular [25] pentesting tools – Nmap, Nikto, and OWASP ZAP – which we described above in Section II-B. We applied those tools to exhaustively search ports and find vulnerabilites present in an instance of the DVWA (Damn Vulnerable Web Application), which is a PHP and MySQL implemented web application that is vulnerable by nature. DVWA is commonly used by the security community to research, learn and practice penetration testing tools [26]. In this case, it provided a useful set of output data during prototyping.

Each of the common tools – Nmap, Nikto, and OWASP Zap – offer a variety of output formats including XML, JSON, HTML, and text files (among others). We found that they did not, however, all offer the same standardized output:

- Nmap ⇒ XML
- Nikto ⇒ XML
- OWASP ZAP ⇒ JSON

Extracting the relevant data from these tools across multiple formats can be very time consuming. Further, for presentation to a Business Intelligence System (BIS), it is useful to have this schema output to a standardized input datatype because each entry corresponds to risk intelligence associated with each of the organizations' computer devices and the data those devices store and process. Furthermore, Yusof et al. found there is an overwhelming data integration issue across BIS that inhibits data-driven decision making [27]. By standardizing these datatypes, our prototype allows data-driven decision making associated with information technology and cybersecurity risk.

In the design selection for our schema, we considered machine-readability as a critical path requirement because our intention was to automate pentest outputs for automatic BIS ingestion. Both XML and JSON are valid options for automatic machine reading. The other main design preference was for human readability, which is likely to positively impact technology acceptance especially among senior managers. We also considered speed of ingestion and processing as a secondary design goal.

Among these output formats, JSON contains a more predictable and ultimately more human-usable format than XML [28]. Also, JSON is faster at data exchange than XML, requires fewer bytes in transit, and is supported by many data visualization tools [29]. Furthermore, in 2022, Yusof et al. specifically recommended a JSON-based standardized format as superior to XML for the overwhelming majority of Business Intelligence Systems [27]. Therefore, our prototype schema utilizes a standardized JSON output.

The variables used in our standardized JSON schema – depicted in Figure 2 – each reference a different *dimension* of the penetration test intelligence outputs, as in:

- address – URL
- host - IP address
- port - port number
- quick scan - the scans which provide the layout of the environment - for example, Nikto scans
- id - to differentiate vulnerabilities found
- method - the method used like 'GET' or 'POST'
- alert name - the vulnerability found
- risk_level - critical, high, medium, or low
- instance_count - the number of instances on which this vulnerability has been found on a given particular website
- solution - the remediation
- references - the proof of successful exploit
- affected scope - the IP addresses which are affected by this vulnerability
- other information – in this the risks which do not find any severity level, i.e., less than Low are included

This schema is noninclusive and provides a basic format and structure that can be modified further to accommodate output from additional tools or specific BIS integration requirements. In particular, automated vulnerability scanners such as Nessus provide a structured format that could be converted to this schema. Such an integration from a vulnerability scanner would lack the detail about the `references` that prove compromise exploited systems, but would still generally conform with the other data types that our schema integrates.

C. Step 3: Code

This step required transformation of pentest tool outputs into the standardized schema described above. This required two main steps:

1) **Converting** the sample XML outputs into JSON using a publicly-available tool [30].
2) **Parsing** the tool-generated reports to compile outputs into the new schema using a custom python script, depicted in Figure 3.

D. Step 4: Test

The resultant JSON file integrates the test outputs from Nmap, Nikto, and OWASP Zap, which reflects a functional proof of concept for the novel schema.

E. Step 5: Integrate

Once the schema file is created, it can be fed into any software or application programming interface that will accept data formatted in JSON.

Our proof of concept integration of standardized JSON format visualizes the data using python visualization tool libraries depicted in Figure 4. Though our proof of concept visualization is relatively simplistic, the JSON format allows for multi-dimensional modeling associated with each of the schema variables. In particular, many BIS such as Microsoft's Power BI[1] and Oracle Analytics Database[2] allow for JSON ingestion and follow-on correlation with other elements of BI.

This is useful because recent work by Chen and Siau has proven that integrating Information Technology infrastructure

[1] https://learn.microsoft.com/en-us/ [2] https://docs.oracle.com/en/

```
{
 "address": "",
 "host": "",
 "port": "",

 "open ports":[{"port number": "<port>", "service running":"<service>"}],
 "closed ports": [],
 "filtered ports": [{"port number": "<port>", "service running":"<service>"}],

 "os information": "",

 "quick scan": [{"id":"<id>","method":"<>","description":"<>"}]

 "Detailed scan": [{"id":"<id>", "alert_name":"<>",
 "description":"<>", "risk_level":"<>", "instance_count":"<>", "solution":"<>",
 "reference":"<>", "affected scope":"<>", "other information": "<>"}]
}
```

Fig. 2: Standardized JSON schema

```
import json

myjsonfile = open('nmap_basic.json', 'r')
jsondata = myjsonfile.read()
obj = json.loads(jsondata)
abc={}
abc['address'] = obj['host']['address']['@addr']
abc['host'] = obj['host']['hostnames']['hostname']['@name']
abc['ports'] = {}
abc['ports'] ['open_ports_count'] = int(len(obj['host']['ports']['port']))
abc['ports'] ['close_ports_count'] = int(obj['host']['ports']['extraports']['@count'])
abc['ports'] ['Total_ports_count'] = int(obj['scaninfo']['@numservices'])
abc['open_ports'] = []
for i in range(len(obj['host']['ports']['port'])):
          open_ports = {}
          open_ports['protocol'] = obj['host']['ports']['port'][i]['@protocol']
          open_ports['portid'] = obj['host']['ports']['port'][i]['@portid']
          open_ports['service'] = obj['host']['ports']['port'][i]['service']['@name']
          abc['open_ports'].append(open_ports)

myjsonfile2 = open('nmap_os_version.json', 'r')
jsondata2 = myjsonfile2.read()
obj2 = json.loads(jsondata2)
pqr={}
pqr['os'] = obj2['host']['os']['osmatch']['@name']

# new = {**abc, **pqr}

# print(new)
# object2 = json.dumps(pqr, indent=4)
```

Fig. 3: After XML ⇒ JSON conversion, this parser extracts each information dimension to populate the schema

with BIS positively correlates with organizational agility [31]. An important use case includes the security properties of the information technology systems used to collect, store, analyze, and report business intelligence itself. By integrating pentest output for each device, decision makers will gain a useful data source that would apply a form of meta-cognition to the BIS inputs themselves. For example, if key business decisions rely on systems or data lakes that are vulnerable to data integrity attacks, decision makers should know this and might seek additional confirmatory signals before taking a decision. Some additional integration potentials include compliance tracking, business impact analyses refinement, and private equity portfolio health.

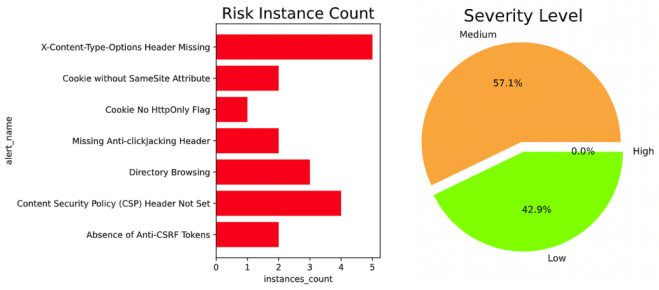


Fig. 4: Proof-of-concept visualization

V. Conclusions and Discussion

Business Intelligence Systems (BIS) enable data-driven insights. However, these insights often lack the cybersecurity risks that penetration tests (pentests) provide. The pentest method provides useful business risk data and this study demonstrates a novel implementation of those outputs to inform BIS. To do this, we analyzed six pentest reports to determine common dimensions and opportunities for automation, developed a prototype schema to display those dimensional data, generated test output from three common pentest tools, converted those outputs into the schema, and demonstrated a proof-of-concept using basic visualization tools.

This standardized schema will help pentesters process and present the results in a rapid and extensible manner. Most importantly, this standardized schema represents the first proof of concept for integrating automatic pentest tool outputs into

a format that is useful for data-driven BIS. Finally, our novel schema reduces pentest outputs into a basis spanning thirteen data elements. This represents the first step toward a usable pentest dimensional model.

VI. Limitations and Future Work

This project deals with just three tools – Nmap, Nikto, and OWASP ZAP – and there are many tools present in the market for performing a penetration test. Hence, there might be a case where the proposed JSON standardized format is missing some of the attributes which should be addressed. Therefore, future work can should consider how other vulnerability scanners and pentesting tools could enrich the format and content of this schema.

Also, this project did not include input from actual BIS users. Gathering feedback about the usability of our schema from the people who will use it will increase the practicality of this project. Furthermore, this feedback might inform dimensionality reduction of our schema to make it more efficient and practical. Hence, follow-on qualitative research should consider the ground-truth of integration of these data into BIS as well as the actual usefulness of each pentest report output field amid the myriad data-driven decisions that managers face.

VII. Acknowledgements

This research was funded in part by ongoing activities in the Eaton Cybersecurity SAFE lab at Rochester Institute of Technology's ESL Global Cybersecurity Institute. S. Sharma would like to express gratitude for the encouragement from her friends and family during this work. J.M. Pelletier would also like to acknowledge the ongoing support he receives from the *Ordo Praedicatorum*.

References

[1] P. Vats, M. Mandot, and A. Gosain, "A comprehensive literature review of penetration testing & its applications," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2020, pp. 674–680.

[2] C. Scully and P. Wang, "Router security penetration testing in a virtual environment," in *Information Technology-New Generations: 14th International Conference on Information Technology*. Springer, 2018, pp. 119–124.

[3] R. Croft, D. Newlands, Z. Chen, and M. A. Babar, "An empirical study of rule-based and learning-based approaches for static application security testing," in *Proceedings of the 15th ACM/IEEE international symposium on empirical software engineering and measurement (ESEM)*, 2021, pp. 1–12.

[4] N. Munaiah and A. Meneely, "Data-driven insights from vulnerability discovery metrics," in *2019 IEEE/ACM Joint 4th International Workshop on Rapid Continuous Software Engineering and 1st International Workshop on Data-Driven Decisions, Experimentation and Evolution (RCoSE/DDrEE)*. IEEE, 2019, pp. 1–7.

[5] H. Hanif, M. H. N. M. Nasir, M. F. Ab Razak, A. Firdaus, and N. B. Anuar, "The rise of software vulnerability: Taxonomy of software vulnerabilities detection and machine learning approaches," *Journal of Network and Computer Applications*, vol. 179, p. 103009, 2021.

[6] P. Morrison, K. Herzig, B. Murphy, and L. Williams, "Challenges with applying vulnerability prediction models," in *Proceedings of the 2015 Symposium and Bootcamp on the Science of Security*, 2015, pp. 1–9.

[7] T. Rangnau, R. v. Buijtenen, F. Fransen, and F. Turkmen, "Continuous security testing: A case study on integrating dynamic security testing tools in ci/cd pipelines," in *2020 IEEE 24th International Enterprise Distributed Object Computing Conference (EDOC)*. IEEE, 2020, pp. 145–154.

[8] R. Pandey, V. Jyothindar, and U. K. Chopra, "Vulnerability assessment and penetration testing: a portable solution implementation," in *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2020, pp. 398–402.

[9] M. Alharbi, "Writing a penetration testing report," United States: SANS Institute., 2010.

[10] M. N. Zakaria, P. A. Phin, N. Mohmad, S. A. Ismail, M. N. Kama, and O. Yusop, "A review of standardization for penetration testing reports and documents," in *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*. IEEE, 2019, pp. 1–5.

[11] A. A. Alghamdi, "Effective penetration testing report writing," in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. IEEE, 2021, pp. 1–5.

[12] K. Barik, A. Abirami, S. Das, K. Konar, and A. Banerjee, "Penetration testing analysis with standardized report generation," in *3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021)*. Atlantis Press, 2021, pp. 365–372.

[13] P. S. Reddy and J. Pelletier, "The pentest method for business intelligence," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2022, pp. 1117–1125.

[14] A. Sinha and P. Das, "Agile methodology vs. traditional waterfall sdlc: A case study on quality assurance process in software industry," in *2021 5th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*. IEEE, 2021, pp. 1–4.

[15] B. Stackpole and D. Johnson, "Cptc-a security competition unlike any other," 2019.

[16] N. Munaiah, A. Rahman, J. Pelletier, L. Williams, and A. Meneely, "Characterizing attacker behavior in a cybersecurity penetration testing competition," in *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. IEEE, 2019, pp. 1–6.

[17] https://github.com/globalcptc/report_examples/blob/master/2020/Finals-A-reportredacted.pdf.

[18] https://github.com/globalcptc/report_examples/blob/master/2020/Finals-C-report-redacted.pdf.

[19] https://github.com/globalcptc/report_examples/blob/master/2020/Finals-O-report-redacted.pdf.

[20] https://github.com/juliocesarfort/public-pentesting-reports/tree/master/RedSiege.

[21] https://go.itpro.tv/pentest-report.

[22] multiple, "Penetration testing execution standard," 2012. [Online]. Available: http://www.pentest-standard.org/index.php/Main_Page

[23] I. Constantiou, A. Shollo, and M. T. Vendelø, "Mobilizing intuitive judgement during organizational decision making: When business intelligence is not the only thing that matters," *Decision Support Systems*, vol. 121, pp. 51–61, 2019.

[24] C. Zhou, A. Stephen, X. Cao, and S. Wang, "A data-driven business intelligence system for large-scale semi-automated logistics facilities," *International Journal of Production Research*, vol. 59, no. 8, pp. 2250–2268, 2021.

[25] M. Shah, S. Ahmed, K. Saeed, M. Junaid, H. Khan *et al.*, "Penetration testing active reconnaissance phase–optimized port scanning with nmap tool," in *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2019, pp. 1–6.

[26] S. K. Oh, N. Stickney, D. Hawthorne, and S. J. Matthews, "Teaching web-attacks on a raspberry pi cyber range," in *Proceedings of the 21st Annual Conference on Information Technology Education*, 2020, pp. 324–329.

[27] M. K. YUSOF, M. MAN, W. A. F. W. HAMZAH, S. SAFEI, and I. ISMAIL, "Native json model for data integration in business intelligent applications," *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 18, 2022.

[28] http://www.json.org/xml.html.

[29] https://www.toptal.com/web/json-vs-xml-part1#:~:text=JSON%20is%20faster%20because%20it,more%20than%20just%20data%20interchange.

[30] https://jsonformatter.org/xml-to-json.

[31] X. Chen and K. Siau, "Business analytics/business intelligence and it infrastructure: impact on organizational agility," *Journal of Organizational and End User Computing (JOEUC)*, vol. 32, no. 4, pp. 138–161, 2020.

Partial SQL Query Assessment

Mario Fabijanić*, Igor Mekterović**
*Algebra University College/Software Engineering, Zagreb, Croatia mario.fabijanic@algebra.hr
**Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia igor.mekterovic@fer.hr

***Abstract* – Automated grading systems in education have been around for sixty years. They have found applications in areas such as online learning systems where virtually an unlimited number of users can test their knowledge that would not be possible to evaluate manually. Implementations within Massive Open Online Courses are a good practice in which users can do the self-testing, and get instant feedback, making the learning process more efficient. Within universities, automated grading systems allow teachers to evaluate solutions and provide feedback for thousands of submissions in a short time. This paper presents an overview of methods used in automatic SQL query evaluation systems, from early implementations when the goal was only to evaluate solutions binary, to today when they enable functionalities like partial and configurable evaluation, rich and customized feedback, learning analytics, learning pattern detection, code quality check, plagiarism detection. These methods are not exclusive, and combining different approaches makes an automated grading system more comprehensive and applicable. Automatic assessment system of SQL queries developed at Algebra University College will be presented as an example of a system which uses dynamic and static analysis, awards partial points, and gives feedback to students based on the wrong parts of their solutions.**

Keywords – SQL; assessment; fully-automated; partial marks; the configurable assessment process

I. INTRODUCTION

The rising demand to educate IT professionals increases the burden on teachers who, during the educational process, often must quickly and fairly evaluate students' solutions in the form of programming language code. The problem was recognized more than 60 years ago [1] and is present nowadays more than ever in IT colleges and universities. Online assessment systems in education have been used since 1961 when one of the first systems was used at Stanford University for the evaluation of submissions written in ALGOL [2].

In database-related university courses, the main problem is the amount of time required to evaluate student solutions. The assessment process imposes a cognitive load that lasts for days and can result in delays and inconsistencies [3]. For instance, teachers included in research [3] spent 22 hours grading 1533 student SQL solutions, and the assessment process was not finished, since students successfully challenged given grades (66% of given grades were changed). Furthermore, only 33% of teachers finished the assessment process before the deadline defined by their university. This problem is often mitigated by involving additional teachers in the assessment process, but that may lead to diminished consistency and favoritism [4]. This additional inconsistency when more teachers are included in the grading process is more emphasized especially when partial grades are assigned because teachers can have different grading criteria [5].

Manual feedback from the teacher helps student performance [6][7]. Thus, automatic feedback becomes an essential part of the education process, because often during the assessment process, an insufficient number of students get feedback that would help them fix the score next time [8]. This precise feedback is often not provided or is not available to most students. Automatic instant feedback increases students' participation and helps them resolve assignments in introductory courses [9], with improved performance and final grades [10]. Automatic instant feedback can also be implemented in a system used to learn how to construct SQL queries, providing tips during writing queries [11]. It also helps teachers to increase teaching efficiency, with a reduced workload [12]. As will be shown in the remainder of the paper, automatic grading and feedback generation are closely related.

II. GRADING SQL STATEMENTS

This chapter focuses on techniques and methods currently used or proposed by researchers in the field for evaluating SQL (Structured Query Language) code used in relational database systems. Semi-automated systems provide only part of the whole grade, and teachers do the final grading, while in case of fully automated systems, teachers do not need to evaluate individuals' solution, but they need to setup assessment system. Many semi-automated or fully-automated assessment systems were introduced during the years [15][16][17][18][19][20][21][22]. Those systems are based on dynamic analysis, which was the first method used in evaluation process with automated assessment systems.

A. Dynamic analysis

Dynamic analysis is performed by executing student's solution and by executing the correct solution assigned by the teacher. Both solutions use the same input datasets to be able to find the differences in results and give a grade that depends on the result sets' similarity. Dynamic assessment then concludes if the student's solution is fully correct or not. Typically, dynamic assessment implementation produces a binary decision, although, partial gradings are possible as with ActiveSQL but are of questionable application. The first generation of assessment systems implemented dynamic analysis with binary grading, without the possibility of giving partial points [15][24].

Dynamic assessment is very sensitive to differences in the result sets. For instance, a small, irrelevant error like an extra column in a result set causes "wrong solution" outcome. Depending on the database state and the question,

sometimes wrong solutions can return the same result set as the correct solution, so we must have a test database and we must pay attention to its content when creating questions.

Authors in [25] used information from the physical query plan of the RDBMS to generate data that will cover each node from the physical query plan represented as an ordered list or relational algebra operations. To have covering data means to have the data adequate to test different situations in the queries. All the predicates from the query should be evaluated as TRUE to get solution data as a result set for the query testing. The goal is to get data that satisfies a coverage criterion and to populate tables with such test data. To achieve this, random search may be used to explore a set of random data, biased random search to explore a set of random data, but with seeded values from the query constants or from the seeding strategy specific for join operations, and the genetic algorithm may use a fitness function for the data from physical query plan that measure how close candidate solution is from covering query target. The idea of generating query-tailored data will improve automated assessment systems making them even more reliable during grading student solutions. Regarding feedback, dynamic analysis cannot provide the reason why the solution is wrong. To overcome that imperfection, some systems are giving feedback based on the comparison of other parameter values like row count [14].

In conclusion, dynamic analysis is a very practical, popular, and widely used technique that is relatively simple to implement. On the other side, it does not provide partial grading facilities and can only provide rudimentary feedback. In a try to overcome that, some of the proposed models provide partial scoring feature comparing students' solution result sets with the expected result set from the correct solution. Partial grade is then calculated as a proportion of correct cell values of the total cell values of the correct result set. However, this is a rather dubious approach with limited applicability [21s]. It should be noted that dynamic assessment attempts to (dis)prove equivalence of two queries by comparing their effects (outputs). Why not compare the queries themselves which is our problem in the first place? The following section describes such approach.

B. Queries equivalence

The problem of recognizing SQL query equivalences is not a simple task because of its' undecidability [26]: two SQL queries can be syntactically different, but semantically/logically the same and both will produce the same result set for the same inputs. A group of authors proposed an automated prover which can determine the equivalence of SQL queries, named Cosette [27]. Cosette consists of two components: first component translates SQL queries into logic formulas, and then searches for the examples of input data which would show that the two queries are not the same, and the second component translates queries into K-relations, which are then validated for the equivalence. K-relations are the representation of relations in the form of mathematical functions which receive a tuple and return input tuples' multiplicity, which can be checked to proof the equivalency [27].

Cosette has several drawbacks due to the limitations in HoTT (Homotopy Type Theory) library [35] for Coq proof tool, on which it is based on [26]. For instance, Cosette does not support foreign keys because they are difficult to model with HoTT library in Coq [26]. To overcome this limitation, authors of Cosette proposed new algebraic structure - unbounded semiring (U-semiring). They used it as new formalism for SQL queries. To prove that two queries are semantically equal, SQL queries are converted into U-semiring expressions (U-expressions). In the next step, U-expressions are compared using UDP (U-expression Decision Procedure) algorithm [26]. U-semiring may model some of the SQL features like integrity constraints, which was one of the main drawbacks in Cosette. UDP supports more different SQL queries then Cosette, with more powerful automated proof search [26]. Although these methods brought new algorithms for checking the equality of two SQL queries, and results were promising, to the best of our knowledge, their development for wider range of SQL features didn't happen, probably due to their overall complexity.

C. Static analysis

Rather than proving that two SQL queries are equivalent, which has proven to be too difficult [26], many researchers turned to static analysis of SQL statements. It resulted in the second-generation automated assessment systems, which use static analysis method [13]. In situations when the result sets from student solution and correct solution are equal, without static analysis system cannot conclude if students' solution fully satisfies the expected solution. For instance, in case when the assignment (and solution) explicitly calls for using joins, and student provides a logically equivalent solution which uses subqueries.

Static analysis is the analysis of the code without executing it. In our context, it can be used to analyze structure of the SQL code, or its' abstract syntax representation which is known as the abstract syntax tree (AST). By comparing parts of AST, the similarity of SQL queries may be determined [28]. AST represents the code written in a programming language and adjusted to contain only nodes that matter for the comparison, without elements like comments, semicolons, parenthesis, aliases, or any other elements without an impact on the functionality of the solution.

Before creating AST, SQL codes that should be compared are converted into canonical forms. Canonical form of SQL query is again a SQL query, but without parts that are irrelevant for the comparison with another query [29]. Canonical forms of SQL queries minimize the syntactic differences between two queries that need to be compared. After the process of canonicalization, abstract trees are created from canonical forms.

i) Canonicalization process

The process of canonicalization can be divided into syntactical and semantic canonicalization. Examples of syntactical canonicalization challenges:

1. When the solution uses the BETWEEN predicate, it is replaced with the equivalent relational operators "greater or equal to" and "smaller or equal to", e.g.,

```
A BETWEEN 5 AND 10
```

is replaced with

`A>=5 AND A<=10` [29].

2. When the solution uses NOT, it is converted into a statement without NOT, e.g., NOT (A>B) is replaced with A<=B. In addition, predicates that contain greater than are replaced with smaller then, e.g., A>B is replaced with B<A [29].
3. When the solution uses a subquery with the ORDER BY clause and without the TOP clause, the ORDER BY clause is removed because it does not change the result set.
4. When the solution uses INNER JOIN, the query optimizer chooses optimal order during query execution [30], and it can be different in the correct solution than the order in the student solution. For instance, when comparing two statements:

```
FROM City as c INNER JOIN Student as s
 ON c.CityID=s.CityID
```

and

```
FROM Student as s INNER JOIN City as c
 ON c.CityID=s.CityID
```

Regarding functionality, those two statements are equal, but their abstract trees are different, and that is why those kinds of nodes are sorted before calculating the similarity of statements.

Examples of semantic canonicalization challenges:

1. When the solution uses the DISTINCT clause to remove duplicates, and they do not exist, the DISTINCT clause can be removed from the solution. The case of the primary key attribute or any other unique attribute in the SELECT list and as a join predicate, meaning the DISTINCT clause can be removed [29].
2. When a solution contains redundant joins, they can be removed, e.g., if the solution is:

```
SELECT * FROM student
LEFT OUTER JOIN department
ON student.dept_id=department.id
WHERE student.dept_name = 'Biology'
```

The condition in the WHERE clause fails if student_dept_name is null, and the solution can be rewritten to use INNER JOIN instead.

Left or right outer join can be replaced with an inner join if (i) the attributes used in the join include all foreign key references from one side of a join to the other side of a join and (ii) foreign key values used in the join are not null [29].

3. When a solution contains the predicates which are replaceable, lexicographically least attribute will be used in all clauses, e.g., if the query is:

```
SELECT student.dept_name FROM student
INNER JOIN department
ON student.dept_name=department.dept_name
WHERE student.dept_name LIKE 'English%'
```

And because of the student.dept_name=department.dept_name, the department.dept_name attribute will be used in the SELECT clause too. In addition, if constant is a part of the equality condition, it is considered as lexicographically least attribute [29].

4. When a solution contains an ORDER BY clause with more than one attribute, and if the second attribute is functionally dependent on the first attribute, the second attribute is ordered the same way as the first attribute, and therefore, irrelevant, and can be removed from the solution. Similarly, when a solution contains GROUP BY, e.g.,

```
SELECT id, COUNT(*) FROM student
INNER JOIN takes ON student.id=takes.id
GROUP BY id, name
```

and if student.id is the primary key, the attribute name can be removed from the GROUP BY clause [29].

ii) Using ASTs

After the canonicalization process, ASTs are used to find the differences between students' solution and correct solution. One possible approach is to calculate tree edit distance between students' solution tree and correct solution tree [28]. It is calculated as number of modifications to the student solution tree until it becomes equivalent to the correct solution tree. This makes partial grading based on the number of needed modifications possible, but the challenge regarding this approach is the need to test the equivalence after each modification of the student solution. There is also a possibility to try all the edits possible at each step, again testing the equivalence at each step, making the number of edits exponential [29]. During tree similarity calculation, ASTs may be decomposed into smaller subtrees which are more suitable for the tree edit distance algorithms.

First tree edit algorithms proposed in the late seventies used a hard-coded strategy and did not consider the shape of the trees which resulted in big differences during runtime [31][32]. More recent algorithms like RTED (Robust Tree Edit Distance) first calculate the decomposition strategy and then adapt to the input tree, which results in better runtime [33]. It is not perfect in terms of memory consumption for large input trees, since the strategy computation can use twice the memory used for distance computation, limiting its usage on large trees. Newer algorithm, AP-TED (All Path Tree Edit Distance) and AP-TED+ reduce the memory consumption by 2/3 compared with the RTED algorithm [33]. This is achieved by releasing the memory early during decomposition.

In addition, the AP-TED+ algorithm considers the fact that during the distance computation phase majority of subtrees contain up to two nodes, and the algorithm uses different functions to compute the distance when one of the trees is small [33]. Another approach of using ASTs is to check a set of grading rules against key-value pairs for each SQL clause within AST. Grading rules make partial grading possible, and they should be created by teachers, together with the correct solution, before assessment process starts.

iii) Without ASTs

The abstract tree can be built when the SQL solution is syntactically correct. If that is not the case, the similarity between two SQL statements should be calculated differently. For that purpose, text comparison of both SQL statements may be done. Before the comparison, both solutions are normalized: tab, newline, and semicolon characters are replaced with whitespace, leading and trailing whitespaces are removed, and all double spaces are replaced with one whitespace. With those adjustments, the similarity of the two texts may be calculated using e.g. Levenshtein distance metric [28]. Levenshtein distance is calculated as the minimum number of single-character edits needed to transform one text into another [34]. Some researchers suggest value 1 as acceptable Levenshtein's distance when a student can still get the maximum grade, and it is considered as an unintentional typo that does not affect student knowledge and therefore, student grade.

D. Combined Approach

Two main approaches are dynamic and static analysis [28], but combining them increases the quality of the automated assessment systems in general. Putting dynamic and static analysis into relation, the result of the dynamic analysis can be one of those:

a) Candidates for correct statements

b) Statements that cannot be executed

c) Partially correct statements.

The dynamic analysis may discover candidates for correct solutions, if they can be executed and provide the expected result set. The problem is that student could get correct result set, but without the need to use expected SQL clauses, and dynamic analysis cannot discover it. Example of this problem is when the UNION statement is used, instead of expected SQL clauses, and tweaks the result set knowing what correct result set should be.

The static analysis is primarily used for partially correct statements, but it can also be used for students' solutions that dynamic analysis confirmed as potentially correct, preventing the fraud like in previous example. In addition, queries that are different in their syntax can be semantically equal, which makes the problem more complex [33].

III. AUTOMATIC ASSESSMENT SYSTEM OF SQL QUERIES DEVELOPED AT ALGEBRA UNIVERSITY COLLEGE

In this chapter, an automatic assessment system of SQL queries developed at Algebra University College will be presented as an example of a system which uses dynamic and static analysis, and therefore may award partial points and give feedback to students, focused on the wrong parts of their solutions. This system uses an input from the teacher: the correct SQL query solution, and the grading rules with the deducting percentage of maximum points in case of errors. During static analysis, the AST of student solutions are compared with AST of the correct SQL query solution. Figure 1 shows an example of the abstract tree application and Figure 2 shows an example of the feedback to student, both taken from an automatic assessment system used during previous exam terms.

Example of the correct SQL query solution:

```
SELECT YEAR(i.InvoiceDate) as InvYr
 , MONTH(i.InvoiceDate) as InvMnth
 , COUNT(i.IDInvoice) as NumberOfInvoices
FROM Invoice as i
 INNER JOIN Customer as c
 ON i.CustomerID=c.IDCustomer
INNER JOIN City as ct
 ON ct.IDCity=c.CityID
WHERE ct.Name='Rijeka'
GROUP BY YEAR(i.InvoiceDate)
 , MONTH(i.InvoiceDate)
HAVING COUNT(i.IDInvoice)<30
   ORDER BY InvYr DESC, InvMnth
```

Example of students' solution:

```
SELECT YEAR(i.InvoiceDate) as InvYr
 , COUNT(i.IDInvoice) as NumberOfInvoices
FROM Invoice as i
INNER JOIN Customer as c
 ON i.CustomerID=c.CityID
INNER JOIN City as ct
 ON ct.IDCity=c.CityID
WHERE ct.Name='Split'
GROUP BY YEAR(i.InvoiceDate)
 , MONTH(i.InvoiceDate)
HAVING COUNT(i.IDInvoice)<30
   ORDER BY InvYr
```

Examples of the grading rules:

a) Regarding the FROM clause:

   i) The weight in the whole query: 100%

   ii) If valid, deduct 0% of possible points

   iii) If not valid, deduct 100% of possible points

b) Regarding the WHERE clause:

   i) The weight in the whole query: 50%

   ii) If not valid, deduct 50% of possible points

   iii) If valid, but wrong columns are included, deduct 25% per each wrong column

   iv) If valid, with correct columns included, but with wrong values included, deduct 25% per wrong value

c) Regarding SELECT clause:

   i) The weight in the whole query: 30%

   ii) If not valid, deduct 10% per each missing column.

Teacher grading rules used during the grading process can be stored in the configuration file or a database. An example of the set of general rules:

```
{ "max_points": 4.0,
  "clause_points_from": 50,
  "clause_points_where": 15,
  "clause_points_groupby": 30,
  "clause_points_having": 15,
  "clause_points_select": 20,
  "clause_points_orderby": 10,
  "parses_and_compiles": 100,
  "wrong_results": 50 }
```

In this example, the JSON configuration file contains the maximum points for the correct solution and weights for each clause in the query. These weights represent percentages that will be deducted from the max_points if erroneous. Also, the rules state the deduction percentage if the solution does not parse and compile, and if it returns a wrong result set. Note that the latter is evaluated via dynamic assessment.

A set of rules per each clause may also be defined:

```
"from_each_missing_table"
"from_each_wrong_join"
"from_max_joins_allowed"
"from_use_specific_join"
"where_wrong_rows"
"where_partial_correct_column"
"where_use_specific_operator"
"groupby_not_correct"
"groupby_partial_correct_column"
"having_wrong_number_of_rows"
"having_missing_aggregate_function"
"select_missing_regular_column"
"select_missing_function_column"
"select_missing_aggregate_column"
"orderby_each_missing_column"
"orderby_each_wrong_ascdesc"
"orderby_each_wrong_column_order"
```

An example of the rule for the HAVING clause:

```
"having_missing_aggregate_function":
   {  "status": "on",
      "deduct": 100,
      "hard_stop": 0  }
```

For each of the presented grading rules, there may be a subset of settings that enable or disable the rule, the deduction percentage if the rule is not satisfied, or whether to stop or continue with the grading process, like in above example. Additional settings can also be used, e.g.:

- **`specific_join`** in the case that student solution must contain required type of inner join, outer join, full join, or cross join in the FROM clause,
- **`specific_operator`** in the case that specific operator must be used in the WHERE clause,

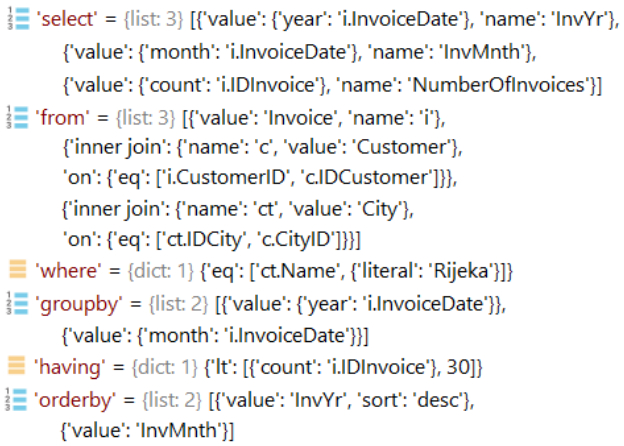


Figure 1. Example of the abstract tree application

```
FROM: Missing join of tables city and customer
WHERE: Incorrect rows filtered
GROUP BY: Incorrect columns
GROUP BY: Partial, column year(invoice.invoicedate) found
SELECT: Missing function invmnth
ORDER BY: Missing column InvMnth
ORDER BY: Column invyr should be sorted desc
ORDER BY: Column 2 should be invmnth but is nonexistent
Task I4Z1 graded, points scored: 1.22/4.0

Total points scored: 1.22/4.0
```

Figure 2. Example of the feedback

- **`num`** as a number for maximum joins or subqueries needed in student solution,
- **`select_subqueries_required`** in the case when subquery must be within the SELECT clause,
- **`where_subqueries_required`** in the case when subquery must be within the WHERE clause.

The precision of the grading rules also has a great influence on the feedback possibilities. Feedback is addressed as one of the most important parts of the whole education process. Rich feedback is a prerequisite that students may become aware of their knowledge gaps and according to those, they can make additional effort in learning [8].

In general, the usability of the feedback is recognized in several areas: (i) info about failure – evaluation logs as the source of that data, (ii) info about failed tests – info about how to proceed, with a hint on how to overcome an issue, (iii) structured report of the evaluation including details, e.g., execution times, resource usage, and tests passed and failed, and (iv) manual report – messages written by the instructor about a submitted solution [13]. With the static analysis implemented, many of the aforementioned ideas for feedback become possible. It is a significant improvement over dynamic analysis, where the feedback was limited to the comments about what is wrong in result set of the query, but without too helpful information about which part of the solution was wrong, and the significance of the error. On top of the partial grading feature, rich feedback is a great improvement in automated assessment systems.

IV. CONCLUSION

Automated grading systems in the educational process are needed more than ever. Their popularity is growing with the growing number of students enrolled in computer science courses and with the desire to improve the educational process. Partial grading of student SQL queries and rich feedback from the automated grading system could have a major impact on achieving this goal. The main problem to be solved is to prove that two SQL queries (exact solution and student solution) are equal. This is an extremely demanding task, and instead of proving the equivalence of SQL queries, two other approaches that use dynamic and static analysis have prevailed. Dynamic analysis is relatively simple to implement but solves only part of the given problem. On the other hand, static analysis significantly contributes to achieving the set goal - partial evaluation of SQL solutions and at the same time rich

feedback for students. Combining both methods greatly increases the capabilities of the system for automatic evaluation of SQL solutions.

REFERENCES

[1] J. Hollingsworth, "Automatic graders for programming classes," Commun. ACM, vol. 3, no. 10, pp. 528–529, Oct. 1960, doi: 10.1145/367415.367422.

[2] S. Wasik, M. Antczak, J. Badura, A. Laskowski, and T. Sternal, "A survey on online judge systems and their applications," ACM Comput. Surv., vol. 51, no. 1, 2018, doi: 10.1145/3143560.

[3] J. Tharmaseelan, K. Manathunga, S. Reyal, D. Kasthurirathna, and T. Thurairasa, "Revisit of automated marking techniques for programming assignments," in IEEE Global Engineering Education Conference, EDUCON, 2021, vol. 2021-April, doi: 10.1109/EDUCON46332.2021.9453889.

[4] D. Fonte, D. Da Cruz, A. L. Gançarski, and P. R. Henriques, "A flexible dynamic system for automatic grading of programming exercises," OpenAccess Ser. Informatics, vol. 29, pp. 129–144, 2013, doi: 10.4230/OASIcs.SLATE.2013.129.

[5] I. Albluwi, "A Closer Look at the Differences between Graders in Introductory Computer Science Exams," IEEE Trans. Educ., vol. 61, no. 3, pp. 253–260, 2018, doi: 10.1109/TE.2018.2805706.

[6] D. Nicol and D. MacFarlane-Dick, "Formative assessment and selfregulated learning: A model and seven principles of good feedback practice," Stud. High. Educ., vol. 31, no. 2, pp. 199–218, 2006, doi: 10.1080/03075070600572090.

[7] J. Hattie and H. Timperley, "The power of feedback," Rev. Educ. Res., vol. 77, no. 1, pp. 81–112, 2007, doi: 10.3102/003465430298487.

[8] A. P. Cavalcanti et al., "Automatic feedback in online learning environments: A systematic literature review," Comput. Educ. Artif. Intell., vol. 2, p. 100027, 2021, doi: 10.1016/j.caeai.2021.100027.

[9] S. Krusche and A. Seitz, "ArTEMiS - An automatic assessment management system for interactive learning," SIGCSE 2018 - Proc. 49th ACM Tech. Symp. Comput. Sci. Educ., vol. 2018-Janua, pp. 284–289, 2018, doi: 10.1145/3159450.3159602.

[10] W. Zhou, Y. Pan, Y. Zhou, and G. Sun, "The framework of a new online judge system for programming education," ACM Int. Conf. Proceeding Ser., pp. 9–14, 2018, doi: 10.1145/3210713.3210721.

[11] M. H. Ying and Y. Hong, "The development of an online SQL learning system with automatic checking mechanism," Proc. - 7th Int. Conf. Networked Comput. Adv. Inf. Manag. NCM 2011, pp. 346–351, 2011.

[12] X. Xie and X. Li, "Research on Personalized Exercises and Teaching Feedback Based on Big Data," ACM Int. Conf. Proceeding Ser., pp. 166–171, 2018, doi: 10.1145/3232116.3232143.

[13] J. C. Paiva, C. I. Tec, and D. C. C. Fcup, "Automated Assessment in Computer Science Education : A State-of-the-Art Review," pp. 1–39, 2022.

[14] I. Mekterovic, L. Brkic, B. Milasinovic, and M. Baranovic, "Building a comprehensive automated programming assessment system," IEEE Access, vol. 8, pp. 81154–81172, 2020, doi: 10.1109/ACCESS.2020.2990980.

[15] A. Kleerekoper and A. Schofield, "SQL tester: An online SQL assessment tool and its impact," Annu. Conf. Innov. Technol. Comput. Sci. Educ. ITiCSE, pp. 87–92, 2018, doi: 10.1145/3197091.3197124.

[16] P. J. Wagner, "The sql file evaluation (sqlfe) tool: A flexible and extendible system for evaluation of sql queries," Annu. Conf. Innov. Technol. Comput. Sci. Educ. ITiCSE, p. 1334, 2020, doi: 10.1145/3328778.3372599.

[17] A. Trongratsameethong, P. Vichianroj, and C. Mai, "ASQLAG - Automated SQL Assignment Grading System for Multiple DBMSs," vol. 1, no. 1, pp. 41–59, 2018, doi: 10.14456/jti.2018.4.

[18] C. Kleiner, C. Tebbe, and F. Heine, "Automated grading and tutoring of SQL statements to improve student learning," ACM Int. Conf. Proceeding Ser., pp. 161–168, 2013, doi: 10.1145/2526968.2526986.

[19] M. De Raadt, S. Dekeyser, and T. Y. Lee, "Do students SQLify? Improving learning outcomes with peer review and enhanced computer assisted assessment of querying skills," ACM Int. Conf. Proceeding Ser., vol. 276, pp. 101–108, 2006, doi: 10.1145/1315803.1315821.

[20] J. Castelein, M. Aniche, M. Soltani, A. Panichella, and A. Van Deursen, "Search-based test data generation for SQL queries," Proceedings - International Conference on Software Engineering. pp. 1220–1230, 2018, doi: 10.1145/3180155.3180202.

[21] A. Cumming and G. Russell, "Automatic Checking of SQL: Computerised Grading," Int. J. Learn. Annu. Rev., vol. 12, no. 3, pp. 127–134, 2006, doi: 10.18848/1447-9494/cgp/v12i03/46714.

[22] J. C. Prior, "Online Assessment of SQL Query Formulation Skills," Ace '03, vol. 20, pp. 247–256, 2003, [Online]. Available: http://delivery.acm.org/10.1145/860000/858433/p247-prior.pdf?ip=201.5.254.119&id=858433&acc=PUBLIC&key=4D4702B0C3E38B35.4D4702B0C3E38B35.6DB1341FA924E6B8.4D4702B0C3E38B35&CFID=525851233&CFTOKEN=96683423&__acm__=1436149768_230e7f22b7329beb08e8473b5e50b453.

[23] P. Brusilovsky, M. V. Yudelson, S. Sosnovsky, V. Zadorozhny, D. H. Lee, and X. Zhou, "An open integrated exploratorium for database courses," Proc. Conf. Integr. Technol. into Comput. Sci. Educ. ITiCSE, vol. 1, no. 412, pp. 22–26, 2008, doi: 10.1145/1384271.1384280.

[24] S. Sadiq, M. Orlowska, W. Sadiq, and J. Lin, "SQLator - An online SQL learning workbench," SIGCSE Bull. (Association Comput. Mach. Spec. Interes. Gr. Comput. Sci. Educ., vol. 36, no. 3, pp. 223–227, 2004, doi: 10.1145/1026487.1008055.

[25] J. Tuya, M. J. Suárez-Cabal, and C. De La Riva, "Full predicate coverage for testing SQL database queries," Actas las 16th Jornadas Ing. del Softw. y Bases Datos, JISBD 2011, no. January, pp. 683–684, 2011, doi: 10.1002/stvr.

[26] S. Chu, B. Murphy, J. Roesch, A. Cheung, and D. Suciu, "Axiomatic foundations and algorithms for deciding semantic equivalences of SQL queries," Proc. VLDB Endow., vol. 11, no. 11, pp. 1482–1495, 2018, doi: 10.14778/3236187.3236200.

[27] S. Chu, C. Wang, K. Weitz, and A. Cheung, "Cosette: An automated prover for SQL," CIDR 2017 - 8th Bienn. Conf. Innov. Data Syst. Res., 2017.

[28] J. Wang, Y. Zhao, Z. Tang, and Z. Xing, "Combining dynamic and static analysis for automated grading SQL statements," J. Netw. Intell., vol. 5, no. 4, pp. 179–190, 2020.

[29] B. Chandra, A. Banerjee, U. Hazra, M. Joseph, and S. Sudarshan, "Automated grading of SQL queries," Proc. - Int. Conf. Data Eng., vol. 2019-April, pp. 1630–1633, 2019, doi: 10.1109/ICDE.2019.00159.

[30] L. Giakoumakis and C. Galindo-legaria, "Testing SQL Server ' s Query Optimizer : Challenges , Techniques and Experiences," Bull. IEEE Comput. Soc. Tech. Comm. Data Eng., pp. 1–8, 2008.

[31] E. D. Demaine, S. Mozes, B. Rossman, and O. Weimann, "An optimal decomposition algorithm for tree edit distance," ACM Trans. Algorithms, vol. 6, no. 1, pp. 146–157, 2009, doi: 10.1145/1644015.1644017.

[32] K. H. Lee, Y. C. Choy, and S. B. Cho, "An efficient algorithm to compute differences between structured documents," IEEE Trans. Knowl. Data Eng., vol. 16, no. 8, pp. 965–979, 2004, doi: 10.1109/TKDE.2004.19.

[33] M. Pawlik and N. Augsten, "Tree edit distance: Robust and memory-efficient," Inf. Syst., vol. 56, pp. 157–173, 2016, doi: 10.1016/j.is.2015.08.004.

[34] L. Yujian and L. Bo, "A normalized Levenshtein distance metric," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 6, pp. 1091–1095, 2007, doi: 10.1109/TPAMI.2007.1078.

[35] J. Gross, M. Shulman, A. Bauer, P. L. Lumsdaine, A. Mahboubi, and B. Spitters. The HoTT libary in Coq. https://github.com/HoTT/HoTT.

DE-DS

International Conference on

DIGITAL ECONOMY AND DIGITAL SOCIETY

Steering Committee

Chairs: Edvard Tijan, University of Rijeka, Croatia
Mladen Mauher, Zagreb University of Applied Sciences, Croatia

Members: Dragan Čišić, University of Rijeka, Croatia
Alen Jugović, University of Rijeka, Croatia
Đani Juričić, Jožef Stefan Institute, Ljubljana, Slovenia
Anđelko Milardović, Institute for Migration and Ethnic Studies, Zagreb, Croatia
Ana Perić Hadžić, University of Rijeka, Croatia
Davor Salamon, Zagreb University of Applied Sciences, Croatia
Vjeran Strahonja, University of Zagreb, Varaždin, Croatia
Marijana Vidas-Bubanja, Belgrade Business School, Serbia

Program Committee

Members: Saša Aksentijević, Aksentijević Forensics and Consulting, Ltd., Viškovo, Croatia
Mladen Mauher, Zagreb University of Applied Sciences, Croatia
Edvard Tijan, University of Rijeka, Croatia

PAPERS

Risk Management in Development of Croatian Maritime Cargo Single Window System

S. Aksentijević*, E. Tijan**, K. Nikolozo*** and A. Perić Hadžić**
* Aksentijević Forensics and Consulting, Ltd., Viškovo, Croatia
sasa.aksentijevic@uniri.hr
** University of Rijeka, Faculty of Maritime Studies/Department of Maritime Logistics and Management, Rijeka, Croatia
edvard.tijan@uniri.hr, ana.peric@uniri.hr
*** Bartolići 6, Zagreb, Croatia
kristijan.nikolozo@gmail.com

***Abstract* - The aim of this paper is to research and describe the effectiveness of the risk management approach used during implementation of the Croatian national maritime Single Window system related to cargo, or Port Community System, initially in the port of Rijeka. The primary stakeholders and their internal systems, as well as formal development requirements, are defined. Project risk is identified within the project management framework's criteria, and the project's execution and important milestones, including risk occurrences during project execution, are recognized and discussed. Finally, modifications to the approach are suggested, as well as future study opportunities.**

Keywords – Single Window System, Port Community System, Port of Rijeka, risk management

I. INTRODUCTION

The major purpose of the Croatian maritime cargo Single Window system (or Port Community System - PCS) implementation project in the Port of Rijeka is to connect all port stakeholders, including government and commercial entities and their digital information systems to the Croatian Integrated Maritime Information System through a single digital platform (CIMIS) [1] [2]. The construction of the Croatian Port Community System began in late 2017, when the Port of Rijeka Authority received funding from the INEA [3], in the form of the Connecting European Facility – POR2CORE – Port Community System project [4]. The project was overseen by the Croatian Ministry of the Sea, Transport, and Infrastructure [5] and managed by the Port of Rijeka Authority. The main goal of the project was to align the implementation activities with other ongoing digital technology developments in the maritime sector, and to create a system that could be used in other Croatian cargo ports after the initial pilot project in Rijeka is successfully completed.

With the project's ambitious scale, numerous interested parties, limited finance, and rigorous contracting conditions and deadlines, it was clear from the start that there would be several major groups of risk sources. For this reason, work began on identifying a risk framework that would be appropriate for tracking the evolution of such a complex system with numerous internal and external stakeholders and risk types. After internal analysis, it was decided to adopt traditional enterprise risk management approach and monitor its performance and capabilities. This approach was used over the course of the project's four years of execution.

The risk management procedures during time and money constrained complicated cargo marine Single Window projects are scarcely and incidentally studied and have not been well addressed. The majority of prior research has focused on general and top-level risk management. The tracking of project risks is frequently entrusted to senior management, who are responsible for monitoring project risks and mitigation activities, which must be consistent with expected timetables [6]. Based on prior experience, some authors argue that the costs of implementation may be larger than the benefits in some circumstances, rendering whole implementation initiatives worthless from a pure financial standpoint [7]. The majority of risk assessments for maritime cargo Single Windows are centred on their economic feasibility. Some writers acknowledge that there is a lack of maritime cybersecurity awareness, as well as the significance of a comprehensive strategy based on hazards and maritime cyber risk valuations linked with maritime authorities and identification of critical assets in this sector [8]. Researchers that are more technically oriented in their studies adopt a different approach, and their perspective of hazards during project execution is restricted to features of the underlying technology, digital components, and cybersecurity. They acknowledge that, until recently, ports were primarily concerned with physical security. However, cyber assaults pose the greatest threat today [9], so there is pronounced need for approach change.

As a result of previous research, it appears that during the development of complex cargo maritime Single Window systems, there is a disjointed approach to risk analysis and management methodology, where different stakeholders, based on their previous experience with separate segments of the system, take a different approach to individual risk types and mitigation measures, rather than an integral model approach. Quantitative advantages may be assessed using traditional

project management methodologies as well as financial indicators and approaches, but qualitative benefits are best realized through strategic analysis.

II. PCS Development Requirements

Interoperability with existing IT systems of individual stakeholders of the port community, the CIMIS NMSW system, and the Republic of Croatia's eCustoms system, was the primary requirement that PCS had to accomplish in order to achieve the project's intention [10]. It was created with the following implementation goal in mind: to enhance and ease the optimal flow of information between participants in integrated maritime and land transportation systems employing integrated components. The most important needs are the simplicity and optimization of business operations among port community members, single data entry, and data confidentiality, integrity, and availability. It had to be adjusted to the port of Rijeka's design and operative construction, which has four distinctive, independent port basins [11].

Port Authority, Harbor Master's Office, Ministry of the Sea, Transport and Infrastructure, Ministry of the Interior (border police), Customs authorities, maritime carriers, port concessionaires - terminal operators, maritime agencies, shipping companies and transport organizers, land carriers (rail and road), sanitary inspection, phytosanitary inspection, veterinary inspection, and other administrative services are among the most important stakeholders in the Rijeka PCS development project. Therefore, it had to be designed to accommodate their needs, but also additional stakeholders in the port community. Stakeholders in the PCS system are involved in goods transportation and associated operations, although their scope varies, and it is dependent on each company's profile and activity. Each future stakeholder must be able to secure its own data, with the ability to manage the data for which they have been granted permission and approved access.

Figure 1 shows the orchestration of various stakeholders via interconnected systems and PCS modules.

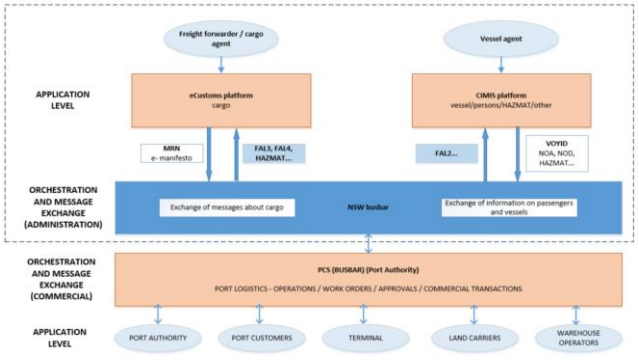


Figure 1: PCS connections orchestration

Identified PCS modules use the following codification:

1. D1 - official procedures of maritime administration and exchange of the information with the NMSW CIMIS system;
2. D2- ECS (Export Control System) and ICS (Import Control System) modules – exchange within customs procedures prescribed by the Ministry of Finance;
3. D3 - port container terminals operations;
4. D4 - NCTS (New Computerized Transit System) module – exchange of transit customs procedures data;
5. D5 - customs module – exchange with the eCustoms system, local risk control and integration with user applications;
6. D6 - port coordination and task planning;
7. D7 - access to the port area - control of vehicles and persons in the port area under ISPS rules and connection with truck announcement system (PORTUNUS);
8. D8 - conventional cargo warehouse operations – exchange with conventional cargo warehousing systems, delivery and dispatch of goods by sea and land, handling of containers;
9. D9 - hinterland container terminal operations;
10. D10 - CFS (Container Freight Station) module - container manipulation orders at the container terminal, storage and warehousing and
11. D11 - AGCT (container terminal operator) Rail module for rail container traffic - railway wagon operations.

Therefore, PCS has to be connected to the following varied information systems utilized by port operations stakeholders in order for all these modules to be effectively constructed and placed into operation: CIMIS, eCustoms, TOS F4B, TOS COMBIS, TOS NAVIS, TOS AGCT CFS and PORTUNUS. These systems are developed by the Ministry of the Sea, Transport and Infrastructure, Croatian Customs office, and cargo concessionaires to facilitate various aspects of the cargo operations in the port.

III. Project Execution

It is worth noting that, according to the original contract annexes, project plans do not foresee critical path milestones; instead, project execution was atomized in order to maximize available financial resources and deploy as many modules and capabilities as feasible, departing from the original plan. However, this appears to be a logical result of differing levels of participation in the project by diverse parties, particularly concessionaires. According to the initial project plan, critical milestones were (1) the creation of module D1 in the supplier's environment, (2) the installation of hardware and network components, and (3) the development of the overall functional specification. The first and second milestones were met; the third milestone

was not, but they proved ultimately not to be important for the overall project's success.

Each step of the project execution was defined by the occurrence of a number of risks, which resulted in annexes to the original contract, scope extensions, and the dynamics of project development. The first significant milestone after formally commencing project development and integration in May 2019 was the preparation of a functional specification for the complete PCS system and development of the CIMIS interconnection module in the supplier's test environment. Due to challenges with the initial project kick-off and supplier team organization and cohesion, minor delays in hardware and integration services delivery, shifts in the focus of different stakeholders (concessionaires, customs), and the fact that the national CIMIS system was not ready for bidirectional communication, it was clear that a change in the project plan would be required. As a result, the service provider was unable to provide a complete functional specification for the system. The first appendix [12] to the original contract was stipulated in May 2020, formalizing the necessary adjustments. Functional requirements for individual modules shall be given simultaneously with module acceptances, according to this annex. The hardware was delivered and installed at the actual locations in the summer of 2020.

There was still no bidirectional connectivity with the national CIMIS system by the middle of 2020, which was a requirement for future integration. While some concessionaires have begun internal studies to assure conformity with the PCS, others have not. For the integrated modules, the provider has begun offering maintenance and helpdesk services. There was an open opportunity to extend the project for one more year, so late in 2020, a second annex [13] to the original contract was stipulated, and the financial value was increased by 17% due to the operational costs of the system (licensing, data links, and other costs) for one more year of development, as well as the need to develop three additional software modules, The year 2021 saw significant developments, and various modules were completed during that time, including the access control module (D7). After the Ministry's development services provider delivered the new version of the national CIMIS system to production in mid-2021, with bidirectional communication capabilities, related module D1 was completed. Customs-related modules (NCTS, ECS/ICS), as well as the planning and control portion of the D6 module for Luka Rijeka j.s.c concessionaire, were also released into production exploitation.

Further delays have been caused by the late development of interconnected systems by other parties, as well as the lack of concessionaires' availability for testing and development. To address this situation, the third annex [14] to the original contract was stipulated at the end of 2021, and with the approval of the INEA agency, the project execution was extended for one more year, until the end of 2022, and itemized prices were negotiated with the supplier in order to ensure development services coverage during that period.

The PCS system was placed into production in January 2022, with Port Control Center (LKC) [15] being the first user of the system in production mode and plans to immediately roll out modules D6 and D7 for concessionaire Luka Rijeka j.s.c.'s planning and coordination. Furthermore, project technical assistance team has created a set of public procurement documents for future maintenance services that was published on the Republic of Croatia's public procurement web early in 2022 [16]. This marks the official commencement of the national PCS's production operation.

IV. PROJECT RISK MANAGEMENT

As a part of bidding documentation, only a few risks were anticipated, thus it was suggested that the supplier provides and explains the model for managing any negative project impacts that it expects to encounter during the project's execution, and that the client's acceptance be considered. By managing potential adverse effects during project implementation, the supplier should have ensured timely detection and effective management of potentially critical external or internal impacts and events, as well as propose effective preventive methodology to eliminate, transfer, or reduce the level of impact of adverse effects on the PCS development project. The technique of communicating with the customer, harmonization of diagnostics, degrees of risk, and harmonization when planning the treatment of undesirable effects are all part of the process of controlling probable adverse effects on particular project components.

Initial potential negative impact identification, as well as the likelihood of occurrence, impact on the project, and strategies to remove, transfer, or decrease the degree and severity of the negative impact have been identified. The risks were classified into four categories: 1) strategic, 2) operational, 3) technical, and 4) regulatory risks. This internal risk registry was eventually utilized as a source of data for the ongoing risk management.

The purpose of mapping identified risks was to make a connection between the PCS implementation project's risk management approach and enterprise risk management goals by emphasizing positive occurrences while avoiding or controlling negative ones [17]. Initial risk management matrix employed a proprietary enterprise project management risk registry. All risks identified and controlled during the project execution until the production work of the first module are processed using the same framework.

Initially, the project risk management strategy was based on standard enterprise risk management methodology as defined by the PMBoK (Project

Management Body of Knowledge) [18], with the goal of identifying not just hazards but also opportunities during project execution. Traditional project management approaches group risk sources according to different criteria with a focus on the risks' inherent qualities (regulatory, technological) and levels on which they exert their effect, without clearly delineating the difference between internal and external forces affecting project risk (strategic, operational). This method is plainly inadequate because it does not adequately depict risks that may occur with other stakeholders and their systems, as the PCS is a system that is connected to many other systems. These hazards that arise in the project's environment are frequently overlooked, yet they have a significant impact on the project execution. The following sources of risk in form of the residual risk were not identified, but had impact on the project until the completion:

1. In the production system, data flow and message exchange between PCS and MNSW CIMIS must be synchronized, as CIMIS does not provide for an up-to-date test environment.

2. As the functional definition and development are in the early phases, there is a possibility of not being able to provide modules D8 and D9 for concessionaire Port of Rijeka, j.s.c., until the determined end of the project.

3. There are multiple parallel project activities and module developments, putting pressure on the supplier's capabilities and endangering module completion and integration, as well as ultimate product quality. This inherent risk is also reduced by the project's prolongation.

4. Given that the comprehensive transition of module D3 to production necessitates additional comprehensive testing, the deadline for the transition of this module to production can be more easily adjusted to the new framework of the final contract extension, considering the concessionaire's and other stakeholders' availability.

5. The risk of module D10 development has been carried forward to the start of production. The project's extension allows for the implementation of the D10 module. It will be necessary to achieve an agreement on whether the concessionaire will allow D3 module production without CFS or insist on simultaneous production of both modules. Because the deadline for continued collaboration on the CFS module's integration with PCS, as well as the expected timeframe for its transfer to production, is uncertain whether the D10-CFS module will be implemented prior to the end of Q4 2022.

6. Customs office has not yet taken a decision on whether or not to keep control lists in the customs system, and the PCS continues to function under present regulations, which means that control lists previously established in the customs IT system will be manually inserted into the PCS system. These features were successfully evaluated as part of module D3's rigorous internal User Acceptance testing.

7. Customs services for NCTS and HRAIS are currently being developed. Although PCS can function without these services, MRN services cannot. They will work with the latest versions of NCTS5 [19] and HRAIS2 [20].

Therefore, it was apparent that the enterprise risk management approach did not clearly identify all barriers, drivers, and success factors of the PCS development.

V. CONCLUSION

Internal and external stakeholders were involved in the process of building a national PCS model in Croatia, including the subject Ministry, Port Authority, integration and development services provider, and technical assistance team, as well as concessionaires, customs, police, port captains, ship and cargo agents, port control centre, and various digital systems that will exchange data with the new PCS. Under the auspices of the CEF EU financing, a project was launched in Croatia in mid-2017 to begin the construction of a national model of Port Community System, which was initially deployed in the Port of Rijeka and began production work in early 2022. The Port Community System, a digital platform for exchanging cargo messages and data among the port cluster's many stakeholders, should be adopted in other Croatian cargo ports in the near future.

The project's execution was substantially influenced by the recognized and residual risk. While the traditional strategy was effective for managing and treating recognized risk, it was insufficient for reducing residual risk to a minimal feasible level. The fact that the final risk identification revealed a more thorough risk occurrence registry, as well as additional drivers, success factors, and obstacles, clearly demonstrates this fact. The supplier proposed project risk management, which was handled in the same way as the traditional enterprise risk management. The project risk was broken down into numerous areas, including strategic, operational, technological, and regulatory aspects, with risk likelihood, impact, and mitigation strategies identified. This approach is appropriate for projects and environments where categorical risk is high (for example, caused by the complexity of the used hardware, network, and system support), but it lacks refinement in situations where multiple stakeholders operate their own information systems that must be integrated and communicate with the PCS.

These environments are often external to the PCS and contain undisclosed disk categories that are difficult to prepare ahead of time, as those planning and administering the project typically have very limited visibility of these complicated external systems and much

less options for mitigation at the time of their occurrence. The use of the described approach in analysing and reengineering of additional processes (mostly commercial processes, not administrative processes) associated to the introduction of integral business information systems best suited for seaport clusters might lead to further study.

REFERENCES

[1] Port Authority Rijeka, European projects, CEF – POR2CORE – PCS [Online]. Available: https://www.portauthority.hr/europskiprojekti/975/.

[2] Republic of Croatia, Ministry of the Sea, transport and Infrastructure, CIMIS – Croatian Integrated Maritime Information System [Online]. Available: https://mmpi.gov.hr/more-86/vts-croatia-114/cimis-hrvatski-integriranipomorski-informacijski-sustav/16553.

[3] The European Union, European Commission, The Innovation and Networks Executive Agency (INEA) [Online]. Available: https://ec.europa.eu/inea/en/welcome-to-innovation-networksexecutive-agency.

[4] The European Union, European Commission, Connecting Europe Facility, CEF Transport, Upgrade of the Rijeka Port infrastructure – Port Community System (POR2CORE-PCS) [Online]. Available: https://ec.europa.eu/inea/en/connectingeurope-facility/cef-transport/2016-hr-tmc-0082-s.

[5] Republic of Croatia, Ministry of the Sea, transport and Infrastructure [Online]. Available: https://mmpi.gov.hr/en.

[6] World Customs Organization, Building a single window environment, Compendium V.1, 2020 [Online]. Available: http://www.wcoomd.org//media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/tools/single-window/compendium/swcompendiumvol1all-parts.pdf.

[7] Kapidani, N., Tijan, E., Jović, M. and Kočan, E. (2020) "National Maritime Single Window – Cost-Benefit Analysis of Montenegro Case Study", Promet – Traffic&Transportation, 32(4), pp. 543-557. doi: 10.7307/ptt.v32i4.3422.

[8] Interreg Italy-Croatia Transpogood project, Drivers and Barriers identification, Final Version of 20/09/2019., Deliverable Number: D.5.1.1. [Online]. Available: https://www.italycroatia.eu/documents/120312/1577990/D5.1.1.+Drivers+and+Barriers+identification.pdf/f0579b91-b7ba-32af-01bdcce7dd67caa7?t=1591009546900

[9] Dr. Nils Meyer-Larsen, Project Manager, The Institute of Shipping Economics and Logistics (ISL), Better Cybersecurity for Port Community Systems, Edition 76: winter 2017., PTI Edition 76: Cyber Risk & Security [Online]. Available: https://www.porttechnology.org/editions/cyber_risk_security/

[10] Republic of Croatia, Ministry of Finance, Customs Administration, e-Customs [Online]. Available: https://carina.gov.hr/e-customs/6678

[11] Actual I.T. d.d., Funkcionalna specifikacija: PORT2CORE PCS ICT sustav, version 8.0, Port Community System 2020., 2021.

[12] First annex to the contract nr. 1 11-18 L.U. for delivery and installation of the PCS system, Rijeka, Croatia, 4th May 2020.

[13] Second annex to the contract nr. 1 11-18 L.U. for delivery and installation of the PCS system, 26th November 2020., Rijeka, Croatia

[14] Third annex to the contract nr. 1 11-18 L.U. for delivery and installation of the PCS system, 29th December 2021., Rijeka, Croatia

[15] Port Authority Rijeka, Port Control Centre [Online]. Available: https://www.portauthority.hr/en/rijeka-traffic-24-h-port-control centre/

[16] The public procurement portal of the Republic of Croatia „Narodne novine", Port Authority Rijeka, Delivery and installation of PCS systems [Online]. Available: https://eojn.nn.hr/SPIN/APPLICATION/IPN/DocumentManagement/DokumentPodaciFrm.aspx?id=5912617

[17] Gregg Witte, Huntington Ingalls Industries Inc., „Implementing an enterprise risk management framework", 12.10.2021., TechTarget [Online]. Available: https://www.techtarget.com/searchcio/feature/Implementing-an enterprise-risk-management-framework

[18] Goodman, R. (2005). The ascent of risk: risk and the Guide to the project management body of knowledge (PMBOK guide), 1987 1996-2000-2004. Paper presented at PMI® Global Congress 2005—Asia Pacific, Singapore. Newtown Square, PA: Project Management Institute [Online]. Available: https://www.pmi.org/learning/library/ascent-risk-pmbokguide7618

[19] Republic of Croatia, Ministry of Finance, Customs Administration, NCTS phase 5 [Online]. Available: https://carina.gov.hr/istaknute-teme/e-carina/ncts-faza-5/9845

[20] Republic of Croatia, Ministry of Finance, Customs Administration, HRAIS2 [Online]. Available: https://carina.gov.hr/istaknute-teme/e-carina/hrais2/8267

E-Accounting System as a Service (E-ASaaS): A Conceptual Overview

N. Vlahović
Faculty of Economics and Business University of Zagreb, Croatia
nvlahovic@efzg.hr

***Abstract* - In this paper we present an innovative concept of a public information system based on several digital technologies and approaches for outsourcing accounting as a public service. Accounting is one of the fundamental business functions of any business entity, that is highly regulated in all the economic systems. Over the recent decades accounting function in business organizations has evolved due to dramatic changes and advancements in information and communication technologies. These technologies-initiated changes in form of digital documentation, online or cloud accounting and similar alternatives to traditional organization of accounting. The goal of this paper is to draw attention to available technologies that may provide even more drastic innovation in organizing accounting function of business entities in a common administrative space such as counties, or entire countries and nation-wide economics. Here we propose a model of a public access E-accounting system that employs the functionality of cloud-based services, data pod concepts and blockchain technologies in order to automate processes of bookkeeping but also reporting on the side of business entities, while at the same time provide autonomous control and audit as defined by law and competent institutions in various segments of supervision (i.e. tax supervision, statistics reporting, etc).**

Keywords* - *Accounting Information Systems (AIS), Accounting, Public Electronic Services, Blockchain, Data pods, Cloud Computing

I. INTRODUCTION

Accounting as the basic of financial reporting is the key element of any economic system in contemporary societies. It is the first business function that was supported through the means of electronic computers, and the gateway to the rise of information and communication technologies (ICT). Over the past decades overall process of digitalization has changed many aspects within the accounting function as well as various related aspects in the business environment. In order to cater for these changes Accounting Information systems (AIS) have evolved and transformed into crucial part of the overall business information systems (BIS) used by both small and medium companies as well as large domestic and international corporations. Yet, beside the initial innovation functions of the average AIS remain the same in their core. At the same time new technologies have been developed that enable new capabilities that core function of the AIS may transform further to achieve new goals and functions that were not attainable before the raise of digitalization and digital economy.

The intention of this paper is to analyze, select and present several key technologies that hold the potential for transformation of AIS. Using the electronic services as the fundamental building block for AIS instead of traditional software, transition from software solutions to service-oriented AIS is available. Electronic services can also provide additional functionalities and capabilities for more effective, efficient, and more comprehensive outlook on accounting and business transactions that take place in the digital economy.

The goal of this paper is to present a conceptual model of a public service that is inspired by currently available electronic services provided through the concept of electronic government. The concept of electronic accounting system-as-a-service (i.e. e-ASaaS) is based on the principle of regarding transactions in digital economy based on digital documents, digital payments, and digital exchange not only of information and finances but in certain situations of goods and services. The system itself relies on digitalization of public sector records that are the legislative basis for any economic activity in any economy. Only by taking the nature of digitalization into account a system with additional information controls and automation capabilities that provide higher level of trust in financial reporting, contracting, investment and business partnerships can evolve.

The main contribution of this paper is to outline three most important technologies that may enable service-based AIS provided by the public sector based on public data records and existing infrastructure. Additional contribution is in the definition of key structural elements of the proposed conceptual model and the SWOT analysis of the proposed system.

The structure of the remainder of the paper is as follows: In Section II current state of financial accounting and financial reporting is discussed as well as fundamental principles and standards that shape the current implementations of AIS. In the following Section key information and communication technologies are described, outlining the most important aspects for the e-ASaaS model. In Section IV description of the conceptual model e-Accounting System-as-a-Service is given. In Section V a discussion of the described concept is given, while examining the feasibility and potential of the system through the SWOT analysis. Finally, in the last Section conclusions on potential benefits and possible challenges are given as well as remarks for the future work.

II. DIGITALISATION AND FINANCIAL ACCOUNTING AS BUSINESS FUNCTION

Financial accounting is an essential part of business reporting in enterprises. The accounting function is a structured layered system that is dedicated to providing information relevant to decision makers and more specifically, to equity investors, lenders and other creditors in their capacity as capital providers. For this purpose, key elements that are described with accounting information are (1) Assets, (2) Liabilities, (3) Equity, (4) Income and (5) Expenses. In order to preserve high quality of accounting information a rigid set of assumptions, principles and constraints need to be followed. For additional assurance various standards exist either on a national level or international level (such as International Financial Reporting Standards – IFRS).

With the development of Information and Communication Technologies (ICT), first digitalized systems of keeping accounts were introduced. Over the years, they have evolved into a comprehensible integrated information systems where the core of these business information systems (BIS) is the accounting information system (AIS). The impact of AIS on accounting functions in undeniable and comprehensive as they in turn shaped the organizational aspects of the accounting profession [9, 10].

Further digitalization of business entities and governmental bodies have also allowed for digitalization of business transaction elements beyond the boundaries of one business entity. This was made possible by establishing, prescribing and regulating standards in electronic document interchange (EDI) usually using XML schemes and other standardized formats, protocols of interaction and various security elements. One such example in Croatia is the introduction of digitalized infrastructure for electronic business that is used for creation of digital documents such as e-Invoices.

In turn this means that AIS had to adapt to these changes not only to accommodate for new types of business documentation but also to take advantage of digitalization and improve their own effectiveness and performance [5, 8].

Certain aspects of AIS or various subsidiary ledgers have been created so that they allow for automatized entry of transactions as the documentation arrives in real time or as business processes are being executed. The efficiency of this type of systems has also influenced the quality of financial information processes by these systems, but also introduced additional controls to detect mistakes and even fraudulent activities [6].

The increased functionality of AIS that improves quality of financial information also comes at a cost in terms of changes to organizational support to accommodate these new functionalities and also, requirements for the accounting professionals in terms of required knowledge and skills [7].

In order to accommodate for convenience of using such complex systems an alternative to traditional integrative software solutions (such as SAP S/4 HANA, or MS Dynamics NAV), service-based solutions (such as Oracle ERP cloud) are becoming increasingly popular ultimately replacing native application approach.

Based on this trend and the fact that regulators (being a part of the public sector) already record important regulatory, statistical, and financial data, a new approach to organization of the overall accounting function lends itself to the idea of public electronic service. In order to take advantage of existing infrastructure service-based solution should define standards and protocols that will certify adequate security and efficiency levels that are currently available in commercial cloud-based solutions. In the following Section we will present several key technologies that can be used to achieve this requirement.

III. OVERVIEW OF RELEVANT KEY INFORMATION AND COMMUNICATION TECHNOLOGIES

There are several ICT solutions that hold a great potential for innovation of existing AIS. Some of these include blockchain technology, cloud computing and data pods. All of these technologies resolve a specific requirement for using existing Internet infrastructure and available public resources to support businesses in meeting their legal obligation for establishing Accounting and Financial reporting function.

In the rest of this Section, for each relevant key technology that e-ASaaS concept is based on, a brief description is given, underlining their advantages and disadvantages.

A. Cloud computing

Over the past decade cloud computing has emerged and established itself as one of the most prominent architectures for providing electronic services [11, 12]. Its influence has been so significant that cloud computing challenged software developers by disrupting the traditional business models both their own and those of businesses running their software. Software industry accepted this transformation shift from local and native software solutions to cloud solutions and e-services. Today, electronic services and cloud-based software solutions are one of the most important segments in software industry.

Unlike Grid computing, cloud computing enabled service-based business models for a whole new range of services that were not traditionally deliverable in digital form. Today dependency on cloud solutions is high enough so that cloud computing faces new challenges in sustaining its security solution, privacy of information, continuity managements etc.

Main benefits of cloud computing such as resource sharing, flexibility, scalability and security as well as current trends in software industry favor the solutions based on cloud computing platforms. These solutions generally perform better in terms of cost efficiency and recovery flexibility in comparison to traditional software solutions.

Cloud computing is defined as a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers,

storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [13]. As such, Cloud computing established a layered architecture that also defines the level of service provided. Each tier can be controlled cumulatively either by the service provider or the service user depending on the scope of control and disposition of resources required by the service user.

e-ASaaS model upgrades the organization of the top two tiers of cloud computing (i.e. data tier and application tier), upgrading them using additional solutions based on technologies described in the remainder of this Section.

B. Blockchain technologies

Blockchain technology has been popularized through the raise of cryptocurrencies since 2009 when first Bitcoin was established. Technology behind the Bitcoin was based on peer-to-peer transaction system where intermediary was removed by recording transactions in a distributed data system with built in cryptographic verification used to establish encryption and digital signatures [1]. The data records, organized as blocks with separate timestamps, are chronologically linked into a chain where additional blocks prevent changes of data along the chain [3]. Soon after the initial cryptocurrency many others were founded using the same technology. The intrinsic permanent and secure record of transactions than can be easily verified using the blockchain soon gain popularity as a secure and trustworthy alternative payment method without intermediaries which led to the popularization of benefits and advantages of blockchain model.

Blockchain can be described as a distributed data ledger that is used to permanently store digital transaction data, that cannot be modified or removed from the records, and where high level of security and privacy of the data is achieved using public key cryptography and unmediated access to data. As a distributed record system a network of participating nodes is established where each node maintains a version of the blockchain, either full or partial. There are different consensus mechanisms used to maintain agreement on the true state of the blockchain content. These mechanisms provide incentive through rewards for nodes that maintain genuine record and prevent fraudulent transactions.

Public data records can benefit from using blockchain technology as founding technology for various transaction ledgers or master records and has been proposed as feasible solution in a number of incentives in e-government projects [2, 4].

Possible drawbacks usually pertain to establishing more efficient and environmentally friendly consensus mechanism as mechanisms implemented for cryptocurrencies require high energy consumption and are time consuming. There are suitable alternatives that can be used in e-government implementations [14].

C. Data pods

Data pods, are secured dedicated data storage locations for private data in distributed web networks [15]. They pertain to reorganization of the commercial logic in contemporary Web solutions where private data is treated simply as the result of data processing and service delivery that commercial exploiter and other third-parties use to generate profit. The intent of the private data pods is the reversal of the traditional digital commerce business model and economic value chain, by putting an asset value on the individual's personal data that is of direct financial and other benefit to that individual [15].

Main idea of data pods is that the owner of data contained in the pod can grant access to third parties to a selection of data contained within the data pod in a flexible manner with the option to withdraw at any time. In this way privacy and confidentiality is preserved while empowering the owner of private data to negotiate the use of their information even on commercial terms. Pods can be defined as decentralized data stores that function as secure personal web servers of data [16].

Currently, the idea of data pods and the implementations of data pods are being explored under the more general project Solid which is envisioned as an alternative view of Web 3.0 with redefined commercial roles [16].

For the purpose of providing public electronic services, data pods can be used to store private and confidential data on the public sector infrastructure, where each data pod is associated with a particular user (person or business) using national identification and authentication system (NIAS). Each user can have multiple data pods which is one of the assumptions of the conceptual model of e-ASaaS that will be explained in detail in the following Section.

IV. CONCEPTUAL MODEL OF E-ACCOUNTING SYSTEM AS A SERVICE (E-ASAAS)

The proposed model takes advantage of existing standards and technologies described in the previous sections of the paper so that it further refines the SaaS tier of cloud computing, improving the value of visibility to the end user with additional automatization options that heavily rely on secure and permanent data records and access to constituting electronic services (Figure 1)

The concept further expands elements in the Data and Application level of SaaS cloud architecture. Data level is organized using data pods for various types of data that may contain private and confidential entity information. A part of this data involves public data records that government bodies already have on record and can be accessed on their origin using data pod permissions. The accounting transactions are stored as a portion of the data layer as they contain blockchain ledger of transactions that were used in recording financial transactions of the entity. This blockchain ledger is distributed across the cloud that contains full nodes and master nodes.

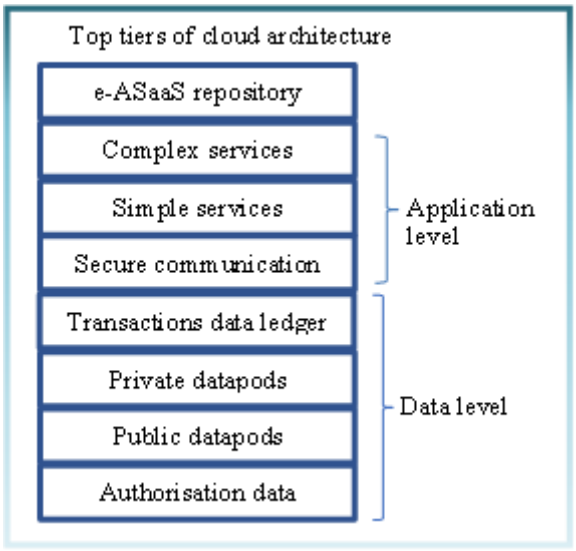


Figure 1. Elements of the e-ASaaS concept model within the cloud architecture

In the transaction data ledger layer, a separate blockchain ledger is kept recording access to the financial information in the general ledger of the entity. A separate blockchain logs data on access to the financial reports of the entity where various regulators and agencies have usually temporary or limited legal authorization for access. Here majority of blockchain consists of partial nodes containing only hash values with references to data that is kept usually by the entity in their separate data record systems or e-archive of the entity.

Finally, application level of the architecture contains several elements that pertain to establishing secure communication between entities, and permission for accessing simple and complex services that are available as basic building blocks of the accounting information system in the form of business logic. In Figure 2, we can see fundamental elements that this business logic contains.

The organization of data and functions is coupled from various data pods and existing services within the public sector cloud. While majority of the e-ASaaS system is contained in the public sector cloud, some services originate in connected information systems such as e-Archive systems, civil registries etc.

Each user entity is granted permanent access to the system where they can access and manage their access policies that regulate access of their employees to General ledger data on one hand, and access to their financial reporting data that can be granted to various agencies.

The general ledger is based on the blockchain and covers five basic elements of the financial reporting framework. Each blockchain transaction contains complete posting information from the journal that are validated for correct credit and debit balance before they are recorded in the general ledger. The chart of accounts and posting models for various typical business transactions are available in the form of business logic within the e-ASaaS system. They are aligned with updated and current legal regulation as it becomes available. Posting procedure itself can be semi-automated (requiring verification by person responsible) or manual using the

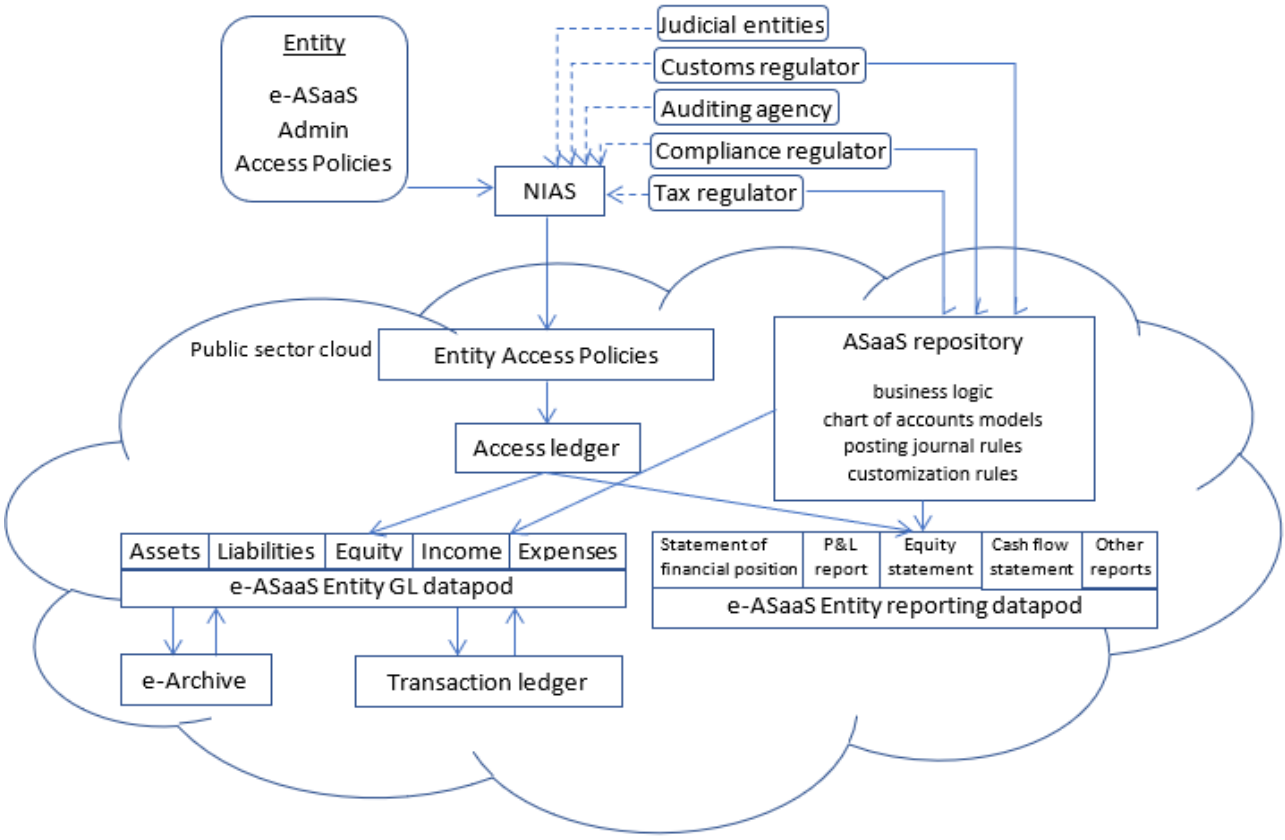


Figure 2. Overview of the e-ASaaS concept model structure

services for creating record provided by the systems business logic.

One example is the use of e-Invoices where the procedure for incoming invoices can be automated when the trading partner is recorded by the system.

Basic financial reports can be generated based on the information contained in the General Ledger and transferred to the reporting data pod of the entity. Each authorized regulator or agency can access data in this data pod and reassemble the data in the form that they prescribe for their intended purpose.

V. DISCUSSION

Public sector in most developed and developing countries has established a level of information infrastructure that can provide higher level of e-services. These e-services can cater for more complex and comprehensive needs of citizens and businesses. There is an opportunity to offer new e-services that are not traditionally offered through this platform. One such service is the proposed conceptual model of electronic accounting information system as a digital service, e-ASaaS. A more detailed overview of proposed concept characteristics is provided through the SWOT analysis (Figure 3).



Figure 3. SWOT analysis matrix for the e-ASaaS conceptual model

A. *Strenghts of the e-ASaaS concept*

Keeping financial books and records is obligation of business entities prescribed by law, and as such proactive approach of e-government can support businesses by providing a public service to fulfil their obligation.

Business entities that would use e-ASaaS system will always use the updated most recent legislative solutions that prescribe various changes in regulations of any aspects of bookkeeping and accounting as they are available with the business logic of the system itself.

There is a potential of instant access to current insight into statistical results of national industry and GDP by government agencies, tax collection, customs, and import and export indicators by government agencies.

If public-private partnership is established, a new service market can be initiated potentially creating additional impulse to business activity as it was the case of previous infrastructural digitalization projects (such as NIAS, e-Invoicing, etc…)

B. *Weaknesses of the e-ASaaS concept*

In case of technical failure or failure of the system there would be a temporary delay in providing this service which may disrupt regular business especially in some crucial moments (i.e., payroll accounting for example). In order to overcome this issue contingency plans, need to be established.

Various supplementary records have a high variety and dependencies on organization circumstances of each entity. It is expected that standardization to provide various requirements and needs of different business entities may prove to be very hard or even unattainable. These records however need to be included in the general ledger at some point and if they are not included or connected to the e-ASaaS system they would have to be included manually or imported from external source.

A rule-based system that assures the compliance with current regulation may have the effect of limiting the options of accounting practices that would otherwise take advantage of loopholes in standards and regulation. This can be considered a weakness of the system because the “creative accounting” practices are indicators for the regulator of the quality of accounting standards and laws, and more generally tax regulation and other governmental policies, driving political and governmental agencies to improve the legal framework.

C. *Opportunities provided by the e-ASaaS concept*

The e-ASaaS provides valuable functionalities that can be further connected with other relevant entities such as commercial banks that have high level of digitalization. Better overall integration of the nationwide economy can be promoted.

There are several factors that could impact higher volume in international trade. Better competitiveness of the domestic businesses, readily available business information, incentive for international entrepreneurs to invest in national business incentive are just a few to name. All of these incentives can eventually further promote business activities.

D. *Threats that may deminish benefits of the e-ASaaS concept*

As continuity of business entities is one of the most important assumptions for accounting, there are several possible threats that could jeopardize confidentiality of entity data. One such development is the advancements in the field of quantum computer where current cryptographical solutions for secured communication would become obsolete and privacy of data would be disrupted.

The distributed architecture of the system may provide high level of privacy and security but may also increase the effects of the system on ecology in terms of power

consumption and similar cost inefficient effects on public spending.

There is still a lack of standards for particular elements of the underlaying technologies that need to be established before implementation of all of the system elements. Further investment on behalf of academic and commercial sector is required in order to overcome these challenges and establish a viable, sustainable and effective organization of overall system.

In the long run, automation in accounting practices may result in changes in demand for experts in the field of accounting. Demand for less specialized and traditional positions may significantly decrease, forcing employees to seek specialization for more complex accounting roles. There is also possibility of increased demand for higher expertise in the field of accounting and related areas that offer better job conditions but also higher responsibilities (such as managerial accounting positions, consultant positions, audit positions etc.). In each case, accounting profession will require updated and more comprehensive IT skills.

VI. CONCLUSION

In this paper we presented an innovative concept of a public information system for outsourcing accounting as a public service. The model is based on available public sector infrastructure and takes advantage of a few rising digital technologies The main premise is that the existing public data, resources, infrastructure, and other public assets can be used to provide novel, high complexity, comprehensive services.

The goal of the paper was to present a conceptual model of the proposed system that may improve cost effectiveness to business entities while achieving additional goals of public interest for all social stakeholders. We presented key characteristics of cloud computing, blockchain and data pods that make the fundamental elements of proposed concept.

The model incorporates new technologies in current cloud computing layered model that enable system to automate various accounting tasks while maintaining secure, private, confidential, and permanent record of business transactions. At the same time the system can provide means of autonomous control and audit as defined by law and competent institutions in various segments of supervision (i.e. tax supervision, statistics reporting, etc).

In order to consider and analyze potential of the proposed model and functionalities blockchain technology and data pod specification bring to the public sector a SWOT analysis is performed. There are significant strengths of the proposed model such as the potential to redefine and re-establish entrepreneurial environment and provide competitive advantage to businesses on a national level. Yet, there is a number of challenges that need to be addressed before implementation of this system to minimize potential security and downtime risks.

REFERENCES

[1] Elisa, N., Yang, L., Li, H., Chao, F., Naik, N. (2019) Consortium blockchain for security and privacy-preserving in E-government Systems. In Proceedings of the International Conference on Electronic Business (ICEB), Newcastle Upon Tyne, UK, 8–12 December 2019; pp. 99–107.

[2] Batubara, F.R., Ubacht, J., Janssen, M. (2019) Unraveling Transparency and Accountability in Blockchain. In Proceedings of the 20th Annual International Conference on Digital Government Research, DG.O 2019, Dubai, United Arab Emirates, 18–20 June 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 204–213. DOI: 10.1145/3325112.3325262

[3] Casino, F., Dasaklis, T. K., Patsakis, C. (2019) A systematic literature review of blockchain-based applications: Current status, classification and open issues, Telematics and Informatics, Volume 36, 2019, Pages 55-81, https://doi.org/10.1016/j.tele.2018.11.006.

[4] Swan, M. (2015) Blockchain Blueprint for a New Economy. O'Reilly Media Inc., Sebastopol, CA, USA.

[5] Bredmar, K., Ask, U., Frisk, E., Magnusson, J. (2014), "Accounting Information Systems Implementation and Management Accounting Change", Business Systems Research, Vol. 5, No. 2, pp. 125-138, DOI: 10.2478/bsrj-2014-0015

[6] Boylan, D. H., Hull, J. E.(2022) "Have Accounting Information Systems Significantly Helped in Detecting Fraudulent Activities in Accounting?", The Journal of Applied Business and Economics; Vol. 24, Iss. 3, Thunder Bay, pp. 45-56.

[7] Al Nasrallah,W.; Saleem, F. (2022) "Determinants of the Digitalization of Accounting in an Emerging Market: The Roles of Organizational Support and Job Relevance." Sustainability, No. 14, 6483. https://doi.org/10.3390/su14116483.

[8] O'Donnell, J. B. (2019). "Are Accounting Information Systems Programs Evolving to Meet the Needs of the Accounting Profession? An Analysis of Accounting Information Systems Programs in 2005 and 2019". Journal of Applied Business and Economics, 21(8). https://doi.org/10.33423/jabe.v21i8.2591

[9] Mardi, M., Perdana, P. N., Suparno, S., & Munandar, I. A. (2020). Effect of Accounting Information Systems, Teamwork, and Internal Control on Financial Reporting Timeliness. The Journal of Asian Finance, Economics and Business, 7(12), 809–818. https://doi.org/10.13106/JAFEB.2020.VOL7.NO12.809

[10] Awosejo, OJ; Kekwaletswe, R M; Pretorius, P; Zuva, T (2013) "The Effect of Accounting Information Systems in Accounting". International Journal of Advanced Computer Research, Vol. 3, Iss. 3. Accent Social and Welfare Society, pp. 142-150.

[11] Cheng, M., Qu, Y., Jiang, Ch., Zhao, Ch. (2022) "Is cloud computing the digital solution to the future of banking?", Journal of Financial Stability, Volume 63, 101073,ISSN 1572-3089, https://doi.org/10.1016/j.jfs.2022.101073.

[12] Hassan, A., Bhatti, S. H., Shujaat, S., Hwang, Y. (2022) "To adopt or not to adopt? The determinants of cloud computing adoption in information technology sector", Decision Analytics Journal, Volume 5, 100138, ISSN 2772-6622, https://doi.org/10.1016/j.dajour.2022.100138.

[13] Zhang, Q., Cheng, L. & Boutaba, R. Cloud computing: state-of-the-art and research challenges. J Internet Serv Appl 1, 7–18 (2010). https://doi.org/10.1007/s13174-010-0007-6

[14] Lykidis, I., Drosatos, G., Rantos, K. (2021). The Use of Blockchain Technology in e-Government Services. Computers 2021, 10, 168.

[15] Castell S. (2022) "Private Datapods: Web 3.0 does not Automatically Mean Blockchain Decentralization". Studies in Social Science Research, Vol. 3., No. 2. Scholink inc., pp. 90-118

[16] Solid, Availabe at https://solidproject.org/about [02-02-2023]

A Preliminary Survey into the Use of Business Process Tools in Croatian Companies

Ljubica Milanović Glavan
Faculty of Economic and Business, University of Zagreb, Croatia
ljmilanovic@efzg.hr

***Abstract* - This paper presents the analysis of process-oriented information technology. Process information technology can be defined as all information technology that a company uses in order to reach the highest possible level of process orientation. It includes information technology for managing business processes and information technology for performance measurement. The paper identifies and describes process information technology and it examines the influence based on actual research done in Croatian businesses, information technology has on levels of business process orientation. Information technology has a favourable effect on process orientation, according to the research. It has been established that information technology is the least valued aspect of process orientation and that corporate informatization in the Republic of Croatia is still insufficient. The adoption and usage of process information technology in Croatian enterprises are crucial, according to the aforementioned data.**

Key words - process-oriented information technology; process management; Croatia; empirical research.

I. INTRODUCTION

The analysis of information technology is the foundation of this study. Every company is defined with different business processes that describe the way the company operates. A series of tasks that make up a business process are logically related and utilize firm resources with the overarching objective of meeting customer demands for goods and services of a suitable quality and cost in a sufficient amount of time while also generating some value [1]. Business processes gather data about the situation and requirements of the company's clients and employees, then process that data in order to take actions that can meet those requirements. They are, in essence, the brains and nervous system of every business [2], so managing them is crucial. There are several important components to business process management (BPM). This idea is understood in science and practice in a variety of ways, from a narrow focus on information technology to a comprehensive view of process management. Business process management initially placed a strong emphasis on the technical, i.e., the application of information technology in business processes and process design [2]. It is only in recent years that researchers have begun to look at using business process management as a comprehensive strategy that goes beyond information technology. A comprehensive picture of business process management is presented by [3]. They view business process management as a novel way of conducting business in an organization in contrast to the conventional functional, hierarchical method. The organizational components of process management are included in a holistic approach, which views processes as the central means and such operations are carried out as long as they are supported by the staff of the organization. That is why business process management integrates a managerial strategy with the right technology.

Organizations today are compelled to connect their business processes via functional units as a result. Organizations employ a process-oriented approach to managing their activities as a result of the requirement for such cross-functional connections. Process orientation (also known as BPO) has several meanings in the literature, but one that claims that an organization's process orientation is determined by how much attention it gives to its primary business processes may be the most accurate [4]. Process orientation in business refers to efforts made by organizations to build organizational structure and strategic planning upon business processes [2].; it emphasizes processes as the antithesis of hierarchy in the organization and it increases organizational efficiency by coordinating organizational activities in a system that is based on overall processes. Process orientation is very closely related to information technology and informatization of business. A suitable degree of orientation of business processes degree is a requirement for informatization of business processes. Informatizing chaotic, unconnected, and uncoordinated corporate processes is pointless. In contrast, information technology supports and encourages many procedures to be performed [5].

The research questions that arise are: what types of technology belong to process information technology and what is its influence on the business process orientation in a company. The first chapter defines and describes the concepts essential for this topic. The second chapter defines information technology, classifies it and emphasizes its importance. In the third chapter, the informatization of the Croatian companies is empirically examined and the information technology's effect on process orientation is tested. Finally, the fourth chapter gives concluding remarks of this paper, explains the scientific contribution and shortcomings of the completed research, as well as providing recommendations for future research.

II. PROCESS INFORMATION TECHNOLOGY

Information technology (IT) can be used to develop process orientation, according to many writers [6;7]. According to [8] IT can affect business processes in 3 complimentary, but different ways. IT adds worth to the entire organization through these impacts on business operations. First, automated impacts refer to value produced from IT's position as a capital asset that replaces labour from the efficiency perspective. Value within this dimension primarily comes from effects like increased productivity, labour cost savings, and cost reduction. Second, IT's capacity to of information collecting, storing and processing is the main source of information effects. Following these outcomes, adding value entails enhancing decision quality, empowering individuals, conserving resources, enhancing organizational effectiveness, and enhancing quality. Third, transformational effects describe the benefits derived from IT's ability to support transformation and innovation of processes. A shorter time period and improved goods and services are how the business value related to these outcomes is seen. According to [7] using information technology to support business process execution has many advantages. He claims that information technology makes it possible to automate business processes, shorten their duration, collect, store, and deliver information more easily through the creation of business process repositories, supervise and monitor business processes, make interventions in real-time, analyse data, connect business process activities, and transfer information more quickly between geographically distant areas. Better process execution is made possible by information technology, but it does not by itself bring value. However, the administration of business processes is a required and inevitable support for information technology. Information technology and processes are closely related, therefore process-oriented businesses must inevitably use technologies to manage their business operations automatically.

Information and communication technologies advanced greatly near the close of the 20th century as they started to be employed more and more in the workplace. This advancement, along with advances in computing power and transmission speed, made it possible to digitize business processes within the firm. The growth of technology brought about by process changes was sparked by the growing interest in enhancing corporate processes as well as technological advancements. This includes the usage of integrated information systems and BPM tools for process integration. Over the past 20 years, business process management systems have seen constant change. First, everything started with process automation, management of business processes and creation of software solutions for connecting different information systems within the company. At the end of the 1990s, the focus shifted to connecting with information systems of business partners, developing and connecting systems for electronic business, and connecting systems to supply chains. Such unification of a traditional information system for managing business flows, a system for interactive communication of groups of users, a system for BPM and information technology for measuring business success resulted in a new technology called process information technology [9].

Information and communication technologies advanced greatly near the close of the 20th century as they started to be employed more and more in the workplace. This advancement, along with advances in computing power and transmission speed, made it possible to digitize business processes within the firm. The growth of technology brought about by process changes was sparked by the growing interest in enhancing corporate processes as well as technological advancements. This includes the usage of integrated information systems and BPM tools for process integration. Over the past 20 years, business process management systems have seen constant change. First, everything started with process automation, management of business processes and creation of software solutions for connecting different information systems within the company. At the end of the 1990s, the focus shifted to connecting with information systems of business partners, developing and connecting systems for electronic business, and connecting systems to supply chains. Such unification of a traditional information system for managing business flows, a system for interactive communication of groups of users, a system for BPM and information technology for measuring business success resulted in a new technology called process information technology [9]. Process information technology can be defined as all information technology that a company uses in order to reach the highest possible level of process orientation. It includes information technology for managing business processes and information technology for performance measurement [6].

A. Information technology for business process management

According to [7] business process management from the aspect of information technology includes five components:

- Business processes are modelled using common notations, called process modelling;
- Process monitoring, which refers to the capability of determining how well a process is performing;
- Measuring the process's effectiveness: evaluating the process's effectiveness based on performance indicators;
- Management of business rules: the objective is to extract business rules from conventional software applications, store them, and manage them using independent components referred to as the business rule engine;
- Process engines are employed as parts of process-based applications and are information technology systems. They ensure that procedures are carried out in accordance with their instructions.

In the beginning, information technology was developed in the direction of support in modelling business processes. But business process modelling is only one phase of BPM life cycle. It is the basis of management because if there is no clearly defined model of how business is conducted, management is not possible

either. Information technology must support all stages of the BPM life cycle [10]. Therefore, software tools oriented to business processes can be divided into two basic groups [11]:

- Tools for modelling and analysing business processes;
- Tools for business processes management.

Business process management tools' primary goal is to automate and regulate the execution of business processes, whereas business process modelling and analysis tools' primary goal is to record, evaluate, and provide suggestions for bettering business processes. Tools for process modelling and BPM tools and have some fundamental differences and these are is that modelling tools are only a modelling environment, while business process management tools are an execution environment. However, these two categories of tools are complementary and it is necessary to combine them when implementing a business process change project. Although there are quite a few software tools available that fall outside of the two categories mentioned, it is not necessary to pay attention to these tools separately because they are frequently either a crucial component or a supplementary module of software tools from the aforementioned categories [12].

B. Information technology for business process managment

The literature has shown that many approaches to business performance measurement proposed in the last two decades emphasize the measurement dimensions and performance indicators that must be considered. In contrast, issues related to information technology that support performance measurement have been largely neglected. [13] points out that building a performance measurement system does not necessarily include information technology. However, for such a system to be successful, information technology is necessary. Information technology enables: monitoring of business operations, ongoing data gathering on business performance, comparison of historical, achieved, and target values of collected data, information extraction from primary data concerning new business trends or anomalies, and dissemination of obtained results to business participants. In other words, it can be said that information technology has various roles in performance measurement: collection, analysis, reporting and presentation roles. The question that arises is what kind of information technology can serve as support for methods and tools for measuring performance. The analysis of the literature shows that for the purpose of measuring the success of a company's operations, business intelligence tools [12] and integral information systems [14].

Software tools of business intelligence can be considered various application software supports created for reporting, analysis and presentation of data [15]. Business operations must be closely watched at all times. To determine the effectiveness of the processes that have been put in place, they must be measured. Tools for business intelligence can be employed to manage business operations. Starting with static process control, which keeps track of process stability using a wide range of dynamic tools like simulation, what-if analysis, and business cost analysis, it progresses to gauging how satisfied internal and external customers, process owners, employees, and suppliers are with the process. Future-focused businesses are starting to understand that gathering information on business processes can result in large operating cost reductions and quicker returns on technology expenditures.

One of the prerequisites for the success of a company's operations is the support of an integral information system (Enterprise Resource Planning, ERP), which is used to collect and distribute information throughout the organization, in order to connect different organizational parts and create a basis for quality decision-making [11]. An ERP system connects all organizational business operations, supports and combines the work of all services and functions, and encompasses the complete business model of the company, as well as external business processes related to business partners [15]. ERP uses a single database and performs the functions of a transactional and managerial information system and integrates elements of a decision support system. It enables document tracking and process level reporting, employee teamwork, and planning, monitoring, and use of all company resources. The basic idea of ERP is the connection between the organization and the process, which enables the recording of all business events so that the data can be used and analysed within the entire organization at the process level. The results of monitoring and measuring the success of business processes serve the management for comparison with competitors, for creating strategic maps of the company and for evaluating the effectiveness of the company's strategy's implementation. The ERP approach to automation and integration of business processes is based on the idea that many key business processes are similar in different organizations and an information system for automating such similar processes can therefore be developed. Business processes integration is achieved when nearly all business processes within an organization are performed within the boundaries of an integral information system. Practice has shown that ERP systems are imperative, but that an ERP system in a company is not enough by itself. According to [16] "for a successful ERP system, the customer is king, technology is the key, information is a competitive weapon, business process management software tools and business intelligence software tools lead the way." This means that CRM must be integrated within ERP, technologies must be improved, especially those related to electronic business, and timely and accurate information about what is happening inside and outside the organization must be available to ERP systems. However, without following two trends recognized in practice today [14], namely: integration of ERP and business intelligence tools; integration of ERP and business process management tools.

III. EMPIRICAL RESEARCH

Companies in Croatia must evaluate the practice of using process information technology in order to be able to become and remain competitive. Therefore, an empirical investigation into how process information

technology is used in large, medium-sized, and small businesses in the Republic of Croatia was conducted.

A. Methodological framework of the research

An electronic questionnaire served as the primary tool used in this study to gather primary data. The questionnaire used examined the process orientation of companies in the Republic of Croatia. Process information technology is one of the aspects of BPO. The practice of using such technology in Croatian companies was examined through this dimension, which contained a total of 11 questions. The other 8 dimensions of BPO are not relevant to the topic of this paper. Since in most cases the questions (statements) are of a closed type with the offered answers of intensity, a discrete 7-point Likert scale was used, where it applies: 1=not true at all,…, 7=completely true, x=don't know. The degree to which respondents approved or disapproved of the declarations was gauged on a scale of 1 to 7. The questionnaire and accompanying letter were sent to the electronic addresses of 1,200 large, medium and small companies in the Republic of Croatia, and information about them was obtained using the register of the Croatian Chamber of Commerce (the e-mail contacts that were available were used). The research sample consists of 127 companies, which means that the response rate is 10.58%.

To determine the general characteristics of the organization, questions about the main activity of the organization and the average number of employees were used. In order to demonstrate the heterogeneity of the sample, the volume of annual turnover was also observed. To determine the activities of the organizations that participated in the questionnaire in response to the question about their main activity, 19 options were offered according to the categorization of the State Bureau of Statistics. The largest number of companies from the sample are from other activities (16.53%), professional scientific and other technical activities (15.75%), followed by companies from processing activities (11.81%) and from trade and repair of motor vehicles (11.02%). Not a single company from education or cultural, entertainment and recreational activities participated in the survey, and one company from the sample did not respond. 33.9% of large, 34.6% of medium and 31.5% of small companies took part in the survey according to the number of employees. 35.7% of large, 27.7% of medium and 36.6% of small companies took part in the survey according to the criterion of annual turnover. As a result of the research's participation by businesses from various industries, the data sample adequately represented all business sectors. Furthermore, the sample is a good representative of the population of large, medium-sized, and small Croatian businesses.

A. Results and discussion

By analysing the responses of all companies from the sample, both the average ratings of the state of process orientation of Croatian companies and the average ratings of its individual dimensions were obtained. The investigation revealed that the worst-rated aspect of process orientation is information technology, which has an average score of 4.1827, which indicates that in the practice of Croatian companies, the informatization of business is insufficient. The issue is caused by the fact that information technology rather than business processes, is based primarily on business functions. Processes must be the foundation of business so that information technology may automate them. Companies gave the lowest ratings of this dimension precisely in the aspects related to the described process information technology: the use of software tools for documenting and analysing business processes (3.62), the use of software tools for managing business processes (3.44), the use of software business intelligence tools for measuring business performance (3.56) and the introduction and use of an integral information system (3.65). Furthermore, these parts (questions) have the lowest marks of all the parts of the entire questionnaire, which is devastating (Table I).

TABLE I. PROCESS INFORMATION TECHNOLOGY QUESTION GRADES

Process information technology	**Average grade**
Business informatization is based on processes (not on business functions)	4.88
Our information system provides the management with relevant information on business processes effectiveness.	4.88
Our information system is flexible and adjustable to business process change.	4.66
An organization uses special programme solutions for customer relationship management.	4.03
An organization uses some aspect of e-business.	3.86
An organization has document management system.	4.62
An organization uses specialised business process modelling and analysis programme tools.	3.62
An organization uses specialised business process management tools.	3.44
An organization applies business process management system.	3.56
An organization uses business intelligence programme tools for measuring business performance.	3.65
An organization applies an integral system for business process management.	4.25

In the future, the use of such tools will have to be intensified, because they allow managers a mechanism to efficiently coordinate all the human and technological resources required to carry out certain business processes, adding value to the firm by harmonizing its operations. Even among companies that have tools for modelling and

documenting business processes, 54.9% occasionally use them, compared to only 45.1% who use them consistently. The situation is no better with business process management tools. They are only utilized routinely by 48% of businesses, and sporadically or infrequently by 52% of businesses. Businesses most commonly employ ARIS Toolset (5 mentions), SharePoint (3 mentions), IBM BPM (2 mentions), and Microsoft Visio among the software tools for modelling and documenting business processes (2 cases). ARIS, MS Dynamics NAV, SAP BPM, and IBM WebSphere Business Integration Suite are the software programs that businesses most frequently employ to manage business processes (listed in four of the cases) (2 cases). Some companies give inadequate answers to questions about specialized software tools for modelling and documenting business processes and to questions about business process management tools they use, from which it can be concluded that these companies are not familiar with the concept and terms of business processes. As for business intelligence software tools for measuring business performance, 66.9% of the companies in the sample use tools for business reporting, 52.8% perform analytical data processing, 63% perform ad hoc queries and analyses, 59.1% statistical analysis and data mining, and only 32.3% alarm systems and early notification systems.

All of the above indicates that the implementation of process information technology in Croatian companies must become imperative. It is information technology that is mentioned in the literature as a factor that facilitates the restructuring of business in the direction of the process paradigm [14; 17], therefore, in the conducted research, we wanted to empirically test whether there really is a connection between information technology and the process orientation of the company. For this purpose, IT and BPO were the subjects of a correlation study. The p-value is equal to 0.000, and the Spearman correlation coefficient is 0.702 in value. According to these findings, Information technology adoption and business process orientation are statistically significantly and positively correlated. At the significance level of 1%, and according to Gulford's table, this relationship is on the cusp of being a moderate to high connection (Table II).

TABLE II. BPO AND IT CORRELATION

		IT	**BPO**
IT	Correlation	1	0,702
	P-value		0,000
	N	127	127
BPO	Correlation	0,702	1
	P-value	0,000	
	N	127	127

As stated previously, many authors in their works classify integral information systems and business intelligence software tools as elements of information technology for performance measurement, but none of them tried to prove it empirically. Given this fact and the fact that today many companies use business intelligence software tools, and all successful companies have integral information systems, the question arises whether business intelligence software tools and integral information systems truly support the management and design of business processes. The conducted research therefore tested whether the use of integral information systems is positively related to the process orientation of the organization and whether the use of business intelligence software tools is positively related to the process orientation of the organization. Spearman's correlation coefficient was used for testing. The use of integral information systems is positively correlated with the process orientation of the organization at a significance level of 1%, according to the Spearman coefficient between the variables of use of integral information systems and the variable of process orientation, which is 0.385 with a p-value of 0.000. The usage of business intelligence software tools is positively correlated with the organization's process orientation, according to the Spearman correlation between the variable use of business intelligence software tools and the variable process orientation, which is 0.495 with a p-value of 0.000. Based on the conducted statistical analysis and all of the above, it can be concluded that the use of information technology is positively related to the process orientation of the company. Given that the connection has been empirically proven, it was additionally wanted to examine whether there is also an influence of the use of information technology on the process orientation of the company. Regression analysis is used to examine the impact. The regression analysis in this case was carried out in such a way that process orientation was a dependent variable, and information technology was an independent variable. The performed analysis indicates that the regression model explained 46.8% of all deviations of the dependent variable. The magnitude of the Durbin-Watson test is 1.852, which means that there is no autocorrelation problem. Information technology as an independent variable is significant at the 1% significance level, and the calculated parameter's value of 0.441 suggests that information technology has a favourable effect on the company's process orientation, meaning that using information technology causes the company to have higher levels of process orientation (Table III).

TABLE III. THE INFLUENCE OF IT ON BPO

Coefficients [a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	p-value	95,0% confidence interval	
		$\hat{\beta}$	Standard error	$\hat{\beta}$			Lower bound	Upper bound
1	Constant	2,994	,186		16,072	,000	2,625	3,362
	IT	,441	,042	,684	10,490	,000	,358	,524

a. Dependent variable: BPO

IV. CONCLUSION

The review of the pertinent literature revealed how crucial it is for businesses to increase their process orientation by utilizing information technology. In order for such a claim not to remain only on a theoretical basis, the conducted research examined the connection of information technology with the process orientation of the company. The paper first defined and classified the so-called process information technology, and then its influence on process orientation was empirically tested. Information technology's effects on businesses' process orientation and the connection between the use of business intelligence tools and integral information systems with process orientation have been statistically proven. In other words, the importance of information technology for the process orientation of the company was empirically proven, which reflects the scientific contribution of this work. In the literature so far, there are only theoretical claims about the existence of the mentioned influence and connection, but there are no empirical tests. The conducted research also examined the practice of using process information technology in Croatian companies. It has been proven that the informatization of business in the Republic of Croatia is still inadequate and that information technology is the least rated dimension of process orientation. All of the aforementioned evidence led to the conclusion that process information technology adoption and use in Croatian businesses are essential.

One of the main limitations of the research in this work is related to the size of the cause on which it was conducted. Namely, although the representative sample consisted of 127 large, medium and small companies of the Republic of Croatia, a larger sample would have provided much more detailed and concrete results. The research, despite using a professionally crafted survey questionnaire, cannot overcome all the drawbacks of this approach, chief among which is the respondents' incapacity to provide objective answers to the questions and their lack of understanding of specific issues. Therefore, increasing the sample with the additional use of the interview method would largely eliminate the limitation related to the shortcomings of the survey method. Using an interview would avoid misunderstanding of the questions by the respondents, as well as their subjectivity when answering.

Finally, empirical research was conducted only in the Republic of Croatia. Future study could be expanded and conducted concurrently in a number of European Union member states in order to determine the standard level of process orientation for European companies. This would enable a transparent comparison of samples from different countries, as well as the development of a methodology for the use of various types of process information technology necessary for businesses to move toward greater levels of process focus.

References

[1] P. Mikalef, J. Krogstie, "Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities", European Journal of Information Systems, Vol. 29(3), pp. 260-287, 2020.

[2] A. Van Looy, G. Poels, M. Snoeck, M., "Evaluating Business Process Maturity Models" Journal of the Association for Information Systems, Vol. 18(6), pp. 461–486, 2019.

[3] W. Aalst, "Business process management: a comprehensive survey", ISRN Softw. Eng., pp. 1-37, 2013.

[4] K. McCormack, W. Johnson, "Business Process Orientation: Gaining the E-Business Competitive Advantage", New York: St. Lucie Press, 2011.

[5] F. Rahimi, C. Møller, L. Hvam, L., "Business process management and IT management: The missing integration", International Journal of Information Management, Vol. 36(1), pp. 142–154, 2016.

[6] Lj, Milanović Glavan, V, Bosilj Vukšić, N. Vlahović,, "Decision tree learning for detecting turning points in business process orientation: a case of Croatian companies", Croatian Operational Research Review, Vol. 6(1), pp. 207-224, 2017.

[7] S. Lahajnar, A. Rožanec, "The evaluation framework for business process management methodologies", Management: journal of contemporary management issues, Vol. 21(1), pp. 47-69, 2016.

[8] D. Ivandić Vidović, V. Bosilj Vukšić, J. Kereta, "Analiza performansi poslovnih procesa za postizanje poslovne izvrsnosti: studija slučaja Hrvatskog osiguravajućeg društva", Poslovna izvrsnost Zagreb, Vol. 1, pp. 113-130, 2009.

[9] P. Dmitry, L. Mayilyan, Z. Zelentsov, "Information technology integration processes for the sustainable development of construction", E3S Web of Conferences, Vol. 258 (2), pp. 123-153, 2021.

[10] N. Hasan, S. Miah, Y. Bao, M. Hoque, "Factors affecting postimplementation success of enterprise resource planning systems: a perspective of business process performance. Enterprise Information Systems", Vol. 13(9), pp. 1217-1244, 2019.

[11] M. Pejić Bach, V. Bosilj Vukšić, D. Suša Vugec, A. Stjepić, "BPM and BI in SMEs: The role of BPM/BI alignment in organizational performance", International Journal of Engineering Business Management, Vol. 11, pp. 34-41, 2019.

[12] R. Kesner, B. Russell, B., "Enabling Business Processes through Information Management and IT Systems: The FastFit and Winter

Gear Distributors Case Studies", Journal of Information Systems Education, Vo. 20, 401-405, 2020.

[13] J. Mendling, B. T. Pentland, J. Recker, "Building a complementary agenda for business process management and digital innovation", European Journal of Information Systems, Vol. 29(3), pp. 208-219, 2020.

[14] J. Ruey-Jer, K., Daekwan, "Information technology and organizational performance within international business to business relationships: A review and an integrated conceptual framework" International Marketing Review, Vol. 25, pp. 563-583, 2018.

[15] D. Suša Vugec, A. Stjepić, L. Sušac, L., "Business Process Management Software Functionality Analysis: Supporting Social Computing and Digital Transformation", Proceedings of FEB Zagreb 10th International Odyssey Conference, pp. 547-558, 2019.

[16] BPTrends, retrieved from: https://www.bptrends.com/ [09.11.2022.]

[17] S. Chavva, S. Ravi, B. Rao, "A Survey on Business Intelligence Tools for Marketing, Financial, and Transportation Services: Proceedings of the Second International Conference on SCI", Vol.2, 48-53, 2020.

The importance of ITIL4 adoption for IT service management in insurance companies

Ajla Ćerimagić Hasibović*, Anel Tanović**, Aida Granulo***
* Faculty of Electrical Engineering/Department of Computer Science, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
acerimagic@etf.unsa.ba
** Faculty of Electrical Engineering/Department of Computer Science, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
atanovic@etf.unsa.ba
*** Faculty of Electrical Engineering/Department of Computer Science, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
agranulo@etf.unsa.ba

Abstract - **In post pandemic era where companies already adopted digital agendas in their everyday business, conservative businesses as insurance companies must intensify activities in creating new values and use of the potential innovations. Insurance companies have to follow new age users, developments in society and new economic laws with new demands for insurers. As ITIL4 describes an operating model for the delivery of tech-enabled products and services, the importance of its adoption significantly increases. The way this adoption helps insurance companies is described in this paper. Trends in IT operations, such as agile approach are also considered.**

Keywords – ITIL4, Incident Management, change management, Insurance

I. INTRODUCTION

ITIL is a framework for service management and a professionally recognized certification scheme that supports organizations and individuals in gaining optimal value from IT and digital services [1]. Whereas previous versions of ITIL have been perceived to focus overly on IT&SM processes, ITIL4 now provides 34 management practices that incorporate guidance on how best to achieve the desired outcomes and value. These practices provide guidance on architecture, organizational change, projects, workforce and talent, stakeholder relationships, risks, business analysis, IT assets, and others [1], [2]. The process of moving from owning processes to creating a service-focused value chain is an evolution and demands involvement and collaboration of all company's departments including management and stakeholders [3]. In ITIL4 practices are part of the service value system. Service value system (SVS) includes Guiding principles, Governance, Service value chain, Practices and Continual improvement [1]. In the SVS is the service value chain which is a flexible operating model for the creation, delivery, and continuous improvement of services. The service value chain defines six main activities [1], [2]: Plan, Improve, Engage, Design and Transition, Obtain/Build, Deliver and Support, Products and Services.

ITIL4 also indicates seven guiding principles: Focus on value, Start where you are, Progress iteratively with feedback, Collaborate and promote visibility, Think and work holistically, Keep it simple and practical, Optimize and automate [2].

The Internet itself is altering the ways to provide and manage products and services, and the relationships between organizations and customers, which is leading to a redefinition of their value chains [4], but using ITIL as a framework can even increase profits. Transformation in IT was accelerated by COVID pandemic conditions making companies to massively adopt teleworking [5]. That was a start point for all kinds of businesses, but companies had to choose services and implement changes in order they've found adequate. In many cases, customers and employees are not always receptive to what technology has to offer, nor do they recognize its associated value [4]. Changes in customer demands also underpin potentially high degree of awareness among managers about market changes and emphasizes the need for consideration of organizational consequences when new segments are implemented [6]. Although significant, new business models based on technology-impact growth, were slowly applicable to the insurance industry in Bosnia and Herzegovina.

Provisioning of relevant reports to supervisory authorities was the first step in the mandatory adoption of new technologies. Opportunity for additional profit and new value for insurers is provided through using mobile applications and online sale, where the benefit was also for customers – they could customize products they use and buy it online. Marketing adapted to new customers is also an essential step in business changes, especially using social networks. Segmentation of the market and new products like travel insurance policies

with risk of covid19 covered got insurers step forward, as in businesses as in segment-specific strategies and plans. In this paper potential changes in an insurance company from Bosnia and Herzegovina will be discussed. Company taken as an example is company of 500 employees – active system users, where more than 150 are in the back office. System is distributed in 8 regional centers and more than 150 offices. IT consists of 10 members, including the CIO, divided into three divisions – application support, system and technical support, software development. The company hasn't been using ITIL framework on any of the processes.

Section I of the paper is introduction. In section II incident management will be discussed. Section III describes change management and change enablement, and final section IV is about conclusion and future work.

A. *Literature review and motivation for research*

The implementation and adoption of ITIL4 would include an examination of the challenges of implementing ITIL4 and a comparison with other service management frameworks. Even the ITIL4 is in use since 2019., there is a significant lack of recent literature. Using literature mainly related to ITIL3, foundation for further research of ITIL4 can be provided. Documents in used databases can provide us better review: main interest in ITIL research was in year 2012 for incident management, and year 2009 for change management. The analysis of used databases is at follows:

TABLE 1: LITERATURE REVIEW

Search topic	**Database**	**Analyze Result**	**Category of interest**
ITIL incident management	Web of science	79	Computer science - 72
			Engineering - 18
			Telecommunications - 12
			Management - 5
			Business, Economic- 8
	Scopus	168	Computer science - 138
			Decision sciences - 40
			Engineering - 38
			Mathematics 36
			Business - 25
ITIL change management	Web of science	96	Computer science - 93
			Management - 16
			Engineering - 14
	Scopus	171	Computer science - 134
			Engineering - 41
			Business - 33
			Mathematics - 33
			Decision sciences - 28

II. INCIDENT MANAGEMENT

The adoption of ITIL in an organization represents an innovation that is both influencing and influenced by the organizational context. Two of the main reasons for adopting ITIL are to become a service-oriented organization and to deliver IT services that meet business needs [7]. When it comes to small and medium businesses that use the ITIL platform, literature [8] [9] says that the first procedure to be implemented should be the Incident Management Process. ITIL defines an incident as an unplanned interruption to or quality reduction of a service [1]. Some IT examples are the too slow internet, a business application is inaccessible, or a printer not working. Adapted guiding principles can be used as proposed methodology for the process implementation.

A. *Focusing on value*

A value stream is a path that starts with demand and concludes with value generation along the service value chain [9]. First, we have to initialize the project – define areas of improvement and set the goals to be achieved. This is the point where is decided would that change be beneficial to company, and what are the values system company could gain. Some of the values of implementation incident management process are better system of information toward customers, increased hardware reliability, improved monitoring and audit made easier. Stakeholders are to be introduced to changes at this point.

In the system analyzed in this paper, focusing on a value is point where:

- All top managers are aware of the project – they are owners of the process – it is guided by company's vision and financed by management.
- System parts that need to be improved are defined – focus is on system monitoring and mobile application
- There is a detailed plan with defined resources for transforming company into service-focused organization

The mission of IT department here is to implement applications which support business activity and reorganize IT team to better meet company's expectations. Shared vision between IT, HR, finances, marketing and other sectors is mandatory to be established.

B. *Current system state*

Customer expectations are higher than ever and they demand IT support instantly and all time available. Because of availability all kinds of support in systems people use privately, they expect mobile support application or chatbot integrated in business application they use. Knowledge base of the observed system doesn't prevent technicians to do routine tasks repeatedly, because the base is not overall accessible.

Also, specialists could be more effective if they weren't interrupted by tasks that could be automated. Current system could be updated in better logging incidents and managing fixes. The system considered in this paper uses an internally developed help desk, mail communication, MS Office 365 planner, MS Teams and PRTG monitor for managing incidents. That could be considered as a second step of the guidance toward ITIL implementation – start where you are.

Literature overview [3], [4], [5], [10] can provide guidelines for the start point analysis. Inventory conduction is the first step in all analyzed literature, but its outcome should be a document that outlines the conduction result and provides analysis of the current situation, as well as the position of current IT practices compared to ITIL best practices. If profiled first, case organization [11] would be as follows:

TABLE 2: ORGANIZATION PROFILE

Geographic location	Bosnia and Herzegovina
Sector	Private
Size	Medium
Industry sector	Insurance/finance
Business focus	Local
System users	Approximately 450
IT department employees	10
IT structure	Centralized
Server and network infrastructure	System adequate
Trigger/ITIL justification	Processes not formally defined or documented, inefficient tracking of incidents through to changes

Secondly, interview with IT department and minimum 3 other departments should be provided. Goal of this step is to analyze practices of IT department, get the list of equipment, applications and other resources they use, and get user feedback from other departments. One of the user questionnaires should be about user satisfaction. ITIL4 focuses on organizations and people, so this is a point to determine if people have enough skills and knowledge to get the job done, but also if they have the information, they need to run the business properly.

ITIL4 best practices that correlate with the observed system would be: the ability to implement the process with a minimum of activities performed; investing in people with the right skills; selecting the right communication channel; information management integrated into the process; continuously updating the knowledge base; automating incident management.

C. The action plan

The incident management process consists of following steps: logging, prioritization, investigation, communication, resolution, review and closure [1].

In company analyzed, there are average 600 phone calls, 350 mails and 20 tasks in help desk per month toward IT. Not all of them are incidents, but following framework could enable process automation and better communication between parties. The ITIL4 aspects of collaboration, automation, and simplicity reflect the values found in other frameworks such as Agile, DevOps, and Lean methodologies [2] which definitely brings using mobile support apps and automation of processes in focus. Automation is most important fact in process and knowledge base connected to it empower users to solve some problems themselves.

TABLE 3*:* IMPACT ON BUSINESS BY ITIL4 ADOPTION

Steps of Incident management process	**Sector and tool used now**	**Example**	**What ITIL4 can improve**
Logging	IT/Help desk log, mailboxes (i.e., employee opened a help desk task)	Module for Online travel insurance is not working	Automation
Prioritization	IT and other sectors/employees sort tasks on personal experience using knowledge base	Customers can be redirected to store, so priority is set by IT on medium	
Investigation	IT and other sectors/ knowledge base, all used system dedicated applications	Application debugging, checking online status, checking if all APIs are working properly	Diagnostic, knowledge management tools help users to solve some issues themselves
Communication	IT in collaboration with sales support/ Functional and hierarchic escalation	Internal exchange of all states of the incident in IT sector	Detection and correction of automations, ease of accessibility
Resolution	IT/resolution confirmation through Help desk or email	All bugs corrected	
Review	IT, sales support/ knowledge base and shared user manuals	New entry in knowledge base	
Closure	IT and other sectors/Help desk, email, phone	Employee who opened task verified the same	Mean resolution time improved

Final goal of incident management implementation is to improve troubleshooting time response and the time completion of client request, agent motivation, and improved system availability. Investigation step usually takes the longest period of time of every incident. This is the sub process organizations can segment, and get customer feedback sooner in process. That is actually an iterative agile approach, which will enable company to respond to changes in the perception of value. Collaboration iteratively with feedback also continues through next steps.

D. *Collaborate and promote visibility and think and work holistically*

The 4 dimensions of ITIL4 cover value streams and processes, information and technology, organizations and people, and partners and suppliers. People are the collective representatives of the whole system, so close collaboration is critical [1] [2]. In organizations like this, processes, projects, and service value streams can flow across many teams and departments. All these departments have to give their contribution to organizational knowledge pool. Problem solution as for example a new insurance product doesn't calculate premium correctly involves all teams from sales department to development team in collaboration. This step has its mayor impact on following incident management process steps: communication, resolution and review. Choosing the right tool for communication is also of great importance for this step. Recommendation is to use multiple communication channels, and get a software for sorting and categorizing collected data. Sharing regular software and system changes, updates, as keeping even virtual meetings regularly improve user satisfaction and encourage users to solve some problems by themselves. Company's services don't usually start and end within the same department. This is the right point where review of information architecture can be done, as well as redesign documentation and even improvement categorization of causes of incidents – what prevents incident from happening.

E. *Economical use of resources*

Implementing ITIL4 doesn't mean abandoning all previous frameworks and good practices that work for our system and keep it functional. Good managers should focus on achieving the desired result and not waste resources on shortening code or its elegance, but also take the optimal number of steps to achieve the result. Critical success factors factors include [11]: senior management commitment, employee awareness, training, software selection, culture changes, and customer-focused metrics. Performance must be measured at all costs, and companies must measure their progress.

F. *Optimization and automation*

ITIL Incident Management Reporting aims to supply Incident-related information to the other Service Management processes, and to ensure that that improvement potentials are derived from past Incidents [1], but optimization and automation is mostly about people. ITIL4 practices encourage organizations to use human resources more wisely – they should perform actions on incidents that cannot be atomized and not on those which can [10]. Collaboration is also very important so all employees can benefit from organizational knowledge. Development of change culture also has to be stimulated to implement. The mayor benefit is for IT support if processes are automated. Metric to evaluate automation must be adopted. Also executing the improvements in iterative way embraces agile in process.

III. CHANGE MANAGEMENT

After COVID19 taught us how fast people can adapt to new normal, it is important for companies to adopt change management models so they can achieve business success in a safe way at a faster speed than ever before [12]. ITIL4 changed "change management" as known in ITIL3 to "change enablement" and it defines the purpose of the new change enablement practice as: "To maximize the number of successful service and product changes by ensuring that risks have been properly assessed, authorizing changes to proceed, and managing the change schedule" [1]. Collaboration between sectors, which is also recommended in the ITIL4 guidelines, improves the chances of long-term success. ITIL4 changed guidance includes how different levels of complexity affect change needs:



Figure 1: ITIL4 reflects complexity as a critical factor for change [1]

Type of change will determine the level of assessment, authorization, and documentation required. Literature [1] identify three types of change, depending on risk and impact a categorization also recognized applicable for Insurance Company in this paper:

TABLE 4*:* THREE TYPES OF CHANGE

Type of change	**ITIL4 coverage**
Standard change (i.e., user password resets and operating	A low-risk, pre-authorized change that is well understood and fully documented and can be implemented without needing

system updates)	additional authorization. Requires a full risk assessment and authorization only during creation, or modification due to business change or occurrence of an incident.
Normal change (i.e., Insurer application upgrade, hardware upgrades)	A change that needs to be scheduled, assessed, and authorized.
Emergency change (i.e., Regulatory changes the business process)	A change that must be implemented as soon as possible. The process for assessment and authorization is expedited to ensure quick implementation, so scheduling and documentation is not a priority.

Another connection with agile approach is change authority. Decentralized change approval enables working in sprints in different domains related to a change. Change advisory board (CAB) consists of individuals or groups with different knowledge. Of course, it is essential that the correct change authority is assigned to each type of change. Whether CAB is one person or a team, they must communicate broadly within company's departments and involved stakeholders. Here, use of scheduler is suggested at least for planning of changes, communication, avoid conflicts, resource allocation and informing parties about implemented changes. Change advisory board of insurance company, which is the subject of research in this paper, is consisted of IT management including CIO, technical support director, directors of the underwriting department, claims department, policy administration and customer service.

Change enablement guides process to stakeholder's satisfaction – it is customer focused process, which also helps organizations to continuously improve their infrastructure and processes ensuring business to easily implement necessary changes without affecting other business operations. Process steps from ITIL3 [13] can be summarized in ITIL4 into: change submission, change planning, change assessment, change implementation, change review, and change closure.

Change submission is usually ticket generated by service desk – initiating change. Basic information like change type and priority is also collected here. For a Company in this paper suggestion would be using of company's help desk in this step. Predefined templates for data collection are prepared by company. Planning is step where are created documents like back out plans or downtime plans. This is a step needed to be presented to stakeholders. Checklist is the most common approach here. Next step is change assessment. Change plan here needs to be approved by change advisory board. In observed company that should be suggested to top management including CIO, team managers, technical support, and finance and marketing teams. Change implementation involves actors from included sectors and work dispersion through the given tasks. Using project management software can significantly ease change implementation, especially if large-scale changes like digital transformation or moving backup to cloud are in focus. After a post-implementation review is carried out to ensure there are no deviations in the implementation, the change can be closed. Closure can have one of statuses: success, fail, incomplete.

IV. DISCUSSION

This this paper highlights the importance of adopting ITIL4 for IT service management in an insurance company that has not yet implemented it. The reason why this company was chosen is the specifics of insurance caused by the pandemic, and the company has experienced a digital transformation in travel health insurance in a short period of time. The organization has recognized the need to become service-oriented and provide IT services that meet business needs to maintain its position in the market, but framework is still missing. The analysis of the available literature allows the elaboration of recommendations specific to this insurance company, for which this paper was the introduction. As recommended by the literature [8], [9], first process to be implemented is the incident management. Incident management should be followed by change management.

There are many factors that can affect the differences in insurance between Bosnia and Herzegovina and other countries. First factor, already mentioned in this paper, are regulations. The regulatory authority is already divided into two entities, each with its own rules, which can affect the types of insurance available, payment methods, and general terms and conditions of insurance companies. This also affects pricing, either in the country itself or in comparison with other countries. Knowing that there are companies in the country that have their headquarters in the countries of the European Union, the conditions may vary between the insurance companies themselves, which only reinforces the need for structured management, where ITIL4 can help.

V. CONCLUSION

In today's insurance market, companies must demonstrate maximum competitiveness. Without guidelines through processes, this competitiveness only becomes a burden. On the other hand, knowledge based on ITIL4 best practices can greatly facilitate key

processes within an organization. Organizations are finding that in reality they need a set of incident management processes to handle different categories of problems quickly and efficiently using digital service management [1]. This could also be observed in the company considered in this paper.

ITIL4 is based on market requirements, and it emphasize value, so companies can use it not only for IT services, but for all services of business. In this paper is given an initial proposal for an insurance company that has no ITIL implemented – to encourage its adoption and company's market position improvement. This could be provided by enabling all teams to work on new values and the prerequisites for that would be reduction in the execution time of processes, improvement of control and monitoring, increase in clients' satisfaction, cost reduction, decreased system downtime, even the better positioning IT in company, all provided by ITIL4 framework. The fact company has not implemented ITIL before makes it easier to accept idea of implementing the new one.

This paper showed how ITIL4 processes Incident Management and Change Management can improve the operations of any company. Incident Management and Change Management processes when implemented in the business environment of a company must be clearly presented through the presentation of: key goals, key activities, risks, roles, responsibilities, key performance indicators and critical success factors.

Future work should follow ITIL4 implementation in a company and provide all metrics and measuring results for comparison. Analysis which services should be implemented would be a prior step to measuring process.

REFERENCES

[1] Axelos, "Axelos," 20 10 2020. [Online]. Available: https://www.axelos.com/resource-hub/white-paper/itil-4-and-devops-white-paper . [Accessed 1 2023].

[2] S. F. S. F. Lopes, "The importance of the ITIL framework in managing Information and Communication Technology services," *International Journal of Advanced Engineering Research and Science (IJAERS),* vol. 8, no. 5, pp. 292-296, 2021.

[3] Raul Valverde, Raafat George Saade, Malleswara Talla, "ITIL-based IT service support process reengineering," *Intelligent Decision Technologies,* 2014.

[4] Gil-Gómez, Hermenegildo; Oltra-Badenes, Raúl; Adarme-Jaimes, Wilson, "Service quality management based on the application of the ITIL standard," *Engineering Journal Dyna,* vol. 81, no. 186, pp. 51-56, 2014.

[5] J. L. R. Sánchez, "A Quantitative Analysis of Information Systems Management in the Educational Industry," *Analytics,* vol. 1, no. 1, pp. 193-209, 2022.

[6] Ann Højbjerg Clarke, Per Vagn Freytag, "Implementation of new segments in small- and medium-sized enterprises (SMEs)," *Journal of Business & Industrial Marketing,* 2022.

[7] Tom Roar Eikebrokk, Jon Iden, "Enabling a culture for IT services; The role of the IT infrastructure library," *International Journal of Information Technology and Management Forthcoming,* 2014.

[8] L. Lema, J.-A. Calvo-Manzano, R. Colomo-Palacios and M. Arcilla, "ITIL in small to medium-sized enterprises software companies: towards an implementation sequence," *Journal of Software: Evolution and Process,* pp. 528-538, 2015.

[9] Yahya Al-Ashmoery; Hisham Haider; Adnan Haider; Najran Nasser; Mohammed Al-Sarem, "Impact of IT Service Management and ITIL Framework on the Businesses," in *2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI)*, Sana'a, Yemen, 2021.

[10] Mourad EL Baz, Malik Motii Armand, Collins Anong, Belaissaoui Mustapha, "ITIL Implementation in a Moroccan Stat Organization: The case of incident management process," *International Journal of Advanced Engineering Research and Science (IJAERS),* vol. 4, no. 4, pp. 71-78, 2017.

[11] Carol Pollard, Aileen Cater-Steel, "Justifications, strategies and critical success factors in successful ITIL implementtions in U.S. and Australian companies: an explaratory study," *Information Systems Management,* 2009.

[12] Mohammad Reza Taghva, Mohammad Taghi Taghavifard, Foad Fasihi, "Evaluation of IT Service Management Based on ITIL Framework in Banks of Industri and MIne (Service Operation Processes)," *Asian Journal of Information Technology ,* vol. 15, no. 2, pp. 169-177, 2016.

[13] Tariq Rahim Soomro, Amanullah Bashir, "Comparative Study on Change Management," *International Journal of Computer Applications,* vol. 39, no. 1, pp. 32-36, 2012.

[14] Jon Iden, Tom Roar Eikebrokk, "Using the ITIL process reference model for realizing IT Governance: An empirical investigation," *Information Systems Management,* vol. 31, no. 1, pp. 37-58, 2014.

[15] Mauricio Marrone, Francis Gacenga, Aileen Cater-Steel, Lutz Kolbe, "IT Service Management: A Cross-national Study of ITIL Adoption," *Communications of the Association for Information Systems,* vol. 34, no. 49, pp. 865-892, 2014.

[16] Bertrand Verlaine, Ivan Jureta, and Stephane Faulkner, "How Can ITIL and Agile Project Management Coexist?," in *International Conference on Exploring Services Science*, 2016.

[17] Jiqing Cao, Shuhai Zhang, "ITIL Incident Management Process Reengineering in Industry 4.0 Environments," in *2nd International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2016)*, 2016.

[18] Igor Aguilar-Alonso, Mariluisa Pereda Pascal and Cristian Mera Macias, "Applying Business Process Modeling to improve IT Incident Management Processes in a Public Entity in Peru," *IBIMA Publishing - Journal of Software & Systems Development,* 2020.

[19] O. Melnyk, M. Zlotnik, "The essence and basics models of strategic enterprise management," *Economics, entrepreneurship, management,* vol. 7, no. 1, pp. 48-62, 2020.

Classifying Relevant Causes of Employee Absenteeism Using Machine Learning

I. Fosić*, A. Živković*, I. Fosić**
*Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek, Osijek, Croatia
**HEP-Telekomunikacije PS Osijek, Osijek, Croatia
ivana.fosic@efos.hr, ana.zivkovic@efos.hr, igor.fosic@hep.hr

***Abstract*: Employee absenteeism has a significant impact on the organization as it imposes high costs that can be reflected in the organization's economic results. The impact can be seen in productivity, sustainability, competitiveness, profitability, and inter-organizational relations. Considering the wide range of causes of absenteeism, this paper provides a systematic overview of the causes considering different theoretical approaches. A primary survey was conducted (N=420), and the causes of absenteeism were divided into two groups of causes (personal causes and organizational causes). Unsupervised and supervised machine learning methods were applied to the data set. In the unsupervised machine learning phase, the Elbow data grouping method was applied, and the K-Means algorithm was used to assign each record in the dataset to a specific cluster. In supervised machine learning, the data were divided into learning and testing parts in a 70/30 ratio, and the Random Forest (RF), Decision Tree (DT), Support Vector Machines (SVM), and k-Nearest Neighbors (K-NN) classification algorithms were used. The best performance and area under the ROC curve (AUC) values are obtained using the RF classification algorithm. Using the machine learning algorithm RF, the main causes of absenteeism in both groups of causes of absenteeism were identified. From the group of personal causes (personal illness as well as illness of family members), from the group of organizational causes (mobbing, conflict with work colleagues) stand out the most important causes of absenteeism.**

Keywords: causes of absenteeism, employee absenteeism, unsupervised and supervised machine learningrt

I. INTRODUCTION

The absenteeism rate in Croatia, as in the countries of the European Union, varies over time. It depends primarily on the industry and the size of the business entity, as well as on various factors such as the economic situation, legislation and the environment. According to Eurostat data [1] on absenteeism in 2021, Croatia is not behind the European Union average. In Croatia, 10.6% of employees were absent from work in the first quarter (EU=9.4%), 8.1% in the second quarter (EU=7.6%) and 14% in the third quarter (EU=15%). The number of absences in 2021 is approaching the average for the period 2015 to 2019, while 2020 is characterized by numerous measures taken to contain the spread of COVID - 19, which consequently had a significant impact on the increase in absenteeism.

Frequent worker absences negatively impact the cost and productivity of the business system, which can result in the end product or service being uncompetitive [2]. Absenteeism represents a significant cost to employers, and according to Absence Insight [3] , it is the second largest cost after salary costs, suggesting that significant savings in total labor costs can be achieved by addressing this issue, as much as 15%.

The aim of this work is to identify which causes of absenteeism from the group of organizational and personal causes play the most important role in absenteeism, taking into account and highlighting the demographic characteristics of these workers. The above findings can be used in efforts to reduce employee absenteeism to the lowest possible level.

This thesis follows the following concept: after introductory remarks highlighting the aim of the thesis, Chapter 2 provides an overview of the theoretical basis and previous research. Chapter 3 describes the methodology used and the data analysis, including the presentation of the results of the research conducted. Finally, Chapter 4 draws conclusions, identifies the limitations of the research, and provides recommendations for future research.

II. THEORETICAL BASIS AND PREVIOUS RESEARCH

Workplace absenteeism or absenteeism is an important issue for all organizational stakeholders, not just HR professionals. Any justified or unjustified absence of an employee from the workplace is called absenteeism. Absenteeism is the absence of an employee from work, and the main causes are sick leave and abuse of sick leave by employees [2]. The slowness of public health procedures, the lack of an absence management process in organizations, a negative organizational culture, and a rigid labor law framework often lead to unnecessary costs [2]. The author also notes that sick leave is the most prevalent form of absenteeism, but that there are other forms, such as being late to work, taking frequent breaks, manipulating leave, and "apparent attendance." Absenteeism should be looked at from different angles to understand its impact at different levels: from the social to the individual level. In general, absenteeism is a major problem in companies, and a distinction must be made between regulated and unregulated absenteeism [4]. Worker absenteeism has negative consequences for the operation of the company, relationships between colleagues, and the ability to work as a team, and affects workers' families and the environment [5]. Identifying the causes of absenteeism allows for better management, recognition of uncontrolled, partially controlled, or fully controlled forms of absenteeism [5].

Ilić [2] highlights that employers often do not consider the indirect costs of absenteeism, such as the cost of a replacement employee, administrative costs, and costs associated with a drop-in productivity, which are difficult to assess without specialized knowledge and skills. For this reason, the costs of absenteeism are generally explained as unmeasurable, minor, and difficult to manage without affecting the organization and cost of work. In order to make

decisions on reducing absenteeism, data on the prevalence of absenteeism in the organization is needed, which is why it is important to have a system for detecting and recording absenteeism and analyzing trends. To achieve this, it is necessary to develop and implement a system for recording absenteeism, including time measurement [5]. Systematic and objective monitoring of absences can provide the organization with useful data on whether absences are above allowable and acceptable levels and whether they are characteristic of particular jobs and categories of workers. Absence data can provide clues to possible causes and factors of absenteeism. However, in order to take more serious action against absenteeism, the causes and conditions for its occurrence and increase must be accurately identified [6].

According to Čikeš and others, [7] the determinants of absenteeism are divided into several categories depending on their nature: Attitudes, personal, demographic, health, and organizational factors. The authors emphasize that understanding absenteeism behavior begins with understanding its determinants and outcomes.

Different authors have different approaches to classifying the factors that influence absenteeism and its consequences. The basic division of causes of absenteeism for the purposes of this paper is between personal and organizational causes of absenteeism.

Personal causes of absenteeism include absenteeism due to: personal illness, family issues, personal needs and problems, sense of entitlement, stress [8] but also illness of family members, own laziness, own work done by the employee in addition to his/her current job, due to difficulties in going to the workplace, social events during working hours, medical examinations, and problems with alcohol and/or drugs. As mentioned earlier, sick leave is one of the most common causes of absenteeism, while according to Kocakulah et al [9], family problems play an important role in absenteeism. Balancing work and family life can be difficult, especially because of the high cost of child care. One of the most common family problems is that adult offspring have the duty of caring for their elderly parents, which means that they have additional obligations and time constraints to meet their parents' needs, such as doctor's appointments or hospitalizations that require additional time from workers [9]. It can be very difficult to recognise alcohol or drug abuse in the workplace, and alcohol and drug dependence can lead to incapacity and absenteeism [10]. It is desirable for organizations to have a substance abuse policy in place to adequately address this problem [10]. Scheduling a necessary doctor's appointment is often a challenge for employees because they don't want it to interfere with their scheduled workday. This poses a problem for both the employee and the employer because time is lost that cannot be compensated for [9]. Workplace health promotion and education, in addition to laws and regulations, are very important and are achieved through the cooperation of employers, employees and society to improve the health and well-being of workers. In order to achieve better work organization and work environment, it is necessary to promote workers' active participation and personal development. It is important to keep in mind that presenteeism can turn into absenteeism and vice versa [11].

Organizational causes of absenteeism include: Dissatisfaction with work, lack of interest in work, poor working conditions, overtime, dissatisfaction with pay, conflicts with co-workers, conflicts with supervisors, and "bullying" (violent behavior) or mistreatment at work [12] Low employee autonomy, participation and responsibility lead to low satisfaction, which in turn leads to higher absenteeism [13]. When employees are dissatisfied with their work or their motivation wanes, they want to spend as little time as possible at work and avoid going to work whenever possible [2]. Nath Gangai [14] emphasizes that it is better to provide positive incentives to workers to reduce absenteeism than to just impose penalties. Rewards such as additional time off or monetary bonuses are more effective in reducing absenteeism than punishments such as loss of benefits or job. The author also notes that a combination of incentives and penalties, with an emphasis on motivational incentives, is the most effective approach to reducing absenteeism.

It is also necessary to emphasize the importance of demographic characteristics. Choi [15] suggests that there is a relationship between basic worker demographic characteristics such as age, gender, and tenure and various absenteeism reduction measures. It is known that younger workers are more likely to be absent due to lower levels of responsibility, job satisfaction, and poorer working conditions, while older workers are more likely to be absent due to greater job satisfaction and higher status and loyalty to the organization, despite having a higher risk of health problems [16].

Study by Choi [15] wanted to investigate what factors managers should consider to reduce absenteeism among their employees. His research relied on worker demographics as potential causes of absenteeism using machine learning predictions, identifying which worker demographic characteristics might predict absenteeism in the workplace.

III. METHODOLOGY AND DATA ANLYSIS

In the research part of the paper, field primary research was conducted in written form at selected organizations to investigate the opinions and attitudes of the respondents. The test method was applied, more precisely the probing with group test. The test was conducted in the premises of the respondents' work organizations. The prepared questionnaire was anonymous and the respondents filled it out independently. The sample size after subtracting missing values is 420 people. For the research part of the work, the collected demographic data and the data on the causes of absenteeism are used. A good combination of unsupervised and supervised machine learning and their complementarity can be applied to methods in various fields of scientific research such as medicine [17], mathematics [18], banking [19] etc. A very widespread application of these methods is the analysis of absenteeism in various organizations such as

schools [20], industries [21] and in general in business [22]. A proposed approach in school absenteeism research [20] involves using both unsupervised and supervised machine learning techniques to examine data pertaining to school absenteeism. First, group detection is solved using unsupervised techniques. Then, supervised learning methods are used to classify the remaining data. The proposed method is a combination of unsupervised and supervised models, resulting in a model for classifying students into high-risk groups, medium-risk groups, and low-risk groups in terms of truancy. The research [21] compares various ML classifiers and finds that they exhibit a high level of accuracy, making ML techniques a viable option for predicting and analyzing absenteeism. The best ML algorithms are used for forecasting and analyzing school truancy. The study [22] analyzes a situation of employees missing work in a delivery company with the aim of assisting HR executives in forming plans and policies to decrease absenteeism. The data was cleaned and several machine learning methods were utilized to categorize the data (zeroR, J48 based on decision trees, naive Bayes, and KNN).

This paper describes the use of machine learning techniques for the analysis of data related to absenteeism for personal and organizational reasons on the collected data. In the first phase, within the unsupervised machine learning, the Elbow method of data clustering was applied as a basic step to determine the optimal value k, i.e., the number of clusters of the observed data set, and the K-Means algorithm was used to assign each data set to a specific cluster. In supervised machine learning, a classification success test was performed on the previously labeled dataset using four different classification algorithms: RF, DT, SVM, K-NN.

A. Unsupervised machine learning

Unsupervised machine learning algorithms are useful for creating labels in data used in performing supervised learning tasks. Unsupervised algorithms allow intrinsic groupings within unlabeled data and assign a label to each data value [23]. To determine the optimal value of k, i.e., the number of clusters of the observed dataset, the Elbow method was used as a fundamental step for the process of unsupervised machine learning. The Elbow method is used to determine the number of clusters in the dataset by calculating the sum of squared errors (SSE eng. Sum of Squared errors). The number of clusters is determined by observing the graph and the position of the point representing the "elbow" [24]. For the personal and organizational cause datasets, graphs were created using the Elbow method, from which the possible four values for the number of clusters are read: 2, 3, 4, and 5. Since no value has a position at the "elbow point," an analysis

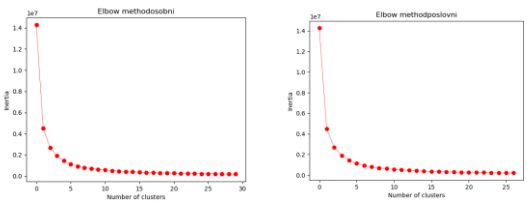


Figure 2. Elbow method applied for both datasets



Figure 1. Silhouette method applied for both datasets

of the Silhouette coefficients was therefore performed for each proposed number of clusters. Determining the silhouette value is a very useful method for finding the optimal number of clusters when the Elbow method does not show a clear "elbow" point. Silhouette coefficient values vary from -1 to 1, with 1 being the best value when samples are perfectly distributed in easily distinguishable clusters. The silhouette coefficient is calculated from the mean distance within the cluster - a - and the mean distance of the nearest cluster - b - for each sample using the following formula [25]:

$$\text{Silhouette score} = \frac{b-a}{\max(a,b)} \qquad (1)$$

According to Figures 1 and 2 and Table I, the optimal number of clusters k = 2 was chosen for both data sets.

TABLE I. SILLHOUETE VALUES FOR BOTH CLUSTERS

Number of clusters	*Dataset - personal cause*	*Dataset - organizational cause*
2	**0.591**	**0.591**
3	0.499	0.499
4	0.481	0.481
5	0.480	0.482

After determining the optimal number of clusters for both datasets, all records in both datasets were assigned to a specific cluster using the K-means algorithm. The K-means algorithm is one of the most well-known algorithms in the field of unsupervised machine learning. This algorithm finds non-overlapping clusters in which each data set is assigned to a particular cluster. This algorithm groups the data by attempting to divide them into n groups with equal variance by appropriately reducing a criterion known as inertia or sum of squares within clusters [26][27]:

$$\sum_{i=0}^{n} \min_{\mu_j \in C}(||x_i - \mu_j||^2) \quad (2)$$

where x_i is a single record in cluster C and μ_j is the mean of the records in the cluster. Using the K-Means algorithm for both datasets, all records were labeled and assigned to one of the clusters. Of the total 420 records in each dataset, 272 records belong to cluster class 0, and 148 records are assigned to cluster class 1. Records in cluster class 1 indicate a more frequent personal or organizational cause of absenteeism, depending on which record is considered. Records labeled class 0 indicate weak absenteeism or some other cause of absenteeism that is not personal or organizational in nature. The distribution of records by the appropriate class in the personal and organizational cause records is shown in Figure 3. Because the datasets are multidimensional, containing 31 and 28 features, respectively, an nD transformation (n is the number of dataset features) was applied to 2D space using the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm to represent the datasets in a 2D coordinate system.

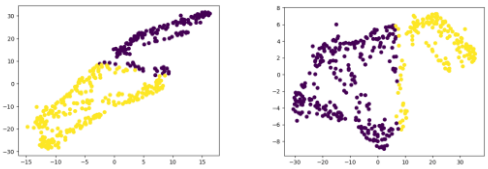


Figure 3. Visualisations of clusters using t-SNE

When applied to high-dimensional but well clustered data, t-SNE shows a visualization with clearly isolated clusters that match well with clusters derived with special clustering algorithms [28].

B. *Supervised machine learning*

The objective of supervised learning is to construct a model from labeled data in order to make predictions on future data. There are two predominant supervised machine learning techniques: classification and regression. Classification involves predicting discrete values such as categories, classes, or labels for new data. Regression, on the other hand, involves predicting the continuous value of a response variable [29]. By applying the classification algorithms to the data clustered in this way from both data sets, it is possible to determine which features contribute most to the classification or to membership in a particular cluster. In this way, the influence of certain questions and answers in the questionnaire used for data collection is indirectly revealed. For classification, four machine learning algorithms were tested and their classification performance was assessed by comparing accuracy, AUC and F1 results. Prior to the actual classification, the data were processed so that they were standardized. Standardization of datasets is a common procedure for many machine learning estimators. Machine learning results could be misinterpreted if the features do not look like normally distributed subplots by default [30]. In practice, the shape of the distribution is often ignored and the data are simply transformed to be centered by removing the mean of each characteristic and then dividing by the standard deviation. The min-max normalization technique involves a linear transformation of the raw data, converting the features in any range to a new range, usually on a scale between [0,1] or [-1,1]. The equation used in this method is:

$$x' = \frac{x_i - x_n}{x_m - x_n} \quad (3)$$

where x_n denotes the minimum value, x_m the maximum value, x_i the input value, and x' the normalized value.

When this normalization method is applied, each feature retains all relational properties in the data [31]. To compare the performance of different classification algorithms, the data for learning and testing must be defined. Prior to the application of machine learning classification, the data was split 70/30, meaning that 70% of the entire dataset is designated for learning and the classification success is tested on the remaining 30% of the data. The split of the dataset into a training and a testing part was done using the train_test_split() function in the Python programming language, which allows random selection of records from the dataset but maintains an equal distribution of classes so that the training and testing data is a representative sample of the original dataset. Classification was performed using four algorithms: RF, DT, SVM, and K-NN , and the results are presented in Table II. for the personal causes of absenteeism dataset and in Table III. for the organizational causes of absenteeism dataset.

TABLE III. CLASSIFICATION PERFORMANCE – DATASET - PERSONAL CAUSE

ML algorithm	*Accuracy*	*F1*	*AUC*
RF	**0.98412**	**0.97777**	**0.98780**
DT	0.97619	0.96551	0.97117
SVM	0.97619	0.96703	0.98170
K-NN	0.93650	0.90476	0.91962

TABLE II. CLASSIFICATION PERFORMANCE – DATASET - ORGANIZATIONAL CAUSE

ML algorithm	*Accuracy*	*F1*	*AUC*
RF	**0.99206**	**0.98876**	**0.99390**
DT	0.97619	0.96551	0.97117
SVM	0.96031	0.94382	0.95898
K-NN	0.93650	0.90476	0.91962

The best results in classifying data with personal and organizational causes of absenteeism are obtained with the algorithm RF, both in terms of AUC, F1, and accuracy. Differences in certain features of the two datasets were examined for their impact on classification results. The popularity of the algorithm RF is based on its ability to make successful predictions. However, it is also important that it provides a complete nonparametric measure of feature importance in prediction/classification [32]. Feature importance allows users to identify those features that play a key role in prediction/classification. In many applications, a good predictive model is only one goal; another, often more important goal, is to identify variables that enable good

prediction/classification. The algorithm RF allows the measurement of the importance of variables, which can be used to rank the variables according to their predictive importance [33]. In this study, the feature_importances_ property and its values were used for each record feature in the algorithm RF. Figure 4 shows the order of importance of the extracted features from both datasets of personal and organizational causes of absenteeism.

When analyzing the data on personal causes of absenteeism, the characteristic of personal illness (A08) was found to be the most important for classification, followed by family responsibilities (A07), illness of family members (A09), and medical examination (A24), while other

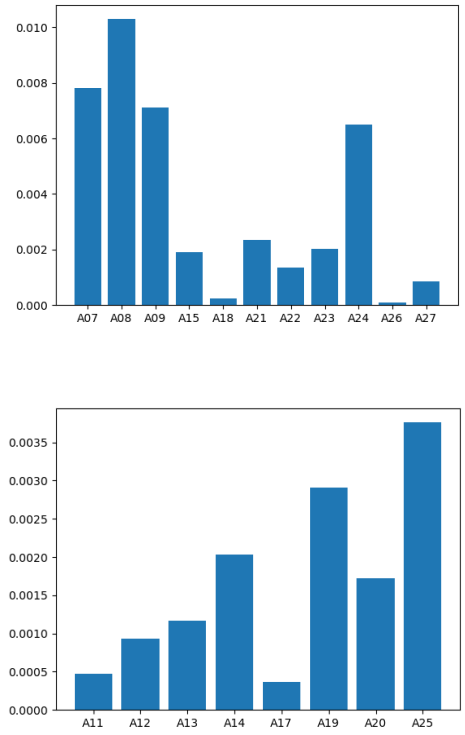


Figure 4. Feature importances

characteristics were less important. For the organizational causes of absenteeism dataset, the most important characteristics affecting classification are bullying (A25), conflicts with work colleagues (A19), overtime (A14), and conflicts with supervisors (A20).

Table IV. shows the order of importance of all characteristics in the classification, including demographic characteristics. In both data sets, respondents' demographic characteristics had the greatest importance in the classification. In the data set of absenteeism for personal reasons, including demographic characteristics, the following characteristics had the greatest importance in classification: total work experience (DM07), work experience in current organization (DM06). In the long case, in the dataset of organizational reasons for absenteeism, along with

TABLE IV. FEATURE IMPORTANCES – BEST DEMOGRAPHIC SCORES

feature importance - demographic scores in personal cause		**feature importance - demographic scores in organizational cause**	
feature	***importance score***	***feature***	***importance score***
DM07	0.367247	DM09a	0.651428
DM06	0.268832	DM10	0.04614
DM02	0.176321	DM12	0.045125

demographic data, the following characteristics had the greatest influence in classification: ownership type of organization (DM09a), level of job in the organization (DM10).

IV. CONCLUSION

In reviewing the literature, numerous classifications of the causes of absenteeism were found, each confirming the other, but also differing from each other, which means that it is necessary to create unique tools both for predicting absenteeism and for predicting the causes that lead to it. Because of the complexity and multifaceted nature of the causes of absenteeism, authors often use certain groups of causes in their research while ignoring others. Using unsupervised and supervised machine learning techniques, it is possible to predict or classify respondents depending on the frequency of their absence from work, which can be personal or organizational in nature. Considering the proposed number of clusters by unsupervised machine learning and the results of the most efficient classification of the Random Forest algorithm of supervised machine learning.

The results show that personal illness and family responsibilities are the most important absences from work for personal reasons. Considering the demographic characteristics of the personal reasons for absenteeism, the most important feature is the total number of years of service and the length of service in the current organization. From the group of organizational causes, bullying and conflicts with work colleagues, as well as the form of ownership in the organization and the level of position in the organization, have the greatest influence on the prediction of absenteeism.

One of the obvious limitations of the conducted research is also the above-mentioned classification that considers only a certain number of organizational and personal causes of absenteeism, which leads to the neglect of other causes that can also have a significant impact on absenteeism in the workplace.

The recommendation for future research is reflected in the increase of the number of respondents, which will contribute to the generality of the obtained results when other possible factors for the cause of absenteeism in workers are included. In the field of machine learning, techniques for reducing and selecting optimal features for prediction or classification can be applied to confirm or improve the adoption of measures to reduce worker absenteeism.

REFERENCES

[1] "Eurostat." Hours of work and absences from work - quarterly statistics - Statistics Explained (europa.eu) (accessed Jan. 30, 2023).

[2] Đ. Ilić, "Measuring absenteeism as a precondition of quality management of absenteeism," Trendovi u poslovanju, vol. 8, no. 1, pp. 66–74, 2020, doi: 10.5937/trendpos2001066i.

[3] "Absence Insight ." https://absenceinsight.eu/ (accessed Jan. 30, 2023).

[4] I. Katić and A. Nešić, "Socio-psychological effects of stress in organizations' absenteeism problems," Work, vol. 66, no. 3, pp. 689–697, 2020, doi: 10.3233/WOR-203211.

[5] D. A. Endovitskiy, "Developing A Classifier Of Relevant Causes Of Absenteeism In An Organisation," in Global Challenges and Prospects of The Modern Economic Development Proceedings of Global Challenges and Prospects of The Modern Economic Development (GCPMED 2020), 15-16 December, 2020, Samara State University of Economics, Samara, Russia, Apr. 2021, vol. 106, pp. 699–706. doi: 10.15405/epsbs.2021.04.02.84.

[6] Đ. Ilić, G. Mrdak, and M. Bojić, "Sociological aspect of labor force absentism," Oditor, vol. 7, no. 1, pp. 195–224, 2021, doi: 10.5937/oditor2101195i.

[7] V. Čikeš, H. M. Ribarić, and K. Črnjar, "The determinants and outcomes of absence behavior: A systematic literature review," Social Sciences, vol. 7, no. 8. MDPI AG, Jul. 24, 2018. doi: 10.3390/socsci7080120.

[8] S. G. Aldana and N. P. Pronk, "Health Promotion Programs, Modifiable Health Risks, and Employee Absenteeism," J Occup Environ Med, vol. 43, no. 1, pp. 36–46, Jan. 2001, doi: 10.1097/00043764-200101000-00009.

[9] M. C. Kocakulah, A. G. Kelley, K. M. Mitchell, and M. P. Ruggieri, "Absenteeism Problems And Costs: Causes, Effects And Cures," International Business & Economics Research Journal (IBER), vol. 15, no. 3, pp. 89–96, May 2016, doi: 10.19030/iber.v15i3.9673.

[10] R. M. Badubi, "A Critical Risk Analysis of Absenteeism in the Work Place," Journal Of International Business Research and Marketing, vol. 2, no. 6, pp. 32–36, 2017, doi: 10.18775/jibrm.1849-8558.2015.26.3004.

[11] H. Brborović and J. Mustajbegović, "Mogućnost prevencije prezentizma i apsentizma zdravstvenih djelatnika," Sigurnost, vol. 58, no. 2, Jun. 2016, doi: 10.31306/s.58.2.2.

[12] A. Cohen and R. Golan, "Predicting absenteeism and turnover intentions by past absenteeism and work attitudes," Career Development International, vol. 12, no. 5, pp. 416–432, Aug. 2007, doi: 10.1108/13620430710773745.

[13] G. Egan, "An Investigation Into The Causes Of Absenteeism In 'Company X.'"

[14] K. Nath Gangai, "Absenteeism At Workplace: What Are The Factors Influencing To It?"

[15] J. W. Choi*, "Prediction of Workplace Absenteeism Time Using Machine Learning," International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 2, pp. 3489–3493, Dec. 2019, doi: 10.35940/ijitee.B6571.129219.

[16] M. Chaudhury and I. Ng, "Absenteeism Predictors: Least Squares, Rank Regression, and Model Selection Results," 1992.

[17] T. Chauhan, S. Rawat, S. Malik, and P. Singh, "Supervised and Unsupervised Machine Learning based Review on Diabetes Care," in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Mar. 2021, pp. 581–585. doi: 10.1109/ICACCS51430.2021.9442021.

[18] J. Shen, W. Li, S. Deng, and T. Zhang, "Supervised and unsupervised learning of directed percolation," Phys Rev E, vol. 103, no. 5, p. 052140, May 2021, doi: 10.1103/PhysRevE.103.052140.

[19] W. Bao, N. Lianju, and K. Yue, "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment," Expert Syst Appl, vol. 128, pp. 301–315, Aug. 2019, doi: 10.1016/j.eswa.2019.02.033.

[20] F. Bowen, C. Gentle-Genitty, J. Siegler, and M. Jackson, "Revealing underlying factors of absenteeism: A machine learning approach," Front Psychol, vol. 13, Dec. 2022, doi: 10.3389/fpsyg.2022.958748.

[21] A. Rista, J. Ajdari, and X. Zenuni, "Predicting and Analyzing Absenteeism at Workplace Using Machine Learning Algorithms," in 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Sep. 2020, pp. 485–490. doi: 10.23919/MIPRO48935.2020.9245118.

[22] M. Skorikov et al., "Prediction of Absenteeism at Work using Data Mining Techniques," in 2020 5th International Conference on Information Technology Research (ICITR), Dec. 2020, pp. 1–6. doi: 10.1109/ICITR51448.2020.9310913.

[23] P. Ebrahimi, M. Basirat, A. Yousefi, Md. Nekmahmud, A. Gholampour, and M. Fekete-Farkas, "Social Networks Marketing and Consumer Purchase Behavior: The Combination of SEM and Unsupervised Machine Learning Approaches," Big Data and Cognitive Computing, vol. 6, no. 2, p. 35, Mar. 2022, doi: 10.3390/bdcc6020035.

[24] H. Humaira and R. Rasyidah, "Determining The Appropiate Cluster Number Using Elbow Method for K-Means Algorithm," in Proceedings of the Proceedings of the 2nd Workshop on Multidisciplinary and Applications (WMA) 2018, 24-25 January 2018, Padang, Indonesia, 2020. doi: 10.4108/eai.24-1-2018.2292388.

[25] "silhouette_score." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html#sklearn.metrics.silhouette_score (accessed Jan. 17, 2023).

[26] "K-means clustering." https://scikit-learn.org/stable/modules/clustering.html#k-means (accessed Jan. 17, 2023).

[27] Md. Zubair, MD. A. Iqbal, A. Shil, M. J. M. Chowdhury, M. A. Moni, and I. H. Sarker, "An Improved K-means Clustering Algorithm Towards an Efficient Data-Driven Modeling," Annals of Data Science, Jun. 2022, doi: 10.1007/s40745-022-00428-2.

[28] D. Kobak and P. Berens, "The art of using t-SNE for single-cell transcriptomics," Nat Commun, vol. 10, no. 1, p. 5416, Nov. 2019, doi: 10.1038/s41467-019-13056-x.

[29] I. Fosic, D. Zagar, and K. Grgic, "Network traffic verification based on a public dataset for IDS systems and machine learning classification algorithms," in 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), May 2022, pp. 1037–1041. doi: 10.23919/MIPRO55190.2022.9803674.

[30] "Standardization, or mean removal and variance scaling." https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing-scaler (accessed Jan. 18, 2023).

[31] G. Aksu, C. O. Güzeller, and M. T. Eser, "The Effect of the Normalization Method Used in Different Sample Sizes on the Success of Artificial Neural Network Model," International Journal of Assessment Tools in Education, pp. 170–192, Apr. 2019, doi: 10.21449/ijate.479404.

[32] H. Ishwaran and M. Lu, "Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival," Stat Med, vol. 38, no. 4, pp. 558–582, Feb. 2019, doi: 10.1002/sim.7803.

[33] F. Degenhardt, S. Seifert, and S. Szymczak, "Evaluation of variable selection methods for random forests and omics data sets," Brief Bioinform, vol. 20, no. 2, pp. 492–503, Mar. 2019, doi: 10.1093/bib/bbx124.

Application of Blockchain Technology and NFTs in a Museum Environment

Matea Bilogrivić*, Helena Stublić**
* Croatian Conservation Institute, Zagreb, Croatia
** University of Zagreb, Faculty of Humanities and Social Sciences / Department of Information and Communication Sciences, Zagreb, Croatia
mbilogrivic@hrz.hr, hstublic@ffzg.hr

Abstract **- Museums as traditional institutions are usually not leaders in the use of the latest and disruptive technologies. However, the Covid-19 pandemic has influenced the growing interest of researchers and museum experts in new solutions for communication and museum operations. That is how blockchain technology and non-fungible tokens (NFT) entered the focus of museums. Although it is most often associated with cryptocurrencies, this technology, due to its main property, the immutability of records on the chain, has great potential for use in various aspects of museum work.**

This paper will show how blockchain and NFTs can be used in museums, from checking the provenance, ownership, and authorship of museum objects, facilitating their borrowing, more efficient digital rights management to building smart museums. It will also focus on how museums have so far experimented with NFTs and blockchain technology and highlight the potential problems of their use. Huge challenges to understanding and adopting this technology is the lack of knowledge base and this paper aims to bring the topic of blockchain technology within the museum field to the table and start a conversation on this fast-growing new technology and possibilities of its implementation.

Ključne riječi – blockchain technology, NFT, museums, galleries, art, cultural heritage

I. INTRODUCTION

Blockchain is a distributed ledger that contains time-stamped records organized into blocks, which are interconnected by repeating a hash code as the first part of one block and the last part of the previous one [1]. In other words, it is a database on a distributed network, a copy of which is kept by all nods on the network. The security of the records on the ledger is one of the main features of this technology and sharing the ledger with all nodes on the network prevents data modification.

Although blockchain technology is usually associated with the emergence of the cryptocurrencies, it was created much earlier. Wanting to investigate trust in information in the digital age, Stuart Haber and Scott Stornetta developed a computer system in 1991 that could digitally time stamp any document through cryptographic coding (hashing) [2]. However, it was not until 2008 that the blockchain gained popularity thanks to the white paper of the Bitcoin, published by an unknown individual or collective under the pseudonym Satoshi Nakamoto [3]. The next big step in the development of the technology was the emergence of the Ethereum blockchain in 2013 and the presentation of the new Ethereum protocol in 2014 [4]. The new blockchain made it possible to create applications and smart contracts which have enabled tokenization, the process of turning an asset into a digital token which represents that asset [5]. Thus, non-fungible tokens were developed on the Ethereum chain, whose popularity skyrocketed in 2021 after the sale of Beeple's NFT *First 5000 days* for $69 million at Christie's auction. Unlike classic cryptocurrencies where every coin is equal, NFT is unique, each has its own value and therefore cannot be exchanged for one another, hence the name non-fungible. NFTs use smart contracts, which Ethereum adapted for use on the blockchain and which enables two parties to perform a secure exchange, without intermediaries [6].

Due to the aforementioned features, blockchain has been increasingly explored in the last few years to exploit its potential in sectors such as finance, commerce, medicine, etc. However, this new technology could also be useful in the heritage sector. Classical heritage institutions are usually not pioneers of new technologies, but despite this, researchers have already begun to study this topic. Although there are not that many published papers, they provide a useful insight into current thinking on blockchain, NFTs and cultural heritage.

Methodologically, this paper is based on literature review and the analysis of external content. The main research question was what are all the possible uses of blockchain and NFTs in the museum environment? From that question arose sub-questions such as: can some museum practices be simplified and improved by using new technology? What are the obstacles in its use? How have museums interacted with this technology so far?

The main part of the paper refers to the application of blockchain and NFTs in museums and is divided into

several sub-chapters, each of which is dedicated to one of the possible uses. In each part, the main conclusions from the literature will be presented, along with projects carried out so far. Also, some problematic aspects of its use are explored, and we raise questions that open up space for new research. The following is a discussion in which we will refer to the context in which research on this topic occurs. Also, we highlight still unexplored or under-researched phenomena related to this technology in museums. We will give our final thoughts on the subject in the conclusion of this paper.

II. Application Of Blockchain Technology And NFTs In A Museum Environment

In the last few years, various research has been conducted on how blockchain technology can be used to improve the work of heritage institutions, including museums. One study from 2022 showed that a search of articles published in the last 10 years on the topic of blockchain technology and cultural heritage highlighted three thematic units: verification of provenance and authenticity, tokenization and fractional capital, and management of intellectual and digital rights [7]. However, a review of the literature revealed that the scope of research is much deeper, and the ideas are more diverse. So far, there are several projects and platforms that have taken advantage of this technology and NFTs to explore their possible benefits in heritage field.

TABLE 1 TECHNOLOGY USED IN DIFFERENT APPLICATIONS

Application in a museum field	Blockchain technology	NFT
A. Provenance, authenticity, authorship, and ownership	✓	✓
B. Digital database of museum objects	✓	
C. Museum rights management	✓	✓
D. Loan of museum objects	✓	
E. Shared ownership and guardianship	✓	✓
F. Smart museum	✓	
G. Making profit	✓	✓

A. Provenance, authenticity, authorship, and ownership

The most common use of NFTs and blockchain technology in research so far is the verification of authorship, authenticity, ownership, and provenance of works of art [1], [8], [9], [10], [11], [12]. Museums and galleries spend a lot of time verifying the provenance of objects from their collections, as well as when buying or selling artworks. A public registry of artworks and museum collections with the mentioned information could therefore be located on the blockchain, which would be accessible to everyone, and every change of owner would be recorded. This would ensure a secure record that is immune to tampering. However, there is a question of the accuracy of the original data entered on the chain, which may be wrong. This is especially the case for, for example, paintings by old masters for which the provenance has not been determined with certainty, so even the record on the blockchain would not be credible [13].

There are several blockchain platforms that deal with issuing certificates of authenticity for artworks and valuables. The first such platform is Verisart, founded in 2015 and managed by Robert Norton. It is a database of artworks through which a certificate of authenticity can be obtained [1], [9], [14], [15], [16]. The biggest flaw of the platform is that anyone can enter the artwork data, which has led to one absurd certificate. In 2018 Terence Eden entered the *Mona Lisa* on the Verisart chain, signed himself as the author, and put the year of creation as 1506 [17]. Artory is a platform on the Ethereum chain that offers collectors certificates of authenticity of artworks, and they cooperate with galleries, museums, auction houses, etc. [1], [16]. This platform, founded by Nanne Dekking in 2016, solved the problem of freely entering data to obtain a certificate. When an individual wants to enter information about an artwork in the register in order to obtain a certificate, experts from art institutions first must verify entered data, such as provenance. The data about artwork enters the blockchain only after approval from the experts, thus ensuring that only verified and accurate information is recorded on the chain. Artory is also the first company that organized an auction through the blockchain, in cooperation with the auction house Christie's. The third platform that deals with the issuance of certificates via the blockchain is Codex Protocol, which enables the verification of the authenticity and provenance of collectible items and the entry of one's own items into the register. To verify the information that is at the beginning of the chain and on which the others are built, they cooperate with auction houses [1], [16].

When we talk about using blockchain technology to authenticate a work of art or a collectible, one question remains open, and that is how to connect the record on the chain and the physical work [13].

B. Digital database of museum objects

Blockchain could be used to create a digital register of museum collections, but also a register of cultural heritage objects of a region or a country [18], [19], [20]. Photographic and written documentation of objects could be kept on the chain, which becomes especially important in case of destruction or loss of a certain object. Given that the data that would be stored on the distributed ledger is large (4k videos, high resolution images, 3D scans) the blockchain should be separated from the files it manages [18]. One of the suggestions for creating a database is to create a hash of the original document that would represent its digital fingerprint and would be stored on the blockchain. Such a record would ensure the accuracy of the data, and anyone could use the hash to verify its authenticity and immutability [19].

Creating a database of museum objects or heritage objects on the blockchain would ensure greater

transparency of the data and make it impossible for them to be modified later. Also, it would enable greater accessibility of data to other professionals and to the public. However, as with the previous use, there is always the problem of the accuracy of the data entered on the chain.

C. Museum rights management

Another use of this technology is management of museum rights and licensing, whether it is the right of ownership, use, display, copyright, or intellectual property right. Museums have an exhibition as well as an educational role and actively participate in the dissemination of knowledge. Therefore, researchers often use museum material to find the necessary information. This process is facilitated by the gradual digitization of museum objects and documentation. The era of Covid-19 pandemic accelerated and increased the digitization of museum collections, as most museums were forced to close their doors and operate exclusively in the digital world. With the digitization of museum collections comes the problem of unauthorized use and sharing of digital content. All this has led to the need for better digital rights management, and the latest solution to this problem is blockchain technology. Three research papers tried to solve the problem with different business models [21], [22], [23]. In order to prevent misuse of museum digital content, first model propose the system of authorizing the use of copyrights based on blockchain, smart contract and cryptography. The research is based on the ECDSA digital signature algorithm [21]. The second paper proposes an algorithm for an authorization mechanism based on blockchain technology that protects the museum's copyrights by using cryptography. This model also includes a cash flow system so that the applicant pays the museum a premium to access the museum's digital content [22]. The third study proposes digital rights management using NFTs and smart contracts. They offer a model in which museums can create NFTs of cultural objects to express their ownership, and the applicant buys an NFT with cryptocurrency to receive a certificate of authorization to use the content [23].

Bhowmik et al. [8] have also made an interesting proposal for digital rights management for lending of digital objects, be it art or archival documents. They propose the creation of a blockchain on which all image transactions will be recorded. The transaction would contain all relevant information about access rules, usage, copyright, etc. An approved transaction receives a hash code, which is then written into the metadata of the JPEG file. Such modified JPEG file is sent to the applicant. The recipient can then verify the authenticity of the file by verifying the record on the blockchain or comparing the signature on the block and in the file they received.

Blockchain technology could also be used to clearly establish certain rights to cultural heritage items, museum objects included, especially those that were dislocated during war or colonization. This division of rights could contribute to a greater willingness to repatriate alienated objects. With the help of tokenization, the rights over the object could be divided so that one party has the ownership right, while the other retains the exhibition right, permanently, or temporarily. Also, tokenization would facilitate the sharing of the monetary income generated by exhibiting the object or some other activities [24]. In order to prevent the unauthorized removal of objects from the country, objects found at archaeological sites could be immediately entered on the blockchain, and thus the ownership rights could not be called into question [24].

D. Loan of museum objects

Another application of blockchain technology concerns the lending of objects among museums [16], [25]. Loans of artworks in museums are regulated by ICOM guidelines and EU provisions, but they often include many other protocols and are often administratively demanding. Loans of museum objects are important both for large institutions and for smaller ones, for whom loans for occasional exhibitions mean a greater influx of visitors and the realization of profits. Therefore, to facilitate the lending process, it is suggested to use blockchain technology. A blockchain can be public, so anyone on the network can see records on it, or private, where only nodes on the private network have access to information. The second option is more suitable for institutions because museums that work together could be connected on a private network. All information about the item, terms of the loan, its time interval, insurance details, value of the item, penalties in case of late return, etc. would be recorded on the chain [16]. Using a smart contract would prevent possible legal problems between museums, because it is automatically executed and charged in the event of a violation of the agreed conditions. Also, other institutions that have access to the network could at any time check where a particular museum object is located and whether it is on loan or in its home institution. This approach would raise the level of security and deepen trust between institutions. The disadvantages of using blockchain technology for loans are on the other hand the high energy consumption, the costs of creating a private chain and the personnel who would maintain it, and the possibility of human error when writing a smart contract. Namely, since smart contracts cannot be changed once they are written on the blockchain, it is necessary to create a new contract to correct incorrect data [16].

This type of loan of museum objects was tested in 2021 among ten museums in Florence. A distributed ledger was created on a private network, which tracked objects that were loaned between museums. Data about each transaction was recorded on the blockchain, including the description of the item, the time interval of the loan, including the start date and end date, insurance details, valuation of the item, verification of the item, the

cost of the loan, if any, and the fee in in case of loss or late return of the borrowed object. Given that no smart contract was used, there was no data on monetary transactions. With this project, in addition to a simpler and safer loan of art, the aim was to raise awareness of the rich inventory kept by the museums in Florence [16].

E. Shared ownership and guardianship

Blockchain technology can serve to tokenize artworks and sell them in parts, leading to shared ownership [1], [15], [16], [25], [26]. An example of shared ownership of art is Andy Warhol's painting *14 Small Electric Chairs*, which was partially sold by the Maecenas platform via the Ethereum chain. Namely, in 2018, the platform divided the ownership of the painting into smaller fragments in the form of art tokens, such as NFTs, and released 49% of the shares of the painting for sale. At the auction, they sold 31.5% of the painting [13]. This left them as the majority owners and allowed people to own a piece of art that they otherwise might not have been able to afford. This kind of shared ownership could be implemented in museums as well. Liddell proposes shared guardianship of digital museum objects, which would deepen the connection between the visitor and the institution [15]. The project led by the National Museum Liverpool and the University of Manchester wanted to explore how blockchain technology can be used to create a meaningful digital museum object, i.e., an object that creates values and connections with the participant. The project took place in three phases from 2018 to 2020. In the first phase, project participants took part in various activities, in which, for example, they described the personal items they had brought with them. The second phase was the organization of an online exhibition called *Crypto Connections: Exploring the Personal*, where the participants' personal items and digital museum items selected at the workshops were exhibited, while the third phase included the creation of digital tokens (cryptocollectibles) of the items exhibited at the exhibition. The results of the research showed that blockchain technology has a positive impact on the creation of connections between participants, museum objects and the museum. By owning a digital token, the participants felt a moral obligation towards the objects and the museum, which is one of the prerequisites for joint guardianship of the same. The token served the participants as a form of memorabilia, which reminded them of the entire process and the experiences gained, thereby adding social value [15], [26].

However, when we talk about shared guardianship, the question of legal ownership and authority over management of objects arises. In the case of a museum, it is difficult to imagine a real sharing of ownership, given that the art always remains in the museum and the authority over it is centralized [27]. Blockchain technology could be a partial solution to this problem. Liddell [15] establishes that by sharing ownership, museums distribute power over the object and in this way, participants become like stakeholders. In order for them to really feel like shareholders, having a material claim would help. Tokens like NFTs could be that material claim, which would be a proof of shared guardianship or ownership and would represent the stakeholder's authority over the museum object.

F. Smart museum

Heritage institutions require certain conditions for the preservation of their materials, such as a certain level of humidity, temperature, security system, etc. To facilitate the monitoring and maintenance of favorable conditions, the Internet of Things (IoT) can be used in combination with the blockchain. IoT represents the connection of smart devices via the Internet and the creation of their network. It provides access to data in real time, collects information about the current state of the building and users, enables collaboration, faster decision-making process, detects critical problems and predicts behaviors or failures. The blockchain, on the other hand, ensures data security and is introduced as a system for monitoring, gathering information about the activity of sensors and systems, and for monitoring the activities of people participating in the process. This approach would prevent unauthorized access to data and would ensure their safe storage and management in real time [28].

G. Making profit

With the rise in popularity of NFTs, many museums saw an opportunity to make money and at least partially recover the revenue they lost due to pandemic closures. So, some of them started digitizing their artworks and sell them as NFTs. Among them were the British Museum, the Uffizi Gallery, the Hermitage, the Whitworth Gallery, the Leopold Museum, etc.

The Leopold Museum in Vienna, for example, made a collection of 24 NFTs of the works of Egon Schiele upon the discovery of one of his long-lost works. They used the money from the sale for the conservation of these works and the restoration of the lost painting [29]. This practice could thereby help heritage institutions, not only museums, to generate income for the restoration and conservation of objects through the sale of NFTs [30].

However, although some of the larger museums have generated significant revenues from such sales, research has shown that only a small number of museums are involved, and only 5% of them generated revenues greater than $250,000 [31]. More and more museums are withdrawing from the sale of NFTs due to the decline of the NFT and cryptocurrency market. Therefore, the use of NFTs for the purpose of providing additional value to visitors has a much greater potential than concentrating only on making a profit.

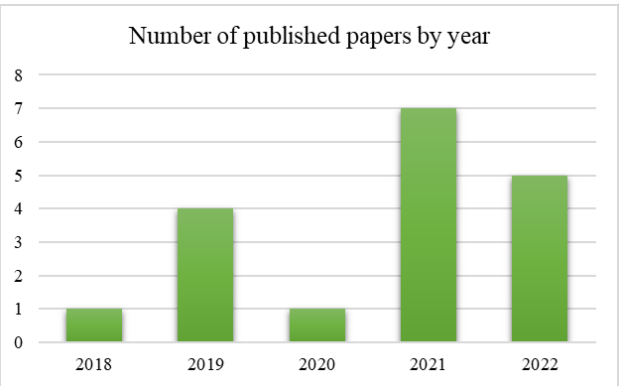
Apart from the profit itself, museums could use the sale of NFTs as an opportunity to attract audiences and reward loyal visitors and customers. Using a smart

contract, the NFT owner could have rights to certain utilities such as museum tickets, a discount on the museum shop or the possibility to participate in the conservation of the museum fund, as was the case in the Leopold Museum. By doing this, museums would improve the connection with visitors, which was already discussed in the chapter on shared ownership and guardianship.

III. DISCUSSION

Papers published so far on the use of blockchain technology and NFTs in the museum environment represent only the beginning of research on this topic. Unlike for example, the archives that began exploring this topic as early as 2016, the first research on application in museums began only in 2018. Covid-19 pandemic and sudden popularity of NFTs have certainly encouraged new research, which is evident from the number of papers published on the topic in 2021 (GRAPH 1). Experts have seen the opportunity in a new technology to solve the problems that museums have faced in terms of content digitization, but also to resolve and simplify some other museum practices such as loan of museum objects or long -term storage of documentation.

GRAPH 1 NUMBER OF PUBLISHED PAPERS ON THE TOPIC OF BLOCKCHAIN TECHNOLOGY AND NFT IN A MUSEUM FIELD BY YEAR



So far, research is mostly based on the potential use of this technology, and only two studies have been conducted in practice and four concrete models for digital rights management have been designed. Further research is needed to better understand blockchain technology and its benefits for the museum sector.

It should be investigated how it is most efficient to connect physical museum objects and their data on the chain, as well as what are the long-term ways of preserving NFTs that become part of museum documentation or collection. Also, it is important to explore the values that NFTs could bring to museums and their audiences, along the lines of social value, which Liddell [15], [26] writes about. Can museums use NFTs and blockchain technology to deepen connections with audiences and achieve greater engagement? Further audience research should answer the question.

Another unaddressed topic in previous research is museum communication through NFTs. Due to the Covid-19 pandemic and the rise in popularity of NFTs, some museums have started experimenting with the new medium and organizing exhibitions dedicated to the new technology, both physically in the museum and in the metaverse. It would therefore be interesting to investigate how the new medium affects museum communication, curatorial practices, and audience experiences in the new environment of the virtual world.

A major obstacle in the implementation of blockchain technology in heritage institutions, including museums, is the lack of knowledge. In most cases, museum experts are not familiar with new technology and the possibilities it provides. It would certainly be useful to do research and create strategies for the education of experts regarding new technologies and their implementation in the museum space. This problem will certainly become less significant over time with the arrival of a new generation of experts who will not be strangers to new technologies, but rather use it as the main tool in their private and professional lives.

IV. CONCLUSION

Blockchain technology and non-fungible tokens have entered a wide social, economic, artistic and many other spheres of life in the last few years. Although it is not a new technology, recently there has been a growing interest in its use. Thus, heritage institutions such as museums, archives and libraries began to explore its benefits. Research conducted so far propose the use of NFTs and blockchain technology in heritage institutions to determine the ownership, authorship, authenticity, and provenance of works of art and documents. Furthermore, they can be used to monitor the loan of objects between institutions, create a register of cultural assets, share ownership of museum objects, etc. Until now, there are several platforms and projects that deal with the mentioned aspects and try to find the best solution for merging new technologies and institutions. Although there are undoubted benefits of using technology, there are also some negative factors associated with it, such as high energy consumption. Also, additional costs for fees, loss of the private key and the risk of incorrectly entered information on the distributed ledger are highlighted.

Many museums saw the sale of NFTs an opportunity to compensate for lost income in the last few years and the possibility of collecting funds for the conservation of museum objects. However, only small number of them made considerable profit. Therefore, museums should shift their perspective and ask themselves how they can use this technology to create additional value to their audience. The first step in doing so is education of museum experts on blockchain technology and NFTs, which was one of the aims of this paper. Blockchain technology and non-fungible tokens in the context of museums is a very broad and interesting topic, and many

research questions remain open and provide opportunities for new research.

REFERENCE

[1] A. Whitaker, "Art and Blockchain. A Primer, History, and Taxonomy of Blockchain Use Cases in the Arts", in Artivate: A Journal of Entrepreneurship in the Arts, vol. 8, no. 2, 2019, pp. 21-46. DOI:10.34053/artivate.8.2.2

[2] S. Haber and W. S. Stornetta, "How To Time-Stamp a Digital Document", in Cryptology vol. 3, 1991., pp. 99-111, https://doi.org/10.1007/BF00196791

[3] S. Nakamoto, Bitcoin: A Peer-To Peer Electronic Cash System, Satoshi Nakamoto Institute, 2008. [Preprint]. Available at: https://nakamotoinstitute.org/bitcoin/ (last accessed 21. 8. 2022.).

[4] V. Buterin, A next generation smart contract & decentralized application platform, 2014, Available at: https://ethereum.org/en/whitepaper/#a-next-generation-smart-contract-and-decentralized-application-platform (last accessed 21. 8. 2022.).

[5] C. Pinto-Gutiérrez, S. Gaitán, D. Jaramillo and S. Velasquez, "The NFT Hype: What Draws Attention to Non-Fungible Tokens?" in Mathematics vol 10, no. 3, 335, 2022. https://doi.org/10.3390/math10030335.

[6] Q. Wang, R. Li, Q. Wang and S. Chen, "Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenge", in arXiv, 2021. https://doi.org/10.48550/arXiv.2105.07447

[7] E. D. Vacchio and F. Bifulco, "Blockchain in Cultural Heritage: Insights from Literature Review", in Sustainability vol. 14, no. 4, 2324, 2022. https://doi.org/10.3390/su14042324

[8] D. Bhowmik, A. Natu, T. Ishikawa, T. Feng and C. Abhayaratne, "The Jpeg-Blockchain Framework for GLAM Services", IEEE International Conference on Multimedia and Expo. San Jose, California, July 2018. http://dx.doi.org/10.1109/ICMEW.2018.8551519

[9] L. Dolan, B. Kavanaugh, K. Korinek and B. Sandler, "Off the Chain: Blockchain Technology—An Information Organization System", in Technical Services Quarterly, vol. 36, no. 3, 2019, pp. 281-295. https://doi.org/10.1080/07317131.2019.1621571

[10] S. Ren, "The Application of Blockchain and its Advantage to the Field of Art", in Journal of Physics: Conference Series, 1437, 012063, 2020. http://dx.doi.org/10.1088/1742-6596/1437/1/012063

[11] A. A. Oyelude, "Trending issues in advancing blockchain technology in libraries, archives and museums", in Library Hi Tech News, vol. 39, no. 6, 2022, pp. 6-7. https://doi.org/10.1108/LHTN-06-2021-0040

[12] D. Ross, E. Cretu and V. Lemieux, "NFTs: Tulip Mania or Digital Renaissance?", IEEE International Conference on Big Data (Big Data), December 2021, Orlando, USA, pp. 2262-2272. http://dx.doi.org/10.1109/BigData52589.2021.9671707

[13] D. Yaga, P. Mell, N. Roby and K. Scarfone, Blockchain Technology Overview, National Institute of Standards and Technology, 2018.

[14] F. Valeonti, A. Bikakis, M. Terras, C. Speed, A. Hudson-Smith and K. Chalkias, "Crypto Collectibles, Museum Funding and OpenGLAM: Challenges, Opportunities and the Potential of Non-Fungible Tokens (NFTs)", in Applied science vol. 11, 9931, 2021. https://doi.org/10.3390/app11219931

[15] F. Liddell, "Building Shared Guardianship through Blockchain Technology and Digital Museum Objects", in Museum & Society, vol. 19, no. 2, 2021, pp. 220-236. http://dx.doi.org/10.29311/mas.v19i2.3495

[16] L. Mucchi, M. Milanesi and C. Becagli, "Blockchain technologies for museum management. The case of the loan of cultural objects", in Current Issues in Tourism, vol. 25, no. 18, 2022, pp. 3042-3056. https://doi.org/10.1080/13683500.2022.2050358

[17] Terence Eden, Mona Lisa, available at: https://verisart.com/works/23f2c64a-08c6-4a42-8013-84ac8422dffb (last accessed 23. 8. 2022.).

[18] D. Trček, "Cultural heritage preservation by using blockchain technologies", in Heritage Science vol. 10, no. 6, 2022. https://doi.org/10.1186/s40494-021-00643-9

[19] A. Anagnostakis, "Towards a Blockchain Architecture for Cultural Heritage Tokens", in Transdisciplinary Multispectral Modeling and Cooperation for the Preservation of Cultural Heritage, First International Conference, TMM_CH 2018, Athens, Greece, October 2018, pp. 541–551. DOI:10.1007/978-3-030-12957-6_38

[20] M. Angelova, "Application of Blockchain Technology in the Cultural and Creative Industries", 2019 II International Conference on High Technology for Sustainable Development (HiTech), Sofia, Bulgaria, October 2019, pp. 1-4. DOI: 10.1109/HiTech48507.2019.9128267.

[21] Y.-C. Wang, C.-L. Chen and Y.-Y. Deng, "Museum-Authorization of Digital Rights: A Sustainable and Traceable Cultural Relics Exhibition Mechanism", in Sustainability vol. 13, 2046, 2021a. https://doi.org/10.3390/su13042046

[22] Y.-C. Wang, C.-L. Chen and Y.-Y. Deng, "Authorization Mechanism Based on Blockchain Technology for Protecting Museum-Digital Property Rights", in Appl. Sci. vol. 11, 1085, 2021b. https://doi.org/10.3390/app11031085

[23] L. Zhao, J. Zhang, H. Jing, "Blockchain-Enabled Digital Rights Management for Museum-Digital Property Rights", in Intelligent Automation & Soft Computing, vol. 34, no. 3, 2022, pp. 1785-1801. DOI:10.32604/iasc.2022.029693

[24] A. Whitaker, A. Bracegirdle, S. de Menil, A. M. Gitlitz and L. Saltos, "Art, antiquities, and blockchain: new approaches to the restitution of cultural heritage", in International Journal of Cultural Policy, vol. 27, no. 3, 2021, pp. 312-329. DOI: 10.1080/10286632.2020.1765163

[25] M. Gallagher and M. G. Fuentes, "Key blockchain projects to improve performance at the British Museum", UCL School of Menagement, 2020, available at: https://www.mgmt.ucl.ac.uk/sites/default/files/upload/pdfs/British%20Museum%20-%20Blockchain%20.pdf (last accessed 9. 5. 2022).

[26] F. Liddell, "The Crypto-Museum: Investigating the impact of blockchain and NFTs on digital ownership, authority, and authenticity in museums", doctoral dissertation, University of Manchester, 2022.

[27] H. Geismar, "Cultural Property, Museums, and the Pacific: Reframing the debates", in International Journal of Cultural Property, vol. 15, no. 2, 2008, pp. 109-122. DOI: 10.1017/S0940739108080089

[28] K. Siountri, E. Skondras, D. D. Vergados, „Towards a Smart Museum using BIM, IoT, Blockchain and Advanced Digital Technologies", ICVISP 2019: 3rd International Conference on Vision, Image and Signal Processing, Vancouver BC Canada, August 2019. DOI: 10.1145/3387168.3387196

[29] TIMELESS REFLECTIONS. THE ORIGINAL EGON SCHIELE NFT COLLECTION, available at: https://www.leopoldmuseum.org/en/press/news/1249/TIMELESS-REFLECTIONS-THE-ORIGINAL-EGON-SCHIELE-NFT-COLLECTION (last accessed 24. 8. 2022).

[30] E. Ertürk, M. Doğan, Ü. Kadiroğlu and E. Karaarslan, "NFT based Fundraising System for Preserving Cultural Heritage: Heirloom", 2021 6th International Conference on Computer Science and Engineering (UBMK), Ankara, Turkey, October 2021, pp. 699-702. DOI: 10.1109/UBMK52708.2021.9559006

[31] Artsy Gallery Insights: 2022 Report, availeble at: https://partners.artsy.net/resource/2022-gallery-insights-report/ (last accessed 24. 8. 2022).

Digital Strategies of Service Sector Enterprises: Promising B2B Decisions

O. Omelianenko*, V. Omelyanenko*, **, ***,, O. Kudrina*, I. Zihunova*, D. Kostyrko*, S. Lytvynenko****
* Sumy State Pedagogical University, Sumy, Ukraine
** Estonian Entrepreneurship University of Applied Sciences, Tallinn, Estonia
*** Institute of Industrial Economics of National Academy of Sciences of Ukraine, Kyiv, Ukraine
**** Sumy State University, Sumy, Ukraine
omvitaliy@gmail.com

***Abstract* – The study is devoted to digital strategies of enterprises in the service sector. The key trends of digitalization of business processes of enterprises in the service sector have been determined. The results of a pilot survey of enterprises in Sumy region of Ukraine are presented. Prospective innovative strategies of enterprises in the service sector are proposed. As part of the process approach, we propose to consider a digital agency within the system of business processes management, the ultimate goal of which is the production or supply of software, or the provision of applied IT services.**

Keywords – service sector, digitalization, business model, strategy

I. INTRODUCTION

Digital transformation is the introduction of modern technologies in the formation and maintenance of the enterprise's business model. From the strategic point of view of creating a business model, it is necessary to understand that digital transformation involves not only the installation of modern equipment or software, but also fundamental changes in management approaches, corporate culture and communications. As a result, productivity and customer satisfaction levels increase, and the company gains a reputation for being progressive and modern.

Digital transformation as a result of the application of information and communication technologies (ICT) is expressed in the creation of unique business models, a qualitative change in the perception of a product or service by customers, a fundamental acceleration and simplification of the operational activities of the participants of the innovation network. It should be taken into account that in modern conditions, the company's success is determined not only by the use of unique technologies, but also by the digitalization strategy. Companies that apply ICT without a previously formed strategy solve only tactical tasks. It is important to remember that ICT is only a means, not an end. In the case of an innovation system, the goal is always a unique product (service) created using such technologies.

Digitization of business processes is relevant not only at the level of individual enterprises. In modern conditions, certain industries choose this path of development as the only opportunity to meet dynamic market conditions. Thanks to this, the digital transformation of the service sector is already changing the life of every person and every company.

Therefore, the article is devoted to the practical issues of development of digital strategies of service sector enterprises based on B2B business model, especially within the cooperation with digital agencies.

II. LITERATURE REVIEW

Authors [1] have identified four types of sources of value creation: increasing efficiency due to reduction of transaction costs, reducing asymmetry of information, increasing the speed of transactions; complementarity: interdependence between products and services for customers (both vertically and horizontally), between online and offline business, between applied technologies and activity strategies; "binding" customers: in the conditions of e-business, the level of competition between enterprises increases, and strategies for creating various loyalty programs for customers, deepening customization, creating a unique offer for the customer as a sales driver are necessary; innovativeness: to achieve leadership, constant (daily) innovations are necessary: in services, content, etc. This idea can be confirmed empirically. As part of the study, e-business models were selected according to the economic logic of business [2]. Research [3] examines the theory of digital transformations and the theory of elements of ecosystems on the example of the development problems of the fintech sector, which is radically changing as a result of digital transformation.

The report [4] notes that digitalization has given rise to a new wave of innovation that will have profound consequences for humanity, changing the relationship between society, the state and business, and will also lead to the transformation of the structure of society and the economy. The pace of economic growth, labor productivity and the development of human potential will be increasingly determined by the level of integration into the digital economy. Indeed, digitalization not only creates new business opportunities, but also create a number of problems and risks. The study [5] examines the case of the Chinese software and services industry. It is shown that it has become a driving factor of economic growth and leading innovations. The study of the structure of an intercity network based on patent cooperation between cities helped to understand the positioning of cities from the point of view of innovations related to service

software (software service-related innovation).

The issue of the development of innovation platforms is very relevant in the framework of the ever-increasing cooperation in the creation of innovation value (innovation value co-creation), which makes it possible to introduce new products (services), increasing the efficiency of business processes and creating new value for customers [6], [7], [8], [9].

A scientific article [10], [11], [12], [13] illustrates the fact that increasing economic attention to knowledge concerns the development of so-called knowledge-intensive business services (KIBS). The researches are aimed to further understand the evolution and growth of KIBS and the nature of networks and markets. It is also argued that changing conditions of competitiveness increase the demand of enterprises in most sectors to innovate and use their core competencies and knowledge. The article [14] investigated the essence of the ICT category, its relationship with such concepts as the Internet of Things, the Industrial Internet of Things, cyber-physical systems, identified and analyzed the factors that determine the implementation of information and digital technologies within the framework of the concept smart manufacturing. Digitization enables the creation of administrative solutions, data protection systems, storage of research data and educational resources, as well as opportunities for better collaboration [15], [16], [17].

The studies [18], [19], [20] analyzed the impact of digitalization on the formation of advantages of new technologies and access to them, as well as determined their impact on organizational transformations in terms of cooperation and competitive advantages. In this context, for a better understanding of "agent-agent" interactions (actor-to-actor interactions), elements of systems thinking, determining the degree of openness of systems and the concept of managing communication flows are used [21], [22], [23].

Research results [24], [25] empirically support the hypothesis that ICT acts as a catalyst for innovation. However, the impact of digitalization on innovation activity depends on the type of ICT innovations, as well as the scale of their application. In particular, a strong positive relationship was found between ICT integrating administrative and industry software and innovative productivity. The identified pattern does not apply to market-oriented ICT, such as e-commerce or customer relationship management software. The results of the study also show that not every combination of ICT is advantageous from the point of view of the innovation process: firms that demonstrate basic use of ICT are not associated with a higher probability of innovation, while firms characterized by active use of ICT are those that who are more likely to innovate.

In the conditions of relying on fundamentally new knowledge and managing it during R&D, virtual formats of interaction can become the most common form of performing promising research and development. At the same time, it is natural to obtain the greatest scientific, technical and commercial effect by integrating these aspects into a single system with a single information and technological support.

III. FOODTECH CASE

The sphere of the restaurant business is not inferior to the sectors mentioned above, the integration of digital technologies in which a whole segment is already forming – Foodtech. The following main areas of Foodtech can be distinguished:

- delivery of food from restaurants and cafes through online ordering;
- food delivery through online ordering;
- the possibility of pre-ordering dishes by a certain time / ordering take-out food;
- availability of a large number of services with a wide database of restaurants and the possibility of ordering meals online.

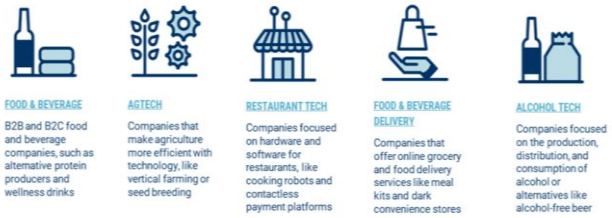


Fig. 1. Foodtech areas (CBINSIGHTS approach)

As a result, the digitalization of the business model of the restaurant business occurs in two directions: communication with the client goes online, and due to the large-scale introduction of automation, the number of personnel involved in service and management is reduced.

According to the research, one of the stable trends in recent years remains the growth in popularity of loyalty programs. Almost all restaurants now have special offers and marketing campaigns, and customers at least look at the restaurant's website, social media posts and read reviews before visiting. In addition, in conditions of growing competition, classic bonus programs are being replaced by more advanced promotions, programs with elements of gamification, etc. Loyalty programs are also moving to a digital format. Instead of plastic cards, mobile applications are appearing that combine a bonus system, the possibility of reservations, an information system, special offers, personal recommendations, food diaries and feedback tools.

Experts of the international IT project on pre-ordering in restaurants and cafes, Foodle, have identified the top 5 problems that inhibit the catering industry:

1. Low rates of development of the Foodtech market and digitalization of the restaurant environment;

2. Deterioration of the quality of food outside the home due to saving time;

3. Lack of effective analytical platforms that provide a wide range of guest behavioral analytics data that allow improving the quality of interaction;

4. The need to save time and increase the speed of serving guests in catering establishments;

5. The need for quick and convenient division of the bill between several guests when visiting the restaurant.

In contrast to the foreign restaurant market, where digitization has currently reached a high level (the share of online service penetration in the food segment in the USA is 3%, in the UK 7%), Foodtech in Ukraine is just beginning to be implemented - the share of online services is no more a few percent.

According to pilot survey of Sumy region of Ukraine, the restaurant business in 2021 has shown some growth, and the field of electronic commerce is also actively developing. Main survey results:

- 82% of surveyed businesses see the future opportunities in the digital technologies application;

- 42% are already using different digital promotion tools;

- 37% are not ready for the expanded use of foodtech technologies due to a number of reasons (lack of financial resources, lack of specialists);

- only 7% use the services of specialized agencies.

These factors indicate the need for active implementation and development of Foodtech, because in the near future the catering market is expected to change. In terms of practical aspects of implementing Foodtech, the main challenge facing the restaurant business will be to make ordering as easy as possible and to innovate to make ordering possible in one or two clicks.

To solve this task, it is necessary to analyze the market and the requests of the target audience. Current trends in the development of the e-commerce market and online orders dictate the need to use advanced analytics about the target audience. As a rule, restaurants do not have such statistics at their disposal and, accordingly, do not have flexible and convenient tools for proactive management of the business model, and therefore sales volumes, sales promotion and revenue growth.

To build a business model and provide quality service, it is necessary to have the most complete information about the target consumer and to personalize data about him as much as possible. This will make it possible to effectively use advertising channels and adjust work processes in the right direction.

In these conditions, the implementation of the B2B model allows obtaining advanced analytics on users in any field, which opens up wide opportunities for business owners. The possibility of personalization, improvement of the quality of services, changes and expansion of the assortment make it possible to significantly improve the quality of the service. With the development of e-commerce and the Foodtech market, advanced analytics will be increasingly necessary in business. In other words, the automation of analytics allows the most effective management of the enterprise. You can accumulate all data on one platform and make decisions based on the real situation in real time. From this point of view, the main task of the restaurant management system is the management of information flows, and business intelligence systems allow storing and processing large amounts of data. Profitability of data-based business model occurs with the beginning of a systematic analysis of the enterprise's business processes for the purpose of strategic improvement. Thus, the preliminary analysis of business processes during the implementation of digital technologies in the Foodtech, in addition to potential conveniences for customers, opens up many new opportunities for the business. Among them, we can note the optimization of work and the possibility of scheduling loading, revenue growth due to reaching a new audience. In these conditions, the services and platforms of specialized business analytics allow to bring work with clients to a new progressive level, as well as to carry out effective marketing management.

IV. B2B-Services Business Model

Gartner experts believe that the success of digitization depends not so much on modern technologies, but on the availability of established production processes and systems. Skills, initiatives and business models are 15 times more important for success in digitalization than discretely implemented technologies.

It should be noted that in the past, in order to successfully enter the online market, it was enough for an entrepreneur to order a typical website, create several pages (landing pages) using one of the online services, purchase a subscription base (in the best case - collect), post several advertising posts on social media networks and understand the basics of SEO promotion. This ensured a sufficient level of profit. But in modern conditions, everything is changing, and Internet technologies are no longer enough for business, traditional advertising only wastes the budget, people's interest quickly fades, one wrong decision – and the subscriber base is no longer effective. At the same time, new digital communication channels are emerging with a wide range of opportunities for promotion, which leads to a change in the business model and shapes the field of digital marketing.

According to survey, the primary task of digital marketing is to identify and attract both traffic and potential customers. Below we have shown the most significant directions of digital marketing for increasing the economic efficiency of the service sector based on B2B business model, especially within the cooperation with digital agencies:

1. Influence Marketing: Companies or people capable of influencing consumer behavior are called opinion leaders. They formed the basis of the concept of digital marketing. It became possible to influence opinion leaders with the help of paid advertising, such as Facebook Advertising, Google Adwords, through complex sCRM systems (social CRM).

The main idea of the approach is that it is easier to deliver advertising content to the consumer who is actively looking for it than to the one to whom the content comes by chance. Analysis of the behavior of Internet users (Online Behavioral Advertising, OBA) consists in collecting information about the online activity of users on various devices and sites in order to deliver advertising messages in accordance with the interests, preferences and needs of recipients.

2. Collaboration: can be established between organizations, technical service providers, digital agencies

to optimize efforts and share resources. In today's conditions, an important factor in creating a marketing strategy is the personalization of advertising space with the help of digital tools.

3. Remarketing: plays a major role in digital marketing. This tactic allows marketers to publish targeted advertising, that is, to show an advertising message to a specific audience, whose representatives are constantly searching for some information on the network, in particular, information about products and services, or visit sites with other specific goals.

Thus, a digital agency is an enterprise capable of providing a restaurant (client) with an expert assessment and ensuring the implementation of ideas in the field of website creation, creativity, strategy, advertising and customer service. Digital agencies are not limited to creating websites. They work comprehensively on creating and strengthening the company's name on the market (and not only on the Ukrainian market), "promoting" it in the media: on forums, partner sites, portals, social networks, Instagram and even on radio and television, using effective viral advertising.

From the point of view of building a business model, it is important that, unlike a regular web studio, the agency's work does not stop when the project promotion ends. Representatives of such an enterprise can continue to develop the project for years, promote it in the media, striving to improve its image and customer loyalty as much as possible. In the competitive restaurant business, it is very important to meet the demands of the modern consumer. According to a number of studies, more than 80% of the restaurant's target audience checks the restaurant's website on a smartphone. Most users in big cities are just looking for a place nearby to grab a bite to eat using their phone.

Thus, one of the components of the success of restaurants is the effective use of marketing opportunities in the Internet space. We consider such practical aspects of building a business model for increasing the economic efficiency of restaurants:

1. Attracting customers with the help of unique Internet positioning. In modern conditions, there are practically no restrictions on creating one's own position, because the Internet provides practically unlimited opportunities for creating the image of the institution, for example, photos of interiors and employees at work or an open kitchen will form a clear image of the institution. It is also possible to make an institution of a certain orientation and collect visitors according to interests.

For example, one of the coffee shops in USA doubled sales by announcing that the establishment has now become a gadget-free zone. Visitors must turn off their devices before ordering. Thanks to this, the coffee shop received not only PR in the network and new visitors. And as practice has shown, turnover accelerated due to the fact that visitors, without being distracted by photos of food or being online, quickly had lunch and went on with their business, freeing up seats for new visitors.

2. Promotion of restaurants and cafes on Instagram. When promoting a cafe on Instagram, there are many opportunities to write about, from the timing of posts to their content, from tags to comments, working with reviews and inviting celebrities.

3. Promotion of the restaurant in Google maps. Promotion in Google maps is a rather voluminous task, and therefore it should be entrusted to specialists.

4. Advertising using geotargeting. After the channels are determined, it is necessary to consider the possibility of advertising the restaurant with geotargeting. It is possible to show advertising of a cafe or restaurant to those who are within a radius of 5 kilometers, for example, this way of advertising is much more effective than advertising for the whole city.

5. Lead generation, discounts and offers for setting up the site, applications and the public in social networks.

6. Positive reviews and crowd technologies. Crowd is a well-thought-out work with feedback covering different platforms. Probably, every city has aggregators and specialized sites where you can get reviews. It is also worth conducting work in Google Local Business, Facebook, posting reviews on your own website, in particular within the framework of the discount system.

Thus, these directions of digital marketing in Foodtech can be most effectively implemented through B2B interaction with specialized digital agencies.

V. CONCLUSIONS

From the point of view of increasing the economic efficiency of the business model, process approach allows:

- to understand the processes, increase their quality and the quality of used resources;
- improve coordination of work between employees, reduce actions duplication;
- impose targeted responsibility for the performance of not just functional duties, but for the implementation of a specific process;
- measure process results through KPIs and stimulate their improvement;
- create an effective employee motivation system.

In a digital agency working to promote a restaurant, a business process can be considered as a set of operations that create a software and information service that has value for the restaurant and its customer. The consumer of the result of the business process can also be another business process of the system. All business processes of an IT company form a network of works performed by structural elements located at different levels of the enterprise's organizational structure.

As part of the process approach, we propose to consider a digital agency as an IT enterprise as a system of business processes, the ultimate goal of which is the production or supply of software, or the provision of IT services. At the same time, one of the key aspects of this approach is to ensure visibility of the management object (IT organization or system) with the help of its accurate, sufficient, concise, convenient for perception and analysis description. Accordingly, we can formulate the following areas of using the process approach to building the digital services business model based on B2B interaction:

1. Features of the organization of the work of a digital agency: increasing the efficiency of work on projects and avoiding problems with overdue deadlines; organization of an effective team; increasing sales in crisis conditions.

2. Possibilities of digital agency IT services: organization of the customer base in CRM and distribution among managers; establishing communication between employees: managers, copywriters, Internet marketers; establishing work with remote specialists; organization of joint work with the client on projects.

In the conditions of relying on fundamentally new knowledge and managing it during R&D, virtual formats of interaction can become the most common form of performing promising research and development. At the same time, it is natural to obtain the greatest scientific, technical and commercial effect by integrating these aspects into a single system with a single information and technological support.

ACKNOWLEDGMENT

Research was funded by Ministry of Education and Science of Ukraine within the projects «Organizational and economic support of the post-war sustainable development of territories based on the infrastructure and service methodology of innovation communities' development», LET EDU, "Territory of innovations: best practices for sustainable development at the local level" and Estonian National scholarship program for foreign students, researchers and lecturers in Estonian Entrepreneurship University of Applied Sciences.

REFERENCES

[1] R. Amit, and C. Zott, "Value Creation in e-Business," Strategic Management Journal, Vol. 22, № 6–7, pp. 493–520, 2001.

[2] L. M. Applegate, "E-Business Models: Making Sense of the Internet Business Landscape," The Academy of Management Journal, Vol. 39, Iss. 1, pp. 216–234, 2006.

[3] K. Boratyńska, "Impact of Digital Transformation on Value Creation in Fintech Services: An Innovative Approach," Journal of Promotion Management, 25:5, pp. 631–639, 2019.

[4] UNCTAD, "Information Economy Report 2017: Digitalization, Trade and Development," United Nations publication. Sales No. E.17.II.D.8. New York, Geneva, 2017.

[5] L. Zhang, Zh. Cao, G. Chen, and Zh. Wang, "A study of China's inter-city networks for innovation cooperation within software and service firms," Eurasian Geography and Economics, Vol. 60, Iss. 5, pp. 582-615, 2019.

[6] A. Aine, T. Björkroth, and A. Koponen, "Horizontal information exchange and innovation in the platform economy – a need to rethink?" European Competition Journal, Vol. 15. Iss. 2-3, pp. 347-371, 2019.

[7] B. Hansen, "The digital revolution – digital entrepreneurship and transformation in Beijing," Small Enterprise Research, 26:1, pp. 36–54, 2019.

[8] L. Melnyk, O. Kubatko, I. Dehtyarova, O. Matsenko, and O. Rozhko, "The effect of industrial revolutions on the transformation of social and economic systems," Problems and Perspectives in Management, Vol. 17 (4), pp. 381-391, 2019.

[9] L. Melnyk, O. Kubatko, V. Piven, K. Klymenko, and L. Rybina, "Digital and economic transformations for sustainable development promotion: A case of OECD countries," Environmental Economics, 12(1), pp. 140–148, 2022.

[10] T. Tirto, Y. Ossik, and V. Omelyanenko, "ICT support for Industry 4.0 innovation networks: education and technology transfer issues," Lecture Notes in Mechanical Engineering. Springer, Cham, 2020.

[11] V. Omelyanenko, "Preconditions analysis of using of technological package concept for development strategy of space metallurgy," Metallurgical and Mining Industry, No. 8, pp. 508–511, 2015.

[12] V. Omelyanenko, "Analysis of strategycal aspects of technology transfer in metallurgy," Metallurgical and Mining Industry, № 12, pp. 394–397, 2015.

[13] Y. Robul, I. Lytovchenko, L. Tchon, O. Khanova, and O. Omelianenko, "Digital marketing tools in the value chain of an innovative product," International Journal of Scientific and Technology Research, 9(4), pp. 158–165, 2020.

[14] M. Ghobakhloo, "Determinants of information and digital technology implementation for smart manufacturing," International Journal of Production Research, 2019.

[15] J. Khalid, B.R. Ram, M. Soliman, A.J. Ali, M. Khaleel, and M.S. Islam, "Promising digital university: a pivotal need for higher education transformation," International Journal of Management in Education, Vol. 12, № 3, pp. 264-275, 2018.

[16] V. Omelyanenko, O. Prokopenko, O. Kudrina, M. Biloshkurskyi, and O. Omelyanenko, "Digital Component of Innovation Landscapes: Con of Sustainable Development at the Local Level," 2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 – Proceedings, pp. 1324–1328, 2021.

[17] S. Muegge, "Platforms, communities and business ecosystems: lessons learned about technology entrepreneurship in an interconnected world," Technology Innovation Management Review,February, pp. 5–15, 2013.

[18] B. Quattrociocchi, M. Calabrese, X. Hysa, and E. Wankowicz, "Technology and innovation for networks," Journal of Organisational Transformation & Social Change, Vol. 14, Iss. 1. pp. 4-20, 2017.

[19] V. Omelyanenko, O. Prokopenko, O. Kudrina, M. Biloshkurskyi, and O. Omelyanenko, "Digital Component of Innovation Landscapes: Con of Sustainable Development at the Local Level," 2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 – Proceedings, pp. 1324–1328, 2021.

[20] T. Eisenmann, "Opening platforms: how, when and why?" Boston: Harvard Business School, 2008.

[21] T. Hughes, and M. Vafeas, "Marketing Agency/Client Service-For-Service Provision in an Age of Digital Transformation," Journal of Business-to-Business Marketing, 26:3-4, pp. 265–280, 2019.

[22] V. Omelyanenko, and D. Volodin, "Nanoinformatics application framework for R&D and industrial analisys," 2017 IEEE 7th International Conference Nanomaterials: Application & Properties (NAP), Odessa, UKraine, pp. 01NNPT03-1-01NNPT03-4, 2017.

[23] O. Prokopenko, and V. Omelyanenko, "Marketing aspect of the innovation communications development," Innovative Marketing, № 14 (2), pp. 41–49, 2018.

[24] P. Santoleri, "Diversity and intensity of information and communication technologies use and product innovation: evidence from Chilean micro-data," Economics of Innovation and New Technology, Vol. 24. Iss. 6, pp. 550-568, 2015.

[25] R. Geissbauer, E. Lubben, S. Schrauf, and S. Pillsbury, "Digital Champions. How industry leaders build integrated operations ecosystems to deliver end-to-end customer solutions," Pricewaterhouse Coopers, 2018.

[26] J. Calof, G. Richards, and J. Smith, "Foresight, Competitive Intelligence and Business Analytics – Tools for Making Industrial Programmes More Efficient," Foresight, vol. 9, no 1, pp. 68–81, 2015.

[27] S. Galloway, and S. Dunlop, "A Critique of Definitions of the Cultural and Creative Industries in Public Policy," International Journal of Cultural Policy, Vol 13, no 1, pp. 17–31, 2007.

[28] R. Hill, and L. W. Johnson, "Understanding creative service: a qualitative study of the advertising problem delineation, communication and response (APDCR) process," International Journal of Advertising, № 23:3, pp. 285–307, 2004.

[29] S. Nambisan, "Why Service Business are not Product Businesses," MIT Sloan Management Review, Vol. 42, № 4, pp. 72–80, 2001.

Factoring economic biases out of tokenomics

T.B. Tommaso Baldo *, M.M. Mauro Migliardi *,
* University of Padova, Padova, Italy
tommasobaldo@gmail.com mauro.migliardi@unipd.it

***Abstract*—Blockchain technology has enabled decentralized applications and peer-to-peer networks to be sustained by an incentive system based on tokens. At the same time, reputation is a paramount component for any limited-trust or no-trust system. However, in many systems, e.g., proof-of-stake-based systems, tokens are a proxy for reputation, therefore there might be confusion between wealth and reputation. In this paper, we focus on the social layer of dApps and present a two-token system that allows a clear-cut separation between wealth and reputation and fits the case of DAOs where user contribution is essential to maintain and sustain the platform, such as social apps where users are asked to vote about newcomers' acceptance or to moderate and filter the content proposed by others. The paper delves into the implementation of such tokenomics, describes how the standard ERC-20 token can be adjusted to become a trustworthy reputation meter, and how the model can be optimized to reduce transaction fees. The value of the proposed work lies in moving away from a wealth-affected reputation in favor of a contribution-based one. Using such a system, dApps can reward users more fairly.**

Keywords*—blockchain, decentralized applications, tokenomics, token, reputation

I. INTRODUCTION

Reward mechanisms have always been a key feature in p2p networks, as they motivate nodes to participate and maintain the stability of the network. The arrival of blockchain technology has enabled the creation of a new generation of incentive systems, based on token incentives and completely free from a central authority. Blockchains themselves are the first examples of such new systems since consensus mechanisms, i.e. the engines of decentralized networks, are based on token-incentive. As an example, Bitcoin and Ethereum motivate nodes in mining or validating new blocks with a consistent token based economic reward.

With regard to the network layer, an incentive schema is often implemented to motivate nodes in sharing bandwidth and storage within the network [1], as in the case of Swarm [2] or Filecoin [3]. The underlying rationale for this concept is that nodes are selfish agents but can be cooperative if an incentive system manages to align the personal interest with the community's one.

Blockchain technology has also generated innovation at the application level seeing as how it has enabled the introduction of a new approach, known as Social Finance (or SocialFi), that makes users able to tokenize their social activities and profit from them [4]. Another noteworthy use case is TCR (Token-curated Registry) where users are in charge of maintaining a high-quality list of items about a certain topic and are rewarded according to their contributions [5]. Rewarding can be leveraged to steer user behavior toward specific goals, as studies underline an already established trend where users are always more motivated to invest time and real money to get a better experience or virtual rewards on online platforms, such as social networks or games [6] [7]. Such systems were already present in platforms like Yahoo! Answer, Stack Exchange, and Quora before the advent of blockchain technology. However, in centralized systems, users collect *social capital*, which is not directly convertible into an economic reward and its scope is limited inside the specific platform [8]. Instead, token-based incentives are directly spendable and they are not limited to a platform. For example, users can trade tokens for other tokens or fiat currencies using Decentralized Exchanges (DEX).

However, token-based incentives are still far from perfect, especially regarding the application layer. For example, several studies highlighted how the inconsistency of token value, which is generally highly influenced by market trends, conditions the user contribution. In other words, users are more active when the economic rewards are higher [9] [10] [11] [12]. Besides, another relevant issue is coin voting. Generally speaking, in DAOs (Decentralized Autonomous System), TCRs, and other mechanisms that involve decentralized community voting, a user's vote is weighted according to his/her wealth. In this way, a platform can efficiently be protected from Sybil attacks but, at the same time, voting power is likely to be condensed into a few wallets [13]. If this happens, the result is more akin to a plutarchy than it is to a democracy. The issue becomes even worse if we consider the fact that a user can buy tokens from external sources, e.g., other users or DEXs, and hence becomes the richest and most influential user without any significant involvement in the platform [14].

The research question we want to investigate is how we can factor economic biases out of reputation but still be able to provide attractive rewards to users. In this paper, we propose a hybrid system where reputation is a trustworthy measure of a user contribution while a token system rewards users according to their reputation. In this way, the reputation value is enhanced because it can offer tangible rewards while the incentive schema is not influenced by an economical factor but is strictly merit-based. The solution we propose is designed for DAOs or DeCMS (Decentralized Content Management System) where users are asked to participate in community activities like voting proposals or moderating the content. However, we think our proposal has a more general value as it can be used in

generic fully-decentralized applications to steer users into contributing both at the networking layer, as nodes, and at the application one, as consumers.
The paper is organized as follows: in section II we explore and compare existing solutions to underline the gap we aim at covering in this paper. Then, in section III we describe our proposed solution and give an insight into the actual implementation. Later, in section IV we analyze the proposed solution in a real-case scenario, in section V, we briefly discuss some potential scenarios and disclose some aspects we would like to delve into as future work.

II. STATE OF THE ART

Since blockchain technology enables a transparent rewarding phase, it has been widely adopted in OSNs (Online Social Network) where rewards are considered a key factor to give value to content and reward the creators [15] [16]. In this section, we will describe several significant examples.

A. Reddit Community Points

As the first example, we may cite Reddit Community Points (RPC). Reddit is one of the most popular Online Social Media (OSM). Its users, also known as *Redditors*, can aggregate into topic-specific communities, namely *subreddits*, where they can discuss and share opinions. Despite the centralized architecture of the platform, Reddit introduced blockchain tokens owned and controlled by the communities themselves. For example, the subreddit about Ethereum has its own implementation of RPC, called DONUTS. RPCs can be spent to buy aesthetic perks, advertising spots, and tipping other users for content creation; furthermore, DONUTS ownership increases a user's vote weight in community polls. There are two ways to earn RPCs: interacting in the subreddit and buying from popular decentralized exchanges like Uniswap or Honeyswap. As a result, the token cannot be considered a trustworthy measure of user involvement in the platform. This claim has been confirmed in [15], where it emerges that users are more likely to sell DONUTS on DEX instead of using them to redeem perks or advertising spots.

B. Steemit

As a further example, we have Steemit [17]. It is a blockchain-based OSN with a complex tokenomics to manage user rewards. It is made of three different tokens:

- STEEM: a liquid cryptocurrency to enable payments between users.
- STEEM Power (SP): tokenization of voting power. A user can earn SP by staking STEEM. SPs are not transferable but can be converted again into STEEM in a 13-week-long process.
- Steem Blockchain Dollar (SBD): a token pegged to the value of the US dollar to provide stability.

Users can earn STEEM thanks to content creation and curation. The latter is based on a staked-based voting system. Indeed, a curator can choose how much voting power is put behind a vote. Obviously, the reward will be proportional to the weight [8]. In such a system, the richest users, namely *whales*, are highly influential. As a result, the original idea has undergone a significant alteration, resulting in a system where often users ask to be paid for voting a certain content or they use bots to automate the curation process and hence earn tokens in an almost passive way [15].

C. Yup

Yup [1] is a social consensus protocol. Its novelty is that it is not a standalone platform, like Steemit, but it is built to work on top of existing OSM. In other words, it is a feed aggregator. In this way, a user can join Yup and, at the same time, enjoy the content he/she likes on whatever platform. The content curation works as follows: a user can like or dislike, even multiple times, every kind of online content. Content must be popular to be ible to be rewarded, i.e. other users have to agree. After the content is rated, the reward is split 50% to the creator and the remaining 50% to the curators. Curators' rewards are again split according to users' influence. The latter is based on a user's token holdings and previous rewards received. To summarize, Yup proposes a staking-based content curation that has the same inherent problems we have described with regard to Steemit.

D. MeritRank

As the last example, we may cite MeritRank [18]. The authors propose a merit-based and Sybil-resistant reward system to motivate user contribution in DAOs. MeritRank is based on feedback aggregation, i.e a peer's reputation is computed on the feedback other nodes provide about him. The proposed model is interesting because it moves away from a staking-based mechanism. However, we argue that feedback aggregation may not be the best fit for DAOs. Indeed, DAOs have a users-to-community model where members are usually asked to vote about new proposals. As a result, user-to-user relationships may be sporadic. Feedback aggregation is a better fit in a scenario like an OSN or a p2p network where users are *prosumers*, i.e. they are producers and consumers, depending on circumstances. Considering these facts, in section III we propose a model that does not adopt feedback aggregation.

E. Other solutions

Other similar solutions are described in the scientific literature. For example, Cai et al. [19] propose an oracle protocol to reward only honest reporting, avoiding in this way the phenomenons like lazy voting or bandwagon effect where users only aim at agreeing with the majority to be awarded. PaySense [20] describes a system where Bitcoin is meant to be a measure of reputation and a reward, simultaneously. We can find other use cases such as content-sharing [21] [22] [23] [24], and educational

[1] https://yup.io/

platforms [25], supply chains [26], and recommender systems [27]. Table I summarizes the main solutions.

III. OUR PROPOSAL: DESIGN AND IMPLEMENTATION

From the literature review, it emerges how creating a single ERC-20 token is a fast and efficient solution to set up a tokenomics. Nevertheless, it is also a limiting model if we aim at building a merit-based platform. A major contradiction is generated when a token is simultaneously signaling user contribution and compensating users because it allows buying and selling reputation. As a result, we have designed a two-token system to separate user contribution from its economic value.

A. Reputation Points

The first token, called *Reputation Points* is based on ERC-20 standard but with a relevant difference: it is non-transferable. In this way, users cannot trade reputation with each other and hence the token can be considered a trustworthy measure of a user contribution. In detail, the ERC-20 contract has a function *beforeTokenTransfer*. It is a hook that is called before any transfer of tokens. As a result, if the condition in the *require* clause is not satisfied, the transfer is blocked. Since we still want the contract to be able of minting and burning tokens, we have written the *require* clause such that it prevents every operation aside from the ones performed from or to the address 0x0, which identifies the contract itself. Any transfer that is not performed from or to the contract address is blocked. In this way, we have created a non-transferable token.
Thinking about a DAO use case, users can earn RPs by:

- creating a proposal.
- voting on a proposal.
- reporting an inappropriate item.
- moderating or filtering content.
- reporting a user who is not acting properly.

In other words, users are rewarded when they are involved in community activities. On the contrary, a user who is negligent with respect to his/her duties has to be punished. As a result, RPs are burned from a wallet when:

- a user doesn't vote.
- a user spams reports or produces a fake report.

The precise amount of RPs minted or burned is strictly related to the use case. Precise tuning is required to be in line with the platform requirements. For example, if a platform requires strict guidelines about the allowed content, it is possible to strengthen the punishment for off-topic posting to reinforce the community rules.

B. Community Token

As the second token, we have the Community Token (CT). It is a standard ERC-20 token and it is spendable to get rewards. The reward system has to be defined according to a use case but generally speaking, we can think of a Reddit-like system where users can buy aesthetic perks and advertising spots. CTs are also meant to be traded outside the platform or sent to other users as tips. The latter functionality could be useful in an OSN to make the creator-consumer relationship profitable for the creator. It is important to notice that CTs are generated from RPs. In this way, we give an intrinsic value also to RPs. Indeed, we have a sort of feedback loop because users are interested in rewards but the only way to redeem them is to earn RPs. Finally, users are motivated to complete activities that earn RPs.

C. Distribution mechanism

The distribution mechanism is responsible for generating CTs from RPs. It is managed by a smart contract, namely *Rewarder*. This design choice adds overhead but it allows to clearly separate the business logic from the token contracts and hence, it increases the upgradability and fault-tolerance of the whole system. We have identified two different ways to mint Community Tokens.

1) Staking-based lazy evaluation: a user stakes RPs on his/her wallet. Every time he/she earns new RPs, interests are computed according to the following formula:

$$CT = RP * r * t$$

where *r* is the interest rate and *t* is the period. Then, the amount of CTs is saved on a local mapping in the *Rewarder* contract. Whenever a user wants to redeem his/her CTs, he/she has to trigger the *Redeem* function and the equivalent amount of CTs will be transferred to his/her wallet. Making the conversion an on-the-fly process optimizes transaction costs because it reduces the number of transactions.

2) Airdrop: it is a periodic distribution. For each user, the amount of CTs he/she will receive is computed according to an allocation policy where the higher the reputation, the more CTs are minted. To optimize the process, it is preferable to select a large period. On top of that, it is important to notice that the function to start the airdrop has to be manually triggered by a user since scheduling a function is not built-in in the Ethereum ecosystem. During the implementation, we have faced another issue: ERC-20 contract associates users to their wallet balances thanks to a mapping. However, in Solidity mappings are not iterable and do not have a *length* property like arrays. As a result, it is not possible to know how many users are holding a certain token. To overcome this issue without changing the standard implementation, we have introduced a subscription mechanism in the *Rewarder* contract. In other words, before being eligible for an airdrop, each user has to register his/her address by calling *subscribeToAirdrop*, which will proceed to add the sender address to an array. When distributing the CTs, the *Rewarder* will iterate over that array. During the testing phase, we have taken into consideration also an alternative implementation, called *Aidrop&Redeem*. The difference is that during the airdrop process, this approach saves the CTs on a local mapping

TABLE I: General overview of incentive mechanisms

Protocol	Distinct Features	Transferable Token	Influence on voting	Notes
DTube [24]	Rewards creation and curation.	Yes	Yes	
Helios [23]	Rewards users according to contributions.	Yes	No	
HiDe [22]	Rewards creators and manage budget allocation.	Yes	Yes	
Innerlight [21]	Has financial bounties to push best responses.	Yes	No	There is no reputation mechanism.
MeritRank [18]	Is sybil-resistant and merit-based.	No	Yes	Feedback aggregation may not be a good fit for DAOs.
PaySense [20]	Uses Bitcoin as reward and reputation measure.	Yes	No	Reputation is easily tradeable because of Bitcoin.
Pranesh [26]	Rewards the supply chain parties.	Yes	No	There is no reputation mechanism.
RPC	Measures user contribution and interaction.	Yes	Yes	
Steemit [17]	Implements tokenization of user economic investment.	Yes	Yes	Direct connection between wealth and voting power.
YUP	Rewards content curation and moderation.	Yes	Yes	

instead of directly minting them. Users have to trigger a function to transfer CTs to their wallets. We wanted to explore if an "on-the-fly" computation could make the airdrop process lighter and hence optimize gas consumption.

IV. EXPERIMENTAL RESULTS

We have set up a local environment to test and compare the three implementations. We have relied on Ganache [2] to create a local and customizable blockchain instance. In detail, we customize the block time to 0.35 s and the gas price to 0.1 gwei, in order to simulate the Arbitrum blockchain and reproduce a real-world scenario. On top of that, Ganache is set to create 100 accounts to interact with smart contracts and test functionalities. The testing phase is managed by Mocha [3], a Javascript framework. We have run a unit test for each function involved in the example. For each execution, we registered the gas consumption and the execution time. In this regard, Table III shows the results obtained. The "user-specific" functions, like *addRPsToAccount* or *redeem* have been executed for each user, i.e. one hundred times. Instead, *airdrop* has been executed only once because it operates on multiple users simultaneously. In this way, we have obtained consistent data to make a comparison.

A. Use Case

Despite our solution being designed to work in general scenarios, we have selected a specific use case for testing. More in detail, we refer to an advertising platform with a voting-based registration mechanism. In other words, registered users have to vote about accepting or rejecting a newcomer every time a new user asks for a subscription. The platform displays user-generated content, so there is a need of moderating and filtering the content proposed by users. Since user participation is essential to make the portal work, we think it is a perfect example of how a well-designed reward system can steer user behavior.
We consider the following example:

1) A user asks for the subscription.
2) Registered users vote.
3) After the deadline is met, the *Rewarder* gives RPs to users who voted and it burns RPs from those who did not.
4) The *Rewarder* distributes CTs according to the policy. For example, Fig.1 shows *Airdrop* distribution process.

In the testing script, we set the interest rate equal to 0.07 and the airdrop period is equal to four weeks.

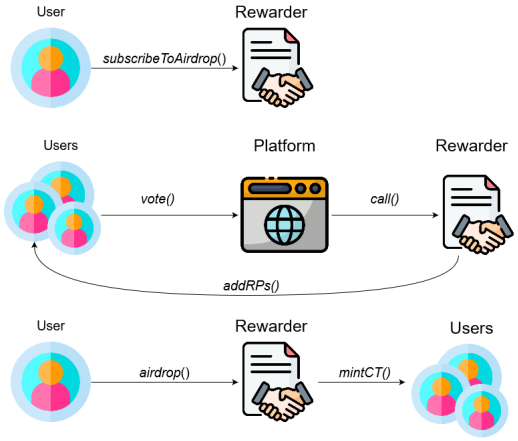


Fig. 1: Rewarding process following the airdrop implementation

B. Contract size

As we can see from Table II, the *Airdrop* implementation is lighter and hence it requires less gas to be deployed. It is worth noting that *Airdrop* has only an array, i.e. the users subscribed, while *staking-*

TABLE II: Deployment Costs

Implementation	Gas Used	Bytecode
Airdrop	590625	4784
Staking-based	638518	5228
Airdrop&Redeem	691868	5726

[2]https://trufflesuite.com/ganache/
[3]https://mochajs.org/

TABLE III: Analysis of costs and time required for each implementation to distribute RPs and CTs to 100 users

Implementation	Transactions	Gas	Time [ms]	Cost per user
Airdrop	*addRPsToAccount*	4170000	110000	75000 = 0.011 USD
	subscribeToAirdrop	4850000	107000	
	airdrop	3360000	5000	
Staking-based	*addRPsToAccount*	7480000	110000	130000 = 0.019 USD
	redeem	5850000	106000	
Airdrop&Redeem	*addRPsToAccount*	4180000	110000	145000 = 0.021 USD
	subscribeToAirdrop	4871000	105000	
	airdrop	3000000	2600	
	redeem	7300000	100000	

based lazy evaluation requires two mappings: one to store the amount of CTs accumulated and one to save the last computation of interests. On top of that, the redeeming process that mints new CTs requires a specific function and hence it increases the contract size. *Aidrop&Redeem* is the heaviest because it has both functions, the one for creating an airdrop and the one for redeeming CTs.

C. Transaction Costs

From the example described above, we consider only two functions: minting new RPs and minting new CTs. Indeed, the voting and registration mechanisms are the duties of the platform. Firstly, we have registered a significant difference between the first time a user receives RPs and further ones. For example, a first distribution requires 86582 instead of 74760 while using the *staking-based* implementation. The reason is that the first distribution includes the cost of initializing the data structures. To be clear, in Table III we have taken into consideration only successive releases. Secondly, Table III shows clearly that the *airdrop* mechanism is the most efficient. Indeed, by removing the interest computation we can halve the gas consumption required by *addRPsToAccount*. *Airdrop&Redeem* is the least sustainable model, we can notice that its alternative version of an airdrop is slightly cheaper but it loses all the convenience because of the cost of *redeem*. Finally, we have computed the gas each user has to spend to complete the process. In this regard, since the airdrop subscription is a one-time transaction we have not considered its cost for calculation. Later, we have converted the gas consumption into US Dollars to give the reader a better understanding. In detail, we have considered the gas price equal to 0.1 gwei and the ETH price equal to 1500 USD, as of the time of writing. The results are included in Table III.

D. Execution Times

Concerning the distribution of RPs, the execution times are similar. However, the airdrop mechanism is more efficient in distributing CTs. Also in terms of scalability, airdrop is a better solution. Indeed, it is a *for* cycle repeated every four weeks, while lazy evaluation is called every time a user earns new RPs. As a result, in use cases such as social networks, where the number of users is high and the interactions between users are even higher, lazy evaluation could be very inconvenient because of transaction costs and the time required. As a final note, *Airdrop&Redeem* halves the airdrop time by removing the minting process.

E. Discussion

Concerning performances, *Airdrop* is the most efficient and scalable among the different options. Indeed, if we consider the case in which a user is already subscribed to the airdrop, the cost of the whole process is almost equal to the cost of adding RPs in the *staking-based* implementation. As a result, the cost per user is halved and, even if the amount of money seems negligible, it is not when we think about a real-case scenario involving millions of users. From an economical point of view, the airdrop mechanism allows arbitrarily selecting the supply. In this way, the platform has control over the coin distribution. For example, it is possible to peg Community Token to a stable currency and hence ensure the tokenomics stability [28]. On top of that, the airdrop process can adopt several distribution policies, e.g. Pareto Distribution or the power law one, and hence it can adapt to different use cases and platforms' requirements. We claim a two-token system with airdrop distribution is a reasonable alternative to coin-based governance where users' wealth is too much influential on community decisions. Indeed, the reputation token we propose is merit-based and it could be easily implemented in a DAO voting mechanism. On a second note, a strength of this proposal is that both tokens are saved on a blockchain. Indeed, solutions like RPC describe a hybrid where information is saved on centralized platforms and then transferred to the blockchain through a minting process. Obviously, a fully-decentralized system like the one we propose is a better fit for decentralized apps. On top of that, it allows interoperability between different services or layers because the information saved on a blockchain is public. In this sense, our solution is similar to *Yup* because it can be implemented on different platforms.

V. CONCLUSION

In this paper, we described a two-token system to motivate users in community activities, e.g. moderating or voting, and reward them according to their contribution to the platform wellness. In comparison with existing systems, we introduce a non-transferable token to measure user contribution. The novelty lies in being able to fully separate economic wealth from actual social contributions, something that is not possible in many DAOs because of coin-based governance. To complete the system, we also added a spendable token to buy rewards, such as aesthetic perks or advertising spots. This token is distributed according to user reputation. In this way, the reward schema lends an intrinsic value to reputation and it generates a feedback loop that keeps users involved. Furthermore, we have studied the implementation by analyzing three different approaches according to deterministic metrics such as contract sizes, transaction costs, and execution time. From these results, we have understood that a monthly distribution, namely *Airdrop*, is an efficient and scalable method to mint and distribute tokens to users. Nonetheless, there still are some unresolved questions that need to be tackled in future works. For example, we could analyze different probability distributions to understand which is the best one for a specific use case. Furthermore, we could extend our analysis to other blockchains to understand which one is the best fit for implementing such a system. Finally, in a larger scope, it would be interesting to examine the consequences of the adoption of a platform capable of using Self-Sovereign Identity for user registration and authentication in terms of privacy and security.

REFERENCES

[1] E. Daniel and F. Tschorsch, "Ipfs and friends: A qualitative comparison of next generation peer-to-peer data networks," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 31–52, 2022.

[2] "Swarm: storage and communication infrastructure for a self-sovereign digital society," 2021. [Online]. Available: \url{https://www.ethswarm.org/swarm-whitepaper.pdf}

[3] "Filecoin: A Decentralized Storage Network," 2017. [Online]. Available: \url{https://filecoin.io/filecoin.pdf}

[4] B. Guidi and A. Michienzi, "Socialfi: Towards the new shape of social media," *SIGWEB Newsl.*, no. Summer, jul 2022. [Online]. Available: https://doi.org/10.1145/3545196.3545201

[5] J. Kaur and B. Visveswaraiah, "A brief survey of token curated registries," in *Emerging Technologies in Data Mining and Information Security*, A. E. Hassanien, S. Bhattacharyya, S. Chakrabati, A. Bhattacharya, and S. Dutta, Eds. Singapore: Springer Singapore, 2021, pp. 189–202.

[6] S. Greengard, "Social games, virtual goods," *Commun. ACM*, vol. 54, no. 4, p. 19–22, apr 2011. [Online]. Available: https://doi.org/10.1145/1924421.1924429

[7] M. Nazir and C. S. M. Lui, "A survey of research in real-money trading (rmt) in virtual world," *International Journal of Virtual Communities and Social Networking (IJVCSN)*, vol. 9, no. 1, pp. 34–53, 2017.

[8] M. Thelwall, "Can social news websites pay for content and curation? the steemit cryptocurrency model," *Journal of Information Science*, vol. 44, no. 6, pp. 736–751, 2018. [Online]. Available: https://doi.org/10.1177/0165551517748290

[9] S. Zhang, J. Park, and R. Ciriello, "The differential effects of cryptocurrency incentives in blockchain social networks," 12 2019.

[10] C. T. Ba, M. Zignani, S. Gaito, and G. P. Rossi, "The effect of cryptocurrency price on a blockchain-based social network," in *Complex Networks & Their Applications IX*, R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, Eds. Cham: Springer International Publishing, 2021, pp. 581–592.

[11] C. T. Ba, M. Zignani, and S. Gaito, "The role of cryptocurrency in the dynamics of blockchain-based social networks: The case of steemit," *PloS one*, vol. 17, no. 6, p. e0267612, 2022.

[12] A. Kiayias, B. Livshits, A. Mosteiro, and O. S. Thyfronitis Litos, "A puff of steem: Security analysis of decentralized content curation," 10 2018.

[13] V. Buterin, "Moving beyond coin voting governance," Aug 2021. [Online]. Available: https://vitalik.ca/general/2021/08/16/voting3.html

[14] B. Guidi, A. Michienzi, and L. Ricci, "A graph-based socioeconomic analysis of steemit," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 365–376, 2021.

[15] B. Guidi, "When blockchain meets online social networks," *Pervasive and Mobile Computing*, vol. 62, p. 101131, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574119220300195

[16] A. Bogliolo, P. Polidori, A. Aldini, W. Moreira, P. Mendes, M. Yildiz, C. Ballester, and J.-M. Seigneur, "Virtual currency and reputation-based cooperation incentives in user-centric networks," in *2012 8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2012, pp. 895–900.

[17] "Steem: An incentivized, blockchain-based, public content platform. whitepaper." August 2017, accessed on 16 January 2023. [Online]. Available: \url{https://steem.com/SteemWhitePaper.pdf}

[18] B. Nasrulin, G. Ishmaev, and J. Pouwelse, "Meritrank: Sybil tolerant reputation for merit-based tokenomics," in *2022 4th Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, 2022, pp. 95–102.

[19] Y. Cai, G. Fragkos, E. E. Tsiropoulou, and A. Veneris, "A truth-inducing sybil resistant decentralized blockchain oracle," in *2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, 2020, pp. 128–135.

[20] S. Delgado-Segura, C. Tanas, and J. Herrera-Joancomartí, "Reputation and reward: Two sides of the same bitcoin," *Sensors*, vol. 16, no. 6, 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/6/776

[21] H. Xu, Y. Cheng, X. Yang, and Z. Li, "Incentives in content-sharing network," *IOP Conference Series: Materials Science and Engineering*, vol. 1043, no. 5, p. 052015, jan 2021. [Online]. Available: https://dx.doi.org/10.1088/1757-899X/1043/5/052015

[22] "Introducing hide - an offline-first decentralized cms." [Online]. Available: \url{https://hide.ac/articles/m7l_Ft2li}

[23] B. Guidi, V. Clemente, T. García, and L. Ricci, "A rewarding model for the next generation social media," in *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*, ser. GoodTechs '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 169–174. [Online]. Available: https://doi.org/10.1145/3411170.3411247

[24] "D.tube: Turning the tables in social media industry." [Online]. Available: \url{https://token.d.tube/whitepaper.pdf}

[25] M. Sharples and J. Domingue, "The blockchain and kudos: A distributed system for educational record, reputation and reward," in *Adaptive and Adaptable Learning*, K. Verbert, M. Sharples, and T. Klobučar, Eds. Cham: Springer International Publishing, 2016, pp. 490–496.

[26] S. Pranesh, V. Kannan V., N. Viswanathan, and M. Vijayalakshmi, "Design and analysis of incentive mechanism for ethereum-based supply chain management systems," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2020, pp. 1–6.

[27] A. Lisi, A. De Salve, P. Mori, L. Ricci, and S. Fabrizi, "Rewarding reviews with tokens: An ethereum-based approach," *Future Generation Computer Systems*, vol. 120, pp. 36–54, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X21000480

[28] M. S. Kim and J. Y. Chung, "Sustainable growth and token economy design: The case of steemit," *Sustainability*, vol. 11, no. 1, 2019. [Online]. Available: https://www.mdpi.com/2071-1050/11/1/167

Adaptation of European Enterprises to COVID-19 Pandemic: Cluster Analysis Findings

M. Pejić Bach*, B. Jaković* and I. Jajić*
* University of Zagreb, Faculty of Economics & Business, Zagreb, Croatia
mpejic@net.efzg.hr; bjakovic@net.efzg.hr; ijajic@net.efzg.hr

Abstract— The Covid 19 pandemic had a significant impact on every aspect of business and personal life. The use of digital technologies was an important lever in overcoming barriers that resulted from the need for physical distance. Businesses made greater use of various digital technologies such as electronic communications, remote use of computers in businesses, and online meetings. In addition, e-commerce has become the primary venue for selling products and services. However, differences between countries became apparent in the response to the Covid 19 pandemic. This paper aims to explore the digital divide between European countries based on their adaptation to the Covid 19 pandemic using a cluster analysis. The Eurostat database is used as a source for the variables measuring the response of firms to the Covid 19 pandemic in terms of communication tools and e-commerce usage. K-Means cluster analysis is used to analyze the homogeneity of countries based on the observed variables. Combined with the LSD post-hoc test, the Anova analysis compares clusters by GDP per capita. The results show that business agility is strongly associated with the country's economic development.

Keywords: Covid-19, pandemic, e-commerce, response, adaptation, Europe, cluster analysis, digital divide

I. INTRODUCTION

Numerous research papers have examined the impact of COVID -19 from a variety of perspectives and geographic locations. In the short term, for example, severe impacts on lost sales, business closures, mass layoffs, and liquidity have been documented in various countries [1]; [2]. The impact on businesses is multifaceted, with sluggish demand most often described as the biggest problem in recent business surveys and becoming more significant over time [3]. While there is considerable uncertainty about the long-term impact of the pandemic, evidence from previous crises suggests the possibility of lasting scarring effects. The impact of the shock on reallocation is expected to persist long after the Covid 19 epidemic has subsided [4]. according to estimates by [5], liquidity shortages occur for firms under various shock scenarios, regardless of age, size, or productivity. The uncertainty could further hamper economic activities. During this epidemic, unprecedented levels of uncertainty have been observed. Companies have been shown to respond by cutting spending on innovation and general management improvements, which is likely to affect future productivity development [6]. However, companies have responded to the crisis with digital transformation [7]; [8], which is already being deployed in various business areas such as human resources [9]; [10], education [11], and smart factories [12].

This paper analyzes the behavior of 23 European countries based on data collected by Eurostat. Variables include companies' remote access to their email system, information and communication technology (ICT) systems other than email systems, meetings, and platforms for selling goods or services. To complement the country comparison on a global level, an economic indicator, gross domestic product (GDP) per capita, was also used [13].

The study uses K-means clustering [14], which has been widely used to study the homogeneity of countries according to various indicators of technological development ([15]; [16]; [17]; [18]; [19]; [20]; [21]). Anova analysis is used to compare the mean values of variables between clusters. In similar studies, Anova analysis is applied to investigate the impact of digitalization, especially remote access in business environment [22]. In addition, an analysis is applied to examine the differences between clusters by GDP per capita, followed by a post-hoc Lease Significant Difference (LSD) test [23]; [24].

The article' is organized as follows: After a brief introduction in the first section, the methodology section details the research technique. The second section consists of the K-Means cluster analysis and the Anova comparison of the variables between the clusters. The third section examines the relationship between country clusters and GDP per capita. The final section provides the concluding observations, limitations, and recommendations for future research.

II. METHODOLOGY

A. Data Sources

The research variables come from the Eurostat database, shown in Table 1, which describes the use of remote access by companies in their daily work environment affected by the Covid 19 pandemic [25]. This study is concerned with four variables that measure the company's remote access to its email system (ENT1), access to ICT systems other than email systems (ENT2), remote meetings (ENT3), and use of online platforms for selling goods or services (ENT4). The dataset consists of 23 European countries, with 2021 being the first year after the Covid 19 pandemic. In addition, the variables are presented in Table 1 in the percentages of firms with at least 10 employees, where 2021 is the period studied. The GDP per capita (GDPpc.)

variable is also included and is expressed in U.S. dollar (USD) currency.

TABLE 1. RESEARCH VARIABLES, % OF ENTERPRISES WITH 10+ EMPLOYEES

Variable description	*Measurement*
ENT1. Companies with an increase in remote access to the company's email system due entirely to the Covid-19	% of enterprises; 10+ employees; year 2021
ENT2. Companies with an increase in remote access to company ICT systems other than email, which was entirely a consequence of the Covid-19	% of enterprises; 10+ employees; year 2021
ENT3. Companies with an increase in the number of remote meetings, which was entirely caused by the Covid-19	% of enterprises; 10+ employees; year 2021
ENT4. Due to the Covid-19, in 2020, the company started or increased its efforts to sell goods or services online.	% of enterprises; 10+ employees; year 2021
GDPpc. GDP per capita (USD)	USD per inhabitant; the year 2021

Source: Authors' work

Eurostat data measuring business-to-business processes and procedures with remote access (Table 2). The variable „Enterprises with an increase in the number of remote meetings exclusively due to the Covid 19 pandemic (ENT3)" has the highest percentage in all countries studied. On the other hand, the variable „Companies with an increase in remote access to the company's email system that is exclusively due to the Covid 19 pandemic (ENT1)" has the lowest percentage of the four variables in all countries in the sample.

TABLE 2. RESEARCH VARIABLES INCLUDED COUNTRIES AND GDP PER CAPITA DATA

Country	*ENT1.*	*ENT2.*	*ENT3.*	*ENT4.*	*GDPpc.*
Austria	21	22	43	19	53.268
Belgium	13	18	46	16	51.768
Bosnia and Herzegovina	7	5	12	11	6.917
Bulgaria	9	9	14	9	11.635
Croatia	7	10	26	8	17.399
Cyprus	17	16	31	23	30.798
Finland	12	13	48	17	53.983
Germany	15	15	30	5	50.802
Hungary	3	3	19	14	18.773
Italy	14	16	27	19	35.551
Latvia	8	9	19	10	20.642
Lithuania	6	7	24	10	23.433
Malta	23	28	38	32	33.257
Montenegro	8	9	20	14	9.367
Netherlands	15	17	48	17	58.061
Norway	5	7	57	12	89.203
Poland	10	14	23	6	17.841
Portugal	9	10	29	21	24.262
Serbia	11	10	21	10	9.215
Slovakia	8	11	19	11	21.088
Slovenia	6	8	28	7	29.201
Sweden	8	12	46	20	60.239
Turkey	9	7	10	14	9.587

Source: Authors' work

B. Statistical analysis

Data refer to firms with more than 10 employees, excluding the financial sector and accounting. To complement the country comparison at the global level, an economic indicator, gross domestic product (GDP) per capita, is also used. GDP is a good economic indicator for this research topic and can therefore be used as an efficient benchmark for the development of a given country in 2021 [26].

To classify the countries, a non-hierarchical statistical cluster analysis is performed. To obtain the initial centroids or estimates, the k-means grouping algorithm and the largest average distance technique are used [27]. After the variables are analyzed, they are assigned to the nearest centroid; new centroid coordinates are generated [28]. The Euclidean squared distance is used to calculate the distance between the center and the 'position of the variables on the map, in order to give more weight to the variables that are farther from the center of the map [29]. This grouping approach has been run a total of fifty times and aims to investigate the homogeneity of the dataset.

In addition to k-means clustering, the authors of this study use an Anova analysis, which is used to study the impact of digitization, especially remote access in the enterprise environment [30]. First, using Anova analysis [31], we compare the mean values of the observed variables that measure the response of countries' to covid pandemics in the different clusters. Second, we compare GDP per capita across clusters using Anova analysis and the LSD test. To examine the assumptions of the Anova analysis, the normality of the distribution of GDP per capita is tested. Also, the homogeneity of the variance of the same variable is examined when comparing the clusters.

All of this is preceded by an individual normalization of the variables using the built-in data normalization function in the Statistica statistical software used for this research analysis. Normalization involves translating the lowest and highest values of the analyzed variables into a predefined range, which increases the accuracy of the grouping algorithm and thus the quality of the clusters created [32]. There are several ways to determine the exact number of clusters, such as the rule of thumb [33], the elbow technique [34], cross-validation, the kernel matrix [35], or the information criterion approach [36]. In this study, we use the V-fold validation and the elbow technique to determine the optimal number of clusters.

III. DISCUSSION

A. Cluster analysis

Figure 1 shows a diagram of the cost sequence using the elbow rule method, which clearly shows that the best solution consists of a total of four clusters. The above diagram shows the error function for distinct cluster solutions [37], i.e., the average distance of observations in the subsamples from the assigned cluster centers. Differences in cluster costs or errors between solutions with two to three clusters are considered significant, i.e., errors decrease by more than 5% compared to a clustered solution with one fewer cluster as the number of clusters increases. The solution with four clusters is the best solution because the error difference between solutions with three or four clusters is less than 5%.

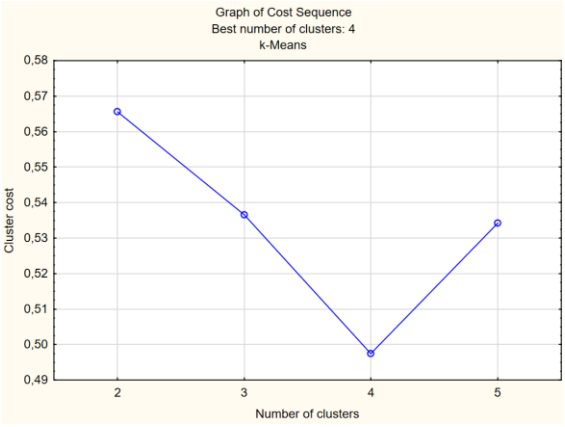


Figure 1. Graph of cost sequence; Source: Authors, Eurostat (2021)

Table 3 shows the results of the Anova analysis, which confirm that the means of the observed variables are statistically different between clusters at the 5% level, indicating that the four-cluster solution is appropriate (note: *** statistically significant below 1% probability).

TABLE 3. ANOVA ANALYSIS OF THE RESEARCH VARIABLES

Research variables	*Between SS*	*df*	*Within SS*	*df*	*F*	*p-value*
ENT1.	451.456	3	102.022	19	28.026	0.000***
ENT2.	587.495	3	136,5055	19	27.258	0.000***
ENT3.	2786.756	3	968.9835	19	18.214	0.000***
ENT4.	424.988	3	461.6209	19	5.831	0.009***

Source: Authors' work

The average values of the research variables by cluster are shown in Table 4. It can be seen that Cluster 1 and Cluster 3 contain the most countries in the research sample. Cluster 4 represents only one country (Norway) with an extremely high value (57%) of variable ENT3, which measures an increase in the number of remote sessions caused solely by the Covid 19 pandemic.

TABLE 4. AVERAGE VALUES OF OBSERVED VARIABLES BY CLUSTERS

Research variables	*C1*	*C2*	*C3*	*C4*
ENT1.	7.77	22.00	13.43	5.00
ENT2.	8.62	25.00	15.29	7.00
ENT3.	20.31	40.50	39.43	57.00
ENT4.	11.15	25.50	16.71	12.00
Number of countries	13	2	7	1
% share	56.52	8.70	30.43	4.35

Source: Authors' work

Table 5. shows the research variables of the countries by clusters together with the distance to the center of gravity.

TABLE 5. COUNTRIES BY CLUSTER AND DISTANCE FROM THE CENTROID

Country	*Cluster*	*Centroid distance*
Bulgaria	1	0.291
Latvia	1	0.097
Lithuania	1	0.274
Hungary	1	0.596
Poland	1	0.575
Portugal	1	0.667
Slovenia	1	0.431
Slovakia	1	0.140
Croatia	1	0.332
Montenegro	1	0.139
Serbia	1	0.274
Turkey	1	0.451
Bosnia and Hercegovina	1	0.366
Malta	2	0.464
Austria	2	0.464
Belgium	3	0.296
Germany	3	0.724
Italia	3	0.406
Cyprus	3	0.619
Netherlands	3	0.340
Finland	3	0.356
Sweden	3	0.664
Norway	4	0.000

Source: Authors' work

13 countries, including Bulgaria, Latvia, Lithuania, Hungary, Poland, Portugal, Slovenia, Slovakia, Croatia, Montenegro, Serbia, Turkey, and Bosnia and Herzegovina, fall under cluster 1. This cluster contains countries belonging to different geographical regions, such as Southeastern Europe (Bulgaria, Bosnia and Herzegovina, Serbia, Montenegro, the eastern part of Turkey), Southern Europe (Portugal), Central Europe (Slovenia, Hungary, Croatia, Poland, Slovakia), and Northern Europe (Latvia, Lithuania). Cluster 1 has the lowest average scores of all research variables compared to the other three clusters. The smallest differences between the countries in cluster 1 were observed for the variable ENT4.

Cluster 2 includes the following two countries: Malta and Austria. By geographic regions, Austria belongs to Central Europe, while Malta belongs to Southern Europe. Both regions fall under the economically developed regions of Europe and are members of the European Union. In contrast to all three other clusters, cluster 2 has the highest average values for the variables ENT1., ENT2. and ENT4, suggesting that the countries in cluster 2 have made the greatest technological progress due to the Covid 19 pandemic.

Cluster 3 includes the following countries: Belgium, Germany, Italy, Cyprus, the Netherlands, Finland, and Sweden. These countries belong to different geographic regions, such as Western Europe (Belgium, Netherlands), Central Europe (Germany), Southern Europe (Italy), Northern Europe (Finland, Sweden), and Southeastern Europe (Cyprus). Countries within this cluster scored the lowest averages for the variable ENT1 and the highest for the variable ENT3. In terms of shares in all variables, the countries in this cluster follow immediately behind Austria and Malta, which form cluster 2.

Cluster 4 contains only one country, Norway. This country is not a member of the European Union and belongs to the Nordic region in Northern Europe. This cluster is specific because it achieves the highest average score in the survey variable ENT3. As for the other three clusters, Norway achieves the lowest average value for the variable ENT1.

It is easy to see that the highest average value in all clusters was recorded for the variable ENT3. This can be easily explained by the fact that in all countries it was necessary to introduce social distancing to prevent the spread of infection. Thus, the physical holding of meetings was replaced by the holding of remote meetings precisely because of the outbreak of the pandemic. On the other hand, the smallest average values for all clusters are observed for the variable ENT1. Considering that the first free email services appeared in the 1990s, it can be assumed that most companies were already using email for their business before the Covid 19 pandemic and that their use is only partly due to the occurrence of the Covid 19 pandemic.

Figure 2. shows the countries included in the clusters. The only country shown is Norway, which represents cluster 4.

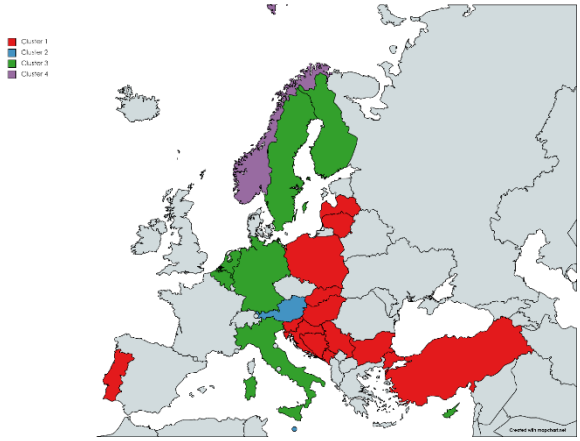


Figure 2. Countries across clusters on the Europe map; Source: Authors, Eurostat (2021); https://mapchart.net/

Table 6. shows the Pearson correlation matrix for the four observed variables. It can be seen that none of the research variables have a higher correlation than 0.7 between the variables included in the data set. The only variables with a higher correlation are variables 1 and 2, and variable 3 and GDP per capita.

TABLE 6. PEARSON CORRELATION MATRIX OF RESEARCH VARIABLES AND GDP PER CAPITA (THE YEAR 2021, USD)

	ENT1.	*ENT2.*	*ENT3.*	*ENT4.*	*GDPpc.*
ENT1.	1	0.940	0.363	0.613	0.260
ENT2.		1	0.496	0.619	0.363
ENT3			1	0.408	0.938
ENT4.				1	0.249
GDPpc.					1

Source: Authors' work

To investigate the relationship between business adaptation to the economy and changes in business methods and procedures and the level of economic growth of selected European countries for the year 2021, we estimated the average values of GDP per capita for each defined cluster. The results of the examined correlations are presented in Table 7. Cluster 4, which consisted only of Norway, achieves the highest average GDP per capita (89,202.80 USD), which was expected considering that Norway has the highest GDP per capita in Europe after Luxembourg, Ireland, and Switzerland [38]. The lowest average GDP per capita is found in cluster 1, which includes the countries of the Western Balkans and the Baltic countries, which, according to [38], are at the bottom of the list of European countries in terms of GDP per capita achieved in 2021.

TABLE 7. GDP PER CAPITA IN 2021 BY CLUSTERS

Cluster	*Average GDP per capita (USD)*	*Number of countries*	*St. Dev.*
C1	16.873,75	13	6.95853
C2	43.262,65	2	14.14956
C3	48.743,14	7	11.22386
C4	89.202,80	1	-
Total	32.012,56	23	21.31979

Source: Authors' work

Figure 3 shows the average value of GDP per capita per cluster. Cluster 4 has the highest value of $89,203, while the lowest value for cluster 1 is $ 16,874.

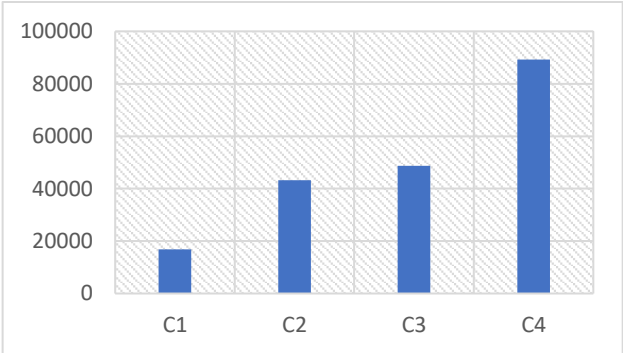


Figure 3. Average values of GDP per capita according to clusters; Source: Authors, Eurostat (2021)

Table 8. shows the results of the Kolmogorov-Smirnov normality test for the distribution of the GDP per capita variable, and the results show that the distribution of GDP per capita follows a normal distribution. Where in a. the normal distribution is tested, in b. the indicators are calculated based on the sample, and in c. the Liliefors correction was used.

TABLE 8. KOLMOGOROV-SMIRNOV TEST OF NORMALITY OF THE DISTRIBUTION OF THE GDP PER CAPITA VARIABLE

One-Sample Kolmogorov-Smirnov Test	*GDP_per_capita*
N	22
Normal Parameters (Mean)[a]	29413.004
Normal Parameters (Std. Deviation)[b]	9223.372
Test Statistic	0.160
Asymp. Sig. (2-tailed)[c]	0.148

Table 9. shows the Levene homogeneity test of the variance of the GDP per capita variable. The test result shows that the assumption of homogeneity of the variance of the GDP per capita variable cannot be rejected.

TABLE 9. LEVENE HOMOGENEITY TEST OF THE VARIANCE OF THE GDP PER CAPITA VARIABLE

Tests of Homogeneity of Variances	*Levene Statistic*	*df1*	*df2*	*Sig.*
GDP_per_capita	1.688	2	19	0.211

Source: Authors' work

The Anova analysis presented in Table 10 shows that the differences in the average values of GDP per capita of the identified clusters are statistically significant, indicating a strong correlation between economic growth and digital transformation in the European countries studied.

TABLE 10. RESULTS OF ANOVA ANALYSIS OF GDP PER CAPITA DIFFERENCES BY CLUSTERS

	Sum of Squares	*df*	*Mean Square*	*F*	*Sig.*
Between Groups	5.04E+09	2	2.5216E+09	31.169	0.000***
Within Groups	1.54E+09	19	8.0901E+07		
Total	6.58E+09	21			

Source: Authors' work; Note: *** statistically significant under 1% probability).

Anova analysis indicates a difference between at least two groups of the observed variable, but does not indicate which groups are statistically significantly different. To test this assumption, a post-hoc LSD analysis is performed. Table 11. shows the post-hoc analysis of GDP per capita between clusters. All clusters show a significant relationship with probability less than 1%, except for the relationship between cluster 2 and cluster 3, and between cluster 3 and cluster 1 (

TABLE 11. POST-HOC LSD ANALYSIS OF THE DIFFERENCE IN GDP PER CAPITA BY CLUSTERS

(I) Cluster	*(J) Cluster*	*Mean Difference (I-J)*	*Std. Error*	*Sig.*
C1	C2	-26388,896	6831,803	0,001***
	C3	-31869,389	4216,681	0,000***
C2	C1	26388,896	6831,803	0,001***
	C3	-5480,493	7211,633	0,457
C3	C1	31869,389	4216,681	0,000***
	C2	5480,493	7211,633	0,457

Source: Authors' work; Note: *** statistically significant under 1% probability).

IV. CONCLUSION

An analysis of the homogeneity of European countries in terms of their adaptation to the changes in business methods and procedures triggered by the Covid 19 pandemic was performed, using non-hierarchical cluster analysis. Using the elbow rule, four clusters were defined as the best fit, which was also tested by an Anova analysis that confirmed the selection of the four defined clusters. Cluster 1 included the following countries: Bulgaria, Latvia, Lithuania, Hungary, Poland, Portugal, Slovenia, Slovakia, Croatia, Montenegro, Serbia, Turkey, and Bosnia and Herzegovina; cluster 2 included Malta and Austria; cluster 3 included Belgium, Germany, Italy, Cyprus, the Netherlands, Finland, and Sweden; and cluster 4 included Norway. All clusters had the highest number of companies that increased the number of remote meetings due to the Covid 19 pandemic, which can be explained by the introduction of social distancing as a method to prevent the spread of the Covid 19 pandemic. The correlation analysis showed the highest positive correlation between companies with an increase in remote access to the company's email system, which was solely a result of the Covid 19 pandemic, and companies with an increase in remote access to the company's ICT systems other than email, which was solely a result of the Covid 19 pandemic.

Similarly, although to a somewhat lesser extent, a strong positive correlation was found between firms that started selling goods or services online or increased their efforts during the pandemic and GDP per capita. Other correlations of the variables are statistically significant and indicate a moderate to good relationship between them. When examining the relationship between business adaptation and changes in business methods and procedures and the level of economic growth of selected European countries for 2021, the estimated average value of GDP per capita was highest in the fourth cluster, i.e. Norway, and lowest in the countries of the Western Balkans and Baltic countries, which are classified in cluster 1. The Anova analysis conducted further showed that there is a strong correlation between economic growth and digital transformation in the European countries studied.

The results of the cluster analysis confirmed significant differences among European countries in terms of business adaptation to the economy and changes in business methods and procedures as a result of the Covid 19 pandemic. In addition, the Anova analysis confirmed a strong relationship between economic growth and digital transformation in the European countries studied, suggesting that more developed countries have adapted more quickly to the Covid 19 pandemic. Since this research is based on only one year's data analysis, this analysis should be repeated for a longer period of time to confirm the obtained results, which also defines the limitation of the conducted analysis.

V. REFERENCES

[1] J.E. Humphries, C. Neilson, and G. Ulyssea, "The Evolving Impacts of COVID-19 on Small Businesses Since the CARES Act.," Cowles Foundation Discussion Papers 2230, April 2020.

[2] A. Adams-Prassl, T. Boneva, M. Golin, and C. Rauh, "Inequality in the impact of the coronavirus shock: Evidence from real time surveys," Journal of Public Economics, vol. 189, 104245, September 2020.

[3] O. A. Younes, and S. Altug, "The COVID-19 Shock: A Bayesian Approach," Journal of Risk and Financial Management, vol. 14, issue 10, 495, October 2021.

[4] D. Christelis, D. Georgarakos, T. Jappelli, and G. Kenny,"The COVID-19 Crisis and Consumption: Survey Evidence from Six EU Countries", SSRN Electronic Journal 3751097, December 2020.

[5] E. Bosio, S. Djankov, E. L. Glaeser, and A. Shleifer, "Public Procurement in Law and Practice," SSRN Electronic Journal, vol. 122, issue 4, pp. 1091-1117, May 2022.

[6] SA Tabish, "Covid-19 Pandemic: Emerging Perspectives and Future Trends," Journal of Public Health Research, vol. 9, issue 1, June 2020.

[7] J. Hunady, P. Pisár, D.S. Vugec, and M.P. Bach, "Digital Transformation in European Union: North is leading, and South is lagging behind," International Journal of Information Systems and Project Management, vol. 10, issue 4, pp. 58-81, November 2022.

[8] M. Tomičić Furjan, K. Tomičić-Pupek, and I. Pihir, "Understanding digital transformation initiatives: Case studies analysis," Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, vol. 11, issue 1, pp. 125-141, 2020.

[9] A.F. Barišić, J.R., Barišić, and I. Miloloža, "Digital transformation: challenges for human resources management," ENTRENOVA-

ENTerprise REsearch InNOVAtion, vol. 7, issue 1, pp. 357-366, March 2022.
[10] A. Trubetskaya, and H. Mullers, "Transforming a global human resource service delivery operating model using Lean Six Sigma," International Journal of Engineering Business Management, vol. 13, November 2021.
[11] I. Pavić, V. Mijušković, and I. Žager, "Which Digital Tools dominate Secondary and Higher Education in Economics: Google, Microsoft or Zoom?," Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy, vol. 13, issue 2, pp. 117-134, December, 2022.
[12] A. Jerman, A. Bertoncelj, and I. Erenda, "The influence of critical factors on business model at a smart factory: A case study," Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, vol. 10, issue 1, pp. 42-52, April 2019.
[13] A. Ishnazarov, N. Kasimova, S. Tosheva, and A. Isaeva, "ICT and Economic Growth: Evidence from Cross-Country Growth Modeling," The 5th International Conference on Future Networks Amp Distributed Systems , pp. 668-671, December 2021.
[14] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means Algorithm: A Comprehensive Survey and Performance Evaluation," Electronics, vol. 9, issue 8, p. 1295, August 2020.
[15] M. Pejić Bach, J. Zoroja, and M. Loupis, "RFID usage in European enterprises and its relation to competitiveness: Cluster analysis approach," International Journal of Engineering Business Management, vol. 8, December 2016.
[16] J. Zoroja, and M. Pejić Bach, "Impact of information and communication technology to the competitiveness of European countries-cluster analysis approach," Journal of theoretical and applied electronic commerce research, vol. 11, issue 1, pp. 1-10, January 2016.
[17] A. Buljan, M. Spremić, "Cluster Analysis of IT Security Risks in Chosen Sectors," ENTRENOVA-ENTerprise REsearch InNOVAtion, vol. 5, issue 1, pp. 489-496, September 2019.
[18] M. Pejić Bach, K. Dumičić, B. Žmuk, A. Bartolić, J. Zoroja, "Use of hierarchical clusters analysis in identification of obstacles to introduction of e-commerce in European countries", IEEE, 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1325-1330, July 2019.
[19] M. Pejić Bach, T. Bertoncel, M. Meško, D. Suša Vugec, and L. Ivančić, "Big data usage in European countries: Cluster analysis approach", Data, vol. 5, session 1, 25, March 2020.
[20] M. Pejić Bach, J. Zoroja, and I. Strugar, "Investment of European enterprises in ICT education of their employees: Relationship with country competitiveness and digital economy development," IEEE, 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 658-663, November 2020.
[21] B. Jaković, T. Ćurlin, I. Miloloža, "Enterprise digital divide: website e-commerce functionalities among European Union Enterprises," Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, vol. 12, issue 1, pp. 197-215, June 2021.
[22] S. Donati, G. Viola, F. Toscano, and S. Zappalà, (2021). "Not All Remote Workers Are Similar: Technology Acceptance, Remote Work Beliefs, and Wellbeing of Remote Workers during the Second Wave of the COVID-19 Pandemic," International Journal of Environmental Research and Public Health, vol. 18, issue 22, p. 12095, January 2021.
[23] A. L. Scutariu, T. Şuşu, C.E. Huidumac-Petrescu, and R.M. Gogonea (2021). "A Cluster Analysis Concerning the Behavior of Enterprises with E-Commerce Activity in the Context of the COVID-19 Pandemic," Journal of Theoretical and Applied Electronic Commerce Research, vol. 17, issue 1, pp. 47–68, March 2022.
[24] T. Alberti, and D. Faranda, "On the uncertainty of real-time predictions of epidemic growths: A COVID-19 case study for China and Italy," Communications in Nonlinear Science and Numerical Simulation, 90, p. 105372, November 2020.
[25] Eurostat, Covid-19 Impact on ICT usage by NACE Rev.2 activity (2022), Available at: https://ec.europa.eu/eurostat/databrowser/view/isoc_e_cvdn2/default/table?lang=en
[26] E. Battisti, S. Alfiero, E., and E. Leonidou, "Remote working and digital transformation during the COVID-19 pandemic: Economic–financial impacts and psychological drivers for employees" Journal of Business Research, 150, pp. 38–50, November 2022.
[27] A. Leogrande, "k-Means Clusterization and Machine Learning Prediction of European Most Cited Scientific Publications", SSRN, 4195846, August 2022.
[28] D. Yu, X. Zhou, Y. Pan, Z. Niu, and H. Sun, "Application of Statistical K-Means Algorithm for University Academic Evaluation", Entropy, vol. 24, issue 7, 1004, 2022.
[29] M. Faisal, E. M., Zamzami, and Sutarman "Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance", Journal of Physics: Conference Series, vol. 1566, issue 1, p. 012112, June 2020.
[30] S. Kalogiannidis, D. Kalfas, F. Chatzitheodoridis, and S. Kontsas, "The Impact of Digitalization in Supporting the Performance of Circular Economy: A Case Study of Greece, " Journal of Risk and Financial Management, vol. 15, issue 8, 349, August 2022.
[31] M. Garaus, and C. Garaus, "The Impact of the Covid-19 Pandemic on Consumers' Intention to Use Shared-Mobility Services in German Cities," Frontiers in Psychology, vol. 12, 646593, March 2021.
[32] I. B. Mohamad, and D. Usman, "Standardization and Its Effects on K-Means Clustering Algorithm," Research Journal of Applied Sciences, Engineering and Technology, vol. 6, issue 17, pp. 3299–3303, September 2013.
[33] F. M. L. Di Lascio, D. Giammusso, and G. Puccetti,"A Clustering Approach and a Rule of Thumb for Risk Aggregation," Journal of Banking & Finance, vol. 96, pp. 236-248, July 2018.
[34] E. Umargono, J. E. Suseno, and S. Vincensius Gunawan, "K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula," Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019), Atlantis Press, pp. 121-129, 2020.
[35] N. Tsapanos, A. Tefas, N. Nikolaidis, and I. Pitas, "Kernel matrix trimming for improved Kernel K-means clustering," IEEE International Conference on Image Processing (ICIP), pp. 2285-2289, 2015
[36] C. A. Sugar, and G. M. James, "Finding the Number of Clusters in a Dataset," Journal of the American Statistical Association, vol. 98, issue 463, pp. 750–763.
[37] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 63, issue 2, pp. 411–423, November 2000.
[38] Statistics Explained. (n.d.). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=GDP_per_capita,_consumption_per_capita_and_price_level_indices

The 5G-supported Unmanned Aerial Vehicles for Emergency Cases Response

Zdravko Paladin[1], Žarko Lukšić[1], Nexhat Kapidani[1], Mario Montagud[2,3], Miguel Fernández-Dasí[2], Soundarya Srinidhi[4], Thomas Wöllert[4], Georgios Boustras[5]

[1]Administration for Maritime Safety and Port Management (AMSPM), Bar, Montenegro

[2] Fundació I2CAT, Barcelona, Spain, [3] Universitat de València, Spain, [4] Atmosphere GmbH, Wessling, Germany

[5] European University of Cyprus (EUC), CERIDES, Nicosia, Cyprus

[1] zdravko.paladin@pomorstvo.me, [2]miguel.fernandez@i2cat.net, [3]mario.montagud@i2cat.net, [4] soundarya.srinidhi@atmosphere.aero, [5]G.Boustras@euc.ac.cy

***Abstract* – In this paper, it is analyzed the general UAVs development and their advancement with the Fifth-generation technology standard for broadband cellular networks (5G network), as well as real-time video streaming and other tools deployed for UAV use cases in the context of EU project RESPOND-A. Practically, the paper provides the case study with simulated risk realization events in which were used UAVs supported by 5G network modems, communication terminals, and 4K 360º video cameras with social viewing tools. Especially for advancing operational and surveillance capacities for SAR missions and environment protection, UAV assets are supported by the recent introduction of the 5G network, which enables a continual flow of information between them and shore-based rescue coordination centers. Such a manner would comprehend the integrated UAV data collecting capabilities, their transmission via the next generation of wireless communications, and seamless video and audio data provision, aiming to raise the level of situational awareness.**

Keywords - UAV; 5G; emergency; SAR; real-time video streaming, social video viewing, immersive video;

I. INTRODUCTION

In extending situational awareness to provide efficient and cost-effective responses to various maritime and land emergency cases, are used versatile Unmanned Aerial Vehicles (UAVs), UxVs applications, and different autonomous assets. Their key importance is reflected in obtaining timely information and persistently monitoring the current situation at the incident area. As an important support asset to maritime and land surveillance, especially when dealing with emergency cases that require persistent monitoring of water areas where an incident occurred, the UAVs are more often used for safety, security, environmental, recognition, surveying, disaster response, and other missions. Therefore, the UAVs are continuously developed in compliance with the latest advancement in mobile networks such as 4G and 5G, which is of crucial importance for mission-critical communications, increased situational awareness, and consequently the efficiency of maritime actions in the required area. The 4G/5G cells support UAV devices and together with satellite communication provide detailed real-time situational pictures which facilitate the decision-making process for responsible maritime and land authorities and first responders on the scene. In such a way, the advanced UAVs operate over 5G networks, as a cost-effective solution for surveillance and real-time actions, bringing significant benefits to national maritime authorities in terms of image accuracy, video streaming, crisis management, damage assessment and evaluation, instant, and timely information flows, as well as support for the decision-making process. The UAVs are also used for visual inspection because of their increasing availability and outstanding camera guiding capabilities, but the primary goal of their applications in Search and Rescue (SAR) operations is to detect, recognize and retrieve subjects in the lowest amount of time feasible [1].

The UAVs used in safety and critical missions for rescuing or surveillance of sea-polluted areas, wildfires, inspection, surveillance, natural disasters, or other security purposes, are mostly designed for long ranges, with a significant number of sensors, and certain payload capacities, depending on the purpose type and conditions at sea or land. Therefore, they are able to fly at great heights and over long distances [2]. Also, the range of capabilities and operations depends on the level of autonomy of UAVs, so there are 4 main types of classification: a remotely controlled system, an automated system (preprogrammed), an autonomous non-learning system (fixed built-in functionalities), and autonomous learning system (able to modify rules of behavior).

The aim of this paper is to provide a review of recent developments of UAVs in the context of their performance and functionalities using 5G network communication. Therefore, the paper unfolds as follows: after introductory considerations, paragraph II provides an analysis of the current development of the 5G network applied in maritime mission communications; paragraph III reviews 5G-based UAVs and their contribution to maritime SAR purposes. Finally, paragraph IV gives a case study of a SAR pilot with testing of UAVs capabilities deployed under the EU RESPOND-A project, while the Conclusion gives overall remarks on the topic.

II. RECENT UAVS DEVELOPMENTS IN THE CONTEXT OF SURVEILLANCE AND SEARCH AND RESCUE ACTIONS

Related to various maritime and land SAR and other emergency missions both at sea and land, in the recent project trials the three models of UAVs proved to be very efficient for required levels of operation, whether it is surveillance from higher layers of the atmosphere or

tactical operation and monitoring in close vicinity from the accident location. These three models are:

(1) the AR3 Net Ray Unmanned Aerial Vehicle (UAV), a fully autonomous UAS with 10 hours of flying capacity and a communications span of up to 80 km, recovered by parachute and airbags or with a net system that is specifically designed for maritime operations;

(2) the AR5 Life Ray Evolution, a medium-altitude, and medium-endurance, fixed-wing UAS conceived and constructed for supporting SAR tasks and long-range intelligence observations (up to 16 hours of flight duration and a communications span of 100 km using Radio Line-of Sight (RLoS) means, or unlimited using satellite communications (SATCOM));

(3) the Zephyr UAS, a High Altitude Pseudo-Satellite (HAPS) that co-operates among satellites, manned aircraft and UASs with an operating altitude of about 21 km [3].

The main strengths of these UAVs are lightness and high seeing, sensing, and communication capacities, as well as the ability to provide better information quality and sharing. The complexity of the UAVs makes them more vulnerable. The significant operational feature is the possibility to approach the areas of high risk for humans, while one of the threats is the collapsing of the UAVs due to severe weather conditions/harsh environments [4]. Another important aspect of UAV development is its improvement in terms of integration with embedded AI modules for detecting, recognizing, and tracing the objects at sea with the ability to instantly transmit the information to ground stations and related data lakes for further processing [5].

In a harsh weather environment, as assumed, the wind can be a big challenge to run the UAV appropriately in SAR and surveillance missions, since it can reach more than half of the UAV's airspeed causing the mission to stop. In order to consider the influence of wind on UAV's kinematics, the Coordinated Turn model has been set up and analyzed in reference [6] complying with IAMSAR manual requirements. This method also involves the probability of containment distribution function over the searching area, with the aim that all SAR facilities be equally and precisely identified.

III. 5G NETWORKS' CURRENT DEVELOPMENT APPLIED IN UAV MISSION-CRITICAL PURPOSES

The 5G New Radio (NR), as a current global generation of advanced mobile network communication, is able to facilitate the following features [7]:

- eMBB (enhanced Mobile Broadband), with a peak data rate from 10 to 20 Gbps, 100 Mbps whenever needed, 10000 times greater traffic, Macro and small cells support, High mobility up to 500 km/h, and network energy savings by 100 times.

- mMTC (massive Machine Type Communications) enables the functioning of a high density of devices (about 2 x 105 - 106/km^2), supports a long-range, has a lower data rate (about 1 to 100 Kbps), leverages the benefits of the ultra-low-cost of machine-to-machine (M2M), offers 10 years of battery life and provides asynchronous access.

- URLLC (Ultra-Reliable Low Latency Communications) provides ultra-responsive connections, 1 ms air interface latency, and 5 ms end-to-end latency between UE (i.e. mobile) and base station, it is ultra-reliable and available 99.9999% of the time, ensures low to medium data rates (about 50 Kbps to 10 Mbps), high-speed mobility.

Considering that 5G also allows usage of non-terrestrial networks, like satellite networks, and device-to-device (D2D) communications without network coverage, i.e. using sidelinks, many other maritime services can be based on 5G networks. To mention a few, these are:

- Search and Rescue missions (getting information from different Internet of Things (IoT) sensors in real time, communication, and accurate localization of ships or people in distress),
- Ship-Shore communications (within Vessel Traffic Monitoring and Information System – VTMIS and for assistance to navigation operations),
- autonomous vessels (5G low latency connectivity is essential for remote operation and management),
- predictive operations/maintenance enabled by artificial intelligence (AI),
- digital twin fed by real-time Big Data coming from heterogeneous sources,
- smart UAVs/drones with high-definition (HD) video streaming and monitoring of the situation in real-time, supported by System Intelligence (SI) and Edge Intelligence (EI) concepts ([3],[4],[8],[9]).

Also, related to network advancements, the 4G Long Term Evolution (LTE) significantly supports all kinds of surveillance aerial users and UAV applications. Its most important performance is latency with a round trip time of 127 ms and a standard deviation of 48 ms for the most negative case scenario. This showed the capability of UAV running in operations using the LTE network with a low altitude of up to 100m.

Furthermore, existing LTE networks used for support of UAV deployment with low density and low altitude are confirmed by Third Generation Partnership Project (3GPP) investigations [10]. In the complex environments, under which maritime, coastal, land operations or any other rescue mission can be considered, the growing number of UAVs deployed in on-site actions, requires the development of new technologies to enable cellular-connected UAVs with air-ground interference avoidance and comprehensive 3D aerial coverage.

Also, the mission-critical requirements in maritime and land SAR actions could be fulfilled using the UAVs together with high-capacity mobile cellular communication systems, based on LTE networks, which enable low latency for all command-and-control devices and ground platforms. In the analysis made by [11], the LTE link is used for critical and high-volume traffic with higher power usage and wide availability of high-quality modems, which enable heterogeneous traffic of data and information with varying requirements. This approach resulted in a resource-guaranteed channel access scheme that provides low latency down to a value of 6 ms, with high reliability.

Furthermore, every UAV integrates several sensors, which provide many functionalities. The most important ones are exact position and height (using a GPS), the ability to find a home and to return (managed by remote control by a human pilot), the transmission of data and

information collected via sensors, mostly from HD, infrared or thermal cameras, which require an appropriate storing capacity, connectivity with the ground station and immediate streaming of the video and information to land operational pilot. The mentioned functionalities as a prerequisite have the interconnection enabled by 4G Long Term Evolution (LTE), 5G (as an advancement of 4G LTE), and Very High Frequency (VHF). In this way, data and information are instantaneously transmitted and represented over pilot command and control interfaces with visualization and management of surveilled situations, providing First Person View (FPV), as it is in real-time [2].

An interesting use case for UAVs in maritime and any other emergency environment is its utilization as support for providing wireless network coverage during a crisis by acting as a flying wireless base station in the case when the ground station is not working. In such a way, many UAVs' use cases can assist the 5G NR to focus on the minimization of path loss, atmospheric attenuation, and enhancement of the data rate.

Especially, the UAV can be of significant support for providing the communication link between the vessel in the deep sea being far away from the coast. Thus, the UAVs can be incorporated into the operational framework for establishing wireless networking communication links between distant base stations and also extending the 5G cellular network. Also, this type of UAV base station can be deployed using an intelligent machine-learning system like the generalized regression neural networks model, evidencing the improvement of quality of Service (QoS) and energy efficiency ([7],[12]).

IV. CASE STUDY: UAV CAPABILITIES DEPLOYED WITHIN PILOTS OF EU PROJECT RESPOND-A

1. UAVs and terminals covered by 5G

Technologies such as UAVs and reliable connectivity amid a disaster scenario are crucial for First Responders to ensure a timely reaction. In the RESPOND-A project, technical partner *Atmosphere* GmbH participated as a technology provider bringing in UAVs, including a connectivity solution called PLANET, to demonstrate the effectiveness of UAVs and 5G connectivity in rescue missions (Fig. 1). *Atmosphere* took part in two pilots of the RESPOND-A project, where one was a simulation of a forest fire, which took place in Cyprus and the other was a simulation of an oil spill caused by the crash of two trucks demonstrated at the port of Valencia, Spain.

PLANET is a connectivity terminal that is flown on the drone and shares onboard data to the ground via a 5G link, which was used for the demonstrations. PLANET provides connectivity from the ground to the UAVs in the air. It is capable of providing both cellular and satellite network services. *Atmosphere*'s drones are a fleet of two octocopters ATMO1 and ATMO2 (Fig. 2), one used as a replacement. ATMO1 is a 92-cm wide, 4.7kg octocopter, that has the capacity to carry a payload of 2-3kg, with a total flight time of 8 minutes. ATMO2 is a 130cm wide, 9 kg octocopter, which has the capacity to carry a higher payload of about 4-5 kg.

The octocopters are self-built, allowing for more flexibility for the integration of the payload on the drone during the integration phases of the project. Customizable drones proved to be of great value as they ease mounting various types of payloads such as cameras and sensors, by making modifications to the dimensions and weight of the drones. Additionally, both octocopters can be flown in manual mode or autonomously with the presence of a licensed pilot.



Figure 1: PLANET connectivity solution flown onboard the UAV (Source: own)

In preparation for the pilot demonstrations for the RESPOND-A project, several flight tests were carried out to test the calibration, and wind speed and to determine the total flight time by increasing the payload gradually.

The UAV and its payload are powered by two batteries. ATMO1 was the choice of UAV that was flown in the Spanish pilot for the demonstration of the rescue mission and an oil spill at the Valencia port. *Atmosphere* has a 5G base station that is capable of providing a 5G network as a stand-alone device, additionally, it is capable of integrating with external 5G cores to set up a 5G network.

In the case of the RESPOND-A project, the 5G base station, which is an Amarisoft callbox mini (Fig. 3) functioned as a 5G radio access network and integrated with a partner's 5G core device.

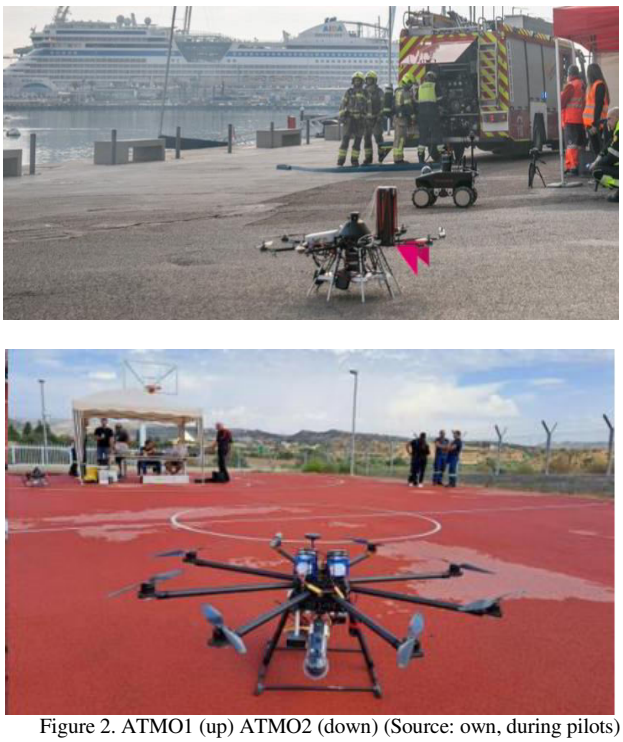


Figure 2. ATMO1 (up) ATMO2 (down) (Source: own, during pilots)

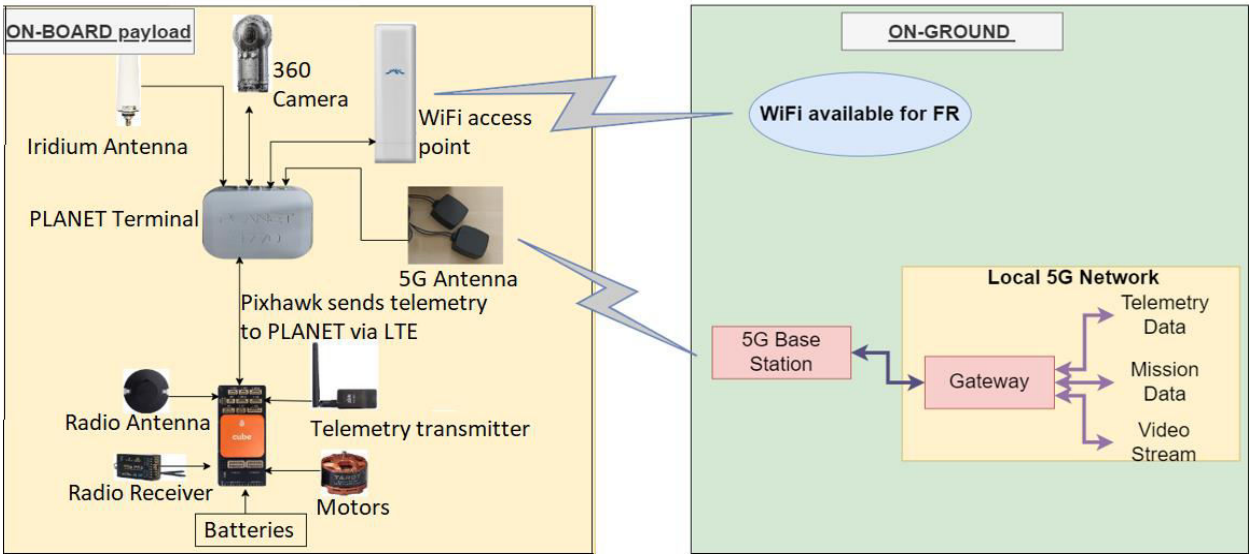


Figure 6. UAV air-to-ground communications (Source: own)

A remote radio head was additionally used to enhance the 5G signal strength, as the main idea of the 5G network deployment was to demonstrate the radiation of a reliable 5G network over >1000m^2 for the use case locations, aiding first responders to be able to reach the 5G network over certain distances around the disaster zone during a rescue mission.

In order to make use of the remote radio head (RRH) and the 5G antenna, certain hardware modifications had to be made, where the Software-defined radio (SDR) card of the Amarisoft callbox mini had to be replaced.



Figure 3. 5G Radio Access Network (Amarisoft callbox mini) (Source: own)

The 5G antenna covers an area of a few hundred meters providing 100+ Mbps of bandwidth, however, these specs are dependent on the characteristics of the area, as certain obstructions can certainly tamper with the RF signal quality of the 5G link radiating from the antenna.

The remote radio head with the 5G antenna is connected to the 5G base station over a long fiber optic cable. Such a setup allows the connection of the 5G network over long distances. In the demonstrations of the pilots at the simulated disaster zones, the 5G base station was set up in the command and control center, with the 5G antenna connected outside the building of the command and control center as visualized in Figures 4 and 5 (in the Cypriot use case).

The UAV flies with *Atmosphere*'s connectivity solution, the PLANET terminal, used as a processing unit for all the UAV onboard components (VR360 camera, Wi-fi access point, telecommunication via 5G link, and satellite link). The PLANET terminal has an integrated 5G modem, which receives a 5G signal through two 5G antennas connected to the terminal. UAV payload air-to-ground architecture is displayed in Figure 6.

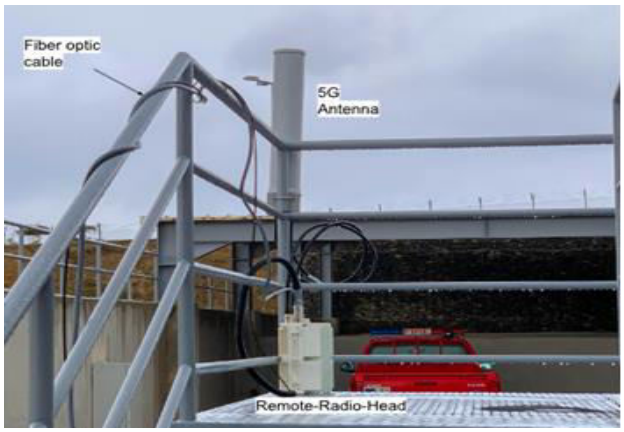


Figure 4. Out-door 5G network architecture (Source: own, during pilots)

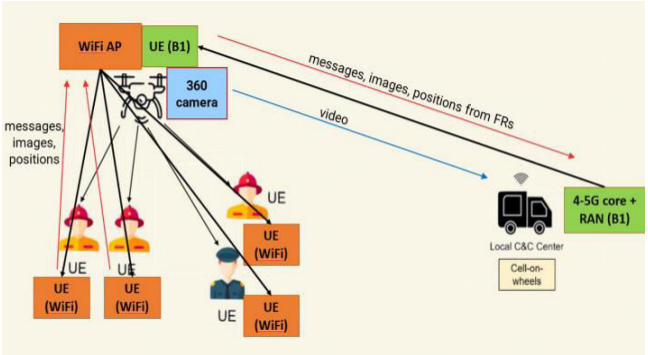


Figure 5. 5G local network infrastructure (Source: own)

On-board data from the UAV, the live video stream, and telemetry data for live localization are sent to the PLANET ground server via the 5G link, and it forwards the data to the command and control tools, to be visualized on the screens (Fig. 6). The camera onboard the UAV connects to the PLANET terminal via an ethernet cable, the camera can then be accessed from the ground via the 5G link. PLANET forwards the video stream from the camera to the processing station on the ground in Command and Control Center (C&C) (Fig.7).

The Wi-Fi access point integrated on the drone, cabled to the PLANET terminal via Ethernet provides Wi-Fi connectivity through the 5G uplink. The main purpose of this feature was to demonstrate to first responders the added communication means to the command and control center, where they are able to send pictures, messages, and other key information under the local Wi-Fi bubble of connectivity provided by the UAV.



Figure 7. Live video stream sent over 5G from the on-board camera (Source: own, during Cyprus pilot)

2. Interactive social 360º viewing for SAR missions

Beyond doubt, the availability of real-time and high-resolution video from emergency and/or surveillance missions, like SAR, can provide huge benefits in terms of safety and decision-making. In this context, 360º videos can provide added value compared to 2D video given relevant facts, like (i) the whole omnidirectional area is captured (and not just a given field-of-view as in directive 2D cameras); and (ii) 360º videos can be watched on Virtual Reality (VR) headsets, allowing for a natural omnidirectional inspection, with high levels of realism and presence (i.e., the feeling of being there, inside the captured scenes).

In this context, recent studies have shed some light on the potential advantages provided by 360º videos compared to traditional 2D videos in education and health sectors for examination of information of interest ([13],[14]).

Real-time and low-latency video streaming from UAVs is currently a technical challenge due to the typically restricted and varying communication (in terms of coverage, and bandwidth availability…) channels between flying UAVs and reception entities on the ground. This challenge is magnified if high-resolution (e.g., >=4K) and immersive video scenes need to be provided. This challenge can be overcome by making use of the latest-generation (e.g., 5G-powered) communication networks and equipment.

Although being provided with 360º video scenes of interest in real-time already provides remarkable benefits, such benefits can be magnified if multiple video feeds can be provided and dynamically selected, and even further if the video feeds are concurrently made available to multiple geographically distributed first responders for collaboratively inspecting the target mission or environment. In this context, recent studies have shown multiple benefits provided by social viewing tools and scenarios, in a variety of use cases, like entertainment and education ([15],[16],[17]).

In response to the existing scientific evidence and gathered user-centric requirements, an interactive social viewing tool for multi-cam 360º videos has been designed and implemented to support SAR missions. In a nutshell, the tool provides a set of technical and interactive features to maximize the benefits in this sector:

- Use of web-based technologies. This maximizes the chances of ensuring cross-platform and cross-device support and eliminates the need for any installation and/or update on the client device to use the tool.
- Low-latency and adaptive streaming capabilities. Provides adapted versions of video streams according to the device being used and the available bandwidth, including to remote consumers via IP networks.
- Free omnidirectional exploration for each deployed 360º camera.
- Dynamic selection of the camera of interest, minimizing switching delays in a synchronized manner (i.e. guaranteeing that video feeds from the different cameras refer to the same time).
- Zoom in / out for each of the video feeds.
- Support for watching on 2D screens (e.g., desktops, laptops, videowalls, smartphones) and 3D VR headsets.
- Availability of text, audio, and video chat channel between the users involved in the mission. Any of these chat modalities can be enabled and disabled, based on needs and preferences.
- Remote users' viewports awareness and selection. Indicates what camera and region of the 360º video the other involved users are watching and is able to select that viewport with just one click (e.g. if another user has detected anything relevant).

Figure 8 shows a UAV equipped with a 4K 360º video camera, connected to a 5G-powered communication terminal. Figure 9 provides a capture of a first responder using that tool to inspect a simulation of an accident in a port area and an associated fuel spill.

A demonstration video of three first responders using that tool in the depicted simulation can be accessed at: https://www.youtube.com/watch?v=hMOAj9Hh1uw. The tool and associated demonstrations have been developed under the umbrella of the EU H2020 Respond-A project (https://respond-a-project.eu/).

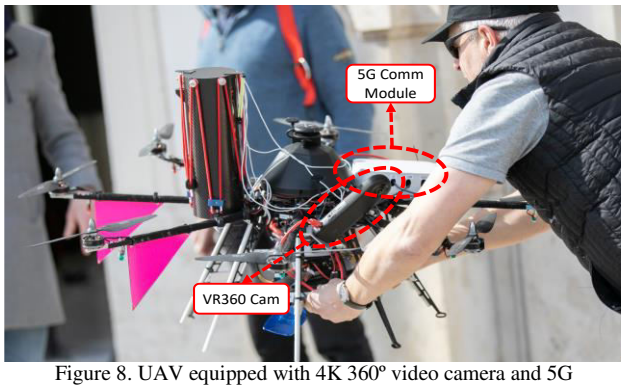


Figure 8. UAV equipped with 4K 360º video camera and 5G communication module (Source: own, during Valencia pilot)

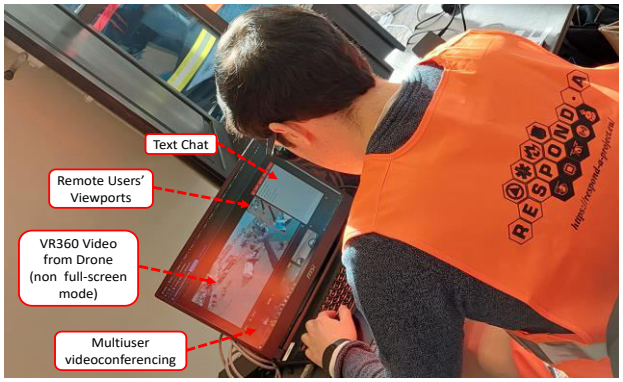


Figure 9. First Responder using a social viewing tool for a collaborative examination of a SAR mission (Source: own, during Valencia pilot)

V. CONCLUSION

The paper shows the overview of the recent developments in UAV capabilities supported with a 5G network, used for enhancing surveillance and control over marine or land areas where disasters, accidents, or environmental risks might occur. The latest advancements in UAVs provide significant support to operating FRs' organizations in SAR actions while the current 4G and 5G mobile networks enable high-quality mission-critical communication, data, video, and relevant information sharing, as well as the use of various types of sensors, devices and equipment, contributing to overall high situational awareness and broader operational picture.

The 5G networks significantly enhance the performance of UAVs making them the most appropriate assets for various emergency use cases. An additional enhancement to overall UAV performance is provided by end-to-end 360° video cameras mounted on the drones with real-time messaging and live streaming protocols ability, interconnected with a 5G powered on-board terminal (PLANET) aiming to show a single 360°, high-resolution panoramic video of surveilled areas. This jointly contributed to efficient communication between involved FRs and C&C centers raising the operational preparedness of all stakeholders. The results of this approach are advanced monitoring and mission-critical communications over 5G with improved response in emergency, safety, security, SAR, and environmental issues, as proved by pilots of the RESPOND-A project.

Finally, further work on this topic will elaborate on the integrated approaches of using the 5G network to cover various applications in maritime environments including mission-critical communications, data sharing, and advanced IoT systems.

ACKNOWLEDGMENT

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 883371 (project RESPOND-A). This article reflects only the authors' views and the Research Executive Agency (REA) is not responsible for any use that may be made of the information it contains. The work of Miguel Fernández has been funded by Secretaria d'Universitats i Recerca de la Generalitat de Catalunya i del Fons Social Europeu (Personal Novel 2021 FI_B_01041). The work of Mario Montagud has been additionally supported by MCIN/AEI/10.13039/501100011033 under Grant RYC2020-030679-I and by ''the European Social Fund (ESF) Investing in Your Future".

REFERENCES

[1] V.S. Ajith, K.G. Jolly, "Unmanned aerial systems in search and rescue applications with their path planning: a review", Journal of Physics: Conference Series, 2021, Vol. 2115, pp. 012020.

[2] R. Fiedler, C. Bosse, et al., "Autonomous Vehicles' Impact on Port Infrastructure Requirements", Fraunhofer Center for Maritime Logistics and Services CML, Hamburg, April 2019.

[3] S. Bauk, N. Kapidani, Ž. Lukšić, F. Rodrigues, L. Sousa, "Review of Unmanned Aerial Systems for the Use as Maritime Surveillance Assets", 24th Conference on Information Technology (IT), 2020.

[4] S. Bauk, N. Kapidani, et al, "Advantages and disadvantages of some UAVs deployed in maritime surveillance", 8th Int. Conf. on Maritime Transport, Barcelona, 2020, pp. 91-102.

[5] Z.Paladin, N. Kapidani, P. Scrima, G. Vosinakis, et al., „Maritime information sharing environment deployment using the advanced multilayered Data Lake capabilities: EFFECTOR project case study", Journal of Maritime Research, Vol. 36(2), pp. 291-304.

[6] F.A. de Alcantara Andrade, A Hovenburg, L. Lima, C. Rodin, T. Johansen, R. Storvold, "Autonomous Unmanned Aerial Vehicles in Search and Rescue Missions using real-time cooperative model prediction control", MDPI Sensors, vol. 19, 4067, 2019.

[7] M. Mahbub, „UAV Assisted 5G Het-Net: A Highly Supportive Technology for 5G NR Network Enhancement", EAI Endorsed Transactions on Internet of Things, Volume 6, Issue 22, 2020.

[8] S. Lins, K. Cardoso, C. Both, et al., "Artificial Intelligence for Enhanced Mobility and 5G Connectivity in UAV-Based Critical Missions", IEEE Access, Vol.9, pp. 111792 – 111801, 2021.

[9] Z. Paladin, E. Fountoulakis, Ž. Lukšić, N. Kapidani, D. Ribar, G. Boustras, „Advanced Mission-Critical Communication in Maritime Search and Rescue Actions", 27th IT Conference, 2023.

[10] Z. Miličević, Z. Bojković, „From The Early Days Of Unmanned Aerial Vehicles To Their Integration Into Wireless Networks", Military technical courier, 2021, Vol. 69, Issue 4, pp.941-962.

[11] Ph. Gorczak, C. Bektas, F. Kurtz, Th. Lubcke, Ch. Wietfeld, „Robust Cellular Communications for Unmanned Aerial Vehicles in Maritime Search and Rescue", 2019 IEEE Symposium SSRR.

[12] R. Alfaia, A. Souto, E. Cardoso, J. Araújo, C. Francês, „Resource Management in 5G Networks Assisted by UAV Base Stations: Machine Learning for Overloaded Macrocell Prediction Based on Users' Temporal and Spatial Flow", Drones 2022, 6, 145.

[13] F. Hekele, J. Spilski, S. Bender, T. Lachmann, "Remote vocational learning opportunities—A comparative eye-tracking investigation of educational 2D videos versus 360° videos for car mechanics". British Journal of Educational Technology, 53(2), 248-268, 2022.

[14] Y. Chao, C. Kang, H. Chuang, et al. "Comparison of the effect of 360° versus two-dimensional VR video on history taking and physical examination skills learning among undergraduate medical students: a randomized controlled trial". Virtual Reality, 2022.

[15] F. Boronat, M. Montagud, P. Salvador, J. Pastor, "Wersync: A web platform for synchronized social viewing enabling interaction and collaboration". Journal of Network and Computer Applications, 175, 102939, 2021.

[16] S. Rothe, A. Schmidt, M. Montagud, D. Buschek, H. Hußmann, "Social viewing in cinematic virtual reality: a design space for social movie applications", Virtual Reality, 25(3), 613-630, 2021.

[17] M. Montagud, J. Li, G. Cernigliaro, et al. "Towards social VR: evaluating a novel technology for watching videos together", Virtual Reality 26, 1593–1613, 2022.

Implementation and Cost-Effectiveness of Blockchain Technology within the IP Ecosystems

P. Karanikić
University of Rijeka Department of Biotechnology, Rijeka, Croatia
pkaranikic@biotech.uniri.hr

***Abstract* - The Fourth Industrial Revolution (Industry 4.0) enables the new digital economy by merging technologies such as blockchain, big data, artificial intelligence, the Internet of Things (IoT), and other emerging technologies through digital transformation. Blockchain technology is among the emerging technologies that are affecting the ways of business operationalization through the automatization of tasks, but also the management of intellectual property rights (IPRs) protection process. The digital economy has an important role in both social and as an engine for economic growth while the IPRs have an important role in the development of the digital economy. The aim of this paper is to investigate and analyze the cost-effectiveness and positive impact of blockchain technologies adoption and implementation in IPRs protection process within the IP ecosystem.**

Key words – Industry 4.0; blockchain technology; economic growth; IPRs protection management; IP ecosystem

I. INTRODUCTION

The Fourth Industrial Revolution (Industry 4.0) gave rise to new emerging and disruptive technologies such as blockchain, big data, artificial intelligence, the Internet of Things (IoT), and other emerging technologies that affect human lives, but also the way how businesses operate. The digital transformation of companies and institutions was accelerated by COVID-19 in adapting to the new digital communication channels. Companies with a higher level of digital transformation adapted faster to the new situation and developed business models that implement blockchain and other emerging technologies.

Intellectual property rights (IPRs) protection and management became more challenging with the rapid development of digitalization and the digital economy where the number of disputes in the IPRs infringement is constantly increasing. The existing technologies used for IPRs protection process are not adequate in the digitalized world to fight against IPRs infringements in a way to monitor, discover, authorize, or forbid the misuse of the IPRs.

Blockchain technology has a transformative and disruptive potential within the IP ecosystems and for this reason, it needs to be properly addressed to fully realize the benefits that blockchain technology offers. It can be used in the management of the IPRs protection process in registration, management, and enforcement of the IPRs.

Practical implementation of blockchain technology within the IP ecosystem can be seen in performing of several activities: providing proof-of-rights, keeping the record of rights, registering the rights, controlling, and tracking the distribution of rights, etc. The idea of a distributed network could be a beneficial solution for IPRs management since IPRs by their nature are distributed across countries and dispersed around the world [1].

The IP system is today faced with problems arising in the digital age where it is very difficult to fight and struggle against the internet who is developing much faster than the laws governing IPRs protection.

The aim of this paper is to investigate and analyze the impact and cost-effectiveness of blockchain technology adoption and implementation for relevant stakeholders in the IPRs protection process within the IP ecosystems. The research question addressed in this paper is whether the implementation and adoption of blockchain technology improve the management and cost-effectiveness of the IPRs protection process within the IP ecosystem, based on the preliminary literature review. It is important to note that the existing literature on this specific topic is scarce and therefore still without sufficient relevant information on the practical implementation of blockchain technology in the IPRs protection process management within the IP ecosystem which presents the research limitation. Therefore, the purpose of this paper is to fill the identified research gap in a way to capture and summarize the available important and relevant information on this research topic.

The paper is organized as follows. In the introduction part, the research topic is introduced and elaborated. In the second part of the paper, the IPRs are defined and elaborated in the economic context, as well as the role of IPRs protection system in economic development. In the third part of the paper is presented and elaborated based on the current literature on how blockchain technology can be implemented within the IP ecosystem. The fourth part of the paper discusses the challenges and opportunities of the potential application of blockchain technology in the management of the IPRs protection process. The final part of the paper, the conclusion, summarizes relevant findings based on the conducted preliminary literature review and proposes the elements for future research on this topic.

II. IMPORTANCE OF IPRS PROTECTION AND MANAGEMENT

Intellectual property rights in a broader sense are the rights that resulted from a creative and intellectual effort of an individual, and that fall into the scientific, literary, and artistic domains. The IPRs can take several forms such as copyright, patents, trademarks, industrial design, trade secrets, and various strategies of IPRs management are implemented accordingly to the specific IPRs form and entering the new or existing market.

The IPRs have all characteristics of any other asset. An asset can be defined as an item, thing, or entity that has potential or actual value to a company or an organization. The value of an asset can be tangible or intangible, financial, or non-financial, and includes consideration of risks and liabilities. While the physical assets refer to the equipment, inventories, and properties owned by a company or an organization and are tangible in its nature, the intangible assets refer to the non-physical assets such as brands, digital assets, use rights, licenses, intellectual property rights, reputation, or agreements [2]. The aim of IPRs protection and management is in transforming intangible assets into tangible ones in a form of products and services, and to commercially exploit them on the market.

From an economic point of view, IPRs can play an important and positive role in economic development and growth. There are two economic objectives of the IPRs system: to promote investments in knowledge and technology creation and business innovation through the exclusivity rights to use and sell the newly developed technologies and to promote the dissemination of new knowledge and technology by encouraging rightsholders to place their inventions on the market. There is a need for a tradeoff between these objectives by policy balancing between the weak and overprotective IPRs system [3].

The IPRs protection system is the institutional system that fosters innovation and economic development. The innovators are enjoying the exclusivity right granted by the government to exclude others from commercial exploitation of their inventions. In return, the innovators are required to disclose the invention to the public and enable the society end economy to fully benefit from the innovation exploitation. The IPRs are having a limited duration of protection and when the protection expires the invention becomes the public good. For this reason, adequate IPRs protection process management needs to be put in place to address all potential difficulties in commercial innovation exploitation.

In the completely opposite approach, an Open Innovation model developed by Chesbrough [4], the companies are encouraged to use external ideas and to combine them with their own ideas to generate synergies for value creation and enable the internal ideas to be successfully applied outside of the company's market. This approach is far away from the traditional IPRs system management that promotes the creation of monopolies and exclusivity by enabling only one company a competitive advantage over all others [5].

Implementation of blockchain technology within the overall IP ecosystem could have a beneficial impact on the IPRs protection process management by making the process easier and smoother and in this sense cost-effective.

III. IMPLEMENTATION OF BLOCKCHAIN TECHNOLOGY IN IP ECOSYSTEMS

The practical implementation of blockchain technology was introduced by Nakamoto. He introduced a new method for value transfer execution between parties that are traceable and reliable [6]. Blockchain technology can be considered as a kind of electronic bookkeeping model since blockchain technology includes distributed data storage, peer-to-peer transmission, consensus mechanism, and encryption algorithm [7].

The IP ecosystem can be defined as a network of different actors (e.g., inventors, companies, organizations, IP offices, enforcement authorities, etc.) that interact and collaborate in the IP environment by using the resources to generate, protect, manage, and commercialize intellectual assets. These interactions present the four phases within the IP value chain [8].

The management of the IPRs protection process is currently well digitalized since the IPRs application process is almost completely digitalized within the relevant national and international IP institutions (i.e., WIPO, EPO, USPTO, etc.). Due to the development and implementation of digital technologies such as blockchain, access to relevant market information, identification of suitable licenses, and automated systems for marketing and licensing are enabled [9].

IPRs and blockchain technology can be observed from two different perspectives: the technology using the blockchain as a basis that can be IP protected if satisfying the necessary requirements, while the blockchain technology can be used in the management of the IPRs protection process [10].

Blockchain technology has many different features such as tamper-evident, timestamp, and traceability possibilities. Another important characteristic that distinguishes blockchain technology is decentralization. This characteristic is very important in the context of IPRs protection process management. The process of IPRs protection is a highly regulative process with strictly defined and complex procedures. Besides the fact that two parties involved in the process of IPRs protection need to verify the authenticity of the information, the participation of the third-party institution is necessary and that is both time and cost-consuming. In addition, the presence of the third-party central institution is very important for controlling the whole process of information exchange and building trust between the two parties involved in the process of IPRs protection [7].

Blockchain technology as an emerging technology has a high potential of changing managerial tasks. There are several important features provided by blockchain technology identified by a number of scholars: an inherent degree of security, records immutability, transactions transparency, digital information integrity

and authenticity, data ownership, and trust substitution in intermediaries by trust in the predefined protocol consensus mechanism [11],[12],[13],[14]. All these aspects are very important for the smooth and successful management of the IPRs protection process.

The potential application of blockchain technology is in the concept of smart contracts. Narayanan et al. emphasize that smart contracts are not enforced by laws or courts of arbitration, but by blockchain technology. The rules and penalties embodied in the smart contracts are defined, and the obligations are automatically enforced [15].

The implementation of blockchain technology in the patent system could result in the reduction of the inefficiencies in recording and efficiently agreeing on the time registrations possibly across several national patent systems, and this could be the first step in the patent application and providing the proof of existence right from the beginning. Besides the fact that the entire process could provide a streamlined and transparent practice by limiting the inefficiencies and paperwork, the implementation of blockchain technology could be an efficient tool to fight against patent trolls, by reducing the costs for companies to timestamp and create track of records for their inventions [16].

The implementation of blockchain technology in a way of skipping the overall traditional process of a patent application for companies in registering their IP through a distributed ledger could reduce the total number of contract disputes and enable the cost-efficiency in reducing the number of interactions with local and national governments. As a result, the effectiveness of the IPRs management system could be improved as well as information distribution between the companies which could have a positive impact on the innovation process [17].

IPRs are traditionally governed and managed by centralized institutions in specific regions where IPRs holders wanted to enforce their rights. However, the globalization and digitalization of the market emphasized a need for a more plausible and novel solution for the management of the IPRs protection process that blockchain technology could offer. One of the potential solutions is a framework model that is based on the consortium blockchain that can potentially solve many practical difficulties that are detected in the IPRs protection process [1].

Another potential use of blockchain technology is as a legal tool in IP law. The implementation of the blockchain technology is capable to affect several processes and provide cost-effectiveness in the IPRs protection process lifecycle such as: registration of the IP rights through blockchain by using mobile applications and replacing the relevant institutions which would minimize or fully reduce the IP registration costs for the IP holders, identification of the rightful IP holders, using smart contracts for IP licensing, reducing the need for third parties in the process, and IPRs protection enforcement against counterfeits [18].

Even though blockchain technology is constantly developing and evolving, when it comes to its practical implementation within the IP ecosystem it is still in the early stage of development.

IV. DISCUSSION

The implementation of blockchain technology could offer new and improved ways of management of physical assets and their digital presentation, value exchange, business operations, and importantly trust mechanism enforcement. Blockchain technology has various implementation that goes beyond its original intent. The implementation of blockchain technology offers an innovative approach to the management of the IPRs protection process.

There is an evident need for both institutional and legislative support for blockchain technology implementation. The legislative support is lagging behind the concrete practical implementation of blockchain technology, especially in the area of IPRs management. The efficient application of blockchain technology is facing several concerns. The first concern is related to the legal binding of digital signatures used in blockchains because not all digital signatures are equally valid. The second concern is the use of blockchain-based evidence in courts. The third concern lies in the need for legislation to support the implementation of blockchain technology for copyright protection. Blockchain technology does not solve the unauthorized use problem, however, it does help creators to track their creations and to capture the value they generate [5].

From the standardization point of view, and in this sense important for the implementation of blockchain technology in the IPRs protection process management, the International Organization for Standardization (ISO) is currently developing standards for Blockchain, and distributed ledger technology called ISO/TC 307 (1). Adoption of such standardization will foster blockchain technology to achieve the relevant required legal status and enable IPRs holders to the creation and save evidence for their IP claims [19].

Implementation of blockchain technology for IPRs management from the operational point of view is faced with several challenges: authenticity problem, provenance problem, and royalty stability problem.

The authenticity aspect is an important and valuable aspect in providing the Proof of Existence of protected work at a certain point in time, as well as Proof of Ownership by including additional information on the IPRs holders the Proof of Ownership since blockchains in its nature are unchangeable. Blockchain technology cannot confirm the authenticity of the registered information per se even though it can store information in an unchangeable manner. Until the authenticity issue is resolved the implementation of blockchain technology in the management of the IPRs protection can only be realized through a permissioned network that requires the validation of the inserted information before it is published on the blockchain [20].

The provenance aspect is related to the question of how to record the transfer of IPRs ownership on the network. It is a very complex task to create a chain of different IP rights holders and it is expected that blockchain is capable to transfer the ownership without the trusted third party involved and to prevent piracy by issuing a digital certificate or a secret key to access the digital information. Until the provenance aspect is resolved, the implementation of blockchain technology in this respect will be by issuing a transferable certificate by a trusted third party such as a relevant government institution [21].

The Royalty Stability aspect presents a big challenge related to the stable payment methods for decentralized royalty management which make the implementation of blockchain technology in royalty payment exposed to a very high level of volatility of cryptocurrencies and respectively difficult to reach a fixed consumable price on the blockchain. For this reason, IPRs holders will have to accept the trade-off between price stability and efficiency [21].

The technical features of blockchain technology have a positive effect on the cost efficiency associated with the authenticity aspect. In addition, the implementation of blockchain technology radically changes the ways of doing business and impacts the changes in the market structures [22].

V. CONCLUSION

Digital transformation is affecting organizations and businesses to adopt their digital transformation strategies. One of the reasons businesses are still not implementing blockchain technology is an economic aspect related to sunk costs and switching costs since money transactions and interactions are changing irreversibly.

Implementation of blockchain technology within the IP ecosystem has great disruptive and transformative potential that needs to be adequately addressed to fully realize its beneficial impact and to properly manage the associated implementation risks.

The implementation of blockchain technology in the IPRs protection process management has a large potential. However, there are some aspects that still need to be resolved to fully capture all the benefits that blockchain technology can bring, especially related to the costs and resources effectiveness. There are still some issues related to security, network size, transaction costs, etc. that need to be resolved and furtherly investigated.

The state-of-the-art related to the implementation and exploration of blockchain technology within the IP ecosystems is still in its beginnings. There is still a lack of understanding of all the benefits of adoption and implementation of blockchain technology can offer.

There is no doubt that the implementation of blockchain technology can improve the IPRs system management that did not follow the rapid technological changes that digitalization brought and the circumstances in which it operates. However, it is arguable if the implementation of blockchain technology can fully replace this system.

There is still a lot of research to be done on the implementation of blockchain technology in the IP system even though blockchain technology was introduced almost two decades ago. Its acceptance and practical implementation within the IPRs system will require further research based on its use cases and accompanying legislation. The results and findings of this preliminary literature review can serve as a basis for the future broader systematic literature review that will encompass all relevant aspects and implications (economic, legal, technical) of the implementation of blockchain technology in the management of the IPRs protection process.

ACKNOWLEDGMENT

The publication of this work has been supported by the University of Rijeka under the project “Smart cities in the function of the development of the national economy” (uniri-drustv-18-255-1424).

REFERENCES

[1] Rambhia, V., Mehta, V., Mehta, R., Shah, R., & Patel, D. (2021). Intellectual Property Rights Management Using Blockchain. In Information and Communication Technology for Competitive Strategies (ICTCS 2020) Intelligent Strategies for ICT (pp. 545-552). Springer Singapore.

[2] ISO 55000:2014(en) Asset management, https://www.iso.org/obp/ui/#iso:std:iso:55000:ed-1:v2:en [Accessed: 12-Feb-2023]

[3] Maskus, K. E. (2000). Intellectual property rights and economic development. Case W. Res. J. Int'l L., 32, 471.

[4] Chesbrough, H. (2006). Open business models: How to thrive in the new innovation landscape. Harvard Business Press.

[5] Rivière, J. M. (2018). Blockchain technology and IP–investigating benefits and acceptance in governments and legislations. Junior Management Science, 3(1), 1-15.

[6] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized business review, 21260.

[7] Luo, L. (2022). Application of Blockchain Technology in Intellectual Property Protection. Mathematical Problems in Engineering, 2022.

[8] WIPO - World Intellectual Property Organization Blockchain Technologies and IP ecosystems: A WIPO white paper WIPO 2022. https://www.wipo.int/cws/en/blockchain-and-ip.html [Accessed: 12-Feb-2023]

[9] Karanikic, P., Bezic, H., and Redzepagic, S. (2021). Digitalization of the University Technology Transfer Process. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 1431-1435). IEEE.

[10] Singh, B. P., & Tripathi, A. K. (2019). Blockchain technology and intellectual property rights.

[11] Beck, R., Avital, M., Rossi, M., & Thatcher, J. B. (2017). Blockchain technology in business and information systems research. Business & information systems engineering, 59, 381-384.

[12] Carvalho, A., Merhout, J. W., Kadiyala, Y., & Bentley II, J. (2021). When good blocks go bad: Managing unwanted blockchain data. International Journal of Information Management, 57, 102263.

[13] Denter, N. M. (2021). Blockchain breeding grounds: Asia's advance over the USA and Europe. World Patent Information, 67, 102082.

[14] Denter, N. M., Seeger, F., & Moehrle, M. G. (2022). How can Blockchain technology support patent management? A systematic literature review. International Journal of Information Management, 102506.

[15] Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press.

[16] Boucher, P. (2017). How blockchain technology could change our lives: In-depth analysis. European Parliament. [Accessed: 13-Feb-2023]

[17] Hancock, M. and Vaizey, E. (2017) Technical Report by the UK Government Chief Scientific Adviser. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf [Accessed: 13-Feb-2023]

[18] Gürkaynak, G., Yılmaz, I., Yeşilaltay, B., & Bengi, B. (2018). Intellectual property law and practice in the blockchain realm. Computer law & security review, 34(4), 847-862.

[19] Cha, H., Lee, W., Choi, Y., Lee, J., & Lee, K. (2019). International standardization on blockchain. Electronics and Telecommunications Trends, 34(2), 110-120.

[20] Ito, K., & O'Dair, M. (2019). A critical examination of the application of blockchain technology to intellectual property management. Business Transformation through Blockchain: Volume II, 317-335.

[21] Asghar, I., Egaji, O. A., & Griffiths, M. G. (2023). An Overview of Blockchain Technology for Intellectual Property Management. Intelligent Sustainable Systems: Selected Papers of WorldS4 2022, Volume 2, 247-266.

[22] Kimani, D., Adams, K., Attah-Boakye, R., Ullah, S., Frecknall-Hughes, J., & Kim, J. (2020). Blockchain, business and the fourth industrial revolution: Whence, whither, wherefore and how?. Technological Forecasting and Social Change, 161, 120254.

Factors Affecting the Use of Online Recommendation Systems in E-Commerce in Croatia

V. Šebek, S. Antolović, C. Kramarić
Algebra University College / Department of Economics, Zagreb, Croatia
vanja.sebek@algebra.hr, stella.antolovic@algebra.hr, claudio.kramaric@racunarstvo.hr

***Abstract* - The concept of utilizing IT for recommendations has existed since the inception of computer science, but it has become more prevalent in recent years as e-commerce has grown significantly. Because recommendation systems are effective at simulating customer behaviour and offering users personalisation, they have helped numerous e-commerce and content providers expand. In order to determine what influences customers' adoption and usage of online purchase recommendation systems, this research explores the factors affecting recommendation systems' adoption and use in e-commerce in Croatia. The structural model of partial least squares and the Unified Theory of Adoption and Use of Technology (UTAUT) were employed to accomplish this goal (PLS). 130 users of online recommendation systems made up the sample. Research has shown that users will only utilize recommendations if they are given supplementary goods that improve their purchasing experience. Users trust recommendation systems (43%), but they have reservations about whether they deliver on their promises which is related to the fact that only a small number of retailers track and evaluate the effectiveness of recommendation systems in their online stores.**

Keywords* - *e-commerce; recommendation systems; artificial intelligence

I. INTRODUCTION

The rapid growth of media, information and network technologies, trading platforms and social networks has generated a massive amount of data that people find difficult to process and comprehend daily. This proliferation of primarily repetitive information places a strain on the human brain and falls short of delivering the desired level of efficiency. The combined effects of this information overload negatively impact productivity and satisfaction with its use. To put it simply, the speed at which users can respond to specific information is slower than the speed at which the information is being transmitted, and the increased availability of information options results in a decreased rate of utilization. This phenomenon is referred to as information overload and is a result of the ongoing and rapidly evolving technological revolution. According to Toffler [1], information overload refers to the difficulties an individual may encounter when trying to comprehend a problem or decide due to the overwhelming presence of information. Information overload places information professionals in a constant state of mental overload or stress, affecting them on a global scale. It has a significant negative effect on the psychological well-being of individuals [2]. Due to a considerable amount of information, the decision-making process in purchasing is under immense pressure [3]. In light of this, recommendation algorithms have a significant impact on the shopping experience in e-commerce. These systems leverage consumer-purchasing patterns to learn about and help users choose the things they should buy, which helps reduce information overload. Furthermore, recommendation systems (RS) lower the time of searching for and choosing products by making recommendations based on the user's profile, and they are critical due to their primary function: predicting whether the user will be satisfied with a particular product [4].

II. THEORY AND PREVIOUS RESEARCH

An online RS is a type of information filtering system that provides personalized recommendations to users based on their preferences, interests and behaviour. It is intended to make pertinent suggestions to consumers while they utilize an online platform, such as an e-commerce website, social-media or an online streaming service, in real-time. RS use a combination of user data, machine learning (ML) algorithms, as well as artificial intelligence (AI), to make personalized recommendations. The data used by these systems includes user profiles, browsing and purchase history, ratings and reviews, social network data, and other behavioural data. The ML algorithms analyse this data to identify patterns, make predictions, and recommend items that are likely to be of interest to the consumer. Online RS come in a variety of forms, including content-based filtering, collaborative filtering, and hybrid methods that integrate both approaches [5]. In recent years, these systems have gained popularity as a means of improving user experience and raising user engagement on online platforms. With the rise of social media, social elements have become important part of RS [5].

The importance of RS in consumer journey are studied in numerous researches, in contrast to prior study that concentrated on technological elements [6]. In context of RS in consumer journey, the focus of

researches varies, from efficiency of RS to consumer satisfaction and trust towards RS [7]. To be used by consumer in the purchase process, RS have to be perceived as trustworthy, as well as relevant. Using RS showed positive impact on the shopping process satisfaction level [8] and those consumer who consulted product suggestions chose suggested products twice as frequently as those who didn't [9]. Perceived RS information quality and system quality, two important characteristics of efficient RS, have a significant effect on consumers' satisfaction and trust [5]. Consumers' perception of the importance of quality and relevant information often leads to neglecting other business related factors, as the relevance of suggested content is top priority focus for RS designers. RS must benefit consumers, but also, they are great way for sellers to increase profitability. Research [10] showed that the RS systems designed by balancing both profit and relevance, compared to those that are oriented to recommending relevant content, generate higher profit, while not influencing the consumer' trust and profit gains were due to balance of consumer's trust and the diversity and relevance of RS. Development of social media increased the social component of RS. Researches [7][11] have shown that consumers most frequently utilize recommendations coming from other consumers. Also, consumers tend to trust more third party RS than RS coming from the seller [7]. Some of the recent researches, such as [12] has investigating connection between the level of psychological ownership (PO) and RS designed preferences, showing that users with high PO prefer user-centric RS approach. Furthermore, consumers learn through their purchase process and apply the weights they ascribe to different product attributes. The RS that learns user preferences through the purchase as even if a recommender system fully understands consumer utility, the highest utility product might not be the greatest advice if the customer needs to discover his or her own preferences [13].

With the development of the technology and more complex consumer journey with information overload as important factor that influence consumer's purchase decision, RS becomes more and more relevant in a practice, as well as in scientific research. With that in mind, and the fact that research of adoption and usage of RS by consumers in Croatia has not been conducted and published in major scientific databases, this research conducted.

III. RESEARCH METHODOLOGY

The research conducted in this paper was influenced by similar research studies [14][16].The methodology used in this research is The Unified Theory of Acceptance and Use of Technology (UTAUT) is one of the most comprehensive technology acceptance models, integrating eight influential acceptance models regarding individual acceptance of newly introduced information technologies. [15]

UTAUT model uses three determinants: performance expectancy (PE), effort expectancy (EE), and social influence (SI). PE refers to the user's expectation that using the system will lead to improved job performance. Research suggested that this is one of the most important predictors of the intention to use technology [15]. EE measures the degree of ease of use or complexity of the technology associated with the use of the technology. SI is the perception of an individual that important others think he should adopt a new system.

UTATU 2 model is often extended by more determinants: facilitating conditions (FC) are the resources and support that a consumer perceives as available to engage with technology; price value (PV) refers to the value that a consumer perceives in relation to the price paid for using the technology and trust (TR) which a crucial variable as it plays a significant role in shaping a person's perception [15].

The UTAUT model was deemed most appropriate for this study, which aimed to assess the behavioural intention and use of RS among customers in Croatia. Data was gathered through an online survey. Snowball methodology of recruiting respondents was applied. N=130 surveys from respondents who indicated they utilize online shopping, made up the analysed sample. A 5-point Likert scale was used to measure all determinants.

Based on previous researcher's findings, UTAUT model was extended and adapted for the purpose of this research (Fig. 1) and hypothesis are formed as follows:

- Hypothesis 1 (H1): PE is expected to have a positive effect on the behavioral intention to use RS in e-commerce.
- Hypothesis 2 (H2) EE positively influences the behavioral intention of RS in e-commerce.
- Hypothesis 3 (H3) SI positively influences the behavioral intention of RS in e-commerce
- Hypothesis 4 (H4) PV positively influences the behavioral intention of RS in e-commerce
- Hypothesis 5 (H5) T positively influences the behavioral intention of RS in e-commerce
- Hypothesis 6 (H6) Behavioral intention of RS in e-commerce positively effects the use

PLS statistical software was used to statistical data analysis.

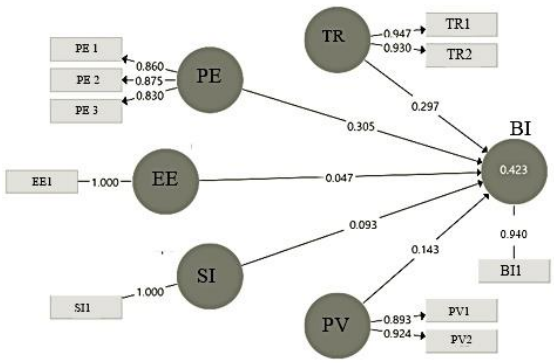


Figure 1. Research model

IV. DISCUSSION

The scales were adopted from the original UTAUT model from [16]. To test the proposed structural model, we used partial least squares–structural equation modelling (PLS-SEM). We checked the absence of measure bias error or common method bias (CMB) as recommended in [17].

In this model, the occurrence of a VIF greater than 3.3 is preferable as an indication of pathological collinearity. To do so, we conducted a full collinearity test in order to rule out CMB. We used data from our original PLS algorithm report. In first step we took results from latent variables and paste them in first free row in our original data set. In a second step, we added a new column called Random V populated with random values that represents all values of that original data set. From that point we created new canvas/model in our structural model. In that model all variance inflation factors (VIFs) extracted by this method must be lower than 3.3 to confirm that the sample has no CMB as in Table 1.

PLS SEM model tries to ensure the reliability and validity of the measurement scales. Factor loading show how well an item represents the underlying construct. Loadings over 0.70 are recommended as in [15]. In our work we eliminated Si3 and Ee3 from the scale.

TABLE I. TEST COMMON METHOD BIAS, VARIANCE INFLATION FACTORS

	Random V
Social influence	1069
Effort expectancy	1505
Performance expectancy	1097
Behavioral intention	1403
Trust	1140
Price value	1008

TABLE II. MEASUREMENT SCALES AND LOADING S

Construct	Items	Loadings
Social influence	Si1	0.804
	Si2	0.903
	Si3	
Effort expectancy	Ee1	
	Ee2	0.89
	Ee3	
Performance expectancy	Pe1	0.86
	Pe2	0.875
	Pe3	0.83
Trust	T1	0.947
	T2	0.93
Price value	Pv1	0.893
	Pv2	0.924
Behavioral intention	Bi	1

The results are presented in Table II.

We analyzed the reliability of the constructs with two indicators: Cronbach's Alpha and composite reliability. The values obtained were over the 0.7 suggested. In terms of convergent validity, the average variance extracted (AVE) was used. The results were over the 0.5 as in Table 3. [15] Fornell–Larcker's test (Table 4.) was used to test discriminant validity of model, which compares the square roots of the AVE (shown in the diagonal in bold) with the correlations of each variable (shown in rows and columns). The first value must be greater than those in their respective rows and columns as in [18]. To check the statistical significance of each of the coefficients or paths (Table 5) we carried out bootstrapping with 10,000 subsamples and used the Standardized Root Mean Square Residual to assess the model's goodness of fit [15]. The

TABLE III. COMPOSITE RELIABILITY AND CONVERGENT VALIDITY

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
PE	0.817	0.823	0.891	0.732
PV	0.79	0.805	0.904	0.826
SI	0.81	0.836	0.912	0.839
Trust	0.865	0.876	0.936	0.88

TABLE IV. FORNELL-LARCKER CRITERION

	EE	PE	PV	SI	Trust	pn
EE	**1**					
PE	0.551	**0.855**				
PV	0.402	0.628	**0.909**			
SI	0.202	0.217	0.228	**0.916**		
Trust	0.524	0.665	0.652	0.299	**0.938**	
pn	0.483	0.578	0.519	0.236	0.584	**1**

TABLE V. STRUCTURAL MODEL

Hyoptesis	Path coefficiens	Rsquare	Fsquare	Validate
EE on BI	0.121		0.031	Supported
PE on BI	0.284		0.041	Supported
PV on BI	0.116		0.019	Supported
SI on BI	0.045		0.004	Not supported
TR on BE	0.228		0.043	Supported
BI		0.437		
srmr				0.051
Q^2predict				0.386

value obtained was 0.051, less than the 0.08, suggesting a good fit of the complete model. The model results in Table 5. also show both the verified hypotheses and the size of their effect. Four of five relationships formulated are significant, of which only the results for expectations–intention to use ratio exceed the minimum 0.2 recommended by [15]. In all cases where the ratios were significant, the effect size (Fsquare) was between 0.015 and 0.35, except in scenarios Si on Bi which were below it. However, the model has a very good fit as the SRMR obtained is only 0.051, well below 0.08. Stone–Geisser's Q2 was calculated to evaluate the predictive capacity of the model. Result is 0.386, value greater than 0 so it is conclude that model has a relevant predictive capacity [15] (Table 5)

From the total number of respondents (N130), there were 70 women and 60 men; 85% or 65% were between the ages of 25 and 54. Of the respondents, 56% declare they have a bachelor's degree or higher, and 41% have completed secondary school. In this context, to the question 'How often do you shop online,' the vast majority, or almost 87% of respondents, answered that they shop one to several times a month. More precisely, 37% of the surveyed shop once or more times a month, and 40% once or more in three months. A significant increase in shopping frequency compared to 2018 research confirms the post-pandemic statements about the rise in sales in e-commerce in Croatia [19].

As we stated previously, the aim was to investigate how the mentioned constructs influence the use of RS in e-commerce. The measurement model that evaluates the validity of the construct and the structural model that evaluates the connection of the variables after application showed a positive result. Survey results showing composite latent variable concordances ranging from 0.860 to 1. The Cronbach's Alphas of the five constructs are displayed in Figure 4. All CAlphas are above the threshold of .70 (Nunnally, 1978), meaning that the reliability is confirmed. Composite reliability indicators for all latent variables must be between the value of 0.7, also called the critical value, and the value of 0.99, representing a very high degree of internal consistency [16]. In this model, the reliable composite values range from the minimum value of 0.890 to the maximum value of 1, in complete accordance with the previously defined interval. It represents a high degree of internal consistency or reliability of the model for the variables. Discriminant validity is a measure that is compared with different but related concepts. It should not be correlated with measures that measure another concept.

When it comes to trust 'in' and expectations 'from' RS - two factors that are incredibly high on the list of users - respondents say that they will use them only when the solutions offered by the systems make sense to them, i.e., when they are shown to them complementary products, similar products but with a better price, products that complement and enrich their shopping experience. According to the answers, 66% of the respondents agree that they can trust the system's recommendations, but almost 30% of the respondents are neutral on this issue (neither sure nor unsure that RS should be trusted). On the next question, their skepticism stands out. Namely, less than half, or 43%, agree that the systems 'fulfill what they claim' and 29% are undecided. In short, users trust recommendation systems but are skeptical about the statement that they live up to what they claim (that is, in what they do, which we could relate to the fact that a tiny number of retailers monitor and analyze data on the performance of RS. Namely, RS often display or link completely non-complementary/irrelevant products and the like). RS are part of our everyday life, and it is almost impossible to avoid them. Accordingly, most respondents, 48% of them, declare that they will use them to the same extent as before, and even 30% that they will use them more often. Only a more minor number declare that they will use them less (4%) or that they will stop using them (1.5%), which, in today's context of the general spread of the Internet and various smart devices, would be an almost impossible mission.

V. CONCLUSION

When an individual is surrounded by constantly changing mass media and new information contexts, their ability to make accurate predictions decreases significantly. To restore their prediction accuracy to an average level, they must rapidly process more data. Recommender systems are a specialized type of information filtering system that attempt to anticipate a user's evaluation or liking for a specific item, product, or service. This prediction is then displayed to the user as recommendations and suggestions based on the results. In this context, the user will be satisfied if three parameters are met: how much the recommendation is focused on his needs (personalized), how interesting it is and how useful it is to him. The primary goal of any recommendation system is to build a quality and objective mathematical model that will be able to measure these three. Their acceptance is wide because they facilitate the processes in the store, but also more broadly: it is used by customers who find it easier and faster to shop, and by merchants in e-stores to increase the value of their basket, make it easier for customers and make shopping as high-quality as possible. After the collected data, the test was carried out through the UTAUT model in the already mentioned SmartPLS 3.0 software tool, which determines the success of the research using various calculation measures. Based on the results all hypotheses from the test are above the limit values and were accepted. The high predictive ability of the model was established for most variables. All selected constructs have a positive influence and the greatest role on the behavioral intention to use the RS in e-commerce. All except the construct: Social impact (what others think about the use), which was to be expected. Namely, recommendation systems are practically impossible to avoid, they do not represent a novelty, and the interest is about how much someone uses it or what opinion they have about the use of the recommendation system - for other users, it is a less important question. Users and customers in e-commerce believe in what the systems recommend to them but express skepticism when they answer whether these systems fulfill what they do. However, almost 80% also state that they intend to use the recommendation systems equally or even more than now. That is important information for store owners and managers: customers and users recognize the importance

of systems in everyday life, are aware of them, and use them almost equally, regardless of whether they buy utilitarian or hedonistic products. On the other hand, e-merchants should pay more attention to the monitoring and analysis/analytics of referral systems because it is proved that these systems have a very measurable impact on revenue growth, customer satisfaction, purchase intentions, time spent on the site, which are all key indicators of success when it comes to e-commerce.

REFERENCES

[1] S. Galloway, Post Corona, Penguin Random House, 2020

[2] W. D. Hoyer, “An examination of consumer decision making for a common repeat purchase product,” Journal of consumer research, vol. 11(3), 1984, pp. 822-829.

[3] N. Zeldes, D. Sward, and S. Louchheim, Infomania: Why we can't afford to ignore it any longer. First Monday, 2007

[4] T. Manos, T. H. King, S. Kallumadi, V. Murdock, and M. de Rijke. "Challenges and research opportunities in ecommerce search and recommendations," ACM Sigir Forum, vol. 54, no. 1, 2021, pp. 1-23.

[5] X. Yang. "Determinants of consumers’ continuance intention to use social recommender systems: A self-regulation perspective." Technology in Society, vol. 64:101464, 2021

[6] C. Yin, Y. Guo, J. Yang and X. Ren, "A new recommendation system on the basis of consumer initiative decision based on an associative classification approach." Industrial Management & Data Systems, vol. 1118, 2018, pp. 188-203.

[7] E. Lepkowska - White, "Are they listening? Designing online recommendations for today's consumers." Journal of Research in Interactive Marketing, vol. 7, 2013, pp. 182-200

[8] G.N. Punj, R. Moore, “Smart versus knowledgeable online recommendation agents,” Journal of Interactive Marketing, Vol. 21, 2007, pp. 46-60.

[9] S. Senecal, J. Nantel. "The influence of online product recommendations on consumers’ online choices," Journal of retailing, vol. 80, 2004, pp. 159-169.

[10] U. Panniello, H. Shawndra, and M. Gorgoglione. "The impact of profit incentives on the relevance of online recommendations." Electronic Commerce Research and Applications, vol. 20, 2016, pp. 87-104.

[11] Y.F. Chen, "Herd behavior in purchasing books online," Computers in Human Behavior, vol. 24, 2008, pp. 1977-1992.

[12] B.G. Seo, D.H. Park, "The effective recommendation approaches depending on user’s psychological ownership in online content service: user-centric versus content-centric recommendations," Behaviour & Information Technology, 2022, pp. 1-13.

[13] D. Dzyabura, J. R. Hauser, "Recommending products when consumers learn their preference weights." Marketing Science vol. 38, 2019, pp. 417-441.

[14] P. Silva,. “Davis' technology acceptance model (TAM)” Information seeking behavior and technology adoption: Theories and trends, 1989, pp.205-219.

[15] J.P. Cabrera-Sánchez, I. Ramos-de-Luna, E. Carvajal-Trujillo, A. F. Villarejo-Ramos, “Online recommendation systems: Factors influencing use in e-commerce,” Sustainability 12, Vol. 21:8888, 2020

[16] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, “User Acceptance of Information Technology: Toward a Unified View,” MIS Quaterly., vol. 27, 2003, pp. 425–478.

[17] N. Kock, “Common Method Bias in PLS-SEM: A Full Collinearity Assesment Approach”. International Journal of e-Collaboration (ijec), vol 11, 2015, pp. 1-10

[18] D. Barclay, R. Thompson, C. Higgins, ”The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Uses as an Ilustration” Technol. Stud. 1995, vol. 2, pp. 285–309.

[19] C. Kramarić, V. Šebek and N. Kaurin Knežević, “Croatian E-commerce Market: Overview, Potentials and Challenges”, Oeconomicus, vol. 13, 2020, pp. 22-30

Co-creating Community-level Indicators: Involving Communities in the Digitalization of Energy for Empowering Energy Citizenship

A. Kumar, A. Wolff, B. Naqvi
Software Engineering, LENS, LUT University, Lappeenranta, Finland
ajesh.kumar@lut.fi

***Abstract* - Energy citizenship refers to the participation of individuals or communities in activities that support the green transition. Whilst individuals may set their own targets for greener living, several initiatives are also being designed and encouraged at the city and regional level, e.g., renewable energy cooperatives and increased electric mobility. Enacting change is a long-term process, requiring monitoring and possible adaptation along the way and initiatives may fail without buy-in from the local community. Digitalization of energy and energy-related systems can produce large quantities of data that the has potential for helping to achieve green targets and monitor progress against them. However, when the initiatives and the data collected to monitor them are defined in a purely top-down manner it risks missing the priorities of local communities, which may affect willingness to participate. This paper explores a co-creation process through which local communities define transition goals and ways to monitor the progress that reflects their own priorities in the form of community-level level indicators (CLIs), as well as the metrics needed to track them. It will explore a common template for defining CLIs and describe their use within several green initiatives across Europe.**

Keywords - digitalization, energy systems, community-level indicators, empowerment, society 5x

I. INTRODUCTION

Energy is a vital resource for communities, and its availability and cost can greatly affect a community's economic and social well-being [1]. Perceptions of energy can vary greatly among different communities with some placing a greater emphasis on environmental concerns and others prioritizing economic development [1]. Additionally, low-income and marginalized communities often have limited access to energy services and may be disproportionately impacted by energy-related issues [2]. It is therefore vital that policymakers and energy providers consider different communities' diverse perspectives and needs when developing and implementing energy policies and programs.

When it comes to the inclusion of multiple perspectives in the design and development of new policies and programs, the co-design approach is often utilized [3]. Co-design is a participatory design approach that involves the active involvement of stakeholders, including community members, in the design process of a project or initiative. It has also been found to be effective in community settings for creating sustainable and equitable energy systems and energy practices [3].

Furthermore, for global goals like decarbonization and green transition where multiple communities and stakeholders get involved, there is a need for metrics that enable policymakers, planners, and community leaders to make informed decisions based on historical and current trends, as well as predictions for future outcomes. A Community-Level Indicator (CLI), also known as a community indicator is a metric that enables multiple stakeholders to make informed decisions and predictions. CLIs are extensively employed in various fields of study, including health, sustainability, environment, climate, energy, and urban planning. The utilization of CLI can provide insights into the performance of policy implementation across social, economic, and environmental dimensions [3]. By combining indicators, they generate a picture of what is happening in a community and can provide information about past trends, current realities, and future directions to support decision-making.

In the field of energy, Community Level Indicators (CLIs) are crucial for promoting energy sustainability and striving toward an equitable transition. These indicators encompass a range of factors, such as access to electricity, availability of clean fuels and technology for cooking and heating, and the proportion of renewable energy in use [4].

This paper discusses the utilization of the co-design approach within the context of energy communities and outlines the steps required to conduct co-design workshops across different communities, either online or face-to-face. Furthermore, the paper presents the outcomes of conducting co-design workshops within four different energy communities in Europe. The workshops were aimed at developing Community Level Indicators (CLIs) that reveal how each community prefers to measure their progress towards objectives and decarbonization targets of local initiatives towards a green transition.

The remainder of the paper is organized as follows. Section II presents the background of the topic. Section III presents the details of developing the co-design framework, Section IV presents the results. Section V presents the discussion, and Section VI concludes the paper.

II. BACKGROUND

This section presents details about (1) community-level indicators (CLI), and (2) the co-design approach. The details are presented in the following subsequent sub-sections.

A. Community Level Indicators

A distinguishing characteristic of CLI is that they are derived from community-level information rather than individual-level data, e.g., indicators focusing on aspects such as the presence of no-smoking areas in restaurants or smoking cessation clinics within a local community to measure progress on improving the health of a community, rather than focusing on individual smoking behaviors and whether they go up or down.

In the existing literature [5], the authors have outlined five main applications for community or neighborhood indicators including, (1) s*ituation Analysis*: analyzing whether situations in a community are getting better or worse., (2) p*olicy analysis and planning*: providing evidence for formulating policy., (3) p*erformance management and evaluation*: monitoring the performance of an intervention., (4) *education and engagement*: empowering communities to act on problems, by presenting data (e.g., crime statistics) in ways that are easy to understand and interpret., and (5) n*eighborhood research*: identifying patterns of cause and effect, such as the effect of neighborhood conditions on individual outcomes.

In the existing literature [6], the authors created a framework for monitoring the sustainability of local fisheries using Community Level Indicators, which combined expert and local knowledge through focus groups. The indicators were categorized based on various aspects of sustainability that were relevant to the local context, including community, ecological, institutional, and socio-economic factors. The authors [6] suggest that a community-level indicator (CLI) approach can incorporate diverse expertise through a combination of top-down and bottom-up processes, as seen in their own case study.

In the existing literature [7], the author identified two approaches for developing CLIs, (1) outcome-oriented, and (2) context-oriented. The former focuses on finding differences in outcomes within the community area, while the latter focuses on the community as an environment that may affect result in positive or negative ways, which in turn may reflect in higher or lower rates of expected outcomes.

Furthermore, Participatory Sensing (PS) data can be useful for individuals with limited experience in data analysis by co-creating CLIs through participatory workshops [8]. Participatory Sensing enables communities to gather data to support their cause, such as recording noise pollution levels to lobby local authorities [8]. In [8] researchers engaged local communities in workshops to identify CLIs that were meaningful. Moreover, in [8], it was found that the co-creation of community-level indicators increased the awareness of the problem and the need to find solutions. This is consistent with the existing work [5] perspective, which suggests that indicators can empower communities to act on issues of interest, particularly when presented in an easily understandable format for a broad audience.

To summarize, the following aspects are the key to be considered in the development of CLIs:

- Familiarization with the problem
- Collaborative goal setting.
- Mapping goals to indicators.

B. Co-design Methodology

The term co-design encompasses different participatory methods that aim to engage users or customers in a collaborative and innovative design process. Co-design offers several benefits, including the ability to produce outcomes that are better aligned with the needs of users or customers, making them more likely to accept and adopt the results of the process [9].

The co-design process involves iteratively refining initially ambiguous ideas. In [10] researchers identified four common phases of co-design, (1) starting with pre-design activities to prepare participants for the process, (2) the generative phase for exploring the design space and producing ideas using tailored tools and methods, (3) evaluative phase to assess the outcomes, and (4) post-design phase for research to determine how well the design meets the needs of people. The generative phase typically employs tools and methods based on making, telling, enacting, or combining these approaches. Making involves creative acts, supported by probes, toolkits, or prototypes. Telling is more dialogue-based, with techniques like World Café, fishbowl, diaries, or prompts. Enacting involves demonstrating through action, such as body storming, where a story is acted out to illustrate a particular design.

To summarize, the key aspects to consider while planning and executing co-design are:

- The four stages can be employed to structure the co-design of CLIs.
- When selecting activities for the generative phase of the co-design, it is advisable to concentrate on making and telling. Enacting does not correspond to the need to support both online and offline co-design, nor is it suitable for the design task.

III. DEVELOPING THE CO-DESIGN PROTOCOL

By drawing from various domain-specific literature we proposed a co-design protocol that includes a 4-stage setup such as pre-design, generative, evaluative, and post-design stages. Table I outlines the stages which can be used for co-designing community-level indicators (CLIs) and furthermore these co-design approaches were implemented for identifying CLIs. The co-design stages provide suggestions for actions or tasks that may be included and describe the rationale considerations and the involvement of stakeholders at each stage. These stages are discussed below along with the type of stakeholders or participants.

TABLE I. CO-DESIGN STAGES FOR IDENTIFYING CLI ACROSS ENERGY COMMUNITIES

Co-design stage	Actions	Description	Task Participants
Pre-design	Identify project goals	Identify the purpose of the CLIs, especially so that these goals can be communicated clearly as part of the co-design process [5]. Define the scope of involvement in utilizing CLIs beyond their ideation, taking into consideration factors such as the availability of data, requirements for specialist tools, and similar.	Researchers
	Conduct a literature review	A literature review allows researchers to produce an initial set of indicators to 'seed' the co-design and provide informative examples. This is common literature and not related to any individual energy community. These CLIs may have an outcome or a contextual orientation[7].	Researchers
	Filter CLIs for local context	Conduct an initial filtering of CLIs to reduce the CLIs from the literature to a small set relevant to the case study.	Researchers
	Identify key stakeholders	The stakeholders represented in the co-design activity may play a significant role in deciding the relevant CLIs.It is thus important to consider how to effectively combine top-down and bottom-up processes to maximize the benefits of community involvement. [6].	Researchers
	Define a recruitment strategy	Define a recruitment strategy for attracting identified stakeholders to participate in the workshop(s). The gains from participation should be identified and communicated.	Researchers, technologists, problem-owners, general public
	Planning for codesign workshop(s)	Select a co-design methodology to follow, e.g., Dialogue Lab, World Café, or another. Select appropriate tools to facilitate ideation within the chosen co-design process, considering activities may be either online or offline – for example making or telling but not enacting [10]. Define how space is set up [11], and in these days, especially whether it will be online, face-to-face, hybrid/synchronous, or asynchronous. Ensure adherence to ethical principles and use of consent forms as well as data management. Identification of materials (including technologies and software) needed to support co-design activities.	Researchers, technologists, problem-owners, general public
Generative	Introduction and familiarisation with the problem	Utilize a subset of CLIs as examples of indicators that are more relevant to the community. In this step, CLIs are explained in detail with examples to show how the CLI is used to measure some community parameters and data required	Researchers, technologists, problem-owners, general public
	Defining the community boundary	Based on geography, or some other definition, boundaries may be defined to mark the scope of the co-design activity[7]	Researchers, technologists, problem-owners, general public
	Collaborative community goal setting	Setting up community goals collaboratively allows participants to identify common goals as well as help reach a consensus on which goals to prioritize and the expected timeline.	Researchers, technologists, problem-owners, general public
	Ideating new indicators within the framing of goals and CLI dimensions	Whether to define thresholds, create CLI repositories, levels of evaluation, etc. How to overcome data literacy, technical, political, and privacy issues? Or problems with 'small' data? [13]. How will CLIs be measured?	Researchers, technologists, problem-owners, general public
Evaluative	Evaluating CLIs and making the final selection	Evaluate CLIs according to aspects, such as relevance, completeness, availability, measurability, reliability, familiarity, nonredundancy, and independence.	Researchers, technologists, problem-owners, general public
	Closing activities	E.g., 'harvest' or sharing the final outputs and capturing the results, debriefing, and filling questionnaires.	Researchers, technologists, problem-owners, general public
Post-Design	Utilizing CLIs within communities as part of defining CTPs	Identifying the availability of required data. Analysing available data in the context of the specific goal. Longitudinal tracking of co-design participants to understand the effects of participation?	Researchers, technologists, problem-owners, general public

As depicted in Table I, during the pre-design stage project goals are identified. The literature review should be conducted to produce an initial set of indicators to 'seed' the co-design and provide informative examples. Once CLIs are identified they should be filtered for local context the next steps are to identify key stakeholders who could take part in the workshop activity and invite the potential participants explaining the goal for the workshop along

with dates, duration, and the location that has been identified for the workshop. Once workshop participants are recruited the planning for codesign workshop(s) must be created.

In the generative stage, workshops are conducted by first giving an introduction about the community and participants are made to familiarize themselves with the problem. Moreover, defining the community boundary which could be based on geography or other factors could help determine the scope of the workshop and CLI generated. In the next step, participants may indulge in collaborative goal setting for the community and ideating new indicators within the framing of goals and CLI dimensions.

In the evaluation stage, CLIs are evaluated according to aspects, such as relevance, completeness, availability, measurability, reliability, familiarity, non-redundancy, and independence into a final list. In closing activities, workshop facilitators share the final outputs and the results, debrief the participants, and may ask them to fill out questionnaires.

In the post-design stage, the generated CLIs can be utilized to map transition pathways to achieve the sustainability goals of the community.

The materials and equipment that could facilitate the face-to-face co-design process may include consent forms, refreshments, a large screen for a group presentation, A3-sized sheets, whiteboards or flip charts, tables and chairs, sticky notes, medium-sized markers and pencils, pin boards or thick cardboard, A5-sized paper or card, and glue.

IV. Results

As stated in section III, the co-design protocol was executed by involving four different energy communities, energy researchers, policymakers, energy executives, and the general public. The workshops aimed to share the needs and experiences of the community on the topic of energy, environment, and climate and to reflect together on what possible actions can be put in place to achieve the decarbonization of their neighborhoods and cities.

The recruitment strategy [14] was incorporated to recruit participants for the workshops. To ensure the successful implementation of CLI workshops, effective recruitment, and involvement of key stakeholders are crucial, and gatekeepers play an important role in this process. Gatekeepers act as coordinators of information between different groups of stakeholders and represent individuals who are heard within their contexts and have easier access to different parties, thus facilitating communication and bridging gaps.

The workshop aimed to identify the manifestations of energy citizenship at various levels such as local, regional, and national levels. The workshop enabled participants to explore various aspects, such as how individuals can reduce energy consumption in their daily lives, how the community can contribute to reducing energy consumption in the neighborhood, and how the community can collectively take action to have a global impact. This allowed participants to gain a better understanding of energy consumption, generation, and energy business as well as local energy policies. This increased the overall understanding of the topic, guided by facilitators, and allowed participants to explore different goals types like short-term, mid-term, and long-term and listed different actions that could lead to the achievement of goals.

The various activities carried out during the workshops led to the discovery of several community-level indicators (CLIs) presented in Table II. Though there were many indicators that were co-created, not all of them were relevant or measurable. These indicators can also be classified into different types environmental, technical, economic, and social. The in-depth discussions during the workshops resulted in well-refined CLIs. All the CLIs explored are outside the scope of this paper.

V. Discussion

Based on the findings, it can be concluded that the co-design approach is effective in producing a diverse range of community-level indicators (CLIs) that addressed various goals and belonged to different categories, including environmental, social, technical, and economic. The workshop plan template was also useful for organizing the workshops into manageable tasks and creating a plan.

In the co-creation of community-level indicators, organizing a workshop can be challenging, especially in situations like a pandemic. Therefore, workshops need to incorporate mitigation strategies and be adaptable. A workshop protocol was found to be useful and adaptable, as both face-to-face and online workshops were able to successfully follow it and tailor it to their own needs. The results showed that the CLIs identified belonged to different categories such as Environmental, Technical, Economic, and Social, broadly. However, a particular workshop may have defined its own set of categories for the indicators.

The next step after listing the community-level indicators is to evaluate them based on specific questions, such as how to measure the indicator, where is the data, and whether new data sources require investment. By selecting some of the proposed indicators, it is possible to analyze the potential for data collection to track improvements in energy citizenship in the area. Data can be collected from various sources, including technical and non-technical data

TABLE II LIST OF CLI CO-DESIGNED DURING WORKSHOPS

Communities	Description	Indicator type	Indicators
Bologna, Italy	The mixed-use district in Bologna comprises two neighborhoods,	Environmental	Percentage of municipal budget directed to environmental-themed workshops.

Communities	Description	Indicator type	Indicators
	namely Pilastro (residential) and Roveri (industrial). Pilastro being planned as a self-sufficient neighborhood with mixed services and activities, it has become a mono-functional residential area with socio-economic challenges such as energy poverty. In contrast, Roveri is home to diverse companies across multiple sectors and is actively involved in establishing the first energy community in the city, leveraging the presence of local industrial partners with large photovoltaic plants.		The number of new energy/environmental-themed associations.
		Technical	The number of people installing Smart Meters or similar devices.
			The number of buildings monitored by Smart Meters.
		Economic	Increase in public incentives to finance photovoltaics.
			The number of municipal energy incentives for the area concerned
		Social	The number of people willing to do the proposed energy improvements.
			The number of events organized by schools to raise awareness and the number of families involved in the events.
UR BEROA, Spain	UR BEROA is a cooperative organization in the Bera Bera neighborhood of San Sebastian that specializes in local energy efficiency. Its main objective is to provide a sustainable energy plan that consists of three natural gas boilers, a cogeneration engine, a biomass boiler, and solar panels that produce hot water, which can be remotely managed based on the specific needs of each area.	Environmental	Number of speeches promoting environmental attitude in the yearly meetings
			Number of articles promoting environmental attitude in the bulletin of the cooperative
		Technical	Electricity generated by photovoltaics(KW)
			Share of rooftop surface covered by PV panels
		Economic	Reduction of the heating and hot water costs
			Subsidies achieved for housing energy efficiency improvements
		Social	Number of new members
			The ratio of subscriptions/de-subscription of cooperative membership
Coopérnico - Portugal	Coopérnico is a renewable energy cooperative in Portugal with over 1,000 members, including individuals, small and medium-sized businesses, and municipalities. Its objective is to engage its members in transforming the energy sector into a more renewable, socially equitable, and collaborative one.	Environmental	"Carbon Handprint" (complementary to "Carbon Footprint") measures the actions (political, economic, educational...) carried out by Coopérnico to encourage the reduction of the carbon footprint (among its members, in society in general).
			To reduce the carbon footprint in photovoltaic projects funded by Coopérnico(CO_2 emissions before and after investment).
		Technical	kWh marketed by Coopérnico.
			kWh produced in Coopérnico projects.
		Economic	No. of photovoltaic systems co-financed by co-members.
			No. and ease of access to public and private financial support for such energy efficiency interventions
		Social	No. of cooperative members.
			Level of compliance with cooperative principles (establishing for each one a metric of their own).
Virtual community-Germany	The Earnest App is an online application that aims to inform and educate its users about energy consumption and CO_2 emissions through interactive features, such as quizzes and challenges. Its goal is to prompt users to reflect on their energy and mobility habits, including the use of sustainable transportation options and reducing long-distance travel, with the hope that this understanding will lead to positive lifestyle changes related to energy efficiency.	Environmental	Assess and increase information on sustainable lifestyles available in the city of Darmstadt
			Collecting data on the diversity of community members
		Technical	Assess internet/app user behavior (numbers, time spent on the website, engaged with what information, etc.)
		Economic	Assess public funding opportunities for various different communities
			Increase funding programs for sustainability education in each region
		Social	Establish organizational headquarters
			Change school curriculum

and additional tools such as guided questionnaires, action research, school training projects, participant observation, administrator or policymaker interviews, and open data portals can be used to gather data. Feasibility is another critical aspect that needs to be considered when shortlisting CLIs, which requires careful consideration of what to measure, which unit is suitable for the measurement, where the indicator can be measured, what the data source could be, and what possible challenges might arise.

VI. Conclusion

The replicable workshop protocol effectively guided the co-design workshops and facilitated the creation of community-level indicators (CLIs). The template involved

activities and discussions that stimulated in-depth conversation, common goals, and actionable points. The workshop facilitators played a vital role in generating ideas and keeping the discussion on track. The co-design workshop is dynamic and may require facilitators to redirect the conversation if it goes off-topic. To overcome such challenges, the workshop plan established a timeline for objectives and actions and discussed energy sustainability and community involvement. Together, they identified goals and actions needed to achieve them, and subsequently, determined how to measure their progress. Challenges may also arise during the workshop, such as in one community, where the heat forced participants to move outside during the workshop as the workshop location was unusually hot.

ACKNOWLEDGMENT

The GRETA project (www.projectgreta.eu) has received funding from the European Union's Horizon 2020 Research & Innovation program under grant agreement No 101022317

REFERENCES

[1] McDonald, N. C., & Pearce, J. M. (2013). Community voices: Perspectives on renewable energy in Nunavut. Arctic, 94-104.

[2] Action, P. (2018). *Poor people's energy outlook 2018: achieving inclusive energy access at scale*. Rugby: Practical Action Publishing.

[3] Phillips, R. (2003). *Community indicators*. Chicago, IL: American Planning Association.

[4] Yuan, M. H., & Lo, S. L. (2020). Developing indicators for the monitoring of the sustainability of food, energy, and water. Renewable and Sustainable Energy Reviews, 119, 109565.

[5] Kingsley, G. T., Coulton, C. J., & Pettit, K. L. (2014). Strengthening communities with neighborhood data. Washington, DC: Urban Institute.

[6] Boyd, H., & Charles, A. (2006). Creating community-based indicators to monitor sustainability of local fisheries. Ocean & Coastal Management, 49(5-6), 237-258.

[7] Coulton, C. J. (1995). Using community-level indicators of children's well-being in comprehensive community initiatives. New approaches to evaluating community initiatives: Concepts, methods, and contexts, 1, 173-199.

[8] Coulson, S., Woods, M., Scott, M., Hemment, D., & Balestrini, M. (2018, June). Stop the noise! enhancing meaningfulness in participatory sensing with community level indicators. In Proceedings of the 2018 designing interactive systems conference (pp. 1183-1192).

[9] Steen, M. (2013). Co-design as a process of joint inquiry and imagination. Design issues, 29(2), 16-28.

[10] Sanders, E. B. N., & Stappers, P. J. (2014). Probes, toolkits and prototypes: three approaches to making in codesigning. CoDesign, 10(1), 5-14.

[11] Lucero, A., Vaajakallio, K., & Dalsgaard, P. (2012). The dialogue-labs method: process, space and materials as structuring elements to spark dialogue in co-design events. *CoDesign*, *8*(1), 1-23.

[12] Angelakoglou, K., Nikolopoulos, N., Giourka, P., Svensson, I. L., Tsarchopoulos, P., Tryferidis, A., & Tzovaras, D. (2019). A methodological framework for the selection of key performance indicators to assess smart city solutions. Smart Cities, 2(2), 269-306.

[13] Murphey, D. A. (1999). Presenting community-level data in an" outcomes and indicators" framework: Lessons from Vermont's experience. Public Administration Review, 76-82.

[14] Wolff, A., Klein, L., Kumar, A. (2022). A replicable workshop design template for co-creating Community Level Indicators.D2.2 of the Horizon 2020 project GRETA, EC grant agreement no.101022317, Lappeenranta, Finland

Exploring Regional Innovation Dynamics Using Directed Acyclic Graphs

F. Molinari, D. Cisic and B. Kovacic
University of Rijeka, Rijeka, Croatia
mail@francescomolinari.es

***Abstract* – Biannually, the European Commission publishes the Regional Innovation Scoreboard (RIS) as a tool to benchmark the performance of 240 regions, 22 EU Member States (MS) and 4 Associated Countries (AC): Norway, Serbia, Switzerland, and the UK. Only Cyprus, Estonia, Latvia, Luxembourg, and Malta are included at the country level. The RIS consists of a synthetic index that starts from the arithmetic mean of a set of 21 indicators. This allows ranking the various territories of Europe for their innovative performance. EU Regions, MS and AC are expected to take their position on the ranking into account when designing innovation policies. Due to its importance for policymaking, the RIS is subject to recurrent assessments of the extent to which it constitutes a meaningful measure of a territory's innovation performance. This paper contributes to the debate by proposing an approach based on algorithms to discover the causal relations among the RIS variables and visualize them as edges of Directed Acyclic Graphs (DAGs). In so doing, we shed some light on the limitations of the RIS tool to assess the dynamics and performance drivers of the EU national/regional innovation systems.**

Keywords – causal discovery; algorithms; DAG; Bayesian networks; territorial innovation

I. INTRODUCTION

Biannually, the Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs of the European Commission (EC) publishes the Regional Innovation Scoreboard (RIS). This is a composite index constituted by the arithmetic average of 21 indicators, measured at regional level for the largest-sized countries and national level for the others. The higher the average of those (normalized) indicators, the better performing – whatever this means – is a regional or national territory said to be.

Currently, the RIS – or the global ranking generated from its data, not to forget the specific rankings by indicator – is used as a policy support tool to benchmark the innovative performance of 240 European Union's (EU) regions, 22 Member States (MS) and 4 Associated Countries (AC), namely: Norway, Serbia, Switzerland, and the UK. Of the 22 EU MS, only Cyprus, Estonia, Latvia, Luxembourg, and Malta are considered at the country, not at the regional, level. Table 1 lists the indicators in use in the two most recent editions of the RIS (2019 and 2021).

Part of the research supporting this paper was made possible by an Interreg MED 2014-2020 grant to EFRI, the University of Rijeka's Faculty of Economics and Business, under the project named "Social and Creative" (S&C). However, the opinions expressed here are solely of the authors and do not engage in any way the Interreg MED Joint Secretariat or other partners of the S&C project.

TABLE 1 . LIST OF RIS INDICATORS

Short name	*Indicator*
air	Air emissions by fine particulate matter (PM2.5) in the manufacturing sector
bp_inn	SMEs introducing business process innovations as percentage of SMEs
cited_pub	International scientific co-publications per million population
design	Design applications per billion regional GDP
dig_skills	Individuals who have above basic overall digital skills
emp_SMEs	Employment in innovative SMEs
empl_KI	Employment in knowledge-intensive activities (percentage of total employment)
exp_business	R&D expenditures in the business sector as percentage of GDP
exp_public	R&D expenditures in the public sector as percentage of GDP
inn_exp	Non R&D Innovation expenditures
inn_exp_pp	Innovation expenditures per person employed in innovative SMEs
nn_SMEs	Innovative SMEs collaborating with others as percentage of SMEs
int_sci_pub	International scientific co-publications per million population
IT	ICT specialists (as a percentage of total employment)
LLL	Percentage population aged 25-64 participating in lifelong learning
Marketing innovators	SMEs introducing marketing or organisational innovations as percentage of SMEs
most_cited_pub	Scientific publications among the top-10% most cited publications worldwide
patent	PCT patent applications per billion regional GDP
performance	RIS performance indicator
PP_pub	Public-private co-publications per million population
prod_inn	SMEs introducing product innovations as percentage of SMEs
Sales	Sales of new-to-market and new-to-firm product innovations in SMEs as percentage of turnover
SME_collab	Innovative SMEs collaborating with others as percentage of SMEs
SME_in_house	SMEs innovating in-house as percentage of SMEs
tert_edu	Percentage population aged 25-34 having completed tertiary education
trademark	Trademark applications per billion regional GDP
Tert_edu	Percentage population aged 25-34 having completed tertiary education

Based on the RIS evidence, EU regional and national governments are expected to (re)design their innovation (or sectorial) policies, to improve their positioning on the ranking across time. The intuition is that progress in one or more of the indicators constituting the RIS (see Table 1) should be positively associated with the region's (or country's) performance. For instance, the most innovative region in 2021 was Stockholm, followed by Etelä-Suomi (Finland) and Oberbayern (Germany).

Due to its importance for policymaking in Europe, the RIS is subject to recurrent additions and removals of its components, as well as frequent assessments of the extent to which it constitutes a meaningful measure of a territory's innovation performance. On the one hand, if we compare the 2021 with the 2019 edition of the RIS, 4 new indicators have been included: Individuals who have above basic overall digital skills (*dig_skills*), Innovation expenditures per person employed in innovative SMEs (*inn_exp_pp*), Employed ICT specialists as a percentage of total employment (*IT*), and Air emissions of fine particulates (PM2.5) in Industry (*air*) [1]. On the other hand, authors like [2] and [3] have raised significant objections to the methodological approach used to create the RIS dataset and rankings, as summarised here below.

According to [2], the RIS (like its country-level-only companion, the EIS – European Innovation Scoreboard) does not constitute a meaningful measure of a territory's innovation performance. It misses the fundamental relation between inputs (notably financial resources) and outputs of innovation activities. In other words, it embeds a systematic bias towards those regions and countries that are currently "overspending" in innovation support, compared with the EU average, while it should reward the efficiency of the underlying processes, or a territory's capacity of being quite successful even with a limited amount of resources.

In turn, [3] criticize the fact that the RIS comes up as a linear (that is, an unweighted) average of indicators, which implicitly puts them all on the same level of importance. This may not be realistic for diverse territories, where the dynamics of innovation may show slight or even profound variations in their respective evolutionary patterns.

More generally, the underlying (to both RIS and EIS) model of innovation dynamics has been criticized for being too simplistic or dictated by what data is available for a majority of EU regions and countries rather than its effective degree of realism or explanatory power.

For instance, in earlier editions of the RIS such as in 2009 [4], the triad was presented of Innovation Enablers, Firm Activities and Outputs. The Enablers captured the drivers of innovation that are external to a firm (including human resources). Activities included any innovation effort that a firm can make (and for which data exists, such as for patents and R&D spending). Finally, Outputs encompassed the benefits of innovation for society, such as in terms of employment, sales and (in the reverse, reduction of) air emissions.

This triad is no longer present in the latest RIS edition for 2021 [5] which uses a slightly more complex taxonomy, based on four elements: Framework conditions (such as levels of education, science and digital skills), Investments (including both spend and ICT employment), Innovation activities and Impacts. However, a comparison between EIS and RIS in terms of data sources – such as that done in [5] – highlights all the limitations of the taxonomy's explanatory power as a result of the huge diminution of available time series when moving from the national to the regional level. It sounds therefore rather questionable from a purely statistical point of view the statement made in the RIS report, that "The most innovative regions are typically in the most innovative countries" [5].

This paper contributes to the debate by proposing an alternative approach to analysing the innovation dynamics of a region. The approach is based on the use of Artificial Intelligence algorithms to "discover" the causal relations among the RIS variables for which indicators exist, in a fundamentally atheoretical way. By atheoretical, we mean that there is no prior commitment to any specific model of the economy that could predict the direction of influence between two or more RIS indicators.

Instead, the proposed approach situates itself in the domain of the so-called Causal Discovery, an advanced data analytics discipline that has been successfully applied to medicine, genetics, and ecology, but is still in its early stages in terms of economic applications [6].

The essence of Causal Discovery is to start from data collections (such as the RIS database, attributing a value to each normalized indicator in each EU country and region for a particular year) and to implement Machine Learning techniques across a sufficiently high number of iterations, leading to visualize the most likely causal relations between variables as edges of Directed Acyclic Graphs (DAGs), connecting some nodes with one-way (causal) relations.

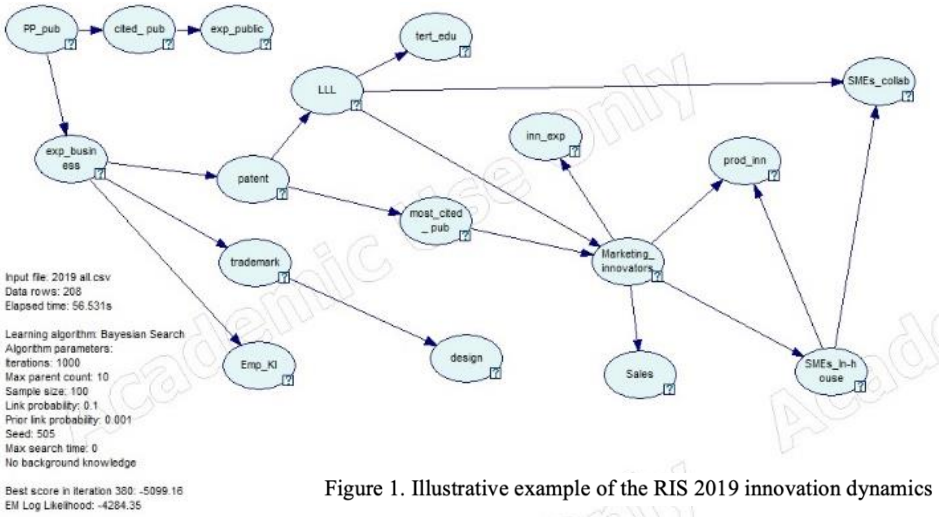


Figure 1. Illustrative example of the RIS 2019 innovation dynamics

An illustrative example of using such technique is visualised in Figure 1, which starts from data of an earlier edition of the RIS, referred to the year 2019.

As one can notice, the DAG starts on the top left with a node without parents and two children. Public-private co-publications per million population (*PP_pub*) as a measure of connectedness between business and academia is shown to directly influence R&D expenditures in the private sector (*exp_business*) and R&D expenditures in the public sector (*exp_public*) only indirectly, i.e. through the intermediate node International scientific citations (*cited_pub*).

This influence looks rather counterintuitive according to both theory and evidence, as publications are normally seen as outputs, instead of enhancers of R&D expenditure. In the case of public administration, scientific publications seem to matter to extent they are cited internationally. A possible explanation can be the importance of reputational factors, deriving from past history of collaborations with the private sector, for academia (as well as enterprises) to receive additional funding for new R&D projects.

Moving one step forward, one can notice that R&D expenditures in the business sector (*exp_business*) have an influence on employment in knowledge-intensive activities (*empl_KI*), on patent applications (*patent*) and on design applications (*design*), although not directly but through trademark applications (*trademark*). Again, this influence is not immediately evident from the experience, unless one accepts that most design innovations in search of protection are in fact those related to a mere aesthetic improvement of some trademark aspects.

Further inspection of the DAG shows an influence of patent applications to lifelong learning (*LLL*), which in turn has influence on the percentage of population with tertiary education (*tert_edu*), the percentage of SMEs collaborating with others (*SME_collab*) and the percentage of SMEs introducing some marketing or organisational innovations (*Marketing innovators*). However, the two latter relations are not exclusive: *SME_collab* are also influenced by the SMEs innovating in-house (*SME_in_house*), which may sound quite reasonable, while *Marketing innovators* are also influenced by the top-10% cited scientific publications (*most_cited_pub*) which has no easy-to-interpret meaning.

Finally, *Marketing innovators* are shown in the DAG to have influence on non-R&D based innovation expenditures (*inn_exp*), on SMEs introducing product innovations (*prod_inn*), on Sales of new-to-market and new-to-firm product innovations in SMEs (*Sales*), as well as on the percentage of SMEs innovating in-house (*SME_in_house*). The latter also has some influence on *prod_inn* together with *Marketing innovators*.

Again, at least part of the above evidence is puzzling. As an example, the direct connection between *LLL* and the propensity of SMEs to collaborate with others or introduce marketing or organisational innovations may sound quite reasonable, but the same cannot be said for the influence of lifelong learning – which pertains to the sphere of vocational training – on an increased percentage of people with tertiary education in a region or country. Same goes for the direct influence of *most_cited_pub* on SMEs acting as *Marketing innovators*, which as we have outlined above, cannot be explained by any promptly available theory.

Without anticipating the conclusions of the paper, we suspect that either a few important indicators are missing from the RIS evidence base, which could be found to act as co-influencers of the dynamics of some variables, or all the heterogeneities that exist within the EU innovation system are too broad to be captured in a single, bird's eye view cutting across the wide diversities of 240 regions and 26 countries of Europe.

The remainder of this paper is structured as follows: in section II we describe in more detail the proposed approach to explore the innovation dynamics of a territory in an atheoretical fashion. In the following three sections we first introduce Bayesian networks as a special category of DAGs and then use them to build an interpretative model of the RIS 2021 database. The model is used for visualisation but also simulation purposes, albeit with some limitations due to unavailable data and lack of computing power. In the last section of the paper, we draw some preliminary conclusions and outline directions for future work.

II. PROPOSED APPROACH

We adopt the Bayesian networks as tools for analysing the dynamics of EU regional innovation sytems. Bayesian networks are a type of probabilistic graphical model used to represent visually and help reflect about uncertain events [6, 7, 8]. They consist of a Directed Acyclic Graph (DAG) that encodes the conditional dependencies between a set of random variables, each associated with a probability table specifying the probability for that variable to have a certain value given the values of the parent nodes in the graph.

The key feature of Bayesian networks is their ability to handle probabilistic inference under uncertainty conditions in a principled and transparent way. They particularly allow for representation and simulation of complex relationships between variables and provide a framework for integrating prior knowledge and data-driven evidence [9]. Bayesian networks are widely used for causal inference and decision-making, as they lead to the identification of the most influential variables in a system and to the calculation and optimisation of alternative decision outcomes despite the uncertainty of available information [10].

Bayesian networks are also used in various thematic fields for modelling and reasoning about complex systems that involve uncertainty and incomplete information. They can be helpful in predicting the likelihood of future events, diagnosing diseases, and making decisions supported by stochastic models. These models are appealing because of their capability to explain intricate processes and provide a well-structured approach for acquiring (solid) knowledge from (noisy) observations [11].

When examining a particular dataset, the objective of Bayesian model selection is to identify the most probable set of connections (causal relations) among the variables that compose it. Unfortunately, if the amount of available data is limited, there could be numerous different models that possess substantial posterior probabilities. Hence, we must use machine learning algorithms that approximate the final model by a repeated number of iterations. While these approaches are computationally compliant, they lack a firm guarantee of the resulting graph's quality. In fact, the graph space is highly "non-convex", and algorithms may become stuck at suboptimal regions.

Additionally, it is known that the exact computation of conditional probabilities in Bayesian belief networks is NP-hard [11], which means that it belongs to a complexity class that is at least as hard as the hardest decision problems that a non-deterministic Turing machine can solve. Thus, an all-too-conservative approach to the visual representation of the 21 indicators now used in the RIS would be very time and resource consuming. However, a variety of algorithms and techniques can be used to simplify the task. These include maximum likelihood estimation, Bayesian search and estimation, and various other methods of machine learning.

The authors have used a Bayesian search algorithm and found out that the results are dependent on the number of iterations and the scoring metric used. The Bayesian search algorithm used for structure learning is Hill Climb with random restarts [12]. The algorithm works by starting with an initial network structure and iteratively making small changes to the structure to improve its score.

One of the key parameters for success is the number of iterations, which determines how many times the algorithm will be run. This number can have a significant impact on the resulting Bayesian model. If the number of iterations is too low, the algorithm may not have enough time to explore all possible network structures and may get stuck in a suboptimal structure. On the other hand, if the number of iterations is too high, the algorithm may overfit the data and create a complex network structure that is not generalizable to new data.

It is also worth noting that the number of iterations is just one of many factors that can affect the quality of the resulting Bayesian model. Other factors include the initial network structure, the quality of the data, and the scoring metric used to evaluate the network.

III. MODEL BUILDING

As mentioned in the Introduction, the two latest datasets RIS 2019 and RIS 2021 do not match each other, because there has been a significant change in indicators and even in the number of EU regions. For instance, compared to the RIS 2019, regional coverage has changed for Croatia from two to four regions, following a revision from the 2016 to the 2021 edition of the NUTS classification. Additionally, the 2019 indicators named "SMEs innovating in-house" and "Marketing or organizational innovators" are no longer in RIS 2021, which introduced four additional ones: Digital skills, IT specialists, Innovation expenditures per person employed, and Air emissions by fine particulates. As the RIS 2019 and RIS 2021 are not fully compatible, we cannot integrate them to get a larger dataset, leaving us with only 240 data rows for 21 indicators. The dataset from RIS 2021 is presented in Figure 2.

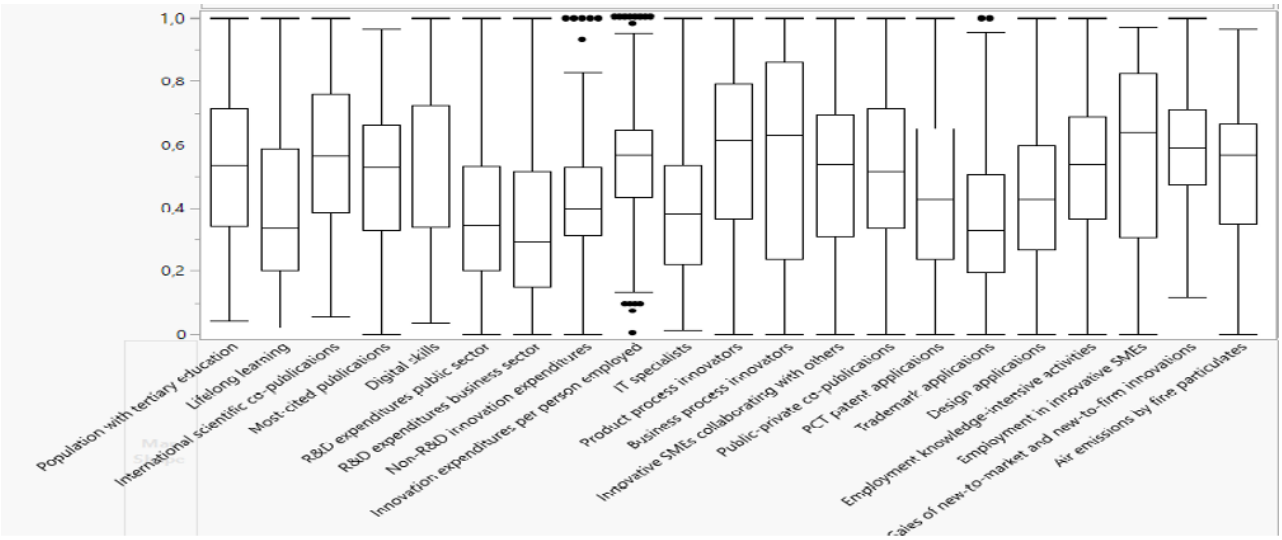


Figure 1. RIS 2021 dataset

To allow Bayesian network modelling, every indicator has been split into 5 equal groups, or quintiles [13], each representing 20% of the given indicator's range. Spreading indicator rankings to quintiles prevents the data from being too thin to be used. This method is normally practiced in economic policy related calculations [6].

For the RIS 2021, we have used the Bayesian Search algorithm with 1000 iterations, max parent count of 10 and 0,001 prior link probability. The resulting network is shown in Figure 3. It was built with the help of the GeNIe modeler [14].

Interestingly, starting point of the DAG is now R&D expenditure in the business sector (*exp_business*), which has influence on patent applications (*patent*), and through them on design applications (*design*), public-private publications (*PP_pub*) and digital skills (*dig_skills*). The latter in turn influence lifelong learning (*LLL*), non-R&D innovations (*inn_exp*), Employment in innovative SMEs (*emp_SMEs*), SMEs introducing business process innovations (*bp_inn*), Air emissions by fine particulate matter (*air*), and top-10% most cited publications (*most_cited_pub*). The latter variable is also influenced by International scientific co-publications (*cited_pub*) and in turn influences Air pollution, which is shown to have some influence on Employment in innovative SMEs. And the description of causal relations between parent and children nodes could continue further, moving along the DAG from left to right.

Globally, the results of this exercise of model building as represented in Figure 3 are rather different from those in Figure 1. This difference cannot be easily explained with

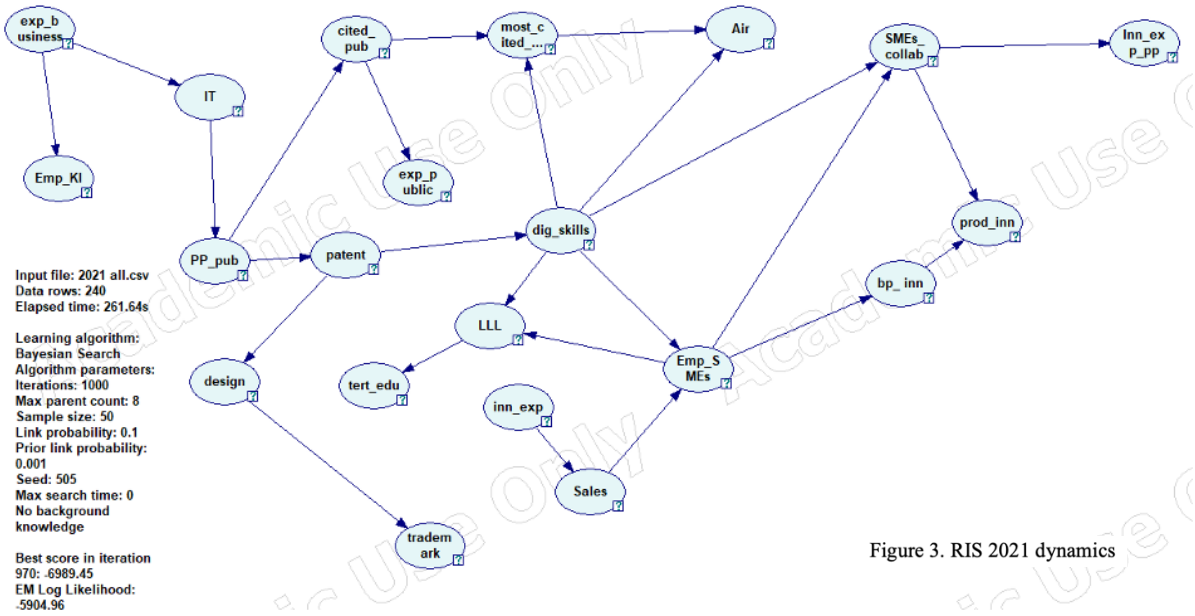


Figure 3. RIS 2021 dynamics

progress of time or with the inclusion/exclusion of some variables in/from the database.

IV. MODEL SIMULATION

When using Bayesian networks for policy definition, the impacts of changes in a node value on another node in the same network are of great interest.

In fact, one of the key reasons for building a Bayesian network is to fit the conditional probabilities of its nodes. This involves estimating the probability distribution of each node given the values of its parents. The process in question involves collecting data and using it to estimate the probability distributions for each node. Estimating the conditional probabilities can be a complex task, especially for networks with a large number of nodes.

Once the conditional probabilities have been estimated, a Bayesian network can be used to make predictions about the relationships between variables. This can be done by propagating the probabilities throughout the network. As it is shown in Figure 4, the RIS2021 Bayesian network model has defined conditional probabilities and predictions that can be calculated as affecting one indicator and then having influence on others.

To exemplify the approach, we have chosen four RIS variables that can be influenced by policy measures. These are Public R&D expenditure, Expenditure on R&D by

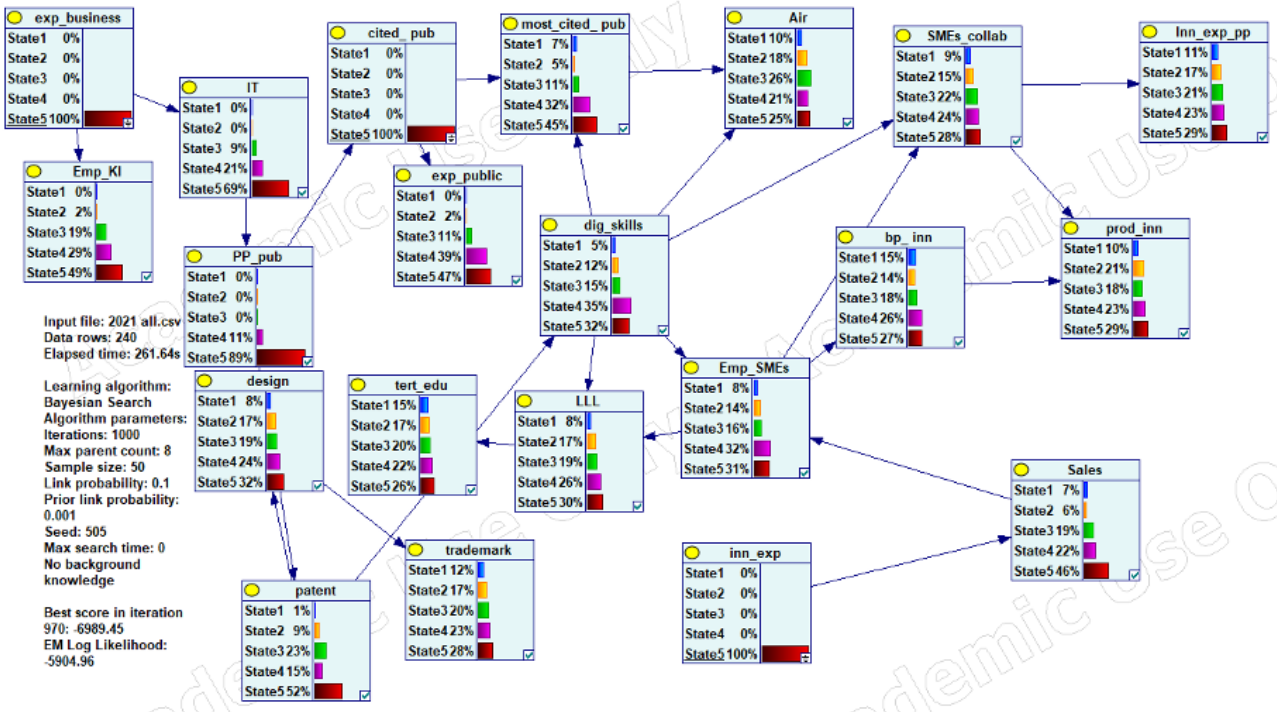


Figure 4 RIS2021 propagating probabilities calculation

business entities, Digital skills and Lifelong learning. Then we simulated the consequences of changing the probability of one indicator and calculated the impact of the change on the other indicators. Some results are displayed in Figure 5.

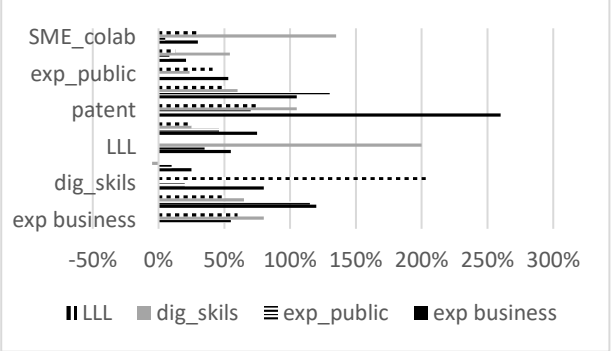


Figure 2. Impacts on indicators when changing another indicator

If one increases public R&D expenditure significantly (e.g. to match the largest expenditure level in the EU), this would increase business R&D expenditure by 55%, patents and publications by 115% and 130%, employment in SMEs by 46%, lifelong learning by 35%, digital skills by 20%, innovation expenditure in SMEs by 8%, employment in knowledge-intensive activities by 10% and collaboration between SMEs by only 5%.

Using policy to boost learning of digital skills would increase public R&D expenditure by 24% and expenditure on R&D in the business sector by 80%. Employment in SMEs would increase by 25%, patent applications by 105%, lifelong learning by 200%, innovation expenditures would grow by 54% and SME collaboration would increase by 135%, while the employment in knowledge-intensive activities would decrease by 5%.

V. Discussion

Bayesian networks seem to be a promising instrument for modelling complex relationships between variables that describe innovation systems. The intuition is to think of Bayesian networks as causal graphs, where every arc (edge) represents a direct causal influence between two connected variables. This view may be too informal and not always mathematically correct, but is widely used by practitioners and can provide a valuable framework for understanding complex systems.

One of the key advantages of using Bayesian networks is that they can help to identify causal factors for certain stochastic variables. A directed arc from X to Y captures the knowledge that X is a causal factor for Y, providing a valuable tool for identifying the most important variables in a system. Additionally, the lack of arcs between pairs of variables can be used to express simple facts about the absence of causal influences between them. This allows for the construction of a data driven model that visualizes all the main causal relationships between variables.

Implementing the proposed approach to the RIS 2019 and RIS 2021 datasets has delivered inconsistent results, both across time and compared with available economic theories and practical experiences. Such inconsistencies are irredeemable, from a purely technical perspective, unless we would consider the possibility of replicating the whole structure learning exercise with the required computing power.

Alternative explanations for the puzzling direction of several causal relations between variables can be manifold: on the one hand, the limited number of observations (only 240 regions compared with 21 indicators) impedes the most obvious corrective action one might consider, that of using RIS data for smaller sized sets of regions – for instance, those belonging to the same geographical areas of Europe or characterised by the same positioning in the ranking (e.g. what the RIS methodology defines “moderate innovators”).

On the other hand, the continuous changes in the data series and metadata definitions do not help consolidate a database with consistent indicators across time, which may be considered as an alternative approach to increasing data size to the required extent. Same goes for moving from a biannual to a yearly publication of RIS statistics, especially for the regional level, which would enable considering time series instead of punctual datasets, and introducing also the time lag dimension, which might be reasonable in a number of cases.

Having more observations and more quality data would be essential to clarify if the heterogeneities that exist within the EU innovation system are too broad to be captured by a single, bird’s eye view cutting across the wide diversities of involved regions and countries.

VI. Conclusion

This paper has proposed an alternative approach to the RIS for analysing the innovation dynamics of a region. The approach is based on the experimental use of Bayesian networks and Machine Learning algorithms to “discover” and visualise causal relations between variables for which RIS indicators exist, without a prior definition of the nature of those variables – such as Innovation Enablers, Firm Activities and Outputs – and therefore in an atheoretical fashion.

To the extent that the proposed approach can be deemed reliable, despite some limitations in computing power and the number of observations, to develop a data driven model predicting the direction of influence between two or more variables, the results of its implementation with both the RIS2019 and RIS2021 datasets has led to counterintuitive results according to both economic theory and experience (or common sense). This raises additional concerns to those expressed in previous literature on the limitations of the RIS tool to assess the dynamics and performance drivers of the EU national/regional innovation systems.

Future work by the authors will include testing again the proposed approach on fewer RIS indicators clustered by typology (for instance, making reference to the triad and quadruplet of elements in earlier and later versions of the methodology). The mapping of results obtained from such exercise would be helpful to shed light on whether a few important datasets are missing, which could be found to act as co-influencers of some variables, or if the direction of some causal relations as (unconvincingly) identified in this preliminary study should instead be reversed.

REFERENCES

[1] European Commission, “Regional Innovation Scoreboard 2021 Methodology Report”. Document date: 14/06/2021. Retrieved online at: https://ec.europa.eu/docsroom/documents/45972

[2] C. Edquist, J. M. Zabala-Iturriagagoitia, J. Barbero, and J. L. Zofío, “On the meaning of innovation performance: Is the synthetic indicator of the Innovation Union Scoreboard flawed?”, Research Evaluation, 27(3), 2018, 196–211.

[3] E.G. Carayannis, Y. Goletsis, and E. Grigoroudis, “Composite innovation metrics: MCDA and the Quadruple Innovation Helix framework”, Technological Forecasting and Social Change 131, 2018, 4-17.

[4] European Commission, “Regional Innovation Scoreboard 2009”. Document date: December 2009. Retrieved online at: https://op.europa.eu/en/publication-detail/-/publication/438811f6-bc27-4049-9872-ad76db87f01e/language-en

[5] European Commission, “Regional Innovation Scoreboard 2021”. Document date: 21 June 2021. Retrieved online at: https://op.europa.eu/en/publication-detail/-/publication/b76f4287-0b94-11ec-adb1-01aa75ed71a1/language-en/format-PDF/source-242412276

[6] A. Darwiche, Modeling and reasoning with Bayesian networks, Cambridge: Cambridge University Press, 2009.

[7] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, “Learning Bayesian networks from data: An information-theory based approach”, Artificial Intelligence 137, 2002, 1-2, 43-90.

[8] N. Friedman and D. Koller, “Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian networks”, Machine Learning 50, 2003, 1-2, 95-125.

[9] J.F. Carriger, M.G. Barron and M.C. Newman, “Bayesian networks improve causal environmental assessments for evidence-based policy”, Environmental Science and Technology 2016, 50, 2.

[10] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, “Using Bayesian networks to analyze expression data”, Journal of Computational Biology 7, 2000, 3/4, 601-620.

[11] P. Dagum and M. Luby, “Approximating probabilistic inference in Bayesian belief networks is NP-hard”, Artificial Intelligence, 60, 1993, 1, 141-153.

[12] D. Heckerman and R. Shachter, “Decision-theoretic foundations for causal reasoning”, Journal of Artificial Intelligence Research, 3, 1995, 405-430.

[13] World Bank, “What are quintiles?” Retrieved online at: https://datahelpdesk.worldbank.org/knowledgebase/articles/1986160-what-are-quintiles, accessed 2.2 2023.

[14] GeNIe Modeler by BayesFusion LLC (academic license). Retrieved online at: www.bayesfusion.com/genie

Methodological Framework for the Comparative Evaluation of the Emission Footprint of Virtual Remote Work and Travel by Autonomous Electric Vehicle to the Workplace

M. Berković, A. Omerhodžić, A. Džananović
Faculty of Traffic and Communications, Sarajevo, Bosnia and Herzegovina
mirza.berkovic@fsk.unsa.ba
adnan.omerhodzic@fsk.unsa.ba
ajdin.dzananovic@fsk.unsa.ba

***Abstract* - The purpose and idea of the paper is to define a methodological framework for the comparative assessment of the carbon footprint of virtual remote work and the footprint of an autonomous electric vehicle for physical mobility to the workplace. The methodology is based on the remote work service, as a typical representative of information and communication solutions with potentially significant opportunities to reduce emissions in the area of physical mobility. On the other hand, autonomous electric vehicles cause less greenhouse gas emissions than diesel cars, even when powered by engines with lower carbon emissions, but we still don't know if it is more environmentally friendly to use digital teleworking services instead of electric autonomous vehicles for trips to the workplace. In the proposed methodology, special attention will be focused on the analysis of emission variables for existing consumption technologies of autonomous vehicles. The originality and value of the work consists in the fact that the results of the work offer an original comparative procedure for determining the value of emission footprint of the physical mobility of an autonomous electric vehicle in relation to the footprint of the virtual mobility of telecommuting.**

Key words - remote work; autonomous electric vehicles; emissions

I. INTRODUCTION

Virtual remote work (telecommuting) is a typical representative of information and communication services, which we define as a model of performing regular work from home, instead of traveling by car to work. In the literature, the terms tele-work or telepresence are also used for this kind of work. Telecommuting as a function of reducing the number of trips has been around for a long time, but the growth of this type of service has been relatively slow since then, except during the pandemic. The indicators of the reduction of traffic emissions from this period show the potential of the service in solving the problem of reducing direct traffic carbon emissions. Different strategies towards less intensive traffic are requiring consideration of potential for carbon reduction by introducing the measure of telecommuting.

On the other hand, vehicles with zero emissions during travel are rapidly developing with consequences for energy consumption and greenhouse gas emissions. Zero-emission vehicles (ZEVs) including hybrid electric vehicles (HEVs), plug-in hybrid electric vehicles (PHEVs), and battery electric vehicles (BEVs) are rapidly developing with varying effects on energy use and emissions. Electric vehicles (EVs) are widely regarded as a promising solution to reducing greenhouse gas emissions and key to the future of low-carbon mobility. It should be noted that BEV has the least amount emissions compared to internal combustion engine (ICE), HEV and PHEV. However, in the paper we focus on autonomous electric vehicles (AEV) 3,4, and 5 levels of autonomy, which have significantly higher efficiency and lower emissions compared to other levels. The vehicles themselves represent a new generation of vehicles that can move independently on roads, that is, with or without any human intervention. Four main drivers that include automated driving, electric powertrains, connectivity and shared mobility can provide a compelling transition to a low-carbon future. The vehicles function using a high-tech system that, among other things, consists of cameras, radar sensors and laser beams, with which the car recognizes the environment in which it is located, as well as other traffic participants. The vehicle continuously collects new amounts of information it receives from objective reality and reacts accordingly. By applying sensor and GPS technologies, vehicles follow road maps, including events, in order to safely reach the desired destination, and certain, more advanced systems have the ability to independently update and optimize the movement path with the help of sensor information. The main advantage of AEVs over conventional systems is the zero emission of harmful gases from the propulsion system during the trip. Also, on the electric vehicle there is no need to change the engine oil, which represents a great potential danger for environmental pollution.

The subject of research in the paper are emissions that can still appear in the process of electricity production in the case of non-carbonized power sources. A specific cause of the negative impact of these vehicles on the environment is the computer system that controls these vehicles and which indirectly generate large amounts of greenhouse gases, i.e. the autonomous driving mode consumes large amounts of energy for processing the autonomy and reliability information of the system. Data centers that house the computing infrastructure used to run applications are notorious for their large carbon footprint.

The methodological framework for researching the operational emissions of AEV computers requires a detailed analysis of the basic consumption factors (360° sensors, the power of the computer on the vehicle, driving hours of each vehicle and the amount of electricity emission that drives all equipment). In addition to the basic energy consumption of electric vehicles, the consumption manifests itself differently depending on additional factors: speed and driving style, topography, road geometry, weather and ambient conditions in the environment, and the electrical load of the vehicle. Achieving long-term energy efficiency of these vehicles means the application of more specialized and energy-efficient hardware designed to run specific driving algorithms as well as the possibility of simple application of new algorithms.

The purpose of this work, in the time of still incompletely decarbonized electricity production, is first of all to look at the indirect emissions of AEVs that are started and run using the energy stored in the battery. On the other hand, authors of the paper are fully aware that an additional trend in the traffic system is towards connected autonomous vehicles (CAEVs), which will play a key role in the new revolution in sustainable mobility with low emissions. Such vehicles can have great potential to operate with greater efficiency, only if they are charged from renewable energy sources that will significantly reduce emissions and dependence on fossil fuel sources. The fact is that such vehicles are complex automotive systems that combine the characteristics of connected vehicles (CV - a vehicle with technology that enables it to communicate with nearby vehicles, infrastructure and objects, but are not automated or electric), autonomous vehicles (AV - a vehicle that, in the broadest sense, capable of driving itself without human intervention) and electric vehicles (EVs).

II. STRUCTURE OF COMPONENTS OF AUTONOMOUS ELECTRIC VEHICLES

The core consumption components of autonomous electric vehicles are: electric motor, battery maintenance systems, various sensors and motor controllers with advanced communication and sensing capabilities and enhanced travel convenience. The system of an electric vehicle with an independent power source is based on the battery, which represents the only source of driving energy, and is charged via a connection from the power grid. E.g. the direct current from the battery through the amplifier and converter is increased and converted into alternating current and, if necessary, sent to electric motors that convert electrical energy into mechanical energy, which is then transferred to the wheels via transmission [1]. The controller manages the operation of the electric motor and the vehicle system through hardware components and software support and monitors the condition of the vehicle, coordinates work and reacts to changes in external driving conditions by issuing appropriate orders to the power unit and other system components. Other parts of electric vehicles are: analog-to-digital signal converter that provides the desired speed information, circuit breaker, fuse or switch, direct current (DC) voltage converter for driving the vehicle's built-in consumers (lights, direction indicators, wipers, sound signal, radio device, etc.), measuring instruments for vehicle management (battery remaining capacity indicator, voltage, current, power, speed), electric battery charger [2].

The energy management system in an electric vehicle aims to increase efficiency and reduce consumption so as to maximize the use of the battery and ensure efficient operation. These include the following control functions:

- *Power supply control:* the energy management system manages the supply of electricity from the battery according to the necessary electrical systems in the vehicle (engine, lights, air conditioning, etc.);
- *Engine speed regulation:* the energy management system controls engine speed to reduce energy consumption and increase energy efficiency;
- *Energy consumption optimization:* the energy management system applies various algorithms to optimize energy consumption, including energy recovery during braking and vehicle speed optimization;
- *Battery management:* the power management system also manages the battery, including its charging and status monitoring;
- *Integration with other systems:* the energy management system is often integrated with other electrical systems in the vehicle, such as the driving control system and air conditioning, to ensure the most efficient use of energy.

Modern user interface concepts of AEVs do not have standard elements such as steering wheel, pedals and gearbox, they are replaced by a large screen. E.g. the architecture of the control system of an autonomous vehicle of the fifth degree of autonomy according to the Society of Automotive Engineers (SAE) classification is divided into four parts consisting of sensor, client, action and user systems [3].

Sensor systems consist of several different sensors that have the task of collecting data from the vehicle's environment in real time. Short range sensors are ultrasonic sensors, capacitive sensors or infrared sensors while long range sensors are LIDAR, RADAR, computer vision and advanced GPS.

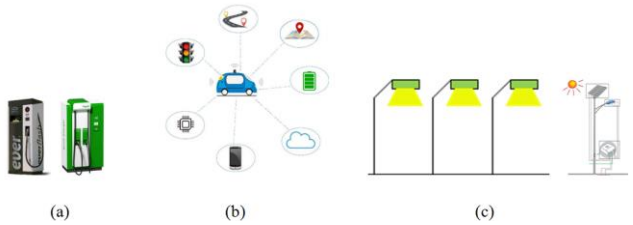
The client system has the task of processing a large amount of collected data and extracting the most important information in order to interpret the objective reality, determine the position of the vehicle in space, and thus decide on the next steps in driving.

Action systems represent the mechanical parts of AEV (e.g. steering, braking and drive systems) that actually execute the commands received from the client system and direct the autonomous electric vehicle.

User systems are a combination of hardware and software that allow the user of AEV to communicate with it in real time. The elements of the user system consist of pointers that provide information to the user about the optimal route of the vehicle or the actions he intends to perform.

III. ENVIRONMENTAL EFFECTS OF APPLICATION OF AUTONOMOUS ELECTRIC VEHICLES

The impacts of the AEVs can mainly be classified as indirect impacts that are a consequence of the current consumption of electricity, where the reasons and consequent connections of these impacts are mostly clear and well understood (Figure 1.).



(a) (b) (c)

Figure 1. (a) Renewable and non-renewable energy sources, (b) consumers of AEVs and (c) secondary consumers

These impacts stem from the goal of using AEVs, i.e. energy efficiency and traffic development based on the energy sector of renewable sources. In this context, it is expected that by 2030, 55% of new vehicles in Europe will be fully electric, and 40% will be hybrids [4]. The introduction of electric vehicles causes the need for global planning of increased energy supply. The energy storage capacity of electric vehicles can be used to equalize demand and mitigate variations.

Other indirect impacts (secondary effects of the production of components for AEVs), which often have more severe consequences for the environment than the first impact, are the production processes and the availability of cobalt, lithium and other materials. In order to reduce this effect and avoid environmental pollution, the recycling of batteries must be encouraged [5].

The third type of impact is cumulative impacts that have additional, multiplying or synergistic effects of the type of construction of additional infrastructure required for AEVs.

The trend is for AEVs to increase the number of kilometers traveled per unit of time, which could contribute to significant environmental improvements given that they are electric and will drive more efficiently. AEVs produce zero emissions when in motion and are less noisy than classic fossil fuel vehicles. Indirect effects occur if the sources for powering the AEVs are non-renewable, e.g. autonomous electric vehicle level 3 does not emit CO_2 directly, but often uses electricity generated with indirect emissions in thermal power plants (Table I.).

TABLE I. EFFECTS OF APPLICATION OF AEVs

Category	Identified effects
Direct influences	None
Primary positive impacts	Reducing the use of diesel and gasoline vehicles.
Secondary positives impacts	Reduction of noise in the room.
	Reduced construction of facilities - pumps.
Primary negative impacts	Increased electricity consumption (from hydro and thermal power plants).
Secondary negative impacts	Expansion of the network of filling stations in the area.
	Increasing demand for batteries and specific materiales.

IV. ANALYSIS OF THE EFFECTS OF VIRTUAL WORK ON THE REDUCTION OF INDIRECT EMISSIONS OF AUTONOMOUS ELECTRIC VEHICLES

The methodology for assessing the effects of remote virtual work on the reduction of emissions of AEVs includes two domains: the first refers to the assessment of emissions that technologies directly cause, and the second to the quantification and valorization of the effects of reducing emissions that arise applying this measure. Both effects must be taken into account to assess the overall utility of the application. The methodological assessment procedure is based on the use of the life cycle approach [6].

In the first step, it is necessary to define the goal of the research, which is the basis for defining the indicator, and in the case of telecommuting, it could be defined as research into the possibility of applying services to reduce emissions in the commuting sector, bearing in mind the increased energy consumption in traffic. The analysis can be performed at the level of office or company, that is, of the employees. The scope of the research involves defining the components of the two systems. An alternative system usually contains a PC, a printer and the necessary infrastructure such as servers and networks. The components of the Baseline Assembly Unit (BAU) system are the components of the AEV defined in chapter II, a private garage and a public parking lot (Table II.).

TABLE II. BAU SYSTEM COMPONENTS

System	Description	System components	
ICT	Virtual remote work (telecommuting)	1. Computer 2. Printer 3. Data center	4. Servers 5. Networks
BAU	Autonomous electric vehicle (relation house - office)	1. Private vehicle 2. Public vehicle	3. Private garage 4. Equipment in the office 5.Public parking

In the next step, it is necessary to look at all positive and negative effects and then select only those that are considered to have a significant impact on the environment. If the effects that are expressed in a relatively short time and at a relatively small scale of implementation (company level) are chosen, the final list of processes that require evaluation is reduced. The reference BAU value is determined, in this case, on the basis of the potentially greatest effect that can be realized [7]. In the case of telecommuting, potential effects are given in Table III.

TABLE III. EFFECTS OF TELECOMMUTING

Category	Identified effects
Direct emission	Emission from telecommuting equipment
Primary positive impacts	Reducing the use of autonomous vehicles.
Secondary positive impacts	Reducing the use of autonomous public transport vehicles.
	Reduced use of public parking areas.
	Reduced construction of additional facilities.
	Reduction of construction of road infrastructure.
Primary negative impacts	Increased energy consumption in the house.
Secondary negative impacts	Expansion of cities in space.
	Increase in trips outside the commuting area.

This means that it is necessary to calculate the level of indirect emissions emitted by the use of a private or public

AEV powered by non-renewable sources per employee (data can be extrapolated for the number of employees in a company), in the house-office/factory relationship (Figure 2.).

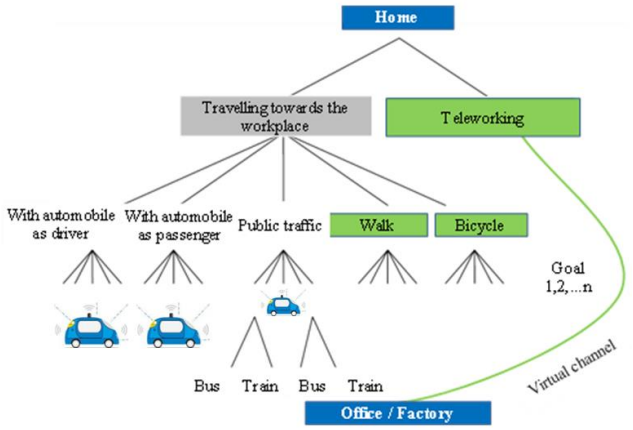


Figure 2. Structure of the physical travel model and the virtual channel of the home-work relationship

Variations in the emissions of the telecommuting components themselves are considered negligible for a shorter period of time, and the operation of some system components such as servers and computer centers are also unaffected because they work independently of where the employees are currently located. In addition, it is assumed that the employees already have the necessary equipment at home (Figure 3.), so the acquisition of new equipment for the needs of virtual work is partially unaffected by the emissions of component production.

The following effects were identified as significant: reduction in the number of trips by AEV, increase in household energy consumption and reduction in the use of equipment and energy in offices and lighting in closed parking lots. All of them represent processes that should be mathematically quantified in the use phase.

The next step is to determine the mean values of the distance from home to work, then determine the percentage of employees who use AVEs and those who use public electric transportation, and use the existing emission factors for the final calculation.

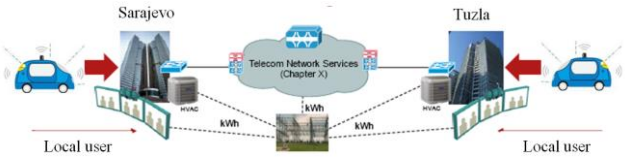


Figure 3. Structure of telecommuting system equipments

V. MODEL OF THE INDIRECT FOOTPRINT OF JOURNEY TO WORK BY AUTONOMOUS ELECTRIC VEHICLE

The emissions footprint parameter of an AEV commute is a combination of a set of characteristics sufficient to determine the behavior of emission values. In the practice of classic vehicles, there are four variables that have an impact on the emission value of the trip: A (total transport activity in the vehicle in km), S (participation of the mode of transport in the road), I (energy intensity) and F (CO_2 content in fuels). In our case, these variables need to be adapted to the AEV technology so that we have: P (drive energy consumption per 100 km traveled, i.e. consumption of AEV activities), D (range of certain AEV models with one charge in km - a higher number of charges causes losses energy), E (CO_2 emissions/km from the power source, with a note that AEVs do not emit CO_2 directly, but use electricity generated with emissions, e.g. in thermal power plants) and S (participation of secondary emissions, such as parking lighting). To determine the savings in CO_2 emissions by replacing travel with virtual work, not only the estimated CO_2 emissions of the vehicle are significant, but also the savings in secondary emissions (e.g. parking lot lighting, congestion in combined traffic, etc.).

To estimate the consumption of AEVs, it is necessary to take an average. With current technology, consumption ranges from 10 to 30 kWh/100 km, while the average consumption is around 17 kWh/100 km, depending on traffic conditions. If we multiply this value by the emission of the thermal power plant as a source of non-renewable energy, we get 9.978,515 gCO_2 for every 100 km. We divide the value by 100 and get a result of 99.78 gCO_2/km. This emission value would be relatively accurate if all the electricity needed for an autonomous vehicle was produced from a thermal power plant, and the emission would represent CO_2 emissions due to the production of electricity from non-renewable sources. If we take the second case that all the required energy is produced from renewable energy sources that generate emissions of up to 8 $gC0_2$/kWh, we would get the value that for every 100 km an AEV has 140 gCO_2. If we divide that value by 100, we get a result of 1.4 gCO_2/km, which is approximately 1.4% of the emissions in relation to the non-renewable source of electricity produced from the thermal power plant.

VI. RESEARCH OF THE MATHEMATICAL MODEL OF MEASURING EMISSION SAVINGS OF JOURNEY TO WORK USING THE SUBSTITUTION MEASURE

For the calculation of emission savings, we propose the PDS (meaning is given below) procedures, as a method for quantifying the carbon emissions of AEVs. The procedure is adapted for the calculation of CO_2 emissions in passenger transport (private AEV, public AEV transport). The total CO_2 emissions of commuting are shown as the product of three factors: energy power consumption (P), travel distance (D) plus the factor of traffic congestion and secondary consumption (S). The elements are of essential importance for assessment and require the definition of the basic labels of their elements:

- E - total CO_2 emissions of all AEVs used by the user on the way from home to work;
- E_i - are the individual CO_2 emissions of the sight (private, public) of the AEV and used by the user;
- k_j - is the type of electrical power source (non-renewable, renewable source) in the form of i (public, private);

- A_i - total vehicle activity in *km* by type *i* (eg number of fillings AEV during the day);
- λ_j - participation of k_i energy sources in the form of *i* (%);
- *j* - energy source in form *i*;
- M_i - participation of type *i* in some area (%);
- I_i - energy intensity of type *i*; and
- E_{ij} - CO_2 emissions from electricity *j* in the form of transport *i* in gCO_2.

According to the proposed model, the emission savings measured by the substitution of an AEV to work can be calculated as:

$$\Delta CO_2 = [(P_{BAU} * D_{BAU}) + S_{BAU}] - (P_m + S_m) \quad (1)$$

where is:

- P_{BAU} - activity of indirect emissions of an AEVs under the BAU scenario,
- D_{BAU} - workplace distance corrected with the congestion factor,
- S_{BAU} - represents the activity of indirect secondary emissions,
- P_m - activity emissions if the measure is taken (virtual work),
- S_m - indirect emissions of work from home.

CO_2 savings from the impact of reducing the activity of AEV after *t* years are calculated as the difference between the annual base value of CO_2 emissions and the annual value of CO_2 emissions in the event in which the virtual remote work activity was undertaken. For each of the components of the PDS procedure, an individual set of defined factors is required to determine the carbon impact of the measure (e.g. with S, detailed data on the type of lighting in the parking lot is required). In the process of measuring CO_2 reduction, there are always secondary effects that influence the result.

Also, an undesirable effect of the application of teleworking is "induced traffic" due to better road conditions (less congestion), when people, in addition to the telecommuting option, want to use AEVs more, traffic is induced and congestion problems are not solved. If several measures are taken at the same time, estimates of the overall impact on emission reduction become difficult.

VII. PROPOSAL OF A GENERIC PROCEDURE OF COMPARATIVE ANALYSIS OF ENVIRONMENTAL IMPACT

The proposed comparative methodology is applicable to cases where telecommuting is used to avoid travel in the following areas: avoiding business trips by AEV, avoiding employe AEV trips to regular work, avoiding personal transportation and using virtual services. The proposed procedure for calculating telecommuting emissions is presented in several steps.

The general approach is to first calculate the direct emissions associated with the provision of data transmission, voice and video access as well as the secondary energy consumption of the end users of the equipment (telephone, video display, or laptop). End-user space includes an employee's home or remote office. After that, through the process, we compare the obtained (positive effects) values with emissions from indirect emissions of AEVs, and in the function of reducing the employee's physical journey to work.

In the paper, we propose a framework of eleven steps for the comparative evaluation and calculation of emissions, i.e. the justification of the application of the BAU alternative, namely:

Step 1: Create an inventory of existing equipment at the first telecommuting location. Then an inventory of endpoints, remote network ICT equipment and company infrastructure. Due to the application of telecommuting services, the company will increase traffic on the network;

Step 2a: Estimate the total energy consumption (phase of use) and related greenhouse emissions of all the company's ICT infrastructure solution and the company's shared network infrastructure;

Step 2b: Calculate the total electricity consumption (phases of use) and associated CO_2 emissions of the end-point telecommuting solution and remote network telecommuting equipment;

Step 3: Estimate the total invested emissions (non-use phase) from telecommuting solutions for enterprise infrastructure and network infrastructure;

Step 4a: Extraction of the total, phase of use, emissions from telecommuting solutions of the company's infrastructure and network infrastructure for the observed period;

Step 4b: Extracting the total, phase of use, CO_2 emissions from the telecommuting solution endpoints for virtual work for a certain period;

Step 5a: Allocation of built-in emissions for all enterprise ICT infrastructure solutions and enterprise network infrastructure above the number of scheduled / actual meetings for that same time;

Step 5b: Allocation of built-in emissions (non-use phase) of telecommuting endpoint solutions in relation to the number of actual working hours;

Step 6: Adjust emissions of one working day, endpoints with the average number of endpoints in the company. Then add the extracts from steps 3 and 4;

Step 7: Estimate the total emissions burden of the service provider for delivering the expected telecommuting solution bit rate for a period;

Step 8: Total emission for all network infrastructure, allocated to usage phases;

Step 9: Individual journey records of journeys for at least 90% of transport will be used to determine greenhouse gas emissions;

Step 10: Assessing how telecommuting solutions affect the reduction of AEV use for travel;

Step 11: Evaluation of alternative modes of travel instead of AEVs, e.g. by bike or by foot.

The results of the comparative analysis procedure between the application of the autonomous electric vehicle and the telecommuting system formally represent the difference in the environmental impact between the AEVs and the telecommuting system:

$$EI_{difference,i} = EI_{AEV,i} - EI_{telecomuting,i} \quad (2)$$

where is:

- *EI = environmental impact*
- *i = i-th comparative category*
- $EI_{difference,i}$ *= i-th secondary environmental impact*
- $EI_{AEV,i}$ *= i-th EI of the autonomous electric vehicle*
- $EI_{telecommuting,i}$ *= i-th EI of the telecommuting system*

By summing the total $EI_{difference}$, i we get the $EI_{difference}$ or the effect of the telecommuting system covered by AEV trips. The equation below shows the formula for calculating the total effect:

$$Total\ EI_{difference} = \sum EI_{difference,i} \quad (3)$$

A positive result (Total $EI_{difference}$ is positive) indicates a positive impact on the environment and a negative value (Total $EI_{difference}$ is negative) represents a negative impact on the environment. Positive secondary effects indicate a reduction in the environmental burden due to the introduction of telecommuting services. Negative secondary effects indicate the opposite.

On the other hand, the achieved effect (x) is expressed in the reduced value (kt) of CO_2 emissions. In the same units, the CO_2 emission that was reduced due to energy saving in commercial buildings/offices (y) is expressed, as well as the emissions caused by additional energy consumption in households (z). The total effect (x+y-z) gives a quantitative value, that is, an answer to the question of the potential usefulness of using telecommuting instead of autonomous electric vehicles to reduce emissions.

VIII. CONCLUSION

The research results are showing new approach of measuring the indirect emission footprint of travel in the electromobility system. In the paper, we determine the role and potential impact of virtual mobility on the reduction of physical electromobility emissions. The methodological framework for the comparative evaluation of the emission footprint of virtual remote work and travel by electric autonomous vehicle to the workplace includes a combination of characteristics of the two systems sufficient to determine a positive impact on the reduction of carbon emissions of commuting. We conclude that the previous comparative assessment methodologies, derived on the basis of positive effects, partially describe the conditions in the electromobility system, and do not include impacts the range of transport activities themselves. For a more comprehensive and detailed approach to this issue, wider research and the availability of field data are needed.

Based on all of the above, we can conclude that in order for autonomous electric vehicles to reach maximum utilization and affect emission reduction, quality integration into electrical energy systems, building of energy-efficient charging infrastructure, production of sustainable batteries and, most importantly, decarbonization of electricity production is necessary. The autonomous electric vehicle system is a very complex system and small changes within the environmental sustainability of one subsystem can have a large effect on the entire system. In order to preserve the overall positive ecological impact of telecommuting on the reduction of indirect emissions of electromobility commuting, it is necessary to have a policy of preferring decarbonized power sources. The increase in demand for electricity to power such cars will present a long-term challenge for electricity suppliers. For this reason, the positive effects on environmental protection could be canceled out by additional emissions from the individual energy sector, if the additional demand for energy is produced from coal-fired plants.

As in any scientific work, within the framework of time and space limitations, a number of additional questions were opened here, which should be treated below, such as the total losses of AEVs in the traffic flow, energy consumption in peak periods of traffic, problems of induced traffic of electromobility etc.

REFERENCE

[1] I. Alajbeg, „Električni automobili i održiv razvoj," Split 2014.

[2] M. Stojkov, D. Gašparović, D. Pelin, H. Glavaš, K. Hornung, N. Mikulandra, „Električni automobil-povijest razvoja i sastavni dijelovi," 2014.

[3] Liu, S., Tang, J., Zhang, Z., „Computer architectures for autonomous driving," IEEE Comput. Archit. Lett., 2017.

[4] Bissel D., Birtchnell T., Hsu E., „Autonomous automobilities: The social impacts of driverless vehicles," Article in Current Sociology, 2018.

[5] Lee, C., „Grabbing the wheel early: Moving forward on cybersecurity and privacy protections for driverless cars," Federal Communications Law, 2017.

[6] GeSI, „Evaluating the carbon-reducing impacts of ICT - An assessment methodology, Global e-Sustainability Initiative Report, February," 2011, Available: www.gesi.org/ReportsPublications/AssessmentMethodology.aspx

[7] P. Nelson, Safirova, and M. Wells, "Telecommuting and environmental policy, "Lessons from the ecommute program," Transportation Research, D 12, pp. 195-207., 2007.

Internet of Behavior – The Transformation of Customer Relationship Management in Logistics

A. Agatić*, E. Tijan**, S. Aksentijević*** and A. Pucihar****
* University of Rijeka, Faculty of Maritime Studies/Department of Maritime and Transportation Technology, Rijeka, Croatia
** University of Rijeka, Faculty of Maritime Studies/Department of Maritime Logistics and Management, Rijeka, Croatia
*** University of Rijeka, Faculty of Maritime Studies/Department of Maritime Logistics and Management, Rijeka, Croatia
**** University of Maribor, Faculty of Organizational Sciences/Department of Informatics, Kranj, Slovenia
adrijana.agatic@pfri.uniri.hr

***Abstract* - The Internet of Behavior (IoB) is an evolution of the Internet of Things (IoT), which is already widely used in logistics. In addition to the technology and data aspect, IoB also includes behavioral science to track and analyze behavioral patterns. By using IoB, logistics companies can provide not only more reliable but also more flexible services, enabling better insight into customer behavior patterns and changing market conditions in real-time. The IoB can enable logistics companies to streamline Customer Relationship Management (CRM). In this paper, the authors research the potential of IoB to transform CRM in logistics. Therefore, the authors analyze the main features of IoB and current trends in logistics CRM. The focus of this paper is to analyze how the application of IoB could transform logistics CRM (from the aspect of predictive services, automation, CRM processes, CRM software and security strategies).**

Key words – Internet of Behavior; Customer Relationship Management; logistics, transformation

I. INTRODUCTION

Logistics companies are increasingly challenged by both demanding customers and the technology improvements and breakthroughs. It is expected of logistics companies to easily adapt to customers' needs to provide personalized services for individual customers. There is a trend of a hyper-personalization of logistics services and the use of disruptive technologies [1].

Quality of customer experience directly impacts the logistics companies' business. Therefore, logistics companies should continuously develop and adopt Customer Relationship Management (CRM) digital solutions [2]. The CRM solution is usually a system that automates and generates customer data about all customer-related activities [3]. In the logistics sector, CRM solutions assist numerous business functions: quote and orders management, rate synchronization, load management, fleet management, warehousing management, sales and marketing management, risk management, reporting, etc [4]. The global CRM market size was estimated at USD 58.82 billion in 2022 and it is expected to grow at a compounded annual growth rate (CAGR) of 13.9% up to 2030 [5].

The Internet of Behavior (IoB) came into greater focus during the COVID-19 pandemic, when companies had to rethink the means of how they communicate with customers. The IoB provides detailed insight into useful information about customers' behaviors, interests, and preferences from the perspective of behavioral psychology. For example, Google and Facebook were among the first global companies which started to use behavioral data so that they can offer and display relevant advertisements and services to their users [6]. The IoB enabled companies not only to connect with their potential audience, but also to be able to track their behaviors. Additionally, technologies like Alexa, OK Google, and Siri are also programmed to study and interpret the data and user behavior i. e. to use the IoB [6].

The global IoB market was valued at USD 472.91 billion in 2022 and it is expected to reach USD 2,143.57 billion by 2030, growing at a CAGR of 20.79% by 2030 [7]. It is also estimated that 40% of people worldwide will be using at least one IoB application by 2025 [8]. This data points to the growing importance of IoB for CRM and logistics services.

In this paper, the authors will analyze the evolution of the IoB and its workflow in logistics operations. In addition, the authors will analyze the emerging trends that will influence and possibly even determine the future of CRM in logistics. The main goal of this paper is to analyze how the IoB can transform CRM and change the provision of logistics services. Therefore, in the 2nd section, the evolution of IoB is analyzed. In the 3rd section, new trends in logistics CRM are identified. In the 4th section, the main part of the paper, authors analyze how IoB could transform CRM in logistics.

II. EVOLUTION OF THE INTERNET OF BEHAVIOR (IOB)

The IoB encompasses the technological aspect of the Internet of Things (IoT) and aspects of behavioral science and psychology, with the aim of gaining deep insights into customer data, i.e., customer behavior, in order to develop better strategies for CRM [9].

The IoB transforms information about customer behavior into knowledge and autonomous decisions based on this knowledge, as presented in Figure 1 [10]. Therefore, the IoB workflow includes the following aspects [10]:

- behavioral tracking: using IoT devices to track users' behavior,
- data collection: collecting valuable data and preparing it for analysis,
- data analysis: using analytics and Machine Learning (ML) to obtain logical and structured information,
- identifying behavioral patterns: using behavioral science and Artificial intelligence (AI) algorithms to identify human behavioral patterns,
- decision-making and predictability: making decisions, predicting user behavior and influencing user behavior.

The IoB data is collected from various IoT sources: smartphones, tablets, PCs, sensors, wearables and embedded systems. These various data sources increase the amount of data collected, which is transformed into meaningful data for businesses using IoB [11].

Some pronounced shortcomings of IoB involve security and privacy issues. The IoB uses a large database about customers and the company, and there is a risk of cybercriminal or unethical use. While various data are collected about customers (e.g., their behavior, likes, and dislikes, etc.), they could assert their privacy rights [12].

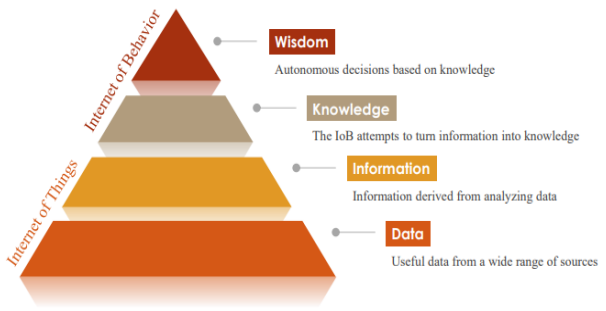


Figure 1. The IoB concept

Despite some shortcomings, the IoB offers more benefits: definition of customer behavior patterns; detailed information about all steps of the customer's journey; notifications and alerts for customers and service providers; predictability of customer behavior and actions; optimization of CRM [13].

III. NEW TRENDS IN LOGISTICS CRM

CRM is the combination of practices, strategies, and technologies that companies use to manage and analyze customer interactions and data [14]. Recently, logistics companies have changed the way they interact with their customers under the influence of technological advances and customer demands. As a result, the quality of customer experience, the accuracy of the information, and transparency are leading to changes in CRM. There are four major challenges that determine the future of CRM in logistics [15], [16]:

- Customer experience: customers are demanding continuous improvement of their experience and personalization, i.e., customized services offered by logistics companies.
- Technological breakthroughs: increased use of technology is inevitable. In the future, CRM will need to integrate various technologies (e.g., voice processing, chatbots, robot process automation, and facial recognition) to develop predictive and real-time solutions for customers.
- Logistics startups: technologies are enabling new business models that focus on one or more aspects of customer service. As a result, startups are introducing new business models based on the use of advanced technologies and they are having an impact on reshaping the existing business models of logistics companies.
- Redefining collaboration: the logistics market consists of different services and different companies providing them, making the service process complex. In order to meet customer requirements, collaboration is expected to increase in the context of digitally networked logistics (data sharing, cooperative freight planning, end-to-end visibility, etc.).

These aspects are evident in the survey conducted by Zendesk (producer of CRM software) on the experiences of 97,000 customers from the transport and logistics sector. The survey results have shown the following [17]:

- 60% of the customers have increased their customer experience expectations in recent years;
- 92% of the customers said they would invest in services from a company that knows the details of their transactions, so they do not have to repeat the process;
- 90% of the customers claim they would spend more on services from companies that offer personalized services;

- 89% of customers prefer companies that offer services online using dedicated customer portals.

The conducted survey shows that customers prefer to work with transport and logistics companies that understand their needs and priorities, i.e., know their behavior patterns. As for the companies, the survey showed that 64% of the companies that responded to their customers' expectations were able to grow their business and retain their customers [17].

IV. IoB AS A TRANSFORMATION FACTOR OF CRM IN LOGISTICS

The previously elaborated changes that are already taking place and will intensify in the future indicate that IoB could be one of the transforming factors of CRM in logistics. Considering that IoB integrates technology, customers, and behavioral science, it has the potential to provide useful inputs to bring about the transformation of CRM in logistics services. The IoB has the potential to become the Enabler of Experience (EoE) in logistics services. The "Experience Enabler Scorecard" identifies the technology and/or strategy strengths and weaknesses to determine if they're able to create digital services and content with an outstanding experience for customers [18]. The IoB could transform CRM in logistics as follows [19], [20]:

- CRM predictive services and automation,
- Redesign of the CRM processes,
- CRM software redevelopment,
- CRM security strategies.

The IoB enables logistics companies to efficiently control customers' behavior and provide 'predictive services' for the customers. The IoB links data from various IoT sources to provide insight into the customer's individual service process – complete ''customer journey''. Logistics companies can then use CRM to create smaller segments of their customer or individuals. This enables logistics companies to offer personalized services, shorter lead times, lower inventory costs, lower shipping costs, fewer delays, etc.

The bottom line is that the highest transformative effect exerted on CRM by IoB is its' automation – a self-adaptive system that provides personalized services with no or minimal human intervention [21]. CRM automation refers to the ability of a CRM system to automate repetitive, manual tasks to streamline work and improve the customer experience [22]. For example, if a customer uses a logistics cargo platform and obtains a quote and books a freight space, the CRM system can use the IoB behavioral pattern to offer a service that fits the customer (type of container, route, price, etc.) without the need for human intervention.

Another simple example is chatbots, which allow companies to automatically respond to questions about their services. There are some positive examples from companies that show IoB could be critical to transform and automate CRM. For example, Transmetrics' AI algorithms recognize patterns that can help predict fluctuations in transportation demand and how customers could react. This enables logistics companies to respond immediately and offer their services more easily for their customers [23].

The data obtained from the IoB can be useful to show necessary changes in the CRM process, i.e., changes in the organization and human resources of logistics companies. Data that provides deep insights into customer behavior and activities can be an indication for logistics companies to redesign and improve the CRM process to fulfill customer expectations. Therefore, IoB data can indicate necessary changes in the CRM process, such as: training of employees on the company's CRM processes; education and training of employees to improve digital skills; possible hiring of behavioral psychology experts to properly analyze customers' activities and behaviors; introduction of new standards and regulations in the company; improving service strategies; reallocation of the budget to invest more in CRM solutions, etc.

It is apparent that the IoB deals with a wide range and quantity of data about customers collected through various sources and technologies. The CRM software should possess appropriate functions to cover all logistics companies' processes. Considering the functions of IoB, it is clear that CRM software should be re-engineered or improved.

Some shortcomings of CRM to fully realize the potential of IoB are the following: low customer satisfaction with user interface and options, lack of alerts and notifications for customers and companies, limited data capacity, limited reporting options, lack of integration with tracking tools, lack of technology solutions and integration with devices, lack of integration with application programming interfaces (APIs), no or weak remote access (e.g., mobile phones), etc. For example, the user interface is the customer's first contact with the CRM. The user interface consists of all the elements that users interact with the following: buttons to activate functions, text to read, graphics, sliders, text input areas, screen layout, transitions, etc. It is responsible for a large part of the user experience [24]. Customers need CRM applications that are rational and logical to navigate and make them feel like they are doing the activities they want to do. The ease or difficulty of interaction with user interface elements determines the user experience in e-logistics services [24].

It is expected that new security protocols and procedures will be introduced to make the use of IoB in CRM more secure. The IoB contains various data about the customers and the logistics company: emails, contact details, names, job titles, addresses, passwords, transaction totals, shipping routes, freight and transport vehicle movements, etc. The introduction of IoB makes data privacy and security more threatened. Customers' experience and behaviour will not be in favour of the logistics companies' digital initiatives if they feel security shortcomings. This will be particularly important in CRM automation. For example, in February 2022, logistics company Expeditors was the victim of a cyberattack that impacted its global operations chain, including customs,

freight, and distribution activities, causing customer shipments to stall and impact the quality of the customer experience [25].

The study, conducted by the Council of Supply Chain Management Professionals, found that 93% of shippers and 98% of 3PL logistics companies view data-driven decisions as crucial for their operations. The study shows that 'predictive service' is no longer an option, but a necessity to improve the customer experience in logistics services. It therefore shows that CRM automation is inevitable and points to the transformative role of IoB in CRM in logistics [23].

V. CONCLUSION

Disruptive technologies are transforming logistics service delivery. The Internet of Behavior, as an extension of the Internet of Things, enables real-time and detailed insight into customer activity and behavior in the delivery of logistics services. Therefore, the CRM of logistics companies is changing, and the IoB is becoming the transformative factor for CRM in logistics as an Enabler of Experience.

The IoB data serves as input for the CRM that logistics companies use to create personalized services. Therefore, IoB transforms the mass perspective of CRM into an individual perspective for each customer. The highest level is CRM automation, meaning that IoB data enable a level of predictability where logistics companies can offer self-created services with minimal human intervention. In addition, IoB-derived data could signal the need to improve CRM processes, i.e., the organizational and human resource aspects of the logistics company, to process large amounts of data and create personalized services for end customers.

Another aspect in which IoB is changing CRM, including logistics CRM, is software re-engineering, meaning the software functions should be adapted (e.g., data capacity, user interface, alerts and notifications) to follow the behavioral patterns of each customer in order to improve customer experience. The security aspect of CRM could change as IoB processes a large amount of customer and logistics company confidential data, and logistics companies should introduce new security features in CRM.

Recent research in the transportation and logistics industry shows the need for CRM transformation as customer experience and data-driven decisions are of paramount importance. Therefore, IoB has the potential to enable this transformation of CRM in logistics.

In terms of future research, there are a few possible directions. First, the aspects of CRM-logistics transformation could be thoroughly analyzed to gain a more detailed insight into the protentional challenges of applying IoB in logistics CRM. In addition, it is possible to investigate whether there are other aspects that should be analyzed. Another direction for future research is to conduct a survey among logistics companies to analyze their capabilities for applying IoB and transforming CRM.

REFERENCES

[1] Tech Mahindra, Hyper-personalization – A New Norm in Customer Service, [Online], Available: https://www.techmahindra.com/en-in/blog/hyper-personalization-a-new-norm-in-customer-service/ [Accessed: 12-Feb-2023]

[2] Cybrosys technology, Importance of CRM in Logistics, [Online], Available: https://www.cybrosys.com/blog/importance-of-crm-in-logistics [Accessed: 12-Feb-2023]

[3] SAP, What is CRM, [Online], Available: https://www.sap.com/insights/what-is-crm.html [Accessed: 12-Feb-2023]

[4] Creatio platform, One platform to automate transportation workflows and CRM with no-code, [Online], Available: https://www.creatio.com/industries/transportation [Accessed: 12-Feb-2023]

[5] Grand View Research, Customer Relationship Management Market Size Report, [Online], Available: https://www.grandviewresearch.com/industry-analysis/customer-relationship-management-crm-market [Accessed: 12-Feb-2023]

[6] Mastek, The Future of the Digital World: Internet of Things (IoT) with Internet of Behaviors (IoB), [Online], Available: https://blog.mastek.com/internet-of-things-iot-with-internet-of-behaviors-iob [Accessed: 12-Feb-2023]

[7] Precedence Research, ICT-Internet of Behaviors (IoB) Market, [Online], Available: https://www.precedenceresearch.com/press-release/internet-of-behaviors-market [Accessed: 12-Feb-2023]

[8] Cprime, A Complete Guide to the Internet of Behaviors (IoB), [Online], Available: https://www.cprime.com/resources/blog/a-complete-guide-to-the-internet-of-behaviors-iob/ [Accessed: 12-Feb-2023]

[9] Softtek, What is the Internet of Behaviour (IoB) and why is it the future?, Available: https://softtek.eu/en/tech-magazine-en/user-experience-en/what-is-the-internet-of-behaviour-iob-and-why-is-it-the-future/ [Accessed: 13-Feb-2023]

[10] J. Sun, W. Gan, H. C., Chao, P. S. Yu and W. Ding, "Internet of Behaviors: A Survey", arXiv:2211.15588v1 doi: https://doi.org/10.48550/arXiv.2211.15588

[11] Redwrite, The Developing IoB Technology and its Applications, [Online], Available: https://readwrite.com/ developing-internet-of-behavior-technology/[Accessed:13-Feb-2023]

[12] Xpanion, Benefits and Risks of Internet of Behavior, [Online], Available: https://www.xpanxion.com/news/news/benefits-and-risks-of-internet-of-behavior-/ [Accessed: 13-Feb-2023]

[13] Experion Technologies, Internet of Behavior (IoB) – A boon for customer relationship management in the coming decade, [Online], Available:https://www.experionglobal.com/internet-of-behavioriob-a-boon-for-customer-relationship-management-in-the-coming-decade/ [Accessed: 13-Feb-2023]

[14] Techtarget, Customer relationship management, [Online], Available:https://www.techtarget.com/searchcustomerexperience/definition/CRM-customer-relationship-management [Accessed: 14-Feb-2023]

[15] TP&P Technology, Customer Relationship Management in Logistic: Why It Matters, [Online], Available: https://www.tpptechnology.com/blog/customer-relationship-management-in-logistic-why-it-matters/ [Accessed: 14-Feb-2023]

[16] Infosys, Whitepaper: CRM – digital transformation driven or digital transformation driver, [Online], Available: https://www.infosys.com/services/salesforce/documents/digital-transformation-driven.pdf [Accessed: 14-Feb-2023]

[17] Faye, Customer Experience Trends in the Transport and Logistics Industry, [Online], Available: https://fayedigital.com/blog/customer-experience-trends-in-the-transport-and-logistics-industry/ [Accessed: 14-Feb-2023]

[18] Medium, The Experience Enabler Scorecard, [Online], Available: https://medium.com/still-day-one/the-experience-enabler-scorecard-6391fb28423 [Accessed: 14-Feb-2023]

[19] CIO, Nine ways to improve your company's CRM system, [Online], Available: https://www.cio.com/article/291103/enterprise-software-9-ways-to-improve-your-company-s-crm-system.html [Accessed: 14-Feb-2023]

[20] M. Javaid, A. Haleem, R. Pratap Singh, S. Rab and R. Suman, ”Internet of Behavior (IoB) and its role in customer services”, Sensors International, Volume 2, 2021, doi: https://doi.org/10.1016/j.sintl.2021.100122

[21] SoftClouds, Customer Relationship Automation will be the future of CRM, [Online], Available: https://softclouds.medium.com/customer-relationship-automation-will-be-the-future-of-crm-4a5735146bf3 [Accessed: 14-Feb-2023]

[22] Select Hub, CRM Automation: Features, Benefits and Top Solutions, [Online], Available: https://www.selecthub.com/customer-relationship-management/automated-crm/ [Accessed: 15-Feb-2023]

[23] Transmetrics, How Predictive Analytics Is Transforming Logistics and Supply Chain, [Online], Available: https://www.transmetrics.ai/blog/predictive-analytics-in-logistics/ [Accessed: 15-Feb-2023]

[24] Quash apps, Logistic Services UI/UX Design: To Enhance User Experience, [Online], Available:

https://squashapps.com/blog/logistic-services-ui-ux-design-to-enhance-user-experience/ [Accessed: 15-Feb-2023]

[25] Kron Technologies, Cybersecurity in the Logistics Industry, [Online] Available: https://krontech.com/cybersecurity-in-the-logistics-industry [Accessed: 15-Feb-2023]

Auto-generated Plans and Decisions as an Instrument of Crisis Management

D. Labaš, M. Galić and M. Pejić Bach
University of Zagreb, Faculty of Economics & Business, Zagreb, Croatia
dlabas@efzg.hr; marta.galic97@gmail.com; mpejic@efzg.hr

***Abstract* – Auto-generated plans and decisions are generated without human input. Machine learning algorithms or other automated decision-making systems develop them based on present criteria and rules. An automated system might create a reaction plan for a natural disaster or cyberattack. The system may assess historical occurrences, discover patterns and trends, and create a reaction plan based on specified parameters. An organisation may employ an automated decision-making technology to choose resources or actions in a crisis. The tool may evaluate data and make a judgement automatically using specified rules and criteria. In urgent circumstances, auto-generated strategies and choices may help. They may also help businesses react consistently to emergencies using objective criteria and data analysis, decreasing human error. This work aims to provide an overview of crisis management and the role of auto-generated plans and decisions in its deployment. The case study of one company is presented to illustrate the usage of auto-generated plans and decisions, taking care that they shouldn't replace human decision-making. To be current and successful, they should be evaluated and updated often.**

Keywords – crisis management, auto-generated plans, machine learning, chatbots

I. INTRODUCTION

Crisis management is a systematic attempt by organisation members to cooperate with stakeholder groups to avoid and prevent or successfully overcome emerging crises [1].

The three fundamental goals of crisis management are [2]: prevention, protection, and mitigation of the crisis. Crisis management often involves a range of companies' activities focused on crisis: from pre-crisis preventive and preparation activities through managing during the crisis and to post-crisis activities of learning new ways of thinking, conducting and revising activities for better future crisis response [3].

Effective crisis management is characterised by the ability to make correct and swift decisions under pressure [4], while attaining organisational tactical and strategic goals [5], and maintaining business continuity along with key competitive advantages [1]. Organisations with effective crisis management are better prepared to anticipate, detect and manage crises [6].

Therefore, organisations should accept crisis prevention and crisis management activities as a formal part of business planning. Accordingly, before a crisis occurs, preventive activities are necessary. The focal point is creating a crisis plan that includes all activities for managing a crisis in the event of its occurrence, emphasising scenario planning in crisis management [7].

This paper aims to present a case study using auto-generated plans and decisions as a crisis management tool in a Croatian company "X". Company "X" is a global leading IT and telecommunications company in omnichannel communication that enables advanced customer engagement and support through use of various communication channels, including chatbots. Data for the case study was collected through an in-depth interview and document analysis. The results showed that the company successfully uses the artificial intelligence tool chatbot as a crisis management instrument.

II. CRISIS MANAGEMENT

A. Traditional approaches to crisis management

Traditional approaches to crisis management are [7]: an event-based (reactive measures) approach that can be defined in comparison to the process-oriented (dynamic approach), which is oriented to improving the efficiency of companies' preparedness for crises. Although the mentioned activities are necessary for crisis management, they are often insufficient for proper crisis response [9]. The traditional event-based and process-based (holistic) approaches that consider crisis preparedness, planning and learning are needed for effective crisis management [10]. Once the crisis has passed, the crisis management team should audit and evaluate the conducted crisis management activities and crisis effects, to increase organisational resilience and crisis preparedness [11].

B. Scenario planning

With crisis management and business continuity, access to crisis preparedness, crisis identification, crisis response, and recovery is easier. Without the abovementioned measures, management will react instinctively in a crisis and often make mistakes instead of a pre-planned strategic response in crisis and conflict situations [6].

Accordingly, scenarios are useful crisis management tools that enable companies to better prepare for the unexpected [12]. Scenario planning helps organisations better anticipate and appropriately prepare for crises and changes and identify threats and opportunities [13]. Crisis management and its core crisis preparedness activity, crisis plan, and within it scenario planning help the organisation improve strategic thinking and external analysis of the environment, as well as better prepare for an uncertain, complex future and provide appropriate, future-oriented strategic thinking and response measures for potential future crises [7].

They are particularly helpful due to the following three features [12]: i) increase organisational preparedness through expanding crisis team members' thinking and insights, ii) help uncover the foreseeable future, iii) enable employees to freely state their opinions and challenge dominant thinking in the organisation.

A potential downside to scenarios is that they might not include all key variables and have a low probability of unforeseen events, such as a pandemic, which can leave organisations exposed to risks. In addition, scenarios should be kept updated with changes in the environment and organisation, and organisations should have at least 4 scenarios in place [12]. Compared to contingency planning, scenario planning has more uncertainties, is more difficult to model, enables employees' creativity, helps identify difficulties and provides several possibilities for interaction and outcome [7]. Using scenario planning improves the detection of organisational and environmental risks and strategic opportunities detection, resilience, crisis preparedness and decision-making [13].

C. Use of chatbots for crisis management

"The conversational agent is a broadly encompassing term for an information system that can communicate with human users through using and processing natural language. It includes chatbots, digital, virtual, or voice assistants" [14]. Chatbot is "a computer program designed to converse with a human being, especially over the internet" [15].

Furthermore, Gartner defines a chatbot as "a domain-specific conversational interface that uses an app, messaging platform, social network or chat solution for its conversation" [16]. Chatbot is software that enables conversation in human form to respond to customers' queries. It can usually help provide routine answers, while human agents are necessary for more complex problem-solving [17]. It has become prevalent in many industries, especially in public relations [18], digital media [19], and human resources management [20] for its digital transformation [21], emergency management [14], [22], widely supporting digital transformation [23], [24].

Key to effective crisis management is timely and thorough crisis preparation, which includes crisis scenario development and communication strategy in times of crisis. In that regard, crisis chatbots can be useful for crisis management. However, they require proper design and training based on certain scenarios. Furthermore, chatbots can also serve as an early warning system through monitoring communication (through analysis of keywords, questions asked, and public sentiment) and enable on-time and adequate crisis communication and response while also alleviating the stress and burden that members of crisis team may have, along with company's customer support agents [25].

IT technologies and AI can advance crisis response, emergency management, and decision-making [14]. Information technologies such as chatbots can improve crisis management, especially in crisis communication and can be used as another communication channel. Their advantages are that they are visible in times of crisis when many people need fast responses to common questions and information [26]. Besides chatbots' benefits for customer service, they enable time savings, cost reduction, 24/7 availability, and better environmental awareness [17].

Chatbots can support emergency management activities through social media, whereas potential activities that chatbots can do in emergency management include [27]: i) Automated responses to FAQ; ii) Additional channels for gathering emergency information; iii) Responding to inquiries regarding ongoing disaster; and iv) Asking for feedback on emergency organisations during the crisis.

Besides the chatbots, several other possible bot types can be applied in emergency management, each with its function: news bot, social bot, monitoring bot and public dissemination bot, whereas social bots can analyse information and respond to inquiries on the crisis, as well as send emergency warnings through social media [27].

D. Auto-generated plans and decisions

Auto-generated plans and decisions are generated automatically, without direct human intervention. They are typically based on predefined criteria and rules and are often generated using machine learning algorithms or other automated decision-making tools like chatbots [28].

For example, an organisation may use an automated system to generate a response plan for a specific type of crisis, such as a natural disaster or cyberattack [28]. The system may analyse data from past incidents, identify patterns and trends, and use this information to generate a response plan based on predefined criteria.

Auto-generated plans and decisions in the form of personal assistant, can be useful when time is critical, and decisions must be made quickly and efficiently [29]. They can also help organisations respond consistently to crises, reducing the risk of human error and ensuring that responses are based on objective criteria and data analysis. One example of such an auto-generated chatbot was developed during the Covid-19 crisis [30].

Similarly, an organisation may use an automated decision-making tool to decide during a crisis, such as

which resources to allocate or actions to take. The tool may use predefined rules and criteria to analyse data and generate a decision automatically, although criticism is expressed concerning such tools [31].

Auto-generated plans and decisions can be useful in crisis management because they can help organisations respond quickly and effectively to unexpected events. Here are a few ways they could be used:

- Rapid response: Time is of the essence when a crisis occurs. Auto-generated plans and decisions can be used to quickly assess the situation and generate a response plan based on predefined criteria. This can help organisations respond faster and more effectively than relying on human decision-making alone.

- Consistency: Decisions must be made quickly and consistently in a crisis. Auto-generated plans and decisions can help ensure that decisions are made consistently, even in high-pressure situations where human decision-making may be affected by stress, emotions, or other factors.

- Scalability: Demand for resources and services can increase rapidly during a crisis. Auto-generated plans and decisions can help organisations scale their response quickly and efficiently without relying solely on human resources.

- Data-driven decision-making: Auto-generated plans and decisions can be based on data analysis and machine learning algorithms that can quickly process large amounts of data. This can enable organisations to make better informed decisions based on real-time data and analysis during a crisis.

- Reducing human error: In high-pressure situations, human decision-making can be affected by stress, fatigue, and other factors. Auto-generated plans and decisions can help reduce the risk of human error, ensuring that decisions are made based on predefined criteria rather than subjective factors.

Overall, auto-generated plans and decisions can be valuable in crisis management, helping organisations respond rapidly, consistently, and effectively in unexpected situations. One example of the application of automated response is autoamted detection of fake news regarding Covid-19 pandemics [32]. However, it's important to note that they should be used to complement human decision-making rather than replace it. Recent surge in use of autogenerative text bot, ChatGPT has revealed that it cannot be used freely, since it is prone to errors [33], and it should be trained for a specific purpose [34].

III. CASE STUDY OF CRISIS MANAGEMENT USING A MODERN APPROACH

A. Application of communication channels for crisis management

Modern companies develop an infrastructure that is highly adaptable to a dynamic market characterised by continuous changes in communication trends, speed and levels of accuracy, reliability and security that the work of today's companies requires. To achieve this, they use the so-called hybrid cloud model, which is coordinated with local services and locally stored data so that they can never run out of resources.

For example, an analysed software company (from now on, company “X”) that uses modern communication channels and, in addition to customer support, provides consulting services, suggests the best forms of integration and offers the most effective solutions and decisions based on best practices, especially if there are unwanted, unknown or unusual activities in everyday business.

The use of such communication channels on the example of one contact centre, whereas for each of the channels advantages of communication with clients are as follows:

• Instagram Business - The Contact Center helps manage all communication on Instagram, one of the main channels for product discovery. It offers the connection through direct messages, comments, mentions in stories and replies with the contact centre on the "cloud" to simultaneously enable a conversation with several clients.

• WhatsApp Business - Offers a connection with customers on a global level, promoting the creation of personalised campaigns and engaging customers with interactive messages and constant reporting on promotions, in-stock items or offers, and must-have activities. This reduces costs and increases agents’ productivity.

• SMS API - Easy integration of SMS messaging into business workflows is encouraged using a single messaging API. Over 650 direct operator connections in 190 countries are available, enabling low-latency, high-reliability interactions that cover every possible case – from SMS alerts on unusual work activity to two-factor authentication solutions. Also, real-time campaign performance analytics and reports are available.

• Web Push - Web push notifications are a quick and easy way to bring site content to visitors without the need to log in or leave personal information.

• Live Chat - Live chat provides continuous customer support with a combination of chatbots and live conversations with an agent, which increases the likelihood that one-time visitors will become valuable long-term clients, given that their questions are answered from the very beginning of their activity and support is provided when it is most needed. By using channels, the time to solve and eliminate problems is reduced, costs are

reduced, productivity is increased, and customer satisfaction is preserved.

• Viber for Business - Enables the delivery of personalised messages "one on one", establishing a relationship with customers based on trust. The main advantage is that if WIFI is unavailable, Viber automatically switches to SMS, ensuring that every message is delivered on time.

Any unusual occurrence in the business process can trigger a crisis in the company. That's why every problem or difficulty in the business chain should be taken seriously and not avoided but instead eliminated. Companies should react appropriately and effectively and learn from the conducted procedures. For companies or certain departments within the company to resolve unusual difficulties in work as painlessly as possible, it is necessary to collect and process all data in an orderly and up-to-date manner and use reports as indicators of the state of the company and business results.

trends, which affect all industries, it is necessary to survive in the market by applying and implementing all available tools.

Chatbots that promptly provide solutions and decisions to emerging crises are an important resource for the company. In addition to providing a competitive advantage, they ensure easier overcoming of crises and simpler implementation of the necessary activities because the system remembers and constantly learns, collects, processes and serves the necessary information and notifications based on which decisions are made. It also suggests an action plan for the implementation of certain crises and challenges that arise.

Plans and decisions must be made in response to the challenges companies face and become a common, but powerful engine that drives various data systems and implements a wide range of content into business, aiming to optimise and implement the strategy.

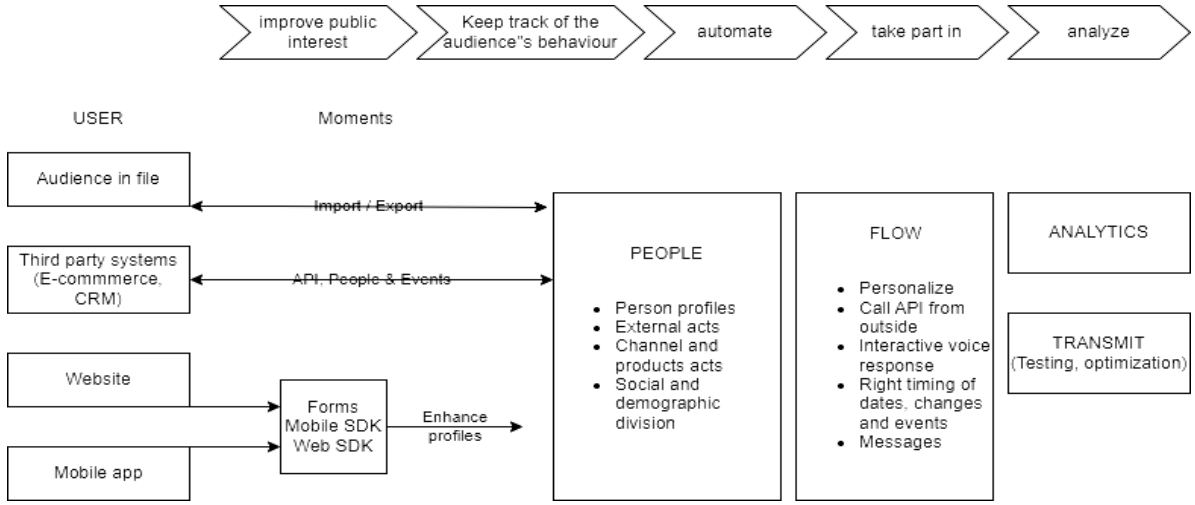


Fig. 1. Platform of analyzed company; Source: Internal materials, adapted by authors

Fig. 1 shows the way the “Platform” of the analysed company “X” works - i.e. the way the client data is collected and processed, how the system reacts to activities by sending an automated message to get a report on the changes that occurred finally, followed with a template on the next actions that need to be taken.

The “Platform” of the analysed company “X” shows the detection of changes in the business process and the sending of auto-messages and responses.

B. Case study of using automated plans and decisions in the form of a chatbot for crisis management

Automating standardised decisions that must be made daily in everyday business activities makes reaching established plans for implementing decisions easier, saving resources and time.

In addition to being exposed to less stress, employees are more agile and productive. Given that today's companies are constantly under the influence of various

Given that auto-generated plans and decisions are based on a large set of data that the system processes and proposes as a solution, it is necessary to adhere to certain principles for the implementation and success of the plan or decision to be valid:

• All proposed decisions should have approximately equal samples and variables.

• If two (or more) decisions are very similar, one can be made more likely by providing more samples for that decision or activity plan, or they can be combined into one and then branched out using attributes (for example, increasing or decreasing the allowable inventory level in the warehouse).

• To achieve maximum utility, decisions should not be similar.

• There must be a certain number of complete sentences that clearly define the intention. The better the bot is trained with complete sentences and data, the better

it will offer solutions and decisions and propose plans with the necessary activities.

• The same keywords or their exact synonyms must not be used in two different decisions.

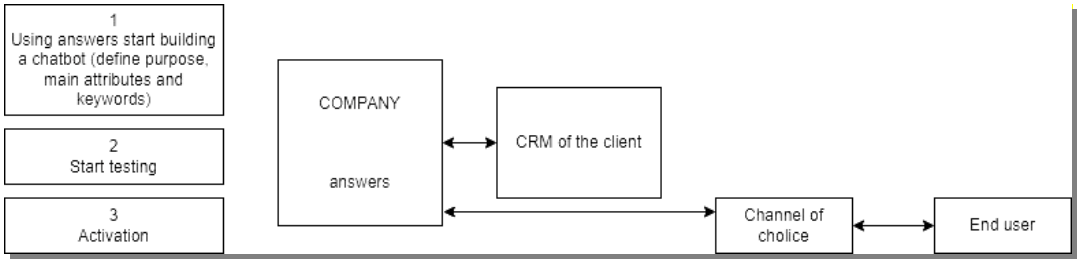


Fig. 2 Chatbot communication with client; Source: Internal materials adapted by authors

The analysed company "X" (provider of cloud communication platform service) has continuous support 24/7 for its clients. To solve clients' problems as efficiently as possible and offer solutions and help them when unusual activities occur in business processes, company "X" uses and follows auto-generated decisions and plans as one of the instruments in preventing crises.

Such a way of conducting business and communicating with clients is facilitated by many platforms, one of which is "Platform" of company "X", which is a platform that is used to activate automatic messages in such a way that the platform is connected to the system. If the system crashes, automatic messages are activated, and supervisors and those responsible are informed about the changes. In this way, it is possible to react to the problem promptly and quickly to eliminate the difficulty/s that hinders the normal functioning of the system and work processes. The "Platform" of company "X" contains all relevant information about clients, which is saved in the database so that all unusual changes and activities can be detected. Communication is based on providing and offering solutions from previous practices; important and special dates, deadlines and other significant moments are monitored. The user is informed by automatic messages by the available data and analytics. Integration is possible through the Call API. Information can be requested from the system or database to prepare decisions and plans for the same or similar business situations more easily and quickly.

To properly implement the plan of necessary business activities when solving problems during crises, it is required to incorporate into the bot words mentioned very often and of great importance. For example, to make a template of an action plan for solving a financial crisis, the word "liquidity" may appear 7/10 times, which makes the word very important for the decision, while the word "In contrast" may appear only once, which is considered to be less important word and a smaller number of activities that are recommended to be undertaken when solving challenges will be directed towards it. For the model used, the word importance weighting scheme does not compare the importance of each word across all decisions and intentions. Still, it assigns importance according to the intention in which it appears.

For the system to send auto-generated messages, i.e. the system for proposing decisions and plans to function, the following three steps must be executed (Fig. 2): Prepare all data and links that the bot will use, Check API integration with the CRM system, and design the dialogue flow. Also, it is necessary to create an intention for each property offered to direct users to the correct report of an error or disturbance, with all available and collected data for processing, thereby directing the system to propose the correct decision or plan.

IV. CONCLUSION

The paper aims to present an overview of auto-generated plans and decisions as an instrument of crisis management. The paper provides an overview of the literature regarding crisis management, considering traditional approaches, scenario planning and chatbots. As literature review and analyzed case study in this paper shows, chatbots are useful for crisis management, especially for advanced communication with clients through various communication channels. In addition, chatbots enable use of previous business practices and can serve as an early warning to prepare decisions and plans simply and swiftly and accordingly help prevent or better manage during crisis. Analyzed case study illustratrates the approach, resulting in easy, fast and customised communication between agents and clients successfully realised, demonstrating the company's ability to provide timely, cost effective and personalised service to its customers as an addition to the development path of the company's digital transformation. However, human intervention and predefined criteria are needed for seamless integration and successful use of automated decision- making tools in crisis management such as chatbots (where it is necessary to define purpose, main attributes and keywords), implying that they are still complementary tools to human decision- making in crisis management.

With expected increase in use of chatbots and other automated decisions making systems, future research could focus on: i) areas and specifics of chatbot use

besides customer service and crisis management, such as healthcare and education; ii) use of AI and natural language understanding in order to improve the existing automated decision making systems; iii) comparative analysis of companies that used chatboes and its effectiveness, as well as evaluation of user satisfaction and areas for improvement.

REFERENCES

[1] Pearson, C., and Clair, J., "Reframing crisis management," Academy of Management Review, vol. 23(1), 1998, pp. 59-76.

[2] Doktor, C., "Strategic crisis management: trends and concepts," Center for Securities Studies (CSS). ETH Zurich, CSS Analysis in Security Policy, vol. 2(23), 2007, pp. 1-3.

[3] Crandall, C., Parnell, J. A., and Spillan, J. E., "A framework for Crisis Management: Crisis management in the new strategy landscape," Sage: Los Angeles, 2010.

[4] Wooten, L. P. and James, E. H., "Linking Crisis Management and Leadership Competencies: The Role of Human Resource Development," Advances in Developing Human Resources, vol. 10(3), 2008, pp. 352-379.

[5] Shaw, G. L., "The competencies required for executive level business crisis and continuity managers," Doctoral dissertation, George Washington University, 2004.

[6] Sapriel, C., "Why do we keep making the same mistakes?" Communication World, vol. 27(4), 2010, pp. 28-30.

[7] Pollard, D., and Hotho, S., "Crises, scenarios and the strategic management process," *Management Decision*, vol. 44(6), 2006, pp. 721-736.

[8] Oscarsson, O., "Crisis-as-practice: Conceptualizing the role of everyday work practices in crisis management," International Journal of Disaster Risk Reduction, vol. 83, 2022, in press.

[9] James, E. H., Wooten, L. P., and Dushek, K., "Crisis Management: Informing a New Leadership Research Agenda," The Academy of Management Annals, vol. 5(1), 2011, pp. 455-493.

[10] Jaques, T., "Embedding Issue Management as a strategic element of Crisis Prevention," Disaster Prevention and Management. Vol. 19(4), 2010, pp. 469-482.

[11] Mikušová, M., and Horváthová, P., "Prepared for a crisis? Basic elements of crisis management in an organisation," Economic Research-Ekonomska Istraživanja, vol. 32(1), 2019, pp. 1844-1868.

[12] Roxburgh, C., "The use and abuse of scenarios," McKinsey Quarterly, November 1, 2009. https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/the-use-and-abuse-of-scenarios

[13] Deloitte, "Scenario planning reduces uncertainty, and increases resilience," June 23, 2022. https://action.deloitte.com/insight/2230/scenario-planning-reduces-uncertainty-increases-resilience

[14] Stieglitz, S., et al., "Design principles for conversational agents to support Emergency Management Agencies," International Journal of Information Management, vol. 63, 2022, in press.

[15] Cambridge Dictionary, "Chatbot," n.d. https://dictionary.cambridge.org/dictionary/english/chatbot

[16] Gartner, "Gartner Glossary: Chatbot," n.d. https://www.gartner.com/en/information-technology/glossary/chatbot

[17] Schweta, and Main, K., "What is a Chatbot? Everything you need to know," Forbes Advisor, August 21, 2022. https://www.forbes.com/advisor/business/software/what-is-a-chatbot/

[18] Kolić Stanić, M., Polak, M., and Togonal, M., "Artificial Intelligence in Communication with Music Fans: An Example from South Korea," ENTRENOVA-ENTerprise REsearch InNOVAtion, vol. 8(1), 2022, pp. 48-63.

[19] Tsioustas, C., et al., "Innovative applications of natural language processing and digital media in theatre and performing arts," ENTRENOVA-ENTerprise REsearch InNOVAtion, vol. 6(1), 2020, pp- 84-96.

[20] Trubetskaya, A., and Mullers, H., "Transforming a global human resource service delivery operating model using Lean Six Sigma," International Journal of Engineering Business Management, vol. 13(9), 2021, pp. 1-16.

[21] Barišić, A. F., Bach, M. P., and Miloloža, I., "Human resources information systems: transactional and strategic paradigm," ENTRENOVA-ENTerprise REsearch InNOVAtion, vol. 4(1), 2018, pp. 224-230.

[22] Barišić, A. F., Barišić, J. R., and Miloloža, I., "Digital transformation: challenges for human resources management," ENTRENOVA-ENTerprise REsearch InNOVAtion, vol. 7(1), 2021, pp. 357-366.

[23] Tomičić Furjan, M., Tomičić-Pupek, K., and Pihir, I., "Understanding digital transformation initiatives: Case studies analysis," Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, vol. 11(1), 2020, pp. 125-141.

[24] Hunady, J., Pisár, P., Vugec, D. S., and Bach, M. P., "Digital Transformation in European Union: North is leading, and South is lagging behind," International Journal of Information Systems and Project Management, vol. 10(4), 2022, pp. 58-81.

[25] Prevency, "5 tips for using chatbots in crisis communication," n.d. https://prevency.com/en/5-tips-for-using-chatbots-in-crisis-communication/

[26] Bonales Daimel. G., and Citalli Martinez Estrella, E., "Using Virtual Assistants and Chatbots for Crisis Communication," Revista Internacional de Investigación en Comunicación aDResearch ESIC, vol. 25(25), 2021, pp.70-90.

[27] Hofeditz, L., Ehnis, C., Bunker, D., Brachten, F. and Stieglitz, S., "Meaningful use of social bots? Possible applications in crisis communication during disasters," In Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Swedenm June 8-14, 2019.

[28] Meinrenken, C. J., "Residential electricity conservation in response to auto-generated, multi-featured, personalized eco-feedback designed for large scale applications with utilities," Energy and Buildings, vol. 232, 2021, 110652.

[29] Hüsson, D., Holland, A., and Arteaga Sánchez, R., "Intelligent Personal Assistant in Business-Context: Key-feature Evaluation for User Acceptance," Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, vol. 11(3), 2020, pp. 147-166.

[30] Li, Y., et al., "Jennifer for COVID-19: An NLP-Powered Chatbot Built for the People and by the People to Combat Misinformation" In ACL 2020 Workshop on Natural Language Processing for COVID-19 (NLP-COVID), Seattle, Washington, USA, July 9-10, 2020.

[31] Araujo, T., Helberger, N., Kruikemeier, S., and De Vreese, C. H., "In AI we trust? Perceptions about automated decision-making by artificial intelligence," AI & society, vol. 35, 2020, pp. 611-623.

[32] Domala, J., et al., "Automated identification of disaster news for crisis management using machine learning and natural language processing," In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) IEEE, July, 2020, pp. 503-508.

[33] Gozalo-Brizuela, R., and Garrido-Merchan, E. C., "ChatGPT is not all you need. A State of the Art Review of large Generative AI models", 2023, arXiv preprint arXiv:2301.04655.

[34] Sallam, M., "The Utility of ChatGPT as an Example of Large Language Models in Healthcare Education, Research and Practice: Systematic Review on the Future Perspectives and Potential Limitations", 2023, medRxiv, 2023-02.

The Legal Status of Digital Nomads

H. Arbutina*, Z. Šinković** and L. Pribisalić**
* University of Zagreb, Faculty of Law/ Chair of Financial Law and Financial Science, Zagreb, Croatia
** University of Split, Faculty of Law/ Chair of Financial Law and Financial Science, Split, Croatia
zoran.sinkovic@pravst.hr

***Abstract* – While the pandemic of the COVID-19 virus took over and changed an increasing part of people's everyday lives, the spread of digital communication and teleworking grew. As a result, there is an increase in interest in a new way of working that is not tied to a physical workplace - digital nomadism. Nomads of the modern, digital age are people who use information and communication technology for work that they can do anywhere in the world, which gives them the luxury of freely choosing their residence. The Republic of Croatia was among the first countries in the European Union to open its doors to digital nomads, just behind Estonia. In the Republic of Croatia, the Foreigners Act regulates and defines the legal status of digital nomads as citizens of a third country (a country that is not a member state of the European Economic Area or the Swiss Confederation) who are employed or perform work through communication technology for a company or their own company that is not registered in the Republic Croatia and does not perform work or provide services to employers in the territory of the Republic of Croatia. Therefore, this paper will analyze the legal status of digital nomads from the administrative and tax aspects (*de lege lata - de lege ferenda*) in the Republic of Croatia and other member states of the European Union.**

***Keywords* – digital nomads, residence, health insurance, health care, taxation**

I. INTRODUCTION

In today's digital world, lifestyle and work without the computer and internet are unthinkable. Digital technologies have brought about huge changes and digital transformation includes more than technologies themselves and relates to changes in ways of thinking. Digital transformation impacts all sectors of the economy and changes our way of life, work, and communication. Existing and often outdated ways of doing business need to be changed, that is, adapted to the needs and challenges of the digital era. There is much room for digital transformation - it is no longer a matter of choice but rather inevitable and essential. Digital transformation is omnipresent in companies' business dealings, the economy, health, education, public administration, and society in general [1].

The concept of digital transformation relates to changes that emerge from the application of digital technologies in all social forms. Digital transformation is a fundamental and accelerated transformation in doing business, processes, capabilities, and models to entirely make use of the possibilities of digital technologies and their impact on society in a strategic and prioritized way [2]. The development of information and communication technologies (ICT) and globalization has opened up the possibility of new flexible work arrangements. In today's world, it is common to work remotely using digital technologies [3].

The concept of "digital nomadism" appeared in popular literature some 25 years ago in a provocative and insightful book “The digital nomad” co-authored by Tsugio Makimoto and David Manners [4].

It is necessary to emphasize how digital nomadism is not the first kind of mobile worker appearing in post-agrarian societies. In medieval Europe, the concept of journeyman (ger. *Wanderjahre*) emerged which marked a period of two or three years in which young merchants set out (ger. *auf der Walz)* waltzing from town to town carrying the tools of their trade, to improve their personal and professional skills. Travelers who practiced this type of lifestyle most often were unattached that is, were not married [5]. Digital nomadism is work that is independent of location. It offers services, the creation of products and marketing, and product distribution or anything offered from a distance via online platforms [6]. Digital nomads are both a product and an example of the ubiquity of mobilities in everyday lives. As our society rapidly transforms itself into a mobile society, in which interactions are also mobilized, the “traditional segmentation of context dissolves so private life can interrupt working life and vice versa” [7]. Digital nomads are primarily young (millennials or Generation Z) individuals who are motivated to explore and combine travel with virtual work [3]. According to Hensellek and Puchala, all definitions of digital nomads have common factors: digital work, flexibility, mobility, identity, and community [8]. Demaj, Hasimja, and Rahimi argue that nomadism is related to digitalization because digital nomads can work while moving from one place to another simply by having a mobile device and an internet connection [9]. Increasing digitalization and workplace flexibility creates a market structure of work where several kinds of tasks are carried out based on casual work contracts, which have brought about the increasing emergence of digital nomads. Today there are more than 35 million digital nomads. Digital nomads work in various areas including: 1. IT: 19%, 2. Creative Services: 10%, 3. Education and Training: 9%, 4. Counseling, Teaching, and Research: 8%, 5. Sales, Marketing, and PR: 8% and 6. Finance and Accounting: 8% [10]. So, we can conclude that a digital nomad is an entrepreneur or an employee who almost exclusively uses digital

technologies to perform his work and is at the same time independent and leads his life in several places. As a rule, digital nomads are engaged in professions that, thanks to the Internet, can perform independently of the place of work, and most often they are highly qualified IT specialists, writers, journalists, designers, and other similar professions. Using the advantages of technology and thanks to the development of technology and the spread of the Internet, a good part of the work is done outside the office and it is not rare that they decide to replace their "office" in a bigger city with another destination. Such individuals freely choose their place of work and life with a laptop computer, mobile devices, and an Internet connection [11].

As stated above, today, there are more than 35 million digital nomads, and almost half of them are citizens of the USA. Their purchasing power amounts to 800 billion dollars a year [12]. Every fifth digital nomad (21%) earns less than 25,000 dollars annually. Nevertheless, 44% earn 75,000 dollars or more. The average monthly earnings of digital nomads amount to 4,500 dollars [10].

Therefore, the legal status of digital nomads in the Republic of Croatia and certain European Union member states (the Federal Republic of Germany, the Republic of Italy, Republic of Estonia) from administrative and taxation aspects will be analyzed in this paper *de lege lata – de lege ferenda*.

II. ADMINISTRATIVE AND TAX STATUS OF DIGITAL NOMADS IN SOME MEMBER STATES OF THE EUROPEAN UNION

Experts predict that by 2035, 1 billion people will be distance working which was considerably increased during the Coronavirus pandemic, and this trend is accelerating. Amongst them will certainly be people who will become digital nomads, those who are not tied to the location of their workplace and who only need digital technology and reliable Internet and access [13] that is, a safe and attractive destination.

This chapter will briefly present the legal status of digital nomads in the member states of the European Union, namely: the Federal Republic of Germany, the Republic of Italy, and the Republic of Estonia.

The Federal Republic of Germany (hereinafter: FR Germany) does not have a visa for digital nomads. However, a person can apply for a freelancer visa (ger. *Freiberufler*), with which he can work remotely for German clients [14]. A citizen of the European Union, the European Economic Area, or Switzerland can ignore the complex bureaucracy and work in FR Germany. In addition, a citizen of the United Kingdom, USA, Australia, Canada, Israel, Japan, New Zealand, or South Korea also does not need a visa, only a residence permit. Some digital nomads who plan to work in FR Germany for less than 90 days in 180 days do not notify the authorities at all and simply enter as tourists. While it is possible to get away with impunity, it should also be noted that this is technically illegal unless the person is a freelancer, translator, tour guide, or model. In these cases, he does not have to apply for a work visa or residence permit, provided that the stay in FR Germany does not exceed 90 days in 180 days [15]. FR Germany wants foreign freelancers to register with the tax authorities. Therefore, it is mandatory to pay several taxes as a digital nomad, namely: 1) Income tax. Income tax depends on earnings. There is a fixed ratio of tax for certain values of income. If a person earns up to EUR 9,744 per year, he is exempt from income tax. Earnings between EUR 9,744 and EUR 57,918 per year make it mandatory to pay income tax, which varies between 14%-42%. If a person earns between EUR 57,918 – EUR 274,612 makes him qualified to pay exactly 42% of the amount. Earnings greater than EUR 274,612 lead to the payment of 45% income tax. Income tax is paid four times a year in March, June, September, and December. At the end of the year, an income tax return is submitted. If the amount paid exceeds the amount stated in the statement, it will be refunded. 2) The value-added tax paid in Germany is 19% in most cases, but it can be up to 9%. This is based on the revenue that the taxpayer generates. These taxes add to the worries of the digital nomad. Health insurance for digital nomads is mandatory in FR Germany [16].

After FR Germany, the legal status of digital nomads in the Republic of Italy will be presented.

In the Republic of Italy (hereinafter: Italy), digital nomad visas are similar to a full-time work permit, insofar as they allow foreign citizens to move to the country and work legally from it with a residence permit. Although the Italian digital nomad visa is yet to be launched at the end of 2022, citizens of countries outside the European Union can still obtain residency and work remotely in Italy. It is usually possible through an Italian self-employment visa. The Italian self-employed visa is similar to the digital nomad visa – both allow "highly qualified" foreign remote workers and freelancers working outside the European Union to move to Italy and work there legally. Although the requirements for the digital nomad visa are still unclear as the program has yet to come into effect, it can be concluded that its requirements are similar to those of the self-employed visa. Eligible self-employed individuals must meet the minimum income requirement set by the Italian government, have a clean criminal record, and have a comprehensive health insurance plan. Those who qualify for a self-employed visa are issued a renewable permit with an initial validity of one year. Once activated, they may be able to renew it provided they continue to comply with the program's requirements. After five years of residence, they can obtain a permanent residence permit. In addition, individuals may include members of their immediate family in their residency applications [17]. Digital nomads who wish to apply for a self-employed visa in Italy will need to ensure that they meet the following conditions before they can obtain a residence permit: they must not be citizens of the European Union or the European Economic Area; hold a clean criminal record; have proof of adequate accommodation in Italy; have an annual gross income of at least EUR 8,500 earned in the year before the visa application or promised by contractual remuneration for services provided to Italian clients or an Italian company; have health insurance to cover hospital or medical expenses in Italy for the first 30 days after entry; attach a certificate that there are no obstacles to the provision of their freelance services and

highly qualified work activities in Italy; provide proof of sufficient financial means to maintain their stay and provide their services [17]. It is important to note that digital nomads will have to pay Italian digital nomad visa taxes or self-employment taxes when they move to the country and comply with social security provisions. While tax residents in Italy would normally have to pay a high tax rate on their annual worldwide income, foreign professionals and remote workers who move to Italy on a self-employed visa will only have to pay tax on 5 percent of their annual income. Moreover, foreign workers who acquire the new digital nomad visa will have a 70 percent tax reduction on their global income if they declare their residency in Italy. If persons were thinking about moving to Italy and working remotely as digital nomads without paying taxes twice, they didn't have to wait until the beginning of 2023. The Self-employed Visa can be used in place of the Italian Digital Nomad Visa, and its benefits are truly unparalleled – whether it's the advantage of double taxation treaties and other tax exemptions, access to better healthcare, or otherwise. By applying for a self-employed visa in Italy, digital nomads will have the right to reside and work remotely in Italy without being considered tax residents there. Also, they will be able to travel to the rest of the European Union without visas within two months from the submission and approval of the application. Moreover, immediate family members can be included in the application and enjoy the benefits of family reunification. All this makes working remotely much more practical [17].

Finally, the legal status of digital nomads in the Republic of Estonia will be presented.

The Republic of Estonia (hereinafter: Estonia) has long been known as a digital country, having first developed an online voting system in 2005. Estonia has a digital nomad visa, which allows remote work for any employer registered outside of Estonia. Estonia offers two types of digital nomad visas. The applicant can choose which one to request and receive depending on the length of stay. There are two types of Estonian digital nomad visas: 1. Estonian digital nomad visa type C. This is a short-stay visa that allows a stay of up to 90 days. 2. Digital nomad visa for Estonia type D. This is a long-term visa that allows a digital nomad to stay in Estonia for one year. To be qualified for a digital nomad visa in Estonia, a person must meet the following requirements: be over 18 years old; can work remotely using telecommunications technology; has an employment contract with a company registered outside Estonia; must prove that he is a freelancer or digital nomad; must prove that he has sufficient financial resources; must have at least EUR 3504 monthly salary [18].

The following documents are required to apply for a digital nomad visa in Estonia: A valid passport. The person must submit a passport that is valid for at least three months after the date of return. It must have two blank pages; Application form. The application form must be completed online. Once filled out, you must print and sign; Two passport-size photos. The Estonian digital nomad visa photo must meet the Schengen visa photo criteria; Contract of employment. A person must prove that he has a contract with an employer registered outside Estonia; a Letter from the employer. A letter from the employer stating that the person wishes to work remotely must be submitted; Proof of sufficient financial resources. The bank statement must be issued within the last six months. It must prove that the person has sufficient financial resources to support himself during his stay in Estonia; Proof of accommodation. A person must prove that he has found accommodation in Estonia. A rental agreement or any reservation can be submitted; Proof of health insurance. The person must have valid health insurance that is valid for the entire stay and that covers EUR 30,000 [18].

If a digital nomad stays in Estonia on a digital nomad visa for less than 183 days, the digital nomad does not need to pay taxes. However, after 183 consecutive days, the digital nomad will be taxed as he will be considered a tax resident. The tax liability in Estonia can be reduced due to the tax that the digital nomad pays in his home country. This depends on whether the country where the digital nomad is resident has a double taxation treaty with Estonia or not [18].

III. ADMINISTRATIVE AND TAX STATUS OF DIGITAL NOMADS IN THE REPUBLIC OF CROATIA

The Republic of Estonia is considered a pioneer in regulating the status of digital nomads, with its "digital nomad visa" program adopted in 2020. Among the more than 30 countries that have followed suit and adopted similar regimes specifically aimed at digital nomads is Croatia, with corresponding legislative changes in force from January 1, 2021 [4]. The Republic of Croatia, as one of the countries for which tourism is one of the most important branches of the economy, has recognized digital nomadism as a form of economy and tourism as one of the future orientations and will promote itself as an excellent destination for digital nomads by establishing a legal framework. It is a new type of tourists who travel and at the same time earn money for their trips, which has become a global trend recently. Namely, digital nomads bring significant value to the destinations where they decide to work, given that they stay there longer even outside the tourist season, and spend the money they earn. Likewise, their importance is also reflected in the sociological sense through the introduction of different values and lifestyles in local communities. The Republic of Croatia is one of the few countries in the world that has regulated the status of digital nomads and made itself a destination for a year-round stay, not just a temporary tourist visit in the summer months [11].

The flexibilization of the labor market and the legal regulation of the residence of citizens of third countries in the Republic of Croatia open up numerous opportunities for foreigners to live and work. The Foreigners Act allows and regulates temporary residence for digital nomads. The same Act introduced the term "digital nomad", which means a foreigner, a citizen of a third country, who is employed or performs tasks through telecommunications technology from the Republic of Croatia for a foreign company or his own company that is not registered in the territory of the Republic of Croatia and does not perform tasks or provide services to employers on the territory of the Republic of Croatia [19]. In Article 3, Paragraph 1,

Point 43, a digital nomad is defined as a citizen of a third country who is employed or performs tasks via communication technology for a company or own company that is not registered in the Republic of Croatia and does not perform tasks or provide services to employers on the territory of the Republic of Croatia [20]. According to Article 3, Paragraph 1, Point 4, a third country is a country that is not a member state of the European Economic Area or the Swiss Confederation, while Article 3, Paragraph 1, Point 3 stipulates that a citizen of a third country is a person who does not have the citizenship of a member state of the European Economic Area or the Swiss Confederation, and has the citizenship of a third country or is a stateless person [20]. Of course, the Foreigners Act, in Article 3, Paragraph 1, Point 2, stipulates that a citizen of a member state of the European Economic Area is a person who has the citizenship of one of the member states of the European Economic Area [20]. According to Article 57, Paragraph 1, Point 11, a digital nomad who is a citizen of a third country who resides or intends to reside is granted temporary residence. According to Article 57, Paragraph 4, a third-country citizen can submit a request for the regulation of temporary residence for other purposes or the purpose of a digital nomad's residence after the expiration of a period of six months from the expiration of the validity of the temporary residence that was approved for other purposes or the purpose of digital nomads residence [20]. Therefore, the temporary residence is approved for one year, and it is not possible to extend the permit, and after the expiration of the period of 6 months from the expiration of the validity of the digital nomad's temporary residence, a request to regulate said residence can be submitted again [11].

A digital nomad must meet the general conditions for approval of temporary residence as a citizen of a third country, and proof that a person performs work via communication technology (statement of the employer or citizen of a third country) is recognized as proof of the purpose of residence as a digital nomad. Thus, it is necessary to attach, for example, a contract on work or performance of work for a foreign employer, or a copy of the registration of one's own company and proof that it carries out the aforementioned tasks through one's own company [11]. Based on the Regulation on the method of calculation and the amount of funds for the maintenance of third-country citizens in the Republic of Croatia, a citizen of a third country who submits a request for temporary residence approval to stay as a digital nomad must have monthly funds for maintenance in the amount of at least 2.5 of the average monthly net paid salary in the last year, and for each additional family member or life partner or informal life partner, the amount is increased by an additional 10% of the average monthly net paid salary in the last year [21]. The following documentation is required as evidence to determine the purpose of the digital nomad's stay in the Republic of Croatia: statements from the employer or a citizen of a third country (as evidence that the work is performed via communication technology): a contract on work or performing work for a foreign employer or a copy of the registration of one's own company and proof that it performs these jobs through one's own company, proof of means of support during one's stay in the Republic of Croatia, a copy of a valid travel document whose validity period must be three months longer than the period validity of the intended stay, proof of health insurance (travel or private health insurance must cover the territory of the Republic of Croatia), proof that he has not been legally convicted of criminal offenses from the home country or the country in which he resided for more than a year immediately before arriving in the Republic of Croatia, address in the Republic of Croatia, with the provision that a temporary address can also be specified (hostel/hotel if accommodation is booked/confirmed) [11]. As the Foreigners Act allows citizens of third countries to stay longer in Croatia, up to one year, the issue of taxation of their receipts during their stay in the Republic of Croatia is raised [11].

This chapter will also discuss the tax status of digital nomads in the Republic of Croatia.

Given that the new Foreigners Act introduced the concept of digital nomads, according to which a digital nomad is a citizen of a third country who is employed or performs tasks via communication technology for a company or own company that is not registered in the Republic of Croatia and does not perform tasks or provide services to employers in the territory of the Republic of Croatia, it was also necessary to define the tax treatment of the receipts that they will realize [11]. In the basic provisions of the Income Tax Act, the scope of receipts that are not considered income has been expanded, among others, for the receipts of natural persons realized on the basis of the acquired status of digital nomads under a special regulation (the Foreigners Act) [22]. At the end of 2020, a new Point 26 was added to Article 9, Paragraph 1, which stipulates that income tax is not payable on the receipts of natural persons realized based on self-employed work or activities for an employer who is not registered in the Republic of Croatia based on the acquired digital nomad status under a special regulation [23]. Therefore, it is for receipts that will be realized by the same thus acquired digital nomad status is prescribed a tax exemption, and there would be no reporting obligation for the receipts thus achieved, and the tax exemption is prescribed to facilitate their decision to choose the Republic of Croatia as their place of residence and work, given that these are highly qualified experts [11]. As a digital nomad is employed or performs tasks for a company or his own company that is not registered in the Republic of Croatia and does not perform tasks or provide services to employers in the territory of the Republic of Croatia, he does not acquire insurance status with the competent institutes (Croatian Pension Insurance Institute and Croatian Health Insurance Fund) based on employment or self-employment and, there is no obligation to pay mandatory contributions [11]. The insurance status of foreigners is regulated under the Compulsory Health Insurance and Health Protection of Foreigners in the Republic of Croatia Act. Article 4 of the Act on Amendments to the Compulsory Health Insurance and Health Care for Foreigners in the Republic of Croatia Act, which entered into force on March 20, 2021, regulates the way to exercise the right to health insurance and health care of digital nomads. This amendment to

Article 6, Paragraph 12 of the integral text of the Compulsory Health Insurance and Health Care of Foreigners in the Republic of Croatia Act stipulates that an application for compulsory health insurance is not required to be submitted by a citizen of a third country on a temporary stay in the Republic of Croatia, who resides in the Republic of Croatia in the purpose of the digital nomad's residence, and, he is obliged to bear the costs of using health care in a health institution, that is, with a health worker in private practice or another provider of health care in the Republic of Croatia [24][25]. With the introduction of the new concept of digital nomads in the Foreigners Act, the conditions of entry, movement, residence, and work of foreigners who are citizens of third countries in the Republic of Croatia, and who carry out work through telecommunications technology from the Republic of Croatia, but not for employers in the territory of the Republic of Croatia, but abroad. Likewise, by defining the tax treatment, i.e. by exempting from taxation the income they will receive based on the acquired digital nomad status, their status was legally regulated to facilitate their decision to choose the Republic of Croatia as their place of residence and work. Therefore, we can conclude that the Republic of Croatia is among the countries that recognized the importance of digital nomads in time and arranged the legislative framework for them, which will influence the attraction of digital nomads in the Republic of Croatia, and consequently have a positive impact on tourism, i.e. on all branches of service activities related to tourism [11].

The Republic of Croatia by regulating the status of digital nomads has become a whole-year destination and not only a temporary tourist summer month jaunt. Research into the NomadList.com portal, a platform that gathers digital nomads from around the world and records their experiences bears witness to the fact that the Republic of Croatia has become the top destination for distance working, and states that monthly between three to five thousand digital nomads come to the Republic of Croatia. A great number of them come from third countries and want to stay and work in the Republic of Croatia longer than 90 days [12].

From the above description, we can conclude that the rules of the Republic of Croatia (both status-legal and tax-legal) are the most suitable for digital nomads of all the analyzed countries.

IV. CONCLUSION

Digital nomads are a new type of tourists who travel and earn money for their travels at the same time. Additionally, they spend their money where they are living, pay value-added tax (VAT), pay for hotels, accommodation, and food, and if they have a family, they pay for kindergartens and schooling, they promote tourism which means that the financial benefits digital nomads bring are great not only for the economy of the Republic of Croatia but also of other states.

The Republic of Croatia has recognized the importance and role of digital nomads in contemporary society, in particular as it is known that the Republic of Croatia is one of the states where tourism is one of the most significant branches of the economy.

Therefore, the Republic of Croatia has brought in regulations that on the one hand simplify or remove the administrative barriers for digital nomads, and on the other, digital nomads' earnings are not classified as income and hence are not taxed. Digital nomads' money earned in FR Germany and the Republic of Italy is taxable. In the Republic of Estonia, digital nomads are exempt from paying tax on income earned provided that have a digital nomad visa for less than 183 days. However, after 183 consecutive days, they will be taxed because they are then considered to be a tax resident of the Republic of Estonia.

The Republic of Croatia, other European Union member states and third countries wanting to attract more and more digital nomads have already regulated the legal status of digital nomads from administrative and taxation aspects, in particular concerning visas, stays, taxation benefits, and tax exemption. On the other hand, digital nomads can live and work in countries that are not European Union member states and are ideal for digital nomads living or working because they can register their company or offshore business, move their property, and so on, which could lead to legal misuse. In such cases, the Republic of Croatia and other European Union member states must always check whether an avoiding double taxation agreement has been formed with those countries so that abuse of laws can be eliminated.

REFERENCES

[1] K. Pandža, Digitalna transformacija na primjeru pružanja javnih usluga – završni rad poslijediplomski specijalistički studij „Organizacija i menagement“, Ekonomski fakultet Sveučilišta J. J. Strossmayera u Osijeku, 2022, p. 10.

[2] S. Golubić, Digitalizacija jedna od temeljnih sastavnica modela razvoja otvorenog informacijskog društva, 2018, http://www.infotrend.hr/clanak/2016/7/digitalizacija-jedna-od-temeljnih-sastavnica-modelarazvojaotvorenog-informacijskog-drustva,88,1251.html, (Accesed 18. 01. 2023).

[3] T. Šímová, A research framework for digital nomadism: a bibliometric study, World Leisure Journal, 2022, pp. 1-17, https://www.tandfonline.com/doi/epdf/10.1080/16078055.2022.2134200?needAccess=true&role=button, (Accesed 18. 01. 2023).

[4] S. Gadžo, Croatia: A New (Tax-Free) Promised Land for Digital Nomads? (Part I), Kluwer International Tax Blog, http://kluwertaxblog.com/2022/02/24/croatia-a-new-tax-free-promised-land-for-digital-nomads-part-i/,(Accesed 19. 01. 2023.).

[5] M. Eddy, Cleaving to the Medieval, Journeymen Ply Their Trades in Europe, The New York Times, 2017, https://www.nytimes.com/2017/08/07/world/europe/europe-journeymen.html, (Accesed 23. 01. 2023).

[6] T. Pancirov, Digitalno nomadstvo je lokacijski neovisan rad, https://www.poslovnifm.com/podcast-emisija/mjesecne-emisije/misliste-mjesecne-emisije/pancirov-digitalno-nomadstvo-je-lokacijski-neovisan-rad/, (Accesed 21. 01. 2023).

[7] O. Hannonen, In search of a digital nomad: defning the phenomenon, Springer, September 2020 (pdf), file:///C:/Users/korisnik/Downloads/In_search_of_a_digital_nomad_defining_the_phenomen.pdf, (Accesed 22. 01.2023).

[8] S. Hensellek and N. Puchala, The Emergence of the Digital Nomad: A Review and Analysis of the Opportunities and Risks of Digital Nomadism. In M. Orel, O. Dvouletý and V. Ratten (Eds.), The Flexible Workplace, Springer International Publishing, 2021, pp. 195-214.

[9] E. Demaj, A. Hasimja and A. Rahimi, Digital Nomadism as a New Flexible Working Approach: Making Tirana the Next European

Hotspot for Digital Nomads. In M. Orel O. Dvouletý and V. Ratten (Eds.), The Flexible Workplace, Springer International Publishing, 2021, pp. 231-257.

[10] Statistika i trendovi za digitalnih nomada u 2023. godini (passport-photo.online), https://passport-photo.online/hr/blog/digitalni-nomadi-statisticki-podaci-i-trendovi/, (Accesed 23. 01.2023).

[11] S. Pezo, Digitalni nomadi i porezni tretman njihovih primitaka u RH, TEB poslovno savjetovanje, Financije, pravo i porezi 4/21, Zagreb, travanj 2021, pp. 25-27.

[12] J. de Jong, „Digitalni nomadi mogu zaustaviti odljev mozgova“, 2022, https://www.nacional.hr/jan-de-jong-digitalni-nomadi-mogu-zaustaviti-odljev-mozgova/, (Accesed 26. 01. 2023).

[13] J. de Jong, Postali smo žarište: Samo lani u Hrvatsku je ušlo 3000 digitalnih nomada; hvale brzi internet, a mjesečno troše i po 10.000 kuna, 2022, https://www.tportal.hr/biznis/clanak/postali-smo-zariste-samo-lani-u-hrvatsku-je-uslo-3000-digitalnih-nomada-hvale-brzi-internet-a-mjesecno-trose-i-po-10-000-kuna-20220604, (Accesed 26. 01. 2023).

[14] https://www.schengenvisainfo.com/digital-nomad-visa/germany/, (Accesed 23. 01.2023).

[15] https://digitalnomads.world/tips/digital-nomad-germany/, (Accesed 23. 01.2023).

[16] https://digitalitinerant.com/europe/germany-digital-nomad-visa-deep-dive/, (Accesed 23. 01.2023).

[17] https://www.globalcitizensolutions.com/italy-digital-nomad-visa/, (Accesed 24. 01.2023).

[18] https://www.schengenvisainfo.com/digital-nomad-visa/estonia/, (Accesed 25. 01.2023).

[19] I. Vidas, Digitalni nomadi i „poslovne“ vize kroz zakonodavni okvir RH, https://www.iusinfo.hr/aktualno/u-sredistu/45843, (Accesed 23. 01.2023).

[20] Zakon o strancima, Narodne novine br. 133/20, 114/22, 151/22 (eng. Foreigners Act).

[21] Uredba o načinu izračuna i visini sredstava za uzdržavanje državljanina treće zemlje u Republici Hrvatskoj Narodne novine br. 14/21 (eng. Regulation on the method of calculation and the amount of funds for the maintenance of third-country citizens in the Republic of Croatia).

[22] S. Pezo, Porezni tretman primitaka digitalnih nomada, Računovodsvo, revizija i financije, Zagreb, svibanj 2021, pp.106-107.

[23] Zakon o porezu na dohodak, Narodne novine br. 115/16, 106/18, 121/19, 32/20, 138/20, 151/22 (eng. Income Tax Act).

[24] Zakon o obveznom zdravstvenom osiguranju i zdravstvenoj zaštiti stranaca u Republici Hrvatskoj Narodne novine br. 80/13, 15/18, 26/21, 46/22 (eng. Compulsory Health Insurance and Health Protection of Foreigners in the Republic of Croatia Act).

[25] Zakon o dopuni Zakona o obveznom zdravstvenom osiguranju i zdravstvenoj zaštiti stranaca u Republici Hrvatskoj Narodne novine br. 26/21 (eng. Act on Amendments to the Compulsory Health Insurance and Health Care for Foreigners in the Republic of Croatia Act).

Comparative Analysis of Machine Learning Algorithms on Data Sets of Different Characteristics for Digital Transformation

D. Oreški*, I. Pihir* and D. Višnjić*
* University of Zagreb / Faculty of Organization and Informatics, Varaždin, Croatia
{dijana.oreski, igor.pihir, dunja.visnjic}@foi.unizg.hr

Abstract **- The application scenarios for machine learning algorithms are getting more complicated as machine learning and real-world situations converge more and more. All fields of study have adopted and benefit from diverse machine learning algorithms implementation. The challenge is to determine which algorithm is best suited to solve a given problem. This problem is especially challenging in social sciences. To tackle that issue, this paper explores a group of machine learning algorithms used for predictive models' development in social science domains of business and education. Several machine learning algorithms are applied here (algorithms of artificial neural networks, k-nearest neighbors, decision tree) along with characteristics of datasets measured by meta-features. In the empirical part of the research, algorithms are compared on the data sets using standard predictive model evaluation metrics. Data sets are extracted from the education and business domain. Research results provide insights into machine learning algorithms' performance depending on their meta-features. Meta-features are significant predictors of machine learning algorithms' performance in both education and business domain. Machine learning-based predictive models developed in this paper are a step forward to the acceleration of digital transformation in the education and business sector.**

Keywords **- comparative analysis; machine learning algorithms; meta-features; digital transformation.**

I. INTRODUCTION

Real-world situations today include digital technology usage on daily basis. In such scenarios, huge amounts of data are generated. Various machine learning algorithms are developed in order to extract knowledge from the data [1]. Every academic discipline, including social sciences, has embraced machine learning and benefited from its use. However, the application of machine learning algorithms is complex and time-consuming. The difficult part is choosing the algorithm that will solve a given problem [2]. The social sciences are particularly challenged by this issue since machine learning for quick and on-demand data analysis, could be a problem or an opportunity to produce new possibilities for new business value and possible digital transformation.

Digital Transformation (DT) is a contemporary paradigm, still growing and emerging in a large number of research papers, used and practiced in business development as well as in research with great interest worldwide [3]. Governments and private companies go for DT to improve business processes, digitally transform their operations, redefine business models and enhance products and services. DT is strategy-oriented and customer-centric, it is based on the introduction of new informational and communicational technology and organizational changes in business models, business processes, and/or products/services [4]. Technological and business changes, as a result of digital transformation, are introduced in new business models. Available data has become the core of interactions in digital business due to its unlimited amount. The digital world has shaped a culture of interaction based on data, and digital transformation leaders are expected to create innovative products and services, as well as increase data readiness to create and share information and deliver seamless digital services [5]. Digital transformation use/reuse technology-related concepts [6] such as Big data, Data analytics, Machine Learning, Artificial Intelligence, Data-Based Decision Making, and Knowledge Management [7], [8], [9] to open possibilities and accelerate digital transformation in business as well as in education and academia.

This paper is structured as follows. Section II overviews related work in the field. Section III describes two datasets used in the empirical research and gives a brief overview of two machine learning algorithms. Section IV provides insights into the characteristics of data measured by meta-features and compares the results of predictive models developed by different machine learning algorithms. Section V concludes the paper by providing guidelines for machine learning algorithms selection in the social science domain and results interpretation in the context of education and business digital transformation.

II. RELATED WORK

Knowledge of artificial intelligence, especially machine learning, is required to process the increasing amount of data that is becoming more accessible in today's digital

environment. A large number of algorithms have been developed in this area that can be used to develop predictive and descriptive models in different domains.

However, since the developed algorithms have different effects on different data, it is necessary to determine which algorithm is best to use in a particular situation, domain, and data set. According to the "no free lunch" theorem there is no best algorithm for all data [10]. Choosing the best algorithm from a large set of available algorithms is a challenging and time-consuming task known in the literature as the algorithm selection problem [11]. Conventional approaches such as trial and error, theoretical analysis, or expert knowledge, have several drawbacks, including high computational costs, difficulty in finding experts, and experts' tendency for personal prejudices and preferences [12], [13]. Considering the shortcomings of traditional approaches to algorithm selection, recent research has focused on automatic algorithm selection [14]. One prominent approach to automated algorithm selection is meta-learning-based algorithm recommendation.

Meta-learning, in the context of machine learning, is the process of learning from previous experience gained by applying different learning algorithms to data sets with different characteristics [15]. There have been numerous successful applications of the meta-learning process in algorithm selection. For instance, Sivakumar et al. [16] compare algorithms in the medical domain for the early detection of cancer. They conduct a comparative analysis of classification algorithms, and the research results in the proposal of a classification algorithm for the medical domain. In the medical domain, meta-learning has also been used to predict covid-19 status based on chest CT scans [17], and [18] proposes MetaPred, a meta-learning framework for predicting clinical risk (e.g. hospital mortality, re-admission to the hospital, etc.). Furthermore, Garcia-Saiz and Zorrilla [19] compare regression algorithms in the educational domain, with the goal of predicting student performance in e-learning courses. Similarly, Romero et al. [20] use meta-learning to recommend a subset of classification algorithms based on Moodle data.

Characteristics extracted from data sets, known as meta-features, play a significant role in the successful application of meta-learning since they can influence the performance of the considered algorithms [21]. They describe different types of data properties that can be used to predict the performance of machine learning algorithms [22], [23]. Prior studies in the domain of classification methods suggest that the characteristics of data sets have a significant effect on the performance of algorithms and demonstrate that the selection of the "best" algorithm depends on the properties of a given data set [24], [25[, [26]. Recent research has not earned much attention to these issues. Kiang suggests that the characteristics of the data set have a significant impact on the performance of classification algorithms [27]. Similarly, Smith's research on the problem of choosing a classification algorithm confirms that it is necessary to understand the characteristics of the data set and that they should be related to the classification algorithm's performance [28].

Many empirical studies [21], [22], [29], [30], [31] proposed different sets of meta-features, but there is no ultimate list that uniformly describes, organizes, and calculates them. Rivolli et al. [30] recently systematized and standardized meta-features into six groups: 1) general; 2) statistical; 3) information-theoretical; 4) model-based measures; 5) landmark measures; and 6) others. This is the most comprehensive list to date.

In this research, we are investigating general and information-theoretical meta-features on datasets from the social sciences domain. Some examples of such usage of different machine learning algorithms and large or open data are explored in various business domains such as agriculture [32], [33], possible use at large and open data in traffic [34], law and government data [35], and used in an education setting [36], [37].

III. DISCUSSION

This section describes two datasets used in the research and introduces two machine learning algorithms applied to the data.

A. Data description and data understanding

Datasets are downloaded from the publicly available repository Kaggle.

The first dataset contains information about research that was conducted in 2008 in two Portuguese schools. The research was conducted for two courses: mathematics and Portuguese language. In this paper, the selected research was on mathematics courses. The number of male and female respondents is approximately equal with 53% female and 47% male. This dataset consists of 395 instances and 33 variables. This dataset is from the education domain and refers to student grade prediction [38].

The second dataset is from the business domain and refers to house sales prediction [39]. The second dataset contains 21 attributes and 21613 instances. Attributes are mostly of numerical type with two categorical ones - date and waterfront. Distributions are of different types, from exponential to multimodal. The most common data distribution is exponential. There is no missing data in the set.

Data preparation consisted of detecting missing values and/or incorrect data. Data understanding included variables description and correlation analysis of numerical variables.

Correlations are performed in the data understanding phase. Results of correlation analysis for the first dataset are presented in the table below. Relations between dependant variable and independent variables are presented.

TABLE 1. CORRELATION BETWEEN INDEPENDENT VARIABLES AND DEPENDENT VARIABLE IN THE EDUCATIONAL DATASET

Variable	by Variable	Correlation	Significance
age	G3	-0.161	p < 0.05
medu	G3	0.217	p < 0.05
fedu	G3	0.152	p < 0.05
Travel time	G3	-0.117	p < 0.05
Study time	G3	0.097	p > 0.05
failures	G3	-0.364	p < 0.05
famrel	G3	0.051	p > 0.05
Free time	G3	0.011	p > 0.05
gocout	G3	-0.132	p < 0.05
Dalc	G3	-0.054	p > 0.05
Walc	G3	-0.051	p > 0.05
health	G3	-0.061	p > 0.05
abscences	G3	0.034	p > 0.05

Most of the correlations are statistically significant. However, there are no high correlations. Table 2 demonstrates a correlation between independent variables and the dependent variable, price, in the case of the business dataset.

TABLE 2. CORRELATION BETWEEN INDEPENDENT VARIABLES AND DEPENDANT VARIABLE IN BUSINESS DATASET

Variable	by Variable	Correlation	Significance
bathrooms	price	0.525134	p > 0.05
bedrooms	price	0.308338	p > 0.05
condition	price	0.036392	p < 0.05
floors	price	0.256786	p > 0.05
grade	price	0.667463	p > 0.05
lat	price	0.306919	p > 0.05
long	price	0.021571	p < 0.05
sqft_above	price	0.605566	p > 0.05
sqft_basement	price	0.323837	p > 0.05
sqft_living	price	0.702044	p > 0.05
sqft_living15	price	0.585374	p > 0.05
sqft_lot	price	0.089655	p < 0.05
sqft_lot15	price	0.082456	p < 0.05
view	price	0.397346	p > 0.05
waterfront	price	0.266331	p > 0.05
yr_built	price	0.053982	p < 0.05
yr_renovated	price	0.126442	p < 0.05
zipcode	price	-0.05317	p < 0.05

Correlation analysis results can determine how the variables are related and that an increase in the construction quality index can cause a price increase, i.e. higher-quality built and designed properties have a higher price than properties of lower construction quality.
In the case of the business dataset, most of the correlations are not statistically significant. Differences in correlation analysis result potentially indicate differences in data characteristics between the two domains of social sciences.

Both datasets are further characterized by meta-features. Simple and information-theoretic meta-features were extracted from the data. Table 3. provides an overview of meta-feature values for those two categories.

TABLE 3. OVERVIEW OF META-FEATURE VALUES

Meta-feature	Educational dataset	Business dataset
General meta-features		
nr_attr	31	19
nr_bin	0	1
nr_inst	395	21613
nr_cat	1	1
nr_num	30	18
attr_to_inst	0.078	0.0008
inst_to_attr	12.742	1137.526
cat_to_num	0.033	0.055
num_to_cat	30	18
nr_outliers	395	386
Information theoretic		
attr_conc.mean	0.456	0.028
attr_conc.sd	0.138	0.056
attr_ent.mean	0.278	3.356
attr_ent.sd	1.549	2.122

The educational dataset has higher dimensionality than the business dataset. The business data set has a higher proportion of numerical attributes in the data. The number of instances is significantly higher in the business dataset. The ratio of outliers is higher in the educational dataset.
Average values of information-theoretic measures show that the educational dataset is more homogenous (measured by attribute concentration), whereas the business dataset has the higher entropy.

B. Machine learning algorithms

The first machine learning algorithm method used in this paper is the decision tree algorithm. A decision tree is a series of nodes connected by branches that extend downward from the root node until they end in a leaf node. The root node is located at the very top of the decision tree graph, and after it, the attributes are tested according to which branches are created by finding new results. Each branch leads to a new decision node or ends with a leaf node.

Decision tree characteristics are:

(i) The decision tree algorithm falls under supervised learning and as such requires the class of the variable being predicted. A training dataset must be applied that provides the algorithm with the values of the variable to be predicted,

(ii) The training data set must have a lot of variation in the result and have a large number of instances, which provides the decision tree algorithm with a large number of branches to classify the data that is added later. A decision tree learns based on examples, if missing examples for a data set, classification, and prediction will be poor or impossible for that data set,

(iii) The attribute to be predicted must be precise, which means that the values of that attribute must be precisely defined so that it is easy to recognize whether an instance belongs to a certain set or not.

The two most used algorithms for decision tree development are the classification and regression trees algorithm (CART) and C4.5 algorithm. Decision trees created employing CART algorithm are binary, which means that each decision has only two branches. The C4.5 algorithm has an advantage over the CART algorithm because it is not limited to binary branches. This algorithm creates decision trees with different branching layouts [40].

Deep learning has become one of the most used machine learning approaches. Artificial neural networks are first representative of the deep learning approach. Neural networks work on the principle of learning from distributed data. The simplest neural network consists of: an input layer, one hidden layer and an output layer.

The process of training a neural network involves the following steps:

(i) defining the structure or architecture of the neural network. This is a very important step, if we create a very extensive network with a large number of neurons then our model will not generalize the data very well.
(ii) select a nonlinear transformation to apply to each link. This transformation controls the efficiency of each neuron in the network.
(iii) decide on the loss function we will use for the output layer. This is true if we have a problem that uses supervised learning,
(iv) to learn the parameters of the neural network, that is, determine the weight values of each connection. The weight values are determined by optimizing the loss function. [41]

IV. RESEARCH RESULTS

The dataset is split into two train and test data, where 70% of the data goes under variables for training, and 30% under variables for testing. In the data preparation phase, outlier detection was performed. In order to perform a neural network model, data are normalized so that only values between 0 and 1 are found in the variables.

A. Predictive models on a business dataset

Optimization of hyperparameters was performed for both machine learning algorithms. Artificial neural network parameters are presented in table 4.

TABLE 4. ARTIFICIAL NEURAL NETWORK PARAMETERS

Artificial neural network parameter	Value
Number of hidden layers	One
Number of hidden nodes	10
Activation function	Sigmoid
Learning rate	0.1

Pruning was performed in decision tree model development to optimize the mode. The smart pruning approach yielded the best model. Models' performance results in terms of error rate are presented in table 5.

TABLE 5. MODELS PERFORMANCE ON THE BUSINESS DATASET

Machine learning algorithm	Error rate
Decision tree	7,55 %
Artificial neural networks	1,6 %

A comparison of decision tree and neural network models indicates the following factors: construction quality index and square footage have proven to be the factors that have the highest impact on price formation according to both models. Furthermore, according to the decision tree model, the third most important factor includes latitude, whereas the year of construction stands out in the neural network model. By adjusting the values of the attributes, significant changes in the price are observed, and thus we can conclude that there is a positive connection between the price of the property, the square footage of the interior space and the quality of construction. A positive correlation between the price of real estate and the number of rooms was also confirmed. Predictive models have shown that the number of rooms has much less influence on the price than expected.

B. Predictive models on an educational dataset

Optimization of hyperparameters was performed for the educational dataset.

TABLE 6. ARTIFICIAL NEURAL NETWORK PARAMETERS

Artificial neural network parameter	Value
Number of hidden layers	One
Number of hidden nodes	15
Activation function	Sigmoid
Learning rate	0.1

Pruning was performed in decision tree model development to optimize the mode. The smart pruning approach yielded the best model. Models' performance results in terms of error rate are presented in the table 7.

TABLE 7. MODELS PERFORMANCE ON EDUCATIONAL DATASET

Machine learning algorithm	Error rate
Decision tree	5,62 %
Artificial neural networks	7,34 %

Based on the model evaluation, we can conclude that there are differences in the performances of different machine learning algorithms on datasets from different domains. In order to establish a connection between algorithms performances and data characteristics, meta-features are examined.

C. Meta-features based meta-model

According to our sample, using data of similar characteristics, an artificial neural network as a machine algorithm should be recommended to do data analysis on a dataset of business domains and a decision tree algorithm in the case of the educational dataset. The educational dataset has higher values for information-

centered meta-feature of attribute concentration. Such measure come from the information-theory field and try to capture the amount of information in the data. The same approach in learning has a decision tree algorithm, which is an information-based machine learning algorithm. Since decision tree-based predictive model yielded more accurate predictions in the case of the educational dataset, inherited information in data provided good basis.

Explanation of the results serves as guidelines for machine learning algorithm selection based on the dataset characteristics measured by meta-features.

V. CONCLUSION

This research examined characteristics of data in education and business. Data characteristics were measured by meta-features. Two machine learning algorithms were applied on the datasets to develop predictive model. Model evaluation led to a conclusion regarding machine learning algorithm selection based on the meta-features. Explanation of the results serves as scientific contributions and guidelines for machine learning algorithm selection based on the dataset characteristics measured by meta-features in favor of the application and research of digital transformation in education and the business world. In the furture research, sample of datasets will be increased, as well as number of extracted meta-features and number of employed machine learning algorithms.

ACKNOWLEDGMENT

This work has been fully supported by Croatian Science Foundation under the project UIP-2020-02-6312.

REFERENCES

[1] Y. Peng, G. Wang, G. Kou, and Y. Shi, (2011). An empirical study of classification algorithm evaluation for financial risk prediction. Applied Soft Computing, 11(2), 2906-2915.

[2] D. Oreski, (2022). Framework of Intelligent System for Machine Learning Algorithm Selection in Social Sciences. J. Softw., 17(1), 21-28.

[3] I. Pihir, K. Tomičić-Pupek, and M. T. Furjan, "Digital Transformation Insights and Trends," *29th CECIIS*, pp. 141–149, 2018.

[4] I. Pihir, "DIGITAL TRANSFORMATION AS UNIVERSITY COURSE - DEVELOPMENT AND IMPLEMENTATION," in *EDULEARN22 Proceedings*, 2022, pp. 7725–7729. doi: 10.21125/edulearn.2022.1799.

[5] L. Hrustek, M. T. Furjan, and I. Pihir, "Enabling Open Data Paradigm for Business Improvement," *56 th Int. Sci. Conf. Econ. Soc. Dev.*, no. May 2021, pp. 174–183, 2020.

[6] Tomičić Furjan, M., Tomičić-Pupek, K., Pihir, I. "Understanding Digital Transformation Initiatives: Case Studies Analysis", Business Systems Research, Vol. 11, No. 1, pp. 125-141, 2020, doi: 10.2478/bsrj-2020-0009

[7] Schwab, K. "The fourth industrial revolution", UK: Portfolio penguin, 2017.

[8] Roedder, N., Dauer, D., Laubis, K., Karaenke, P., Weinhardt, C. (2016), "The digital transformation and smart data analytics: An overview of enabling developments and application areas", IEEE International Conference on Big Data, IEEE, Washington, DC, USA, pp. 2795-2802, doi: 10.1109/BigData.2016.7840927

[9] Gartner inc., "Hype cycle research methodology," gartner.com, 2018, [Online]. Available: https://www.gartner.com/en/research/methodologies/gartner-hype-cycle. [Accessed: Jan. 23, 2023]

[10] D. H. Wolpert, "The Supervised Learning No-Free-Lunch Theorems," in *Soft Computing and Industry*, London: Springer London, 2002, pp. 25–42. doi: 10.1007/978-1-4471-0123-9_3.

[11] R. Ali, A. M. Khatak, F. Chow, and S. Lee, "A case-based meta-learning and reasoning framework for classifiers selection," *ACM Int. Conf. Proceeding Ser.*, pp. 1–6, 2018, doi: 10.1145/3164541.3164601.

[12] G. Wang, Q. Song, and X. Zhu, "An improved data characterization method and its application in classification algorithm recommendation," *Applied Intelligence*, vol. 43, no. 4. pp. 892–912, 2015. doi: 10.1007/s10489-015-0689-3.

[13] R. Ali, S. Lee, and T. C. Chung, "Accurate multi-criteria decision making methodology for recommending machine learning algorithm," *Expert Syst. Appl.*, vol. 71, pp. 257–278, 2017, doi: 10.1016/j.eswa.2016.11.034.

[14] B. M. Franklin, "Automatic Selection of MapReduce Machine Learning Algorithms : A Model Building Approach," p. 237, 2018, [Online]. Available: http://digitalcommons.mtu.edu/etdr/604/

[15] P. Brazdil, C. Giraud-Carrier, C. Soares, and R. Vilalta, *Metalearning*. Cognitive Technologies, 2009. doi: 10.1007/978-3-540-73263-1_3.

[16] S. Sivakumar, S. R. Nayak, S. Vidyanandini, J. A. Kumar, and G. Palai, "An empirical study of supervised learning methods for breast cancer diseases," *Optik (Stuttg).*, vol. 175, pp. 105–114, 2018, doi: 10.1016/j.ijleo.2018.08.112.

[17] E. Güldoğan, İs. O. Yildirim, S. Sevgi, and C. Çolak, "Artificial Intelligence-Assisted Prediction of Covid-19 Status Based on Thorax Ct Scans Using a Proposed Meta-Learning Strategy," *Acta Medica Mediterr.*, vol. 38, no. 3, pp. 1515–1521, 2022, doi: 10.19193/0393-6384_2022_3_228.

[18] X. S. Zhang, F. Tang, H. H. Dodge, J. Zhou, and F. Wang, "MetaPred," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul. 2019, pp. 2487–2495. doi: 10.1145/3292500.3330779.

[19] D. García-Saiz and M. Zorrilla, "A meta-learning based framework for building algorithm recommenders: An application for educational arena," *J. Intell. Fuzzy Syst.*, vol. 32, no. 2, pp. 1449–1459, 2017, doi: 10.3233/JIFS-169141.

[20] C. Romero, J. L. Olmo, and S. Ventura, "A meta-learning

approach for recommending a subset of white-box classification algorithms for Moodle datasets," *Proc. 6th Int. Conf. Educ. Data Mining, EDM 2013*, 2013.

[21] B. Bilalli, A. Abelló, and T. Aluja-Banet, "On the predictive power of meta-features in OpenML," *Int. J. Appl. Math. Comput. Sci.*, vol. 27, no. 4, pp. 697–712, 2017, doi: 10.1515/amcs-2017-0048.

[22] C. Castiello, G. Castellano, and A. M. Fanelli, "Meta-data: Characterization of Input Features for Meta-learning," 2005, pp. 457–468. doi: 10.1007/11526018_45.

[23] I. Khan, X. Zhang, M. Rehman, and R. Ali, "A Literature Survey and Empirical Study of Meta-Learning for Classifier Selection," *IEEE Access*, vol. 8, pp. 10262–10281, 2020, doi: 10.1109/ACCESS.2020.2964726.

[24] N. Dessì and B. Pes, "Similarity of feature selection methods: An empirical study across data intensive classification tasks," *Expert Syst. Appl.*, vol. 42, no. 10, pp. 4632–4642, 2015, doi: 10.1016/j.eswa.2015.01.069.

[25] C. Chen and M. L. Shyu, "Clustering-based binary-class classification for imbalanced data sets," *Proc. 2011 IEEE Int. Conf. Inf. Reuse Integr. IRI 2011*, pp. 384–389, 2011, doi: 10.1109/IRI.2011.6009578.

[26] O. Kwon and J. M. Sim, "Effects of data set features on the performances of classification algorithms," *Expert Syst. Appl.*, vol. 40, no. 5, pp. 1847–1857, 2013, doi: 10.1016/j.eswa.2012.09.017.

[27] M. Y. Kiang, "A comparative assessment of classification methods," *Decis. Support Syst.*, vol. 35, no. 4, pp. 441–454, 2003, doi: 10.1016/S0167-9236(02)00110-0.

[28] S. Ali and K. A. Smith, "On learning algorithm selection for classification," *Appl. Soft Comput. J.*, vol. 6, no. 2, pp. 119–138, 2006, doi: 10.1016/j.asoc.2004.12.002.

[29] A. Filchenkov and A. Pendryak, "Datasets meta-feature description for recommending feature selection algorithm," *Proc. Artif. Intell. Nat. Lang. Inf. Extr. Soc. Media Web Search Fruct Conf. AINL-ISMW Fruct 2015*, no. November 2015, pp. 11–18, 2016, doi: 10.1109/AINL-ISMW-FRUCT.2015.7382962.

[30] A. Rivolli, L. P. F. Garcia, C. Soares, J. Vanschoren, and A. C. P. L. F. de Carvalho, "Meta-features for meta-learning," *Knowledge-Based Syst.*, vol. 240, p. 108101, 2022, doi: 10.1016/j.knosys.2021.108101.

[31] Y. Peng, P. A. Flach, C. Soares, and P. Brazdil, "Improved dataset characterisation for meta-learning," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2534, pp. 779–784, 2002, doi: 10.1007/3-540-36182-0_14.

[32] D. Oreški, I. Pihir, and K. Cajzek, "Smart Agriculture and Digital Transformation on Case of Intelligent System for Wine Quality Prediction", MIPRO 2021, 44th International Convention Proceedings, pp 1565-1570, 2021, doi: 10.23919/MIPRO52101.2021.9596979

[33] D. Oreški, I. Pihir, and N. Kadoić, N, "Smart Agriculture: Machine Learning in Modelling Wine Quality Based on Laboratory or IoT Sensory Analysis", Proceedings of 56th Croatian and 16th International symposium on Agriculture / Rozman, Vlatka ; Antunović, Zvonko - Osijek : Faculty of Agrobiotechnical Sciences Osijek, pp 717-722, 2021.

[34] M. Vujic, L. Dedic, M. Tomicic Furjan, and I. Pihir, "The Benefits of Open Data in Urban Traffic Network". In: Knapčíková, L., Peraković, D., Behúnová, A., Periša, M. (eds) 5th EAI International Conference on Management of Manufacturing Systems. EAI/Springer Innovations in Communication and Computing. Springer, Cham, pp 267–282, 2022, doi: 10.1007/978-3-030-67241-6_22

[35] C. Alexopoulos, I. Pihir, and M. Tomičić Furjan, "Automatic End-to-End Decomposition and Semantic Annotation of Laws Using High-Performance- Computing and Open Data as a Potential Driver for Digital Transformation", 33rd Central European Conference on Information and Intelligent Systems, pp 189-194, 2022, doi: http://archive.ceciis.foi.hr/app/public/conferences/2022/Proceedings/DTCC/DTCC4.pdf

[36] D. Oreški, I. Pihir, and M. Konecki, "CRISP-DM process model in educational setting", Li Yongqiang, Anica Hunjet, Ante Roncevic (ed.), Economic and Social Development, 20th International Scientific Conference on Economic and Social Development, Book of Proceedings. pp 19-28, 2017, doi: https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdPrague_2017_Online.pdf

[37] D. Oreški, M. Konecki, and I. Pihir, "Predictive Modelling of Academic Performance by Means of Bayesian Networks", Konecki, M., Kedmenec, I. & Kuruvilla, A. (ed.), 47th International Scientific Conference on Economic and Social Development, Book of Proceedings, pp 435-441, 2019, doi: https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdPrague2019_Online.pdf

[38] Student grade prediction, https://www.kaggle.com/dipam7/student-grade-prediction, downloaded 01/02/2023.

[39] House sales prediction https://www.kaggle.com/datasets/harlfoxem/housesalesprediction, downloaded 01/02/2023.

[40] D. T. Larose and C. D. Larose, *Data Mining and Predictive Analytics*, 2nd Editio. Wiley, 2015. doi: 10.1016/b978-0-12-800229-2.00003-1.

[41] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Python (A Problem-Solver's Guide to Building Real-World Intelligent Systems)*, 1st ed. ed. Apress, 2018. doi: 10.1007/978-1-4842-3207-1_11.

Clustering of Croatian Software Development Companies Based on Their Financial Indicators

Irena Konecki*, Dijana Oreški * and Nikola Kadoić*
* University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, Varaždin, Croatia
irena.konecki@foi.unizg.hr, dijana.oreski@foi.unizg.hr, nikola.kadoic@foi.unizg.hr

***Abstract* - The main objective of this study is to examine the financial performance of companies in the Croatian software development sector and identify which financial indicators could be predictive of future profitability. The research uses the Financial Agency (FINA) data focusing on pre and post-COVID-19. Cluster analysis of companies is performed based on their financial indicators. Four clusters of companies were identified. Findings indicate that there is no statistically significant difference between the clusters in the profitability growth rate after the COVID-19 crisis. However, the observed period might be too short to measure the long-term effects. The financial analysis of companies in the Croatian software development sector gives insight into how the sector operates and thus provides valuable information for policymakers who would like to support its growth and development.**

Keywords* - *financial indicators, software development sector, Croatia, cluster analysis, COVID-19 crisis

I. INTRODUCTION

Having a strong IT sector contributes to economic growth, employment and productivity. To create the support policies that would suit the software industry, it is necessary to know more about its heterogeneity and dynamic features during the expansion and during the economic crisis. In other words, it is beneficial to know whether there are different subsets of companies in the industry with similar goals, strategies, and financial characteristics. For example, it would be wrong to assume that all the companies in the sector would benefit from the loans aimed at growth since probably not all of them want to grow, and of those that do, maybe not all of them have the financial structure that would benefit from the loan. Thus, financial data can be used to identify clusters of companies with similar financial backgrounds to tailor public policies to their specific needs.

In this paper, the Croatian software development sector is in focus. Section 2 overviews the Croatian software development sector and the belonging financial data. Section 3 defines financial indicators and describes their usage. In Section 4, the methodology used is presented. Section 5 contains the main findings of the clustering procedure, while Section 6 concludes the paper.

This paper has been prepared in the scope of the project SIMON: Intelligent system for automatic selection of machine learning algorithms in social sciences. The main goal of the project SIMON is to develop an intelligent system for the automatic selection of machine learning algorithms in the social sciences that perform better on a given data set, taking into account the specific characteristics of the data. In this paper, we focus on the analysis of business data.

II. CROATIAN SOFTWARE DEVELOPMENT SECTOR

Based on criteria of export orientation, employment capacity and earnings before taxes, interest and amortization, the activity of computer programming, consultancy and related activities (J62.0, NKD 2007) may be considered as the driving force of Croatian industry [1]. It is one of the rare sectors in Croatia that has had growth of gross added value and employment in the period of the highest economic recession [2]. Its profitability measured by gross profit margin indicates convergence to the European Union average [1, 3].

One-third of the revenues of Croatian software companies come from two leading software markets (Northern America and Europe), which allows a better inflow of international capital in situations when domestic capital is scarce [15]. Unlike large companies in the sector, small and medium-sized companies have adopted an agile approach to software development as their main competitive advantage in the industry, characterized by fast development and quick-win strategies in discovering and targeting new markets [4].

The sector predominantly consists of micro-enterprises. In 2012 only the companies that were 16 years old or older had, on average, ten employees [5]. The software industry and data processing activity have the smallest minimum efficient size among ICT services [6], which generally increases the survival chances of enterprises in a particular industry [5]. They also have low entry barriers, capital intensity coefficient [6], and industry concentration levels [7].

Previous research confirmed Gibrat's law in the Croatian software supply and consultancy industry, meaning that the growth of companies in this sector does not depend on their size at the start of observation [8]. That indicates there is no tendency for smaller companies to grow faster in this sector compared to larger companies. This is in accordance with the fact that in the sector, besides the companies that cannot achieve growth due to their internal weaknesses (survivalists) or due to strong competition (low-growth companies), there are a lot of lifestyle entrepreneurs who favor those strategies that do not strive for growth [9]. In the Croatian software

development sector, the relationship between net working capital and profitability is concave quadratic, suggesting an optimal level of net working capital concerning the company's total revenue [10], which the management should strive for.

Research shows that entrepreneurs in this sector do not perceive Croatian institutions as supporting their activities, but on the contrary, they emphasize their repressive function executed through penalty instrumentary [2]. Non-flexibility, slackness, passiveness, standardization, and lack of understanding of the entrepreneurial field are the most common features attached to institutions and their employees. At the same time, the political system is seen as full of illegal favoritism [2]. Inconsistent legislation, high tax burden, and cumbersome and inefficient public administration are reasons why Croatia has not positioned itself as a desirable high-tech investment destination [11].

ICT industries are human capital intensive [12], which is reflected in average salary as a proxy for human capital. Because of this, software companies should pay attention to knowledge management. Research shows that there is a positive attitude towards knowledge sharing in Croatian IT companies [13].

The software industry is a part of the high-tech industry in which noticeable improvements can be made even without substantial financial resources. However, the lack of developers and other IT professionals is impeding more remarkable progress, especially since, after Croatian accession to the European Union, a number of Croatian high-tech professionals have decided to leave the country [11]. One of the possible solutions is engaging more women to pursue ICT professions. The popularization of informatics among women within the educational system should potentially increase their self-efficacy in this area and consequently make ICT occupations more attractive to women [14].

In figures 1-5, we present the sector of Croatian companies classified into area J, class 62.01 Computer Programming (2007 National Classification of Activities) through (the data collected from the Info.biz platform [15]):

- the number of entrepreneurs in the sector in the last five years (Figure 1.),
- the number of employees in the sector in the previous five years (Figure 2.),
- the number of entrepreneurs with profit/loss in the sector in the last five years (Figure 3.),
- total income and expenses in the sector in the previous five years (Figure 4.)
- the trade balance in the sector in the last five years (Figure 5.),

The number of entrepreneurs in the sector grows steadily in the observed period. With the growing number of companies, employees increased by approximately 50%.

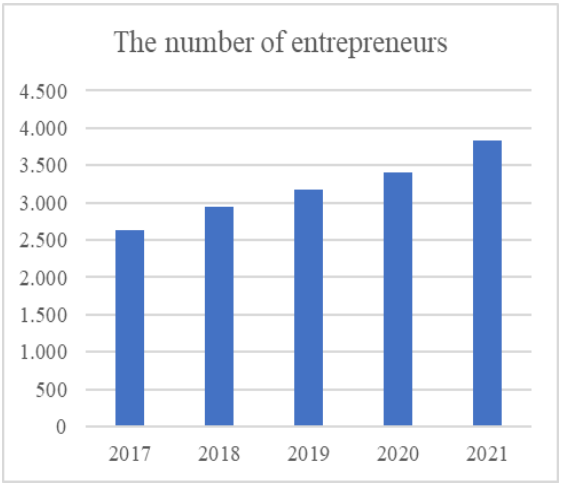


Figure 1. The number of entrepreneurs in the class of Computer Programming in the last five years (authors' work based on [15])

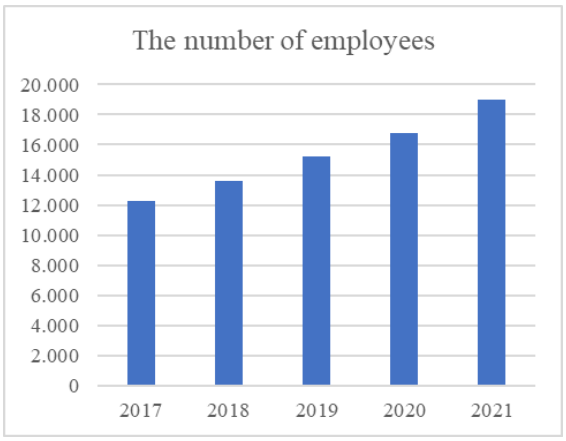


Figure 2. The number of employees in the class of Computer Programming in the last five years (authors' work based on [15])

Most of the companies in the sector make a profit. The difference between total income and expenses increases throughout the observed years, confirming the IT sector's growing profitability and attractiveness.

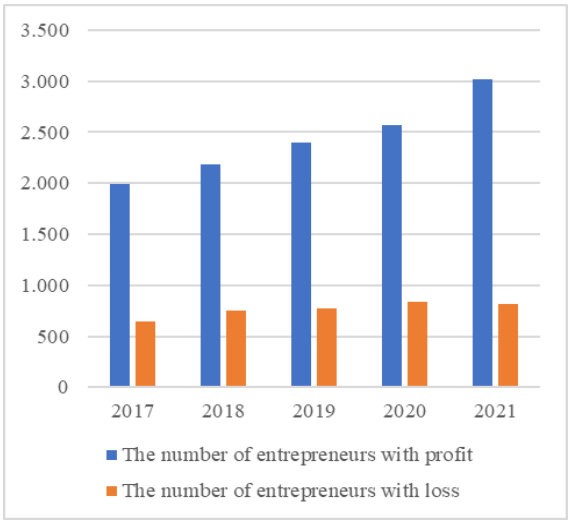


Figure 3. The number of entrepreneurs with profit/loss in the class of Computer Programming in the last five years (authors' work based on [15])

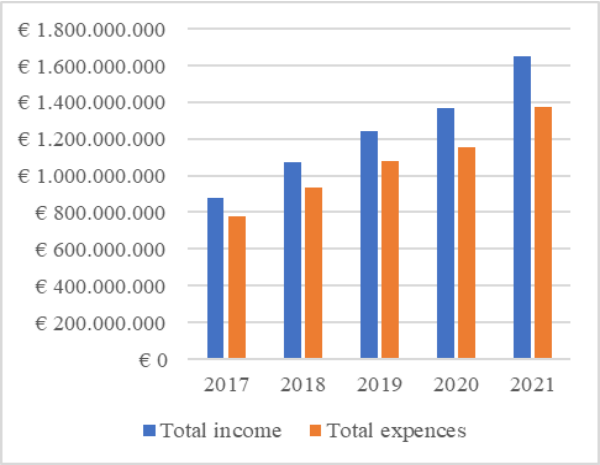


Figure 4. Total income and expenses in the last five years in the sector (authors' work based on [15])

The sector is strongly export-oriented. The trade balance has almost tripled in the observed period.

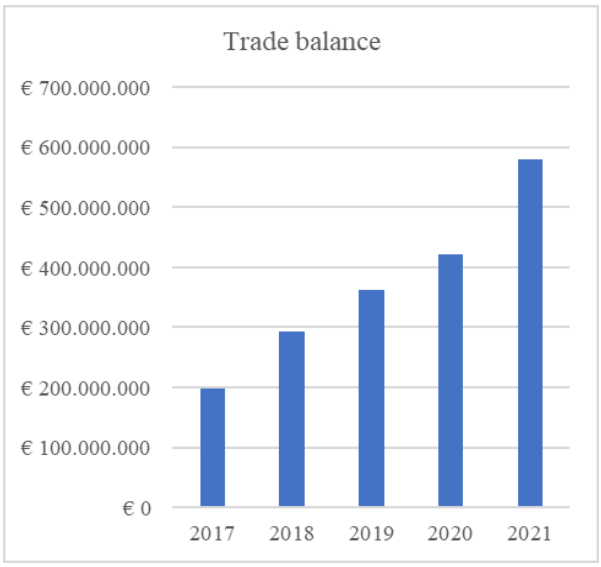


Figure 5. The trade balance in the class of Computer Programming in the last five years (authors' work based on [15])

III. FINANCIAL INDICATORS

Financial statements, mainly balance sheets and profit and loss accounts, are the main source of information when evaluating a company's business activities. Financial statements analysis includes calculating various financial indicators covering several areas of business performance. Such evaluation of business performance is one of the pillars of managerial decision-making. A glance at the financial indicators enables risk assessment and gives information on business security and success. Financial indicators are thus interesting for various stakeholders, internal (owners, managers, employees) and external (partners, creditors, media, etc.).

Financial performance measures are often used because they provide an aggregate view of an organization's performance resulting from its strategic, operating, and tactical management [16]. In that way, financial indicators are "lagging" because they result from former actions [17]. Their objectivity is dependent on the quality of prepared financial statements [18].

Financial indicators are organized into six groups: liquidity indicators, indebtedness indicators, activity indicators, efficiency indicators, profitability indicators, and investment indicators [12].

Business management is assessed based on the security criterion reflected in the liquidity and indebtedness indicators and on the performance criterion for which one looks at efficiency, profitability, and investment indicators. Activity indicators are both security indicators and performance indicators.

Liquidity indicators measure the company's ability to meet its due short-term obligations. They are calculated based on data from the balance sheet. Liquidity indicators include cash ratio, quick ratio, current ratio, and coefficient of financial stability [12].

Indebtedness indicators reflect the structure of liabilities. They show the percentage of assets financed from own capital (principal) and the percentage of the assets funded from external sources (liabilities). Indebtedness indicators are debt-to-assets ratio, equity-to-assets ratio, debt-to-equity ratio, indebtedness factor, level of coverage 1, and level of coverage 2 [19].

Activity ratios measure how efficiently a company performs day-to-day tasks such as collecting receivables and managing inventory [20] and are calculated based on data from the balance sheet and profit and loss account. They include total assets turnover, working capital turnover, receivables turnover, and days of sales outstanding [19].

Efficiency indicators measure the ratio of revenues and expenditures, i.e., they show how much income is generated per unit of expenditure and are calculated based on data from the profit and loss account. They include the efficiency of the overall business, the efficiency of sales, the efficiency of financing, and the efficiency of extraordinary activities [19].

Profitability ratios measure the company's ability to generate profitable sales from its assets [20]. Profitability indicators are net profit margin, gross profit margin, net return on assets (ROA), gross return on assets, and return on equity (ROE) [19]. Investment indicators measure the success of investments in ordinary shares (for joint-stock companies) [19].

Financial indicators have the potential to predict the future financial success of companies. A company's financial success is usually reflected in earnings growth rate before interest, tax, and depreciation (EBITDA). The goal of this paper is to use the cluster analysis based on the financial indicators and to test the following hypothesis:

> *H1: Based on the financial indicators, there is a statistically significant difference in the EBITDA growth rate between the clusters identified.*

IV. METHODOLOGY

The research sample consists of Croatian companies classified into area J, class 62.01 Computer Programming (2007 National Classification of Activities) that existed both in the pre-pandemic year 2019 and 2021. Financial Agency (FINA) database was used for this research [15]. Croatian enterprises have to submit regular financial reports to Financial Agency each year. The companies that disappeared from or entered the market during the observation period were excluded from the initial sample of 4355 companies. The final data set consists of 2069 companies. According to Financial Agency's data, there were three medium companies, 218 small companies, and 1848 micro companies in the sample.

Companies' data included in this research are basic information about companies and quantitative (financial and non-financial) variables, including the percentage of domestic capital, current ratio, debt-to-assets ratio, days of sales outstanding, the efficiency of the overall business, return on assets (ROA), return on equity (ROE), average monthly gross 2 salaries per employee.

The current ratio is calculated as the ratio of short-term assets to short-term liabilities. For the company to be liquid, it should be greater than 2. Debt-to-assets ratio is a percentage of a company's assets financed from external sources. Days of sales outstanding is the average number of days a company takes to receive payment for a sale. The efficiency of the overall business is calculated as the ratio of total income to total expenditure. ROA is the ratio of profit to total assets. ROE is the profit-to-equity ratio [19]. EBITDA growth rate was calculated as a growth rate of earnings before interest, tax, and depreciation from the year 2019 to the year 2021.

The clustering algorithm is applied to group companies. Financial indicators are numerical variables, so a k-means clustering algorithm is applied. The k-means algorithm divides the companies into a certain number of clusters using an iterative algorithm. The k-means algorithm divides data into clusters in order to reduce the distances between cluster centroids. The number of clusters, *k*, must be determined. The outcomes of various *k* values are compared to identify the best number of clusters. The cubic Cluster Criterion (CCC) is applied as a criterion for cluster evaluation. CCC estimates the number of clusters using Ward's minimum variance method, based on minimizing the within-cluster sum of squares. The sum of squared residuals in each clustering approach selects the number of clusters. The clustering model is interpreted by considering mean cluster values for each variable (as algorithms names suggest).

The Levene test and the Welch test were used to determine whether the EBITDA growth rate is statistically significantly different among the companies in different clusters and, consequently, to test the hypothesis.

V. RESEARCH RESULTS

Table I shows the movements of the variables used in the research for the last five available years. The average current ratio of the companies in the sector is above 2 in all observed years, indicating high liquidity. Moreover, liquidity of the sector slightly grows in the observed period. On average, 40% of the assets are financed by debt. Indebtedness has decreased in the last five years. Both average ROA and average ROE are growing from year to year, confirming the sector's attractiveness. What is more, the average EBITDA margin shows the growing efficiency of the sector, while in absolute terms, the total EBITDA of the sector has more than doubled in the last five years. The average monthly gross 2 salary per employee is also growing, but more slowly than profits. It has grown by 25.85% in the last five years. Days of sales outstanding are decreasing in the observed period indicating better business activity.

The k-means algorithm enables clustering by using continuous numerical variables. Five cluster analyses were run separately, specifying two, three, four, five, and six-cluster solutions. The Cubic Cluster Criterion (CCC) was used to determine the optimal number of clusters. When comparing the CCC values for two models, the model with the higher CCC is considered better.

TABLE II CLUSTERS EVALUATION (AUTHORS' WORK)

Number of clusters	CCC	Best
2	-4.62081	
3	-10.837	
4	3.034135	Optimal CCC
5	-9.36535	
6	-0.66089	

The results presented in Table II lead to the selection of four clusters solution. Considering sample sizes in each cluster (table III) and interpretability, the four-cluster solution proved to be the best.

TABLE III CLUSTERS EVALUATION (AUTHORS' WORK)

Cluster	Number of companies
Cluster 1	635
Cluster 2	126
Cluster 3	577
Cluster 4	731

TABLE I THE MOVEMENTS OF THE VARIABLES USED IN THE RESEARCH FOR THE FIVE LAST AVAILABLE YEARS (AUTHORS' WORK)

	2017	**2018**	**2019**	**2020**	**2021**
Current ratio	2,09	2,29	2,21	2,36	2,55
Debt-to-assets ratio	0,43	0,39	0,4	0,38	0,35
ROA	13,30%	15,02%	15,80%	17,64%	19,73%
ROE	25,29%	27,35%	29,40%	31,13%	32,94%
EBITDA margin	16,01	15,83	16,58	18,81	19,90
EBITDA	€ 139.255.585	€ 167.618.066	€ 203.810.154	€ 252.405.263	€ 323.860.871
Average monthly gross 2 salary per employee	€ 1.698	€ 1.858	€ 1.935	€ 2.011	€ 2.137
Days of sales outstanding	64,21	59,71	55,69	55,65	47,74

TABLE IV CLUSTERS` CHARACTERISTICS (AUTHORS' WORK)

Cluster	% of domestic capital	Current ratio	Debt-to-assets ratio	Efficiency of the overall business	ROA	ROE	Average monthly gross 2 salary per employee	Days of sales outstanding
1	99.8837	4.3234	0.4583	114.8768	0.1183	0.2227	935.47	72.3891
2	3.440945	4.3559	0.5148	116.9631	0.2058	0.4509	1996.47	43.4001
3	98.55187	18.342	0.1178	328.7169	0.5741	0.6603	648.24	50.2990
4	99.68144	4.9496	0.2943	155.7054	0.5543	0.8100	1053.44	43.1428

Cluster 4 is the largest cluster comprising most companies, followed by Cluster 1. Cluster 2 is the smallest. Table IV reports on the profiles of the four identified clusters.

Cluster 1 can be characterized as companies of Croatian ownership with the lowest current ratio and the longest days of sales outstanding. This is the cluster with the lowest profitability reflected both in ROA and ROE, with a monthly gross 2 salary per employee at the average level for the sector. Cluster 2 consists of foreign-owned companies with a low current ratio, medium debt-to-asset ratio, and medium profitability. This cluster has the highest average monthly gross 2 salary per employee.

Cluster 3 comprises companies characterized by the highest liquidity and the lowest indebtedness. The efficiency of the overall business is the highest in this cluster. However, this is not reflected in the monthly gross 2 salaries that are below average for the sector. On the other side, ROA and ROE are high. Cluster 4 consists of companies of average liquidity and low indebtedness. This is the cluster with the highest average ROE.

The Levene test results (p=0.000) showed the appropriateness of using the Welch test to test the difference in EBITDA growth rate between the four clusters. The Welch test showed that there is no statistically significant difference in the EBITDA growth rate between the four clusters (p=0.102), rejecting hypothesis H1.

VI. CONCLUSION

Products and services provided by software development companies are important to overall economic growth and development since the software industry has a high level of innovation that spills over to other industries and to research. Strengthening the software industry through institutional support can aid the overall economy since all industries rely on ICT to a greater or lesser degree [4]. It is important to understand the different characteristics of the companies in the sector to create entrepreneurship incentive policies and employment and self-employment policies.

The paper proved that cluster analysis based on financial data is successful in identifying four clusters in the sector of computer programming. This indicates potentially different needs of the companies belonging to different clusters in terms of incentive policies. Future research should complement financial indicators with other quantitative and qualitative variables. It would be interesting to see if the level of internationalization, the level of innovation, and the type of management strategy would improve the clustering model. Also, the more extended observation period of the EBITDA growth rate might yield a different result and should be used in future research.

ACKNOWLEDGMENT

This work has been fully supported by Croatian Science Foundation under the project UIP-2020-02-6312.

REFERENCES

[1] Republika Hrvatska Industrijska strategija Republike Hrvatske 2014. – 2020., Zagreb. Retrieved from http://www.mingo.hr/userdocsimages/industrija/Industrijska_strategija.docx. 2014

[2] T. Šmaguc and K. Vuković. "Institutional field in the Republic of Croatia: Perspective of entrepreneurs pertaining to computer programming activities," In Central European Conference on Information and Intelligent Systems (pp. 145-151). Faculty of Organization and Informatics Varazdin. 2017

[3] Republika Hrvatska. Ekonomska analiza sektora profesionalnih i poslovnih usluga, Zagreb. Retrieved from http://www.cut.hr/resources/publikacije/9/Ekonomska%20analiza%20sektora%20profesionalnih%20i%20poslovnih%20usluga.pdf 2014

[4] N. Vlahović, Lj. Milanović Glavan and A. Franković. "Comparative advantages of software industry in developing countries: Study of structure, market strategies and software development approaches in Croatian software companies," Recent Advances in Computer Science / Xiaodong Zhuang - Zakynthos Island, Greece: CSCC, 2015, 227-233, 19th International Conference on Computers (part of CSCC'15) Zakintos, Grčka, 16.07.2015. - 20.07.2015. 2015

[5] K. Vuković. "Enterprise size distribution in the Croatian software industry," Journal of International Scientific Publications: Economy and Business, 6. 2012

[6] Z. Kovačević and K. Vuković. "Performanse poduzeća u hrvatskom sektoru informacijskokomunikacijske tehnologije (ICT)," Ekonomska misao i praksa, 15(2):217-240, 2006

[7] Ministry of Economy. Industrial Strategy 2014 - 2020, http://www.mingo.hr/default.aspx?id=4980,April 12th 2014

[8] K. Vuković, D. Korent and I. Kedmenec. "Testing Gibrat's Law on software companies in Croatia," In Central European Conference on Information and Intelligent Systems (p. 86). Faculty of Organization and Informatics Varazdin. 2014

[9] Ecorys. FWC Sector competitiveness studies - competitiveness of the EU SMEs in the ICT services industry, http://ec.europa.eu/enterprise/sectors/ict/files/study_report_ict_services_en.pdf, downloaded: March 10th 2014.

[10] D. Korent and S. Orsag. "The impact of working capital management on profitability of Croatian software companies, "Zagreb International Review of Economics & Business, 21(1), 47-65. 2018

[11] G. Dukić, G. Kozina, and D. Dukić. "Croatian high-tech industry: Current state and prospects," Tehnički vjesnik, 26(2), 545-551. 2019

[12] K. Vuković and Z. Kovačević. "Economic impact of information and communications technology - Identifying its restraints and exploring its potential," In M. J. Er (ed.), New Trends in Technologies: Devices, Computer, Communication and Industrial Systems, Sciyo, Rijeka, pp. 119-134, 2010

[13] M. Klačmer Čalopa, J. Horvat and L. Kuzminski. "The role of human resource management practice mediated by knowledge management (Study on companies from ICT sector, Croatia)," TEM Journal, 4 2; 178-186, 2015

[14] T. Šmaguc. "Gender stereotypes in entrepreneurship process: evidence from the Croatian ICT industry," Economic and Social Development: Book of Proceedings, 52-60. 2020

[15] FINA (Financial Agency). "Infobiz platform," https://infobiz.fina.hr/, 2023

[16] R .S. Kaplan and A. A. Atkinson. "Advanced management accounting," (third ed.), 3 Prentice-Hall, USA. 1998

[17] S. Cohen, D. Thiraios and M. Kandilorou. "Performance parameters interrelations from a balanced scorecard perspective: An analysis of Greek companies," Managerial Auditing Journal, Vol. 23 Iss: 5, http://dx.doi.org/10.1108/02686900810875307, 2008.

[18] I. Kotane and I. Kuzmina-Merlino. "Assessment of financial indicators for evaluation of business performance," European integration studies, (6). 2012

[19] K. Žager and L. Žager, L. "Analiza financijskih izvještaja," second edition. MASMEDIA: Zagreb, 2008

[20] T. R. Robinson. "International financial statement analysis," John Wiley & Sons Inc., Hoboken, New Jersey. 2009

Artificial Intelligence in Healthcare Services – Regulation, Implementation and Future Challenges

M. Boban*, M. Klarić**
*University of Split, Faculty of Law, Department of Information Science and Statistics, Split, Croatia
**University of Split, Faculty of Law, Department of Administrative Science, Split, Croatia
marija.boban@pravst.hr

***Abstract* - Digitalization of healthcare services in European Union with implementation of artificial intelligence represents one of the modern forms of healthcare delivery defined in the European digital agenda in order to assure better quality in healthcare services delivery and improve quality of healthcare systems. Authors of the paper analyse the process of digitalization and implementation of artificial intelligence in healthcare service while opening some technical, ethical and legal questions, which need to be better defined. European union is still in process of developing the regulatory framework for implementation and harmonization of digital solutions in healthcare system and the authors give an overview of actual situation and give recommendations concerning the privacy and data protection, data quality and data management as well as harmonized rules on artificial intelligence which should stand as the pillars of the regulatory framework and be the main focus of the future regulation in order to protect the patients data and patients privacy in healthcare services.**

Key words: artificial intelligence, data privacy, data protection, digitalization, European union, healthcare services, personal data protection, regulation

I. INTRODUCTION

Healthcare services are usually part of public services, organized to maintain and manage important social activities in field of public health by public institutions and government bodies in many contemporary states. This approach in providing of healthcare services are usually characteristic for European states, where provision and delivery of healthcare services represents one of the basic human rights, which is usually described as third generation of human rights, which are focused on important social services such as social housing, different types of social insurance and other types of social policy. [1] Organization and functioning of public services is specific, because of significance and importance of those services for daily life of the citizens. [2] Approach to the public services mostly depends on attitude which prevails in social life of community. In some countries can be find opinion that public services are increasingly important for public community, and states create regulatory framework for their provision. In other countries public services have relatively reduced influence on daily activity in society, and their regulation depends on delivery and provision regulated by common framework for performance of all other services in society.

Public services are important part of public administration in contemporary society. They provide many important and additional services, which are important for daily life of citizens in community. They are usually divided to different division: on central (national) and local public services, where central public services deliver type of services important for the state and its activity in society, and local public services, where local government institutions participate in organization and maintaining of this type of services. [3] Public services which are regulated and organized from central government administration and institutions operate for entire state, and their impact is global for national community. Local government services which are regulated and organized from local government institutions and bodies and operate on local level, with limited local influence to fulfil local public needs in community. [4]

Second division is services of general economic and non-economic interests, where services of general economic interests represent important public services with economic or commercial elements, and services of non-economic interests which are also important for community, but without economic character. [5] Division of public services on services of economic and non-economic interests are relatively new and define services which are based on commercial and economic interest, and other which are organized on non-profit principles, with purpose to fulfill some public needs defined as important for community, independently of their costs and organizational complexity. [6] They are usually organized and managed as non-profit institutions financed from the budget, but their activity is important for normal functioning of community. [7]

Criteria for division of public services on commercial or non-commercial services depend on approach to principles of delivery and provisions to the citizens in society. If prevail commercial interest of providers, public services are defined as economic or commercial. [8] If public interest of community is important for availability of some public service, or fulfilling of some public need in society and if those interest or need prevail instead on commercial interest

of possible provider, public services are usually defined as non-commercial. [9]

Healthcare services usually represent important part of public services, with regulatory framework which is similar to the other types of public administration. They can be defining as commercial or non-commercial services, which depends on two different types of interest, commercial or non-commercial, according to the incomes which they realize on the market. [10] If healthcare service is predominantly market oriented, then represents services of general economic interest. If it is non-profit oriented, and main principle of organization and functioning is the principle of intergenerational solidarity, then falls into the category of non-commercial services. Accordingly, health services can be defined as services of general economic interest, when they have organization and delivery based on commercial elements, or services of general interests, when their organization and delivery are primary motivated with fulfilling public needs in public community.

According to the commercial or non-commercial status of operating in community, health services are defined with specific regulatory framework, which includes two main elements: technical aspect of healthcare service providing and social aspect of healthcare service providing. *Technical aspect* defines technical conditions and minimal standards for provision and delivery of healthcare services. Minimal standards and technical conditions for maintaining of healthcare services are usually defined from central government authorities and or professional institutions with public authorities which regulates professional standards for healthcare providing. Government authorities usually prescribe propositions for healthcare services delivery, with minimal technical standards of quality. Additional standards of healthcare services provision quality establish and prescribes professional institutions such as medical chambers or professional medical associations. Those institutions usually have public authorities shared from the central government institutions, to additional regulate delivery of healthcare services.

Social aspect of healthcare service providing has accent on availability of healthcare service. If provision of healthcare services represents social category in society for fulfilling of public needs, then provision of healthcare services is widely accessible. If commercial element prevails in healthcare service organization, then accessibility of healthcare services depends on asset census. Complexity of healthcare services can be described by three different level of public healthcare: primary, secondary, and tertiary. [11] Primary level includes general medical practices, secondary special medical practices with clinical specialists and tertiary hospital healthcare services. General medical practices which operate on the primary level are simplest part of healthcare services, but actually most important because of direct communication with the patients, universal healthcare protection and availability.

In most countries, general medical practice represents main pillar of efficient and affordable healthcare services. [12] Second level of medical practice, with clinical specialists, represents upgrade to general medical practice, and assure efficient and affordable healthcare protection and available healthcare services. Hospital healthcare services are most complex form of healthcare service. This type of protection incorporates various forms of healthcare services with combination of clinical specialists and sub-specialists, which include combination of specialized care and long-term care. Because of complexity, hospital healthcare services are most expensive part of healthcare system in every country and need some type of selection to approach on its services. [13]

II. DIGITAL TRANSFORMATION OF HEALTHCARE SERVICES

Expenses of the health care system are increasing with development of medical technology and population aging. Development of technological solutions, with new diagnostic methods and more complex and comprehensive therapies lead to rising costs of healthcare services. Because of that, contemporary health systems need to use new solutions to reduce costs of the healthcare services and harmonize future development. Digitalization represents one of useful solutions to cut healthcare expenses. European countries are developing healthcare systems with characteristic such as efficiency, high quality, equity, affordability and accessibility. They are trying to balance between those values and their potential conflicts in values and goals. [14]

According to the above written, digitalization represents technological process which facilitates that healthcare services be organized, produced and delivered in new ways. In context of technological changes, digitalization also represents organizational and cultural change. [15] Digitalization of healthcare services includes technical and social component. Technological component describes new technological aspects of digitalization, which effect on various elements of healthcare services, such as structure of healthcare organizations, healthcare culture and functioning of healthcare services, functioning of medical professions, treatments and other effects of healthcare system. [16] Social aspect of digitalization includes social, political and economic impacts of digital transformation, which represents innovation in health technologies and healthcare delivery process. [17] Digital transformation can be describing as a process which introduce to significant changes in in healthcare delivery [18] One of the most important part of digitalization of health services is mobile health service, usually known as mHealth, which includes implementation of digital technologies to maintain health process and conditions in movement. This can be important

in following health conditions of patients such as asthma, heart rate, breading, but also in managing of various healthcare services important for daily life of individuals in community. [19]

Main goals of digitalization health services are effective outcomes, which can be improve, safe harm with adequate healthcare, patient-centered care with involving of patients in process of care, efficient health services with the best value for the money and with equal and available care quality for all users of services in healthcare system [20] Impacts of digital transformation in healthcare are usually visible on operating with medical records and prescriptions which are dematerialized, true therapy which can be adopted by using virtual channels via smartphones and tablets, by using voice interface, which can separate less urgent and complicate request for medical care in real time, digital connection with communication system which reduce communication distance between patients, medical staff and operators. There are several applications that can be develop for the purposes of healthcare system: extension of medical devices, which helps in maintain, diagnostic and therapy; application for health support of patient in prevention or therapy, application which measures health indicators and support healthy behavior of patients; application which gives patients useful specific disease information; application which helps in communication with their physicians and application that provides access to clinical medical information which are stored in digital system. [21]

III. HEALTHCARE SERVICES AND ARTIFICIAL INTELLIGENCE (AI)

Special part of digital transformation in healthcare services can be dedicated to artificial intelligence (AI), which represents next step in development of digital health. [22] AI is expected to radically change workplace and professions. [23] AI is considered as a tool for productivity and progress, and effects the social relations and organizational structure of healthcare system. [24]

A. Development of the ethical and legal framework of artificial intelligence at the level of the European Union

The key foundation for European initiatives to develop artificial intelligence is to create an atmosphere of trust and responsibility around the development and use of artificial intelligence. The EU has a solid and balanced regulatory framework that can serve as a foundation and world standard for sustainable access to this technology. The Union has high standards in terms of product safety and responsibility. The first rules at the EU level on the security of network and information systems and stricter rules on the protection of personal data adopted by Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016 on the protection of individuals in connection with the processing of personal data and on free movement of such data and on the repeal of Directive 95/46/EC (General Data Protection Regulation - hereinafter GDPR), which applies from May 25, 2018. [25] The General Data Protection Regulation ensures a high standard of personal data protection, including technical principles and integrated data protection. It guarantees the free flow of personal data in the Union. It contains provisions on decision-making solely on the basis of automated processing, including profiling. In such cases, respondents have the right to receive meaningful information about the logic of decision-making. [26] The general regulation on data protection provides the individual with the right not to be subject exclusively to automated decision-making, except in certain situations. The Commission closely monitors the application of the Regulation in the context of artificial intelligence, as well as national data protection authorities and the European Data Protection Board, which is the competent authority at the European level.

B. Ethical guidelines for the development of artificial intelligence

The first step in solving ethical issues was the drafting of ethical guidelines for artificial intelligence based on the Charter of Fundamental Rights of the European Union entitled "Ethics Guidelines for Trustworthy Artificial Intelligence by the High-Level Expert Group on Artificial Intelligence (AI HLEG) published on April 8, 2019. [26] The Commission gathered all relevant stakeholders in the European Commission's High Level Independent Expert Group on Artificial Intelligence, which participated in the development of ethical guidelines. The ethical guidelines took into account issues such as the future of work, fairness, protection, security, social inclusion and algorithmic transparency. More broadly, the impact of fundamental rights will be considered, including the principles of privacy, dignity, consumer protection and non-discrimination. [27] In the guidelines, in terms of achieving reliable AI, not only compliance with the law is required, but also one of its three components presented through the three chapters of the guidelines. It is emphasized that laws are not always in step with technological development, sometimes they cannot keep up with ethical norms or they are simply not adequate enough to solve certain problems. For AI to be reliable, they should also be ethical, ensuring compliance with ethical norms.

Based on fundamental rights and ethical principles, the Guidelines list seven key requirements that AI systems should meet in order to be trustworthy:

- Human agency and oversight
- Technical robustness and safety
- Privacy and Data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental well-being
- Accountability

Aiming to operationalize these requirements, the Guidelines present an assessment list that offers guidance on each requirement's practical implementation. [27] This assessment list will undergo a piloting process to which all interested stakeholders can participate, in order to gather feedback for its improvement. In addition, a forum to exchange best practices for the implementation of Trustworthy AI was created.



Figure 1. Seven key requirements that AI systems should meet in order to be trustworthy Ethics Guidelines for Trustworthy AI [28]

C. *Privacy protection and data management*

Closely related to the principle of harm prevention, in accordance with ethical guidelines, is privacy, a fundamental right that is particularly affected by the development of AI systems. Preventing privacy harm also requires appropriate data management that covers the quality and integrity of the data used, its relevance to the domain in which the AI systems will be deployed, its access protocols and the ability to process the data in a way that protects privacy.

AI systems must guarantee privacy and data protection throughout the entire life cycle of the system. This includes information initially provided by the user, as well as data generated about the user during their interaction with the AI system AI generated for certain users or how users responded to certain recommendations. Digital records of human behavior can enable AI systems to infer not only individuals' preferences, but also their sexual orientation, age, gender, religious or political views. In order to enable individuals to have confidence in the data collection process, it must be ensured that data collected about them will not be used to unlawfully or unfairly discriminate against them.

The quality of the data sets used is the most important for the performance of the AI system. When data is collected, it may contain socially constructed biases, inaccuracies, errors, and mistakes. This must be resolved before training with any given data set. [29] In addition, data integrity must be ensured. Feeding malicious data into an AI system can change its behavior, especially with self-learning systems. The processes and datasets used must be tested and documented at every step, such as planning, training, testing and use. This should also apply to AI systems that were not developed internally, but were acquired elsewhere.

In any organization that processes the data of individuals (whether someone is a user of the system or not), it is necessary to establish data protocols that govern access to the data. These protocols should define who can access data and under what circumstances. Only appropriately qualified personnel who are competent and need access to certain data should be allowed to do so. [30]

D. *Harmonised rules on artificial intelligence (Artificial intelligence act)*

The emergence of artificial intelligence, especially the complex system that enables it and the features of autonomous decision-making, requires reflection on the appropriateness of some established rules on security and civil liability. Given the breadth of application of artificial intelligence, it may be necessary to review both horizontal and sectoral rules.

The EU safety framework already covers the intended purpose and (mis)use of products when they are placed on the market. This has led to the development of a robust body of standards in the field of AI-enabled devices that is constantly adapting to technological advances. The further development and promotion of such safety standards and the support of standardization organizations at EU and international level will enable European companies to achieve a competitive advantage and increase consumer confidence. [31]

The Commission is now assessing whether the EU and Member States' security and liability frameworks are appropriate in the context of these new challenges and whether certain shortcomings need to be corrected. A high level of security and an effective mechanism for the legal protection of injured persons in case of compensation helps in building user trust and social acceptance of these technologies. Evaluations of the Product Liability Directive and the Machinery Directive have already been carried out. An initial assessment of existing liability frameworks in the context of artificial intelligence technologies and new technologies was also carried out. In the further analysis of these challenges, the Commission is assisted by an expert group.

Regulation of artificial intelligence in EU countries is next step in process of digital transformation of healthcare services. Main aspect of this regulation are ethical questions

which lead promising technological solutions. Ethical issue is usually related with enactment of healthcare decisions which involve strong and complex judgement and fitting into the social context, which determine decision making process. Those decisions, which are product of human reflection, become part of AI operating, with potentially far-reaching consequences. [32] Accordingly, European Parliament started the discussion regarding the influence and consequences of AI on healthcare services and European Commission proposed to European Parliament and European Council enactment of Artificial Intelligence Act, as a general legal framework for regulation of artificial intelligence in Europa. [33] This act is in legal procedure from 2021., and temporarily is in ordinary procedure in Council of European Union.

According to the proposition, Artificial Intelligence Act needs to regulate different fields of risk, which can be generally divided on unacceptable risk applications, high-risk applications, and low-risk applications. [34] The EU Artificial Intelligence Act could become global standard in Europe (and abroad), which would determine on what implementation and extent of artificial intelligence is more or less accepted or non-accepted in daily life and which application can have more positive than negative effect to the users. Second aspect, which is important in context of this proposal, is question of personal and societal harm, which will arise true implementation of this Act. [35] To reduce potentially harm (individual as societal) proposal of Artificial Intelligence Act prohibited AI practices to the group of individuals who are vulnerable according to their social or economic situation. Second important thing is classification of AI risks with clear definition that divided high risk situations in AI implementation from low-risk situation. Additional question is general purpose of AI system, where Act proposal specifies certain situation where the principles of high-risk AI system are implemented on design of AI system which can be used for many different purposes (general purpose AI). [36] Important part of AI Act proposal is implementation transparency principle according to implementation, development and use of high-risk AI. [37] New added provisions predicted obligation for users of AI emotion recognition system to inform people when they exposed to the influence of the system.

The extensive use of AI-enabled tools healthcare services as in other business-to-consumer transaction should be fair, transparent and compliant with consumer protection legislation. Citizens as users of healthcare services should receive transparent information about the use, features and properties of AI-enabled products. They should be able to control the data generated by using these tools, completly compliant to GDPR regulation, and know whether they are communicating with a machine or another human. When interacting with an automated system, special consideration should be given when users need to be informed about how to get in touch with a human and how to ensure that the system's decisions can be checked and corrected.

IV. CONCLUSION

According to the conclusion of the report on the State of Health in the European Union, only a substantial reflection on health systems can ensure that they are fit for purpose. This means that they must be systems that continue to promote health, prevent disease, and provide patient-centered care that meets the needs of citizens. Data is the key enabler of digital transformation. Health data may be available in various forms; they are not managed in the same way in all member states of the European Union or within the borders of national health systems. They are often not even available to patients themselves or to public authorities, healthcare professionals or researchers to help them develop and deliver better diagnosis, treatment or personalized care. And where they exist, health data often depend on technologies that are not interoperable, preventing their wider use. The basis of this reflection is the conclusion that health and care systems need reforms and innovative solutions in order to become more resilient, accessible and efficient in providing quality care to European citizens.

The paper gives an insight to the potential of applying artificial intelligence in healthcare, but also points out the key risks that must be managed, especially with regard to the protection and security of data in healthcare and protection of privacy of users, as well as the need for a detailed elaboration of the regulatory framework: the security, ethical and legal framework of application and use of artificial intelligence. In this regard, the application of artificial intelligence in healthcare, but also digital solutions in general, if they are purpose-built and implemented in a cost-effective way, can improve the well-being of millions of citizens and radically change the way health and care services are provided. patients.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] T. H Marshall, “Citizenship and Social Class”, Cambridge University Press, 1950., pp. 46 – 75

[3] A. Maziarz, “Services of General Economic Interest: Towards Common Values?”, European State Aid Law Quaterly, Vol. 15, No. 1, 2016., pp. 16 – 30

[4] G. A. Boyne, George, „Sources of Public Service Improvement: A Critical Review and Research Agenda“, Journal of Public Administration Research and Theory, Vol. 13, No. 3, 2003., pp. 367. – 394.

[5] R. Batley, C. McLoughlin, “The Politics of Public Services: A Service Characteristics Approach, World Development”, Vol. 74, 2015., pp. 275 – 285.

[6] Klarić, Mirko: New European Approach to Public Services, Faculty of Law University of Zagreb, Zagreb, 2008.

[7] E. Szyszczak, Erika, “Introduction: Why Do Public Services Challenge the European Union, in Szyszczak, Erika – Davies, Jim – Andeanaes, Mads – Bekkedal, Tarjei: Developments in Services of General Interest, T.M.C. Asser Press, The Hague, 2011., pp. 1–16

[8] S. Van de Walle, “The impact of public service values on services of general interest reform debates”, Public Management Review,

Vol. 8, Issue 2, 2006., pp. 183 – 205 Source: https://doi.org/10.1080/14719030600587422

[9] J. W. Van de Gronden, "Social Services of General Interest and EU Law", in Szyszczak, Erika – Davies, Jim – Andeanaes, Mads – Bekkedal, Tarjei: Developments in Services of General Interest, T.M.C. Asser Press, The Hague, 2011., pp. 123–153

[10] F. Schulz-Nieswandt, "The dynamics of European definition policy of health and social services as services of general (economic interests)", Zeitschrift für öffentliche und gemeinwirtschaftliche Unternehmen: ZögU / Journal for Public and Nonprofit Services, Vol 33., No. 1 (2010), pp. 31-43

[11] A. Cygan, „Public Healthcare in the European Union: Still a Service of General Interest?", The International and Comparative Law Quarterly, Vol. 57, No. 3, pp. 529– 560

[12] M. Klarić, "Administrative and Political Aspects of Healthcare System", Faculty of Law University of Zagreb, Zagreb, 2004., pp. 85 – 89

[13] M. Tahir, R. Amiruddin, S. Palatturi, F. Rivai, M. Saleh, – Lalu, "The relationship between organizing and leadership style and the quality improvement of primary healthcare services", Enfermería Clínica, Vol. 30, Supplement 4, 2020, pp. 39 – 43.

[14] C.J.S. Orte, R.A. Bautista, C.B. Borocco, J.E.C. Neo, A.M. Parico, M.A.S. De Dios, "Comparative study on patient satisfaction on healthcare service delivery in selected private and government hospitals", Enfermería Clínica, Vol. 30, Supplement 7, 2020., pp. 47 – 51.

[15] EXPH (Expert Panel on effective ways of investing in Health), Assessing the impact of digital transformation of health services, 20 November 2018.

[16] R. Koppel, "Great Promises of Healthcare Information Tehnology Deliver Less", in Weaver, Charlotte A. – Ball, Marion J. – Kim, George J. – Kiel, Joan M.: Healthcare Information Management Systems. Cases, Strategies and Solutions, Springer Cham, 2016., pp. 101 – 125.; https://doi.org/10.1007/978-3-319-20765-0

[17] J-I. Ker, Y. Wang, N. Hajli, "Examining the impact of health information systems on healthcare service improvement: The case of reducing in patient-flow delays in a U.S. hospital", Technological Forecasting and Social Change, Vol. 127, 2018., pp. 188. – 198.

[18] M. Mirbabaie, S. Steigliz, N.R.J. Frick, „Hybrid intelligence in hospitals: towards a research agenda for collaboration, Electronic Markets", Vol 31, Issue 2, 2021., pp. 365 – 387.

[19] W. Hong, D. Zhaohua, B. Wang, S. Gupta,"How does service price influence patients' decisions? An examination of the free-market pricing mechanism in online health communities", Electronic Markets, Vol 31, Issue 4, 2021., pp. 877 – 893.

[20] H. Gimpel, T. Romberg-Manner, F. Schmied, T.J. Winkler, "Understanding the Evaluation of mHealth app features based on a cross country Kano analysis", Electronic Markets, Vol 31, Issue 4, 2021., pp. 765 – 794.

[21] E. Maier, U. Reimer, N. Wickramasinghe, "Digital health services in transition", Electronic Markets, Vol 31, Issue 4, 2021., pp. 743 – 746.

[22] Kovačić, M., Mutavdžija, M., & Buntak, K. (2022). e-Health Application, Implementation and Challenges: A Literature Review. Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, 13(1), pp. 1-18

[23] P. Glauner, „Artificial Intelligence in Healthcare: Foundations", Opportunities and Challenges in Glauner, Patrick – Plugmann, Philipp – Lerzynski, Guido: Digitalization in Healthcare, Springer Cham, 2022., pp. 1 – 15.; doi: https://doi.org/10.1007/978-3-030-65896-0.

[24] A. Bohr, K. Memarzadeh, "Current healthcare, big data and machine learning", in Bohr, Adam – Memarzadeh, Kaveh: Artificial Intelligence in Healthcare, Academic Press, pp. 1 – 24. https://doi.org/10.1016/C2018-0-04097-9

[25] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4.5.2016, p. 1–88 (Text with EEA relevance) Available at https://eur-lex.europa.eu/eli/reg/2016/679/oj

[26] J. Connolly, J., A., Staines, A., R. Connolly, A. Boilson, A., P. Davis, D. Weston, D., N. Bloodworth, N., Evaluating Impact of an Emerging Big Health Data Platform: A Logic Model and Q-Methodology Approach. ENTRENOVA-ENTerprise REsearch InNOVAtion, 4(1), 2018. ,pp. 81-87.

[27] A. Boilson, S. Gauttier, R., Connolly, P. Davis, Connolly, J., D. Weston, D., A. Staines, Q-Method Evaluation of a European Health Data Analytic End User Framework. ENTRENOVA-ENTerprise REsearch InNOVAtion, 5(1), 2019., pp. 187-199.

[28] Figure retrieved at source: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1#Human%20agency

[29] M. P. Bach, M. P., D., Cosic, D., Data mining usage in health care management: literature survey and decision tree application. Medicinski glasnik, 5(1), 2008., pp. 57-64.

[30] Ethics Guidelines for Trustworthy Artificial Intelligence (AI) - High-Level Expert Group on Artificial Intelligence (AI HLEG) Available at https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

[31] M. Tomičić Furjan, M., K. Tomičić-Pupek, K., & Pihir, I., Understanding digital transformation initiatives: Case studies analysis. Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy, 11(1), 2020., 125-141.

[32] M. Boban, Primjena umjetne inteligencije u zdravstvu – regulatorni okvir i zaštita podataka, Međunarodni kongres "2. Kongres KOKOZA I 4. Kongres medicinskog prava s međunarodnim sudjelovanjem, Sveučilište u Splitu Pravni fakultet, 2020. pp 67 - 90

[33] Hunady, J., Pisár, P., Vugec, D. S., & Bach, M. P. (2022). Digital Transformation in European Union: North is leading, and South is lagging behind. International Journal of Information Systems and Project Management, 10(4), pp. 58-81.

[34] M. Mirbabaie, F. Brünker, F. Möllmann, R.J.Nicholas, S. Stieglitz, "The rise of artificial intelligence – understanding the AI identity threat at the workplace", Electronic Markets, Vol 32, Issue 1, 2021., pp. 73 – 99.

[35] J. Meszaros, J. Minari, I. Huys, „The future regulation of artificial intelligence systems in healthcare services and medical research in the European Union", Frontiers in Genetics, Vol. 13, 2022. https://doi.org/10.3389/fgene.2022.927721

[36] Panel for the Future of the Science and Technology: Artificial intelligence in healthcare. Application, risks and ethical and societal impacts, European Parliamentary Research Service, 2022.

[37] Procedure 2021/0106/COD, COM (2021) 206: Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative act https://eur-lex.europa.eu/legal-content/EN/HIS/?uri=COM:2021:206:FIN

ICTLAW

International Conference on

INFORMATION AND COMMUNICATION TECHNOLOGY LAW

Steering Committee

Chairs: Tihomir Katulić, University of Zagreb, Croatia
Hrvoje Lisičar, University of Zagreb, Croatia
Lucija Vejmelka, University of Zagreb, Croatia

Members: Nina Gumzej, University of Zagreb, Croatia
Marko Jurić, University of Zagreb, Croatia
Đorđe Krivokapić, University of Belgrade, Serbia
Ashwinee Kumar, VUB, Brussels, Belgium
Ivana Kunda, University of Rijeka, Croatia
Romana Matanovac Vučković, University of Zagreb, Croatia
Hrvoje Stančić, University of Zagreb, Croatia
Goran Vojković, University of Zagreb, Croatia

Program Committee

Members: Marko Jurić, University of Zagreb, Croatia
Tihomir Katulić, University of Zagreb, Croatia
Hrvoje Lisičar, University of Zagreb, Croatia
Lucija Vejmelka, University of Zagreb, Croatia

PAPERS

Comparative Analysis of the AI Regulation of the EU, US and China from a Privacy Perspective

V. Gábor Rádi
Károli Gáspár University of the Reformed Church, Budapest, Hungary
radi.vilmos@gmail.com

Abstract - Artificial intelligence is already part of our everyday lives. We encounter it several times a day, when using a smartphone or social media, when shopping online, or even without any visible signs, such as in case of a secret facial recognition programme.
The area of data protection is closely linked to the evolution of technology, especially wit so-called 'disruptive' technologies raising new data protection issues and risks. Artificial intelligence also creates a new situation because, unlike other technologies, it is difficult to explain how it works and to predict the precise outcome of its use. Regulatory authorities detected the need of proper regulation which ensures both the technical advance and the protection of the natural person's private life.
This paper presents an overview of the major regulatory tools that have emerged in the field so far in the EU, in the US, and finally in China. However, the author of this paper finds that current regulation is rather 'fragmented'; the EU is the first one to come up with a comprehensive and wide-ranging regulation of AI; hopefully the others will follow its path with respect of privacy regulation as well..

Keywords: AI, data protection, privacy, US, China, EDPB.

I. INTRODUCTION

AI is one of the most interesting emerging technologies. As we hear day by day exciting news about the capabilities of artificial intelligence, more and more states or international organizations feel the necessity to create an effective regulation about the technology. My hypothesis is that there is currently no truly effective, comprehensive and wide-ranging legislation on AI that is already in place; or that the current legislation is "fragmented", with sub-areas and functions being the focus of legislators, and other actors stepping in with unified proposals in this area. In order to clarify this issue, I will examine the regulatory instruments that have been put in place so far, concentrating on the legal material of the three most powerful economies of the world: the US, the European Union and China.

II. AI REGULATION IN THE US

The US is probably the country which is taking the lead in the development of artificial intelligence; and several advanced AIs are coming out from US-based companies. Despite this fact, the regulation is still at it's first steps in the US; several pieces of legislation related to AI have been proposed at the federal level, which unfortunately have not yet been finally adopted. For example, the Algorithmic Accountability Act [1] and the Consumer Online Privacy Rights Act [2], which would have dedicated a specific chapter to algorithm-based consumer decision-making. More recent proposals are the Good AI Act of 2022 [3] which would establish an "Artificial Intelligence Hygiene Working Group"; and the Advancing American AI Act [4] intending to regulate the use of AI by governmental agencies. Both regulations were already introduced last year, but since then there has been no significant advancement. It can be said that there is currently no explicit comprehensive federal legislation on AI.

At this stage, it appears that the first major piece of federal-level AI legislation could be the American Data Privacy and Protection Act (ADPPA) [5], which is a „possible" federal-level privacy legislation, although it is also currently in the legislative process and has not been finalised. The draft generally designates the Federal Trade Commission (FTC) as the competent supervisory authority for data protection. The chapter on "Civil Rights and Algorithms" contains the regulation of interest to this paper: within 2 years of the enactment of the prospective legislation, it will require organizations managing large databases that use AI that may cause significant harm to individuals to: first, conduct a privacy impact assessment (based on criteria described in detail in the draft); and second, review the algorithm's architecture, design, structure and inputs, training data, to mitigate risks to the rights of data subjects. Impact assessment and review documentation must be submitted to the FTC within 30 days of their completion. The legislation also requires the FTC to conduct studies on the application of this section of the bill and to publish guidelines on the subject. As we can see, AI is only a minor part of the future ADPPA law.

To address this shortcoming of the lack of federal AI legislation, the White House Office of Science and Technology Policy released the Blueprint for AI Bill of Rights (Blueprint) [6] in October 2022. The Blueprint sets out 5 principles for the use of AI:

(a) building safe and effective systems;

(b) protection against algorithm-based discrimination;

(c) data privacy: concepts similar to GDPR such as "privacy by design", the exercise of data subjects' rights, and protection against abusive data practices are mentioned here;

(d) notice and explanation: data subjects should know when they are using an automated system and should also be informed of the impact of the automated system on them;

(e) human alternatives, and the possibility of consideration and fallback: the data subject should be able to opt-out of AI-based processing or, if they so wish, to request human intervention in the AI decision-making process.

Again, it is important to stress that the Blueprint is not a binding law, but only a recommendation for the development and use of AI.

At the end of the chapter, the author concludes that there is currently no meaningful AI regulation in the US, although significant progress is expected in the near future through ADPPA - however, the author does not expect the implementation of a comprehensive binding regulation as the one in Europe, which will be introduced below. In the US, there are also currently recommendations (the most important of which is the Blueprint) that set the direction for the development and use of AI.

III. REGULATION AND GUIDELINES AT EUROPEAN LEVEL

The first important piece of legislation on data protection is the General Data Protection Regulation (hereinafter GDPR) [7]. Article 5 of the GDPR provides for principles for the processing of personal data, of which the most prominent in the context of AI:

(a) transparency: ”Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject”;

(b) purpose limitation: ”personal data [must be] collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes (…)”;

(c) data minimisation: data must be ”adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed";

(d) accuracy: the data used must be "accurate and, where necessary, kept up to date (...)".

Profiling is defined in Article 4 of the GDPR as "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements ".

The concept of automated decision-making is briefly mentioned in the 71th section of the preamble of GDPR, stating it is a form of a „decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her”. The Article 29 Working Party WP251 Guideline [8] gives us more elaborated details: "Automated decision-making has a different scope and may partially overlap with or result from profiling. Solely automated decision-making is the ability to make decisions by technological means without human involvement. Automated decisions can be based on any type of data, for example:

- data provided directly by the individuals concerned (such as responses to a questionnaire);
- data observed about the individuals (such as location data collected via an application);
- derived or inferred data such as a profile of the individual that has already been created (e.g. a credit score).

Automated decisions can be made with or without profiling; profiling can take place without making automated decisions. However (…) automated decision-making process could become one based on profiling, ()."

The GDPR places a strong emphasis on adequate information, thus, according to Articles 13(2)(f) and 14(2)(g), the data controller is obliged to provide information and according to Article 15(1)(h), the data controller is obliged to provide information to the data subject, at his or her request, on whether automated decision-making, including profiling, is taking place in relation to him or her. The data controller shall also inform the data subject at least of the logic used in such cases and of the comprehensible information relating thereto; and of the significance of such processing and the likely consequences for the data subject.

In order to protect the data subject, the GDPR sets limits on profiling and automated processing in Article 22:

"1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

2. Paragraph 1 shall not apply if the decision:

(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;

(b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

(c) is based on the data subject's explicit consent.

3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9 (1) [of GDPR], unless point (a) or (g) of Article 9 (2) [of GDPR] applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

We can conclude automated functions are well-regulated, but it takes only a smaller piece of the GDPR.

Let’s look at some legal text specifically targeted on AI: On 25 January 2019, the Council of Europe issued its Guidelines on the relationship between artificial intelligence and data protection [9]. The declaration is

divided into 3 main parts: general guidance, advice for developers and advice to legislators and regulators. The Guidelines state that AI development should be in line with the so-called Council of Europe Convention 108 on Data Protection (or its modernised version) and accordingly sets out data protection principles and emphasises the rights of data subjects. For developers, it sets out a number of privacy-by-design principles, draws attention to the importance of ensuring that data subjects are always aware that they are communicating with AI, and calls for developers to establish independent bodies to be consulted on ethical issues in the development of AI. The final chapter of the declaration calls for the drafting of codes of ethics and witness mechanisms and stresses that human intervention should always be allowed. The chapter advocates that national supervisory authorities should have adequate resources to monitor AI developments and that different disciplines should cooperate on this issue; it also calls for developers to consult with supervisory authorities before creating AI that could have a significant impact on the rights of data subjects.

The European Commission have founded the High-level Expert Group on Artificial Intelligence (hereinafter referred to as AI HLEG) which serves as an advisory body to them. On 8 April 2019, the AI HLEG issued the "Ethical Guidelines for Trusted AI" [10], in which it identified 3 elements of trusted AI that must be met throughout the lifecycle of the system: a) it must be lawful; b) it must be ethical, i.e. ensure compliance with ethical principles and values; and c) it must be technically and socially robust.

The Guidelines also identified seven key requirements for ethical AI: human agency and human oversight; technical stability and security; data protection and data management (lawfulness); transparency; diversity, non-discrimination and equity; social and environmental well-being; and accountability. As we can see, the requirements of data protection, transparency and accountability are in parallel with the principles of the GDPR. The AI HLEG subsequently issued a guidance note for stakeholder organisations and individuals (”Assessment List for Trustworthy Artificial Intelligence (ALTAI)”) for self-assessment [11] on 17 July 2020, which can be used to determine the extent to which the AI they are developing or applying can be considered trustworthy. The guide asks questions along the 7 requirements that can be used for self-assessment, so that requirements relevant to data protection are also included.

The White Paper On Artificial Intelligence - A European approach to excellence and trust [12] published by the European Commission summarizes the desired goals, risks and partially repeats the former documents of AI HLEG mentioned above. The White Paper states that its goal to achieve both an ”ecosystem of excellence” and an ”ecosystem of trust” while also highlighting the main problems:

- risks for fundamental rights, including personal data and privacy protection and nondiscrimination and
- risks for safety and the effective functioning of the liability regime.

The White Paper outlines the scope of future compulsory regulation, mentioning the high-risk AI and emphasizes the requirements set up earlier by AI HLEG, and also elevates privacy into the important topics.

The EU binding legal act currently exists only in draft form: the ”Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts” [13] was presented on 21th of April, 2021. The proposed legislation aims to make Europe a global centre for trusted AI. As The draft AI Act requires the below listed criteria to be met for a software to be identified as AI: a) the AI must use specific technologies listed in Annex I of the AI Act; b) it must be able, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations; or c) it must be able to produce outcomes that "influence" the environment.

The draft AI Act classifies AI programmes into several categories based on a risk approach:

The first group includes AI systems that pose unacceptable risks, or otherwise "prohibited", and that clearly threaten the safety, livelihood and rights of people and will therefore be banned. Examples of such programmes include: those that manipulate human behaviour to circumvent the free will of users, or those that allow "social scoring" by governments.

At the other end of the scale are "not high risk" schemes, which can be further broken down into two sub-categories. The first category includes "minimal-risk AI" programmes, the use of which poses almost no risk to the rights and security of users; these are not covered by the AI Act. The second sub-category is the group of "limited-risk AI", where only basic requirements are included (the requirement of transparency: when interacting with a system, users must be aware that they are interacting with a program and can decide whether to continue or discontinue the activity on that basis. Examples of AIs with limited risk are chatbots or deepfake programs.

Most of the draft AI Regulation deals with the last category, so-called "high-risk" AI. The categorisation is set out in the Annexes II. and III. of the AI Act, and – amongst others – includes AI used in the following areas: biometric identification of natural persons; critical infrastructure (e.g. transport) that may endanger the life and health of citizens; education or training (e.g. exam scoring); employment (e.g. CV sorting software); essential private and public services (e.g. assistance assessment, credit scoring); law enforcement (e.g. polygraph AI analysis, personal risk analysis, evidence reliability assessment); asylum and border control (e.g. document authentication); justice and programmes used in democratic processes (e.g. application of laws to specific facts).

For high-risk AI systems, the AI Act sets out a number of requirements: a risk management system must be implemented and risk management measures must be developed in accordance with the criteria described in the draft; the data sets implemented into the software must be of high quality and accurate; technical documentation

must be prepared containing the elements described in the relevant annex of the AI Act, including the information on the AI and its purpose that the authorities need to assess the adequacy of the system; logging of events for subsequent traceability and traceability; clear and adequate information to users about the AI; ensuring adequate human supervision; and finally, the requirement for accuracy, stability and cyber security. As can be seen, the requirements are in line with the expectations of the GDPR in a number of respects, and the GDPR compliance of AI systems will also help to meet the requirements set out in the draft. The AI Act imposes specific obligations on "manufacturers" as well as on importers, distributors and users. The draft imposes a number of detailed rules on the standards, certificates, conformity assessment processes and assessment bodies that will be required in the future, as well as their registration and notification to the public.

The AI Act probably will be an important step from a privacy perspective as well, because it mentions the relationship between the regulation of AI and data protection at several points. In this context, the European Data Protection Supervisor (hereinafter "the Supervisor") is given a number of powers: under Article 53, she or he is empowered to create a regulatory "sandbox" for AI facilitation, for a limited period of time, the development, testing and validation of innovative AI systems before they are put into service. The Supervisor is a permanent member of the European Artificial Intelligence Board under Article 57. Under Article 59, the Supervisor acts as the competent authority for the supervision of the AI aspects of the EU institutions falling within the scope of the AI Act, and as the market surveillance authority under Article 63, and can even impose fines on EU institutions under Article 72.

On 18 June 2021, the European Data Protection Board ("EDPB") and the Supervisor published their joint opinion No 5/2021 ["EDPS - EDPB Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)"] (hereinafter: the "Opinion") [14] on the draft AI Act. The Opinion welcomed the draft, but called for a number of points to be added. The Opinion disapproves of the exclusion of international law enforcement cooperation from the scope of the draft, as this exclusion poses a significant risk of circumvention, for example by third countries or international organisations operating third party high-risk AIs. The EDPS urges that the concept of 'risk to fundamental rights' be aligned with the GDPR and suggests adding a recital stating that the draft does not intend to affect the application of existing EU legislation on the processing of personal data. In relation to risk assessments, the EDPB and the Supervisor propose that the draft AI Act should be amended so that distributors should carry out an initial risk assessment of the AI, taking into account the use cases, and then the user of the AI should carry out a data protection impact assessment, taking into account the specific context in which the AI will operate. The Opinion calls for a complete ban on "social scoring" and "automated recognition of human characteristics in publicly accessible spaces" without exceptions.

According to the EDPS, national data protection authorities should also be designated as national supervisory authorities for the purposes of Article 59 of the draft AI Act, which would ensure a more harmonised regulatory approach and help avoid inconsistencies between Member States of the EU in the implementation. Finally, the Opinion suggests that national data protection authorities should be involved in the preparation and development of harmonised standards and certificates.

On 28 September 2022, a draft Directive [”Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)”] [15] was published, which mainly addresses civil liability actions in front of a national court, and has no direct data protection implications.

The European Parliament's proposal to the Commission [”European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics”] is also worth mentioning [16]. The proposal deals primarily with robots, i.e. AI with a physical appearance; it has a data protection aspect in that it confirms the application of the principles and methods of the GDPR also to robots. As of February of 2023, the proposal seems to have made no progress.

To sum up, the author concludes that Europe is close to achieving a uniform AI regulation, although it may take some years before this is achieved; at present, however, the EU has only issued formal documents on the subject at a declaratory level. The citizens can find detailed privacy rules in data protection authority’s guidelines, instead of EU mandatory law.

IV. THE CURRENT STATE OF REGULATION IN CHINA

China has taken a number of steps over the last few years to develop data protection and IT regulation. Article 24 of the Personal Information Protection Law of the People's Republic of China (PIPL) [17], in force since 1st of November 2021, dedicated it’s Article 24 to automated decision making. Just like the GDPR, the PIPL dedicates only a small portion to AI. According to the PIPL: „Personal information processors using personal information for automated decision making shall ensure the transparency of the decision making and the fairness and impartiality of the results, and may not apply unreasonable differential treatment to individuals in terms of transaction prices and other transaction conditions. Information push and commercial marketing to individuals based on automated decision making shall be simultaneously accompanied by options not specific to their personal characteristics or with convenient means for individuals to refuse. Where a decision that may have a significant impact on an individual's rights and interests is made through automated decision making, the individual shall have the right to request clarification from the

personal information processor and the right to refuse the processor for making the decision only through automated decision making."

The use of AI appears to be of particular importance in China, with a number of official acts issued in the two years; however, it can be noted that there has been only one binding law, but many more recommendations - three in number - have been issued - although it is presumably worthwhile for stakeholders to follow the recommendations closely.

Not surprisingly, the binding legislation focuses on the sub-area of algorithm-based internet recommendations, as has been the case with the first regulatory instruments in other countries and regions. The legislation, which entered into force from 1st of March 2022 and is jointly subscribed by several governmental bodies (but coordinated by the Cyberspace Administration of China). Although it is called "Internet Information Service Algorithmic Recommendation Management Provisions" [18], we can regard it as binding legislation, since it contains a legal liability section which gives the opportunity to issue warnings, decide about suspensions, or impose fines. The legislation requires algorithm-based internet recommendations to be ethical, moral, accountable and transparent. The regulation requires companies to inform users when an algorithm plays a role in determining what information to display to them and to give users the opportunity to opt out of a targeted recommendation based on their personal characteristics. In addition, the regulation prohibits the use of algorithms that use personal data to offer different prices to different consumers for the same product/service.

Three recent recommendations and guidance documents cover:

1. The "White Paper on Trustworthy Artificial Intelligence" [19] issued by the China Academy of Information and Communications Technology together with JD Explore Academy in July 2021, which, among other things, addresses the privacy challenges of AI systems and recommends avoiding or minimizing the use of personal data where it is not necessary to achieve the purpose of AI. The White Paper's policy recommendations include the creation of trustworthy AI-related legislation in China, the development of commercial AI insurance, and a cautious approach to AI research. The document brings up the example that "the frequent use of biometric authentication increases the risk associated with potential data breaches, which could cause the leaking of confidential user data".

According to the White Paper, "the trustworthy characteristics of AI are summarized by five main aspects: transparency, security, fairness, accountability, and privacy." As we can see, the first four aspects are also connected to privacy.

The White Paper even devotes a separate chapter dedicated to AI privacy protection technologies, analyzing types of attacks against anonymized data sets, and protection approaches like the differential privacy.

2. The "Ethical Norms for the New Generation Artificial Intelligence" [20] issued by the National Governance Committee for the New Generation Artificial Intelligence, which preamble mentions six basic ethical requirements: improving human well-being, promoting fairness and justice, protecting privacy and security, ensuring controllability and credibility, strengthening responsibility, and improving ethical literacy. The third subsection of Article 3 is dedicated to the protection of privacy and security: [AI shall] "fully respect the rights of personal information to know, consent, etc., process personal information in accordance with the principles of legality, justice, necessity, and good faith, protect personal privacy and data security, and must not damage the legitimate rights and interests of personal data, and must not illegally collect and use personal information through theft, tampering, or disclosure information without violating personal privacy."

3. "The Guiding Opinions on Strengthening Overall Governance of Internet Information Service Algorithms" [21], which is the work of several different agencies. This document proposes to strengthen the management control systems of AI programs, and deals with management, regulatory, and supervision issues.

It can be stated that no comprehensive regulation on AI exists in China, but rather separet rules concentrating on specific areas; and a number of guidelines and norms. In an interesting development a comprehensive regulation is already included in the "Municipal Ordinance" of Shenzhen City [Shenzhen Special Economic Zone Artificial Intelligence Industry Promotion Regulations] [22]. The city of Shenzhen is important because it was the first so-called special economic zone to be subject to specific rules, and because of this, the zone is home to a number of Chinese technology giants. In addition to supporting R&D projects, the regulation calls for the development of local standards and certification mechanisms, the development of an ethical risk assessment and the establishment of a "Municipal Artificial Intelligence Ethics Committee". The Ethics Committee formulates ethical and security standards; monitors their implementation; analyses the impact of algorithms on the protection of information rights, social ethics, work and employment, etc.; publishes ethical and security practice guidelines, white papers, "good practices"; takes the lead in establishing ethical security management systems for AI.

V. CONCLUSION

The author of the present paper concludes that the need for regulation of AI is a pressing matter. If we see upon the already existing official documents, it is worth contrasting binding regulations with recommendations and guidance documents. Europe is at the forefront of AI regulation, as the GDPR already provides some rules for automated decision-making and profiling, but the draft AI Act and the draft AI Liability Directive are also expected to be adopted in the near future. In the US, there is currently no substantive AI regulation, with changes

expected through ADPPA, which will mandate impact assessments and review of AI systems. In China, binding legislation has only been developed in a sub-area, focusing on algorithm-based internet recommendations; more comprehensive legislation has been developed by the Shenzhen "Municipality" on a territorial basis.

On the other hand, there is a wide range of guidelines, in the EU, guidelines are currently in place rather than comprehensive regulation, see the Council of Europe Guidelines, the European Commission White Paper, the AI HLEG Guidelines. In the US, the Blueprint sets the direction for the development and use of AI, rather than the current lack of binding legislation. In the last two years, China has issued four high-profile official acts on AI, one of which is binding legislation and three recommendations.

In the light of the above, the author of the present paper considers the hypothesis to be well-founded and concludes that there is currently no truly effective, comprehensive and wide-ranging regulation of AI that has already entered into force; and that the current regulation is fragmented, with sub-areas and individual functions being the focus of legislators, and other actors stepping in with unified proposals in this area.

REFERENCES

[1] S.1108 - Algorithmic Accountability Act of 2019, 116th Congress (2019-2020) https://www.congress.gov/bill/116th-congress/senate-bill/1108/text?q=%7B%22search%22%3A%5B%22algorithmic+accountability+act%22%5D%7D&r=2&s=1 [2023. 02. 01.]

[2] S.2968 - Consumer Online Privacy Rights Act, 116th Congress (2019-2020) https://www.congress.gov/bill/116th-congress/senate-bill/2968/text?q=%7B%22search%22%3A%5B%22consumer+online+privacy+rights+act%22%5D%7D&r=1&s=1 [2023. 02. 01.]

[3] H.R.7296 - GOOD AI Act of 2022 https://www.congress.gov/bill/117th-congress/house-bill/7296/actions?s=1&r=2&q=%7B%22search%22%3A%5B%22ai+bill+of+rights%22%5D%7D [2023. 02. 01.]

[4] S.1353 - Advancing American AI Act https://www.congress.gov/bill/117th-congress/senate-bill/1353?s=1&r=4&q=%7B%22search%22%3A%5B%22ai+bill+of+rights%22%5D%7D [2023. 02. 01.]

[5] H.R.8152 - American Data Privacy and Protection Act, 117th Congress (2021-2022) https://www.congress.gov/bill/117th-congress/house-bill/8152/text [2023. 02. 01.]

[6] Office of Science and Technology Policy: Blueprint for an AI Bill of Rights Making Automated Systems Work for the American People https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf [2023. 02. 01.]

[7] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=HU [2023. 02. 01.]

[8] Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01) https://ec.europa.eu/newsroom/article29/items/612053 [2023. 02. 01.]

[9] Council of Europe: Guidelines on Artificial Intelligence and Data Protection. T-PD (2019)01, Stasbourg, 2019. január 25. https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8 [2023. 02. 01.]

[10] High-Level Expert Group on AI: Ethics Guidelines for Trustworthy Artificial Intelligence https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai [2023. 02. 01.]

[11] High-Level Expert Group on Artificial Intelligence: Assessment List for Trustworthy Artificial Intelligence https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment [2023. 02. 01.]

[12] WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf [2023. 02. 01.]

[13] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206 [2023. 02. 01.]

[14] EDPS - EDPB Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) https://edps.europa.eu/node/7140_en [2023. 02. 01.]

[15] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52022PC0496&from=EN [2023. 02. 01.]

[16] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017IP0051 [2023. 02. 01.]

[17] Personal Information Protection Law of the People's Republic of China http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm [2023. 02. 01.]

[18] Internet Information Service Algorithmic Recommendation Management Provisions http://www.gov.cn/zhengce/zhengceku/2022-01/04/content_5666429.htm and https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/ [2023. 02. 01.]

[19] China Academy of Information and Communications Technology together with JD Explore Academy: White Paper on Trustworthy Artificial Intelligence http://www.caict.ac.cn/english/research/whitepapers/202110/P020211014399666967457.pdf [2023. 02. 01.]

[20] National Governance Committee for the New Generation Artificial Intelligence: “Ethical Norms for the New Generation Artificial Intelligence” https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html and https://ai-ethics-and-governance.institute/2021/09/27/the-ethical-norms-for-the-new-generation-artificial-intelligence-china/ [2023. 02. 01.]

[21] Guiding Opinions on Strengthening Overall Governance of Internet Information Service Algorithms http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1779/202109/t20210929_568182.html and https://digichina.stanford.edu/work/translation-guiding-opinions-on-strengthening-overall-governance-of-internet-information-service-algorithms/ [2023. 02. 01.]

[22] Standing Committee of the Seventh Shenzhen Municipal People's Congress: Shenzhen Special Economic Zone Artificial Intelligence Industry Promotion Regulations https://law.pkulaw.com/chinalaw/eb370a7e0d9edd5e8ca8bb1a5fa6a5e7bdfb.html [2023. 02. 01.]

Algorithmic Contracts in European and United States Contract Law - A Comparative Legal Analysis

K.Biczysko-Pudełko*
* University of Opole/Faculty of Law and Administration, Opole, Poland
kbiczysko@uni.opole.pl

Abstract - **Although perhaps still abstract to many, the term 'algorithmic contracts' may soon become as common and more widely known as the - equally abstract-sounding - cryptocurrencies (bitcoin), smart contracts, or blockchain. Algorithmic contracts are a separate category from smart contracts, the so-called digitally enhanced contracts, and are contracts in which an algorithm determines the obligations of the parties - this is how L. Scholz aptly described the essence of these contracts.**

In my presentation, I would like to analyze whether current contract law in the European Union is adequate for algorithmic contracts, i.e., can an algorithm effectively make fundamental determinations about the content of a contract, and will such provisions bind the parties? Will it be possible to invoke an error of declaration of intent in the case of a malfunction of the algorithm caused, for example, by a systems failure, or in a situation where, admittedly, there was no failure, but the algorithm misanalyzed the data?

I would also like to refer my considerations to the US legislation and the views of doctrine representatives presented there and further answer the question of whether the regulatory approach adopted there could also guide possible developments in European law.

Keywords - component; formatting; style; styling; insert (key words)

I. INTRODUCTION

For lawyers dealing with the law of new technologies, it is obvious that it is necessary to seek answers to a number of legal challenges and problems that successively arise in connection with the development of information technologies. Paraphrasing the Latin ubi societas ibi ius, one can say that where there is technology, there, is also law. Thus, in the past, the scientific discourse focused on issues related to the provision of electronic services, cloud computing, personal data protection and, more recently, issues related to Blockchain, smart contracts, the Internet of Things or artificial intelligence. At present, there seems to be a lack - at least on a scale comparable to that concerning artificial intelligence - of discussion concerning algorithmic contracts, although undoubtedly this topic seems to have a high practical relevance. It should be mentioned that, as early as 2021, algorithmic contracts were already used in more than half of the high-frequency trading in the US financial markets [1]. In the consumer market, too, so-called algorithmic contracts are increasingly common, a good example being, for example, the purchase of airline tickets.

II. ALGORITHMIC CONTRACTS – CONCEPT, TYPOLOGY AND CHALLENGES

Algorithmic contracts are contracts in which an algorithm determines the obligations of the parties - this is how Lauren Henry Scholz aptly defined the essence of such contracts in 2017 in one of the first works devoted to this issue [2]. Following the definition cited above, Yasmine Benaich, on the other hand, proposed that algorithmic contracts should be understood as contracts whose terms are defined in whole or in part in a code that allows the automation of an algorithm, which can act not as a tool but as an agent of the parties, depending on the importance, complexity, and predictability of the decisions delegated to it [3]. Importantly, however, these algorithmic actions can take on different forms, making this category of contracting not uniform.

And so, as indicated by L.H.Scholz, algorithmic contracts can be differentiated according to the role to be played by the algorithm - i.e. whether as a tool or as an "agent", further according to the task to be assigned to the algorithm (gap filling or negotiation) and further taking into account the criterion of the algorithm used, i.e. a black box algorithm or a clear box algorithm [4].

And so, an algorithm can act as a "negotiator" before a deal is made, and in doing so the parties specify the conditions to be offered or accepted that the algorithm can apply. An example is the so-called high-frequency trading, which is widely used in financial markets mainly in the United States. Such contracts in practice are concluded by an algorithm, which can react quickly to changes in the financial market and allows the submission and eventual execution of orders at frequencies determined in milliseconds. It is the algorithm that sets individual order parameters, such as the conditions for triggering an order, the timing of its execution, the price or quantity of the instruments subject to the order, or how the order is managed once it is placed, with limited or no human involvement.

Another group is contracts in which the algorithm fills in so-called gaps. Their essence will be that the algorithm fills a gap in a standardized set of conditions. Such a classic example of contracts based on a gap-filling algorithm are those concluded for the purchase of airline tickets, where the price of the offered ticket will be determined by the algorithm based on such data as the time of ticket purchase, the number of flights, etc.

In summary, for a contract to qualify as an algorithmic contract, it must meet certain necessary conditions, viz:

1. These are agreements based on pre-built decision-making models.
2. The role of the man actually comes down to the notion of deciding what the role of the algorithm is, i.e., as a "negotiator", or "agent" of the parties or as a tool to determine and fulfill in the contract those provisions that the parties have not previously agreed on.
3. They are characterized by automaticity

The way an algorithm works or can work directly creates all sorts of challenges that need to be realized. And so the use of algorithms to determine terms in a contract creates the possibility of outcomes that are not and indeed could not have been anticipated by the algorithm's creator. An algorithm can produce a result that is unpredictable. A kind of knowledge gap is created, which is dangerous not only for the entity using the algorithm but also for those who want to enter into such a contract. These concerns are not merely theoretical today. For example, point to the situation where a book on flies was priced at $24 million by pricing algorithms that erroneously adjusted prices based on competitors' expectations [5]. Also worth mentioning is the case of Quoine v B2C2, in which the transaction was conducted through an algorithm. Quoine operated a market where investors could exchange BTC (Bitcoin) and ETH (Etherum). Due to a technical problem, Quoine's market maker program temporarily did not have access to external price data for BTC and ETH3. As a result of the above, B2C2 started offering ETH at a reserve price, which was programmed into B2C2's algorithm as a hedge. As a result, B2C2 bought millions of dollars of BTC at a rate 250 times more preferred than the market rate. When Quoine realised the algorithm was wrong, it took action to undo the transaction citing the error. B2C2 filed a lawsuit, claiming that the transactions should be binding [6].

In light of the above, a number of questions arise almost intuitively on the subject of the adequacy of the current provisions of contract law to the reality created by these contracts, i.e. from those concerning the possibility or effectiveness of making legally binding declarations of intent by the algorithm or the possibility of claiming a declaration of intent error in the event of a disruption in the operation of the algorithm. In this context, in turn, the comparative legal perspective between the legal order of the United States and the legal orders of individual European countries seems particularly interesting, and this is due, among other things, to the fact that in the United States, the scientific discourse seems to be much further advanced.

III. ALGORITHMIC CONTRACTS AND U.S LAW

The issue of using simple algorithms to make a contract has been debated among scholars as far back as the 20th century. In their book,, Can Computers Make Contracts," Tom Allen and Robin Widdison stated that 'neither American nor English law, as they currently stand, would confer legal status on all computer generated agreement' [7]. Currently, however, both representatives of the doctrine and the legislation itself seem to present some concepts different from those presented above.

First of all, as far as legal standards are concerned, it is necessary to point to the Uniform Electronic Transactions Act (1999) (UETA) (*"Uniform Electronic Transactions Act" (PDF). uniformlaws.org. National Conference of Commissioners on Uniform State Laws. 1999. Retrieved April 10, 2022.)* and the Electronic Signatures in Global and National Commerce Act of 2000 (‘E-Sign Act’).

UETA provides uniform rules governing electronic commerce transactions. It establishes a legal foundation for the use of electronic communications in transactions where the parties have agreed to deal electronically.

UETA validates and supports the use of electronic communications and records and places electronic commerce and paper-based commerce on the same legal footing [8].

According to H.L.Scholz, the UETA is the piece of legislation that facilitates the creation of algorithmic contracts in that it equates electronic records and signatures with paper records and handwritten signatures [9]. Moreover, this author concludes that the UETA includes an understanding of algorithms as agents [10]. At the same time, however, attention should also be paid to the voices of those representatives of the doctrine who point to certain doubts that may arise in the context of algorithmic contracts considering the provisions of the UETA. Well, the authors S.Chopra and L. White, referring to Section 9 of the UETA, which states that an electronic record or electronic signature is attributable to a person if it was the act of the person” pointed out that this provision allows for the possibility of avoiding a situation in which responsibility for the mistakes of artificial agents is assigned to the operator or user. However, in the opinion of these authors, this provision is not fully clear. This is because it is difficult to judge based on its content when a particular record or signature will not be the act of a person. This will be especially difficult when the system is malfunctioning in such a way that one might be tempted to conclude that a new act (agent) has occurred [11].

As previously indicated, in the context of algorithmic contracts, the E-Sign Act, which requires that the actions of an artificial agent be "legally assigned" to its user, will also be legally relevant, with the statement that,,A contract. . . may not be denied legal effect, the validity or

enforceability solely because its formation, creation or delivery involved the action of one or more electronic agents so long as the action of any such electronic agent is legally attributable to the person to be bound" [12].

In the context of the above, it seems quite obvious that the consensus also seems to be that it is possible under US law to qualify certain activities of algorithms as electronic agents. Hence, it may not be surprising to see L.H.Scholz's recognition that since algorithms act as humans the law of agency is the appropriate source of law in this regard, but nevertheless, a qualification assuming that the algorithm is not a human being should be included. Ultimately, it is possible to consider an algorithm as an agent without legal personality or as a quasi-agent for the purpose of understanding the legal obligations of their principles [13].

Parallel to the above, which also cannot be fastened, is the evaluation of the algorithm as a mere tool, which only implement the will of their owners without the ability to learn and make decisions. However, as L.H.Scholz rightly points out, already today the acceptance of this position is not fully possible given the capabilities of their operation [14].

IV. ALGORITHMIC CONTRACTS – EUROPEAN PERSPECTIVES

When analyzing the regulatory environment for algorithmic contracts in Europe, it should first be emphasized that there is currently no single piece of European Union law that comprehensively regulates the issues related to these contracts. Furthermore, it should be pointed out that there is currently no single, extensive, and coherent concept of contract law in European Union law. However, there are some European regulations that explicitly address the subject of obligations, such as, for example, directives related to consumer trade or product liability, etc. The only significant achievement in this area is the fact that the European Union has now adopted a single, comprehensive, and coherent concept of contract law. The only significant achievement in the unification of European contract law is the Principles of European Contract Law (PECL), which, and it should be emphasized, act as an instrument unifying contractual practice, and not a strictly normative act. Therefore, the purpose of this work is firstly to signal the challenges that algorithmic contracts may bring to contract law, and secondly to present the concepts and views emerging in individual European countries (in and outside the European Union). These goals are vital since the above-mentioned lack of regulation in the area of European Union law makes a comparison of the law in the European Union and the law of the United States problematic. At the same time, it is not possible to discuss the legal systems of individual European countries in detail in this paper, if only due to editorial limitations, and therefore the analysis will be limited to presenting the views of various doctrines established by different European countries concerning such fundamental issues such as for example, the question of freedom of contract or declaration of intent.

First and foremost, it seems indisputable that the decision of the parties on whether or not to conclude a contract using an algorithm is within the bounds of the principle of freedom of contract, which is deeply rooted in all legal systems of European countries. Also the previously mentioned Principles of European Contract Law – PECL, Article 1:102 states that Parties are free to enter into a contract and to determine its contents. Thus, if the 'effect' of the algorithm is known to the parties in the light of the above mentioned principle, a contract concluded using the algorithm is legally binding. The views of the doctrine which states that a declaration of intent can be made in electronic form, including by automated means, and by doing so not affecting its validity, remains valid.

Even though the above conclusion states that a declaration of intent remains valid, the questions remains open – whether automated declarations of intent fall within the concept of a declaration of intent at all? [15].

The answer to this question was one of the first to be sought after by the representatives of German legal doctrine as early as in the 1960s. In German science, this issue is compared to the analogy for affixing a handwritten signature to a *blank* statement. It is pointed out that a person who signed and gave such a statement to another person will be bound by its contents, although in essence, he had no way of predicting what content would be placed above the signature. Someone who signs a *blank* document must expect that their declaration will be treated as a statement of intent and that there is a risk of the document being filled out contrary to the will of the signatory. According to the concept presented, the source of the imputation of the legal effects of a statement of intent is not the will of the one making the statement, but the confidence induced by the statement in other participants in the binding force of the statement. And this view can also find its application in algorithmic contracts [16].

Establishing that we will be dealing with a statement of intent, moreover, allows us to seek an answer to another question, namely, what if the content of the statement made by the algorithm was not anticipated or was not even foreseeable by the person using the algorithm?

The answer to this question will depend on what caused the operation of the algorithm not to coincide with the statement of intent, i.e., whether the cause is an error inherent in the algorithm itself or at least in the operation of the antivirus.

If there has been a distortion of the statement of intent due to an error inherent in the algorithm, it would seem that entities using algorithms should bear strict liability for legally momentous actions performed through these algorithms, and this is due to at least the following two circumstances:

- The certainty of economic trading as that which prevails over individual interest

- Algorithms often make statements en masse, so potentially evasion can have incomparably more serious consequences than if they were made individually.

Based on the views outlined above, the so-called concept of declaration of intent based on the principle of liability for induced trust has been formed in German doctrine, according to which declarations of intent made in the context of electronic communication will concern those behaviors to be a manifestation of the will of the sender, or in other words – the user of the program [17].

Polish doctrine, on the other hand, emphasizes that when assessing the problem of a possible discrepancy between the content of statements generated in an automated manner and the user's actual will, priority should be given to the protection of the entity acting in reliance on the declarant's conduct and the resulting legitimate expectations [18].

As for Anglo-Saxon doctrine, on the other hand, this doctrine, to some extent analogous to the views presented on the backdrop of U.S. legislation, has developed the concept that protection should be granted to such an interpretation of the programmed machine's actions as would be shared by reasonably judgmental addressees. This concept was developed on the basis of the 1971 English court decision in Thornton v. Shoe Lane Parking, where the court held that a parking machine expressed the will of its owner, and the good faith of customers should be protected [19].

At the same time, however, it should be pointed out that if the other party could have easily noticed that an error had occurred, then such an entity using an algorithm can claim a legally relevant error. Thus, if the entities using algorithms are usually entrepreneurs whose goal is to conclude economically beneficial transactions, the question should be asked whether the mere fact of quoting an obviously low price for a good or service should arouse suspicion of the existence of an error on the part of the submitter. That should certainly be decided in the circumstances of a particular case, and not at the legislative level.

Of course, the above mentioned concepts are not uniform in all EU countries, so I would like to refer to the model law. And so art. 4.110 Principle of European Contract Law states the party may avoid a contract for a mistake of fact or law existing when the contract was concluded if the other party knew or should have known of the error, and leaving the party in error was contrary to the principles of good faith and fair dealing. Also, the Unidroit Rules state that each party should act (including in shaping the content of the contract) in accordance with the principles of good faith and fair dealing.

In conclusion, it should be pointed out that the principles of an algorithmic contract essentially correspond to the content of the contractual provision. Moreover, due to the principle of freedom of contract, it may be concluded that an algorithm can create a binding contract.

However, in a situation where the final content of the contract does not correspond to the intention of the parties, the possibility of evasion will depend on other circumstances. However, it is important to remember that a statement of intent – regardless of its form, technological environment, time and place of submission, and other factors – always remains the same institution of civil law. Any other assumption would have to be considered unhelpful on pragmatic grounds, as it would complicate incredibly business transactions. Importantly, in the case of algorithmic contracts, it is also relevant to refer to the so-called meta-principles, i.e. good faith and fair dealing.

Of course, at this point, it is very important to note that the issues related to algorithmic contracts are not limited to those related to the statement of intent, but as shown above, European doctrines tend to focus around this issue.

V. CONCLUSION

The above considerations lead to the conclusion that although the current theoretical and legal approach to algorithmic contracts in the United States or some European countries seems to focus on other issues, in both cases it seems indisputable that these contracts are fully legally binding and admissible. Moreover, the current regulatory environment makes it possible to assess that contract law is neither archaic nor completely out of step with the realities created by the reality of new technologies, for many views already developed by doctrine representatives can also be successfully applied to algorithmic contracts. Of course, the above does not exclude the need for further debate in this area, the development of certain guiding principles or model rules, or even certain legislative changes, but these are not revolutionary changes. Nor is it the case, contrary to what is commonly assumed about issues related to new technologies, that the emergence and subsequent spread of algorithmic contracts leave several legal questions unanswered. It is therefore entirely possible to agree with M. Olivier, who rightly states that "Thus, the problem for contract law is not so much that these contracts are doctrinally unenforceable, but rather that contract law doctrines need to adapt to explain why these contracts are enforceable" - which both doctrinal representatives from the United States and individual European countries seem to do as demonstrated in this work.

REFERENCES

[1] M. Olivier, Contracting by Artificial Intelligence: open offers, unilateral mistakes, and why algorithms are not agents, Australian National University Journal of Law and Technology 2021, p. 59.

[2] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p.128.

[3] Y. Benaich, 'Hello World!': Towards a New Era of Algorithmic Contracting? Implications for Laws and Regulations, p. 38.

[4] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p. 136

[5] O.Solom, How A Book About Flies Came To Be Priced $24 Million On Amazon, Available at: https://www.wired.com/2011/04/amazon-flies-24-million/).

[6] M. Olivier, Contracting by Artificial Intelligence: open offers, unilateral mistakes, and why algorithms are not agents, Australian National University Journal of Law and Technology 2021, p. 46).

[7] M. Olivier, Contracting by Artificial Intelligence: open offers, unilateral mistakes, and why algorithms are not agents, Australian National University Journal of Law and Technology 2021, p. 55).

[8] S. Norman-Eady, Available at: https://www.cga.ct.gov/2000/rpt/2000-R-1076.htm.

[9] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p.166).

[10] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p.166).

[11] S. Chopra, L. White, Artificial agents, and the contracting problem: a solution via an agency analysis, JOURNAL OF LAW, TECHNOLOGY & POLICY Vol. 2009, no. 2, 2019, p. 388.

[12] S. Chopra, L. White, Artificial agents, and the contracting problem: a solution via an agency analysis, JOURNAL OF LAW, TECHNOLOGY & POLICY Vol. 2009, no 2, 2019, p. 389.

[13] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p.165.

[14] L. H. Scholz, Algorithmic contracts, 20 STAN. TECH. L. REV. 128 2017, p.132.

[15] K.Gajek, Oświadczenie woli w komunikacji elektronicznej, Warsaw 2020, p. 37).

[16] N.Rycko, Zautomatyzowane umowy, pieczęć elektroniczna i podpis biometryczny . Problematyka złożenia oświadczenia woli i formy czynności prawnych, Instytut Wymiaru Sprawiedliwości 2021, p. 11).

[17] K.Gajek, Oświadczenie woli w komunikacji elektronicznej, Warsaw 2020, p. 37).

[18] K.Gajek, Oświadczenie woli w komunikacji elektronicznej, Warsaw 2020, p. 37).

[19] K.Gajek, Oświadczenie woli w komunikacji elektronicznej, Warsaw 2020, p. 37).

Medical Data in the Digital Era - Legal Challenges Related to Providing Information Security, Applying GDPR and Respecting the Professional Secrecy

Krzysztof Świtała*
* Cardinal Stefan Wyszyński University/ Faculty of Law and Administration, Warsaw, Poland
k.switala@uksw.edu.pl

Abstract - **The processing of medical data in electronic form poses many challenges in terms of the security of these resources and ensuring patient privacy. This paper presents new dimensions of healthcare data processing, such as EHR, eHealth, mHealth, IoT or Big Data, in the perspective of the challenges of enforcing legal regulations to ensure the security of such healthcare information resources while also guaranteeing cross-border interoperability of these solutions. The technical and organizational measures applied by the controllers, as well as the legal requirements, should be consistent and comprehensive, considering not only the challenges related to the protection of personal data, but also cybersecurity and the protection of professional secrets in health care. In this context, ensuring the integrity and availability of information is also important as protecting its confidentiality.**

Keywords - eHealth, EHR, data protection, data concerning health, information security, cybersecurity, professional secrecy

I. INTRODUCTION

The protection of information processed in health care cannot be approached selectively. A coherent and comprehensive approach to ensuring the safety of such resources should respect the obligations under the GDPR, the 2022/2555 NIS2 Directive, instruments for legal protection of medical profession secrets and medical records, as well as other related regulations. No less important than ensuring confidentiality of patient information is to guarantee its integrity and availability. The processing of such resources should meet the appropriate standards of business continuity and be characterized by adequate quality guaranteeing that information is up-to-date and complete. It should be remembered that pecuniary or non-pecuniary damage associated with the loss of integrity of medical data used during a medical procedure or the lack of availability of these resources in the event of a life-threatening condition may be at least no less severe than the unauthorized disclosure of such information. Only such an integrated approach to this issue will allow for due respect for the rights and freedoms of patients as data subjects.

I The author acknowledges funding and support from the National Science Centre, Poland (NCN Preludium grant no. 2014/13/N/HS5/03523).

The purpose of this article is to identify the legal challenges associated with the processing of health data and to identify potential methods for increasing the effectiveness of regulations relating to ensuring the security of health data. Only such an integrated approach to this issue will allow for due respect for the rights and freedoms of patients as data subjects. The question whether the legal regulations relating to the protection of information resources in healthcare resulting from the GDPR and the NIS2 directive are sufficient to properly guarantee the information autonomy and patient's privacy should be answered.

II. MEDICAL DATA AND DATA CONCERNING HEALTH

The information on health condition obviously belongs to the matters linked to the private life of an individual [1], constituting the emanation of their rights to information privacy and self-determination, whose enforcement should be appropriately protected by legal, organisational and technical safeguards. Data concerning health are qualified as a particularly protected category of sensitive data laid down in Article 9 GDPR and related directly to the individual's sphere of intimate life. As such, the loss of their confidentiality may cause the feelings of shame, embarrassment, and restraint. It is worth adding that the ISO 29100 standard relating to the framework for the protection of privacy in organizations extends to data related to human health to sensitive personally identifiable information data requiring special precautions [2]. This confirms the need to take special care of such categories of data.

Therefore, any misuse of these special categories of personal data could have more severe consequences on the individual's fundamental rights related to the protection of privacy and information self-determination [3]. It is worth adding that in Poland, as in most other states, such data resources are also covered by the scope of medical professional secrets. It is the experts' opinion on Article 29 Working Party that all EHR data shall be treated as particularly protected [4]. The application of such an approach allows to unify the principles for the resources protection in health records systems, eliminating problems related to the occurrence of weak links in the security

systems and no consistency in the safeguards used. It is a guarantee of proper protection of such resources and data subjects' privacy.

Recommendation CM/Rec (2019) 2 explicitly indicating that health-related data reveal information about a person's health status and this kind of data also includes information on the provision of health services. The explanatory memorandum to the above-mentioned document indicates that the data cover all information relating to the identification of the patient in the care system or the method used for gathering and processing health data, all information obtained during a medical check-up or examination, including biological samples and genome data, all medical information such as an illness, a disability, a risk of illness, clinical, physiological or biomedical information or information concerning medical treatment, irrespective of its source, as well as all data, refer to an individual, generated by professionals practising in the medical welfare sector [5]. This category of information resources should also include data on individual predispositions and health risks related to a specific person [6]. These resources may be associated with a specific health context (such as presence in a region affected with disease) or be the result of a "self check" survey where symptoms specific to a particular disease are given [7]. Conclusions of the CJEU in the Lindqvist [8]. case confirmed the possibility of using such a wider interpretation of the concept of health-related data.

The "data concerning health" term, as known from Article 4(15) GDPR, is also defined in an analogous manner as health-related data. The terminological inconsistency that appears here, referring to similar terms, i.e. health-related data and data concerning health with identical scope, deserves criticism. The notion "personal health information" was also identified in ISO 27799 standard as information about identifiable person that relates to the physical or mental health of the individual [9], which is consistent with health definition proposed by World Health Organization. In this sense, health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity [10]. Thus, this needs to be understood that the analysed meaning of personal health information is to a large extent analogous to the earlier presented normative expression of data concerning health laid down in the GDPR. Besides sensu stricto data concerning health, patient's identification data (a number, symbol, or particulars assigned to an individual to uniquely identify the individual for health purposes) have also been included in this particular category. Attention should also be drawn to the essential link of health information, thus understood, with the data of the person - medical service provider - providing services to a given patient. Furthermore, the personal health information does not include information that, either individually or in combination with other information available to the holder, is anonymized so that the identity of the individual who was the subject of the information cannot be ascertained from the information. Such resources are called sensu largo medical data.

Therefore, it should be considered appropriate to depart from the narrowing of medical data by limiting it only to information about the health status of a specific person. The term also includes resources processed in healthcare other than personal data, such as anonymized research and statistic data, information about health conditions for the population, medical knowledge, as well as know-how. The leading value that determines the protection of data concerning health is the privacy and intimacy of the data subject. With regard to other data processed by healthcare institutions, the security of these resources is primarily provided to ensure appropriate development for medical sciences, healthcare sector and the medical industry in order to maintain and systematically improve the health condition of the population, as well as to guarantee its safety.

Attention should be drawn to inextricable links between medical and genetic data. It is often impossible to put a clear dividing line between their scope. For instance, recital 35 of the GDPR clearly indicates that data concerning health also includes information derived from genetic data [11]. Undoubtedly, such resources qualified in the GDPR for specific categories of data, due to their information potential regarding the ongoing assessment and prediction of the health status of a particular person and their relatives, require care for security appropriate to their nature and specific risks for privacy [12]. Firstly, anonymized genetic data may in the future re-identify a person by recognizing new relationships between them [13]. Secondly, the identification capacity of these resources and their ease of processing in a digital environment can pose potential threats related to identity theft [14]. Finally, the significant value of genetic data in the context of profiling processes may lead to exclusion and discrimination of specific groups of individuals by predicting their personal characteristics and health status.

III. NEW DIMENSIONS OF PERSONAL DATA PROCESSING IN HEALTHCARE

A. The role of a common approach to ensuring information security in healthcare

The processing of medical data takes place not only in an automated way or in filing systems, including health records. Healthcare professionals share oral information about specific patients or make ad hoc notes about the treatment process. These information resources are used not necessarily for purposes related to treatment processes, but also for research aims and health policy management. Furthermore, some entities performing medical activities are operators of essential services, which entails additional obligations related to their securing. The actions implementing them should be complementary to the approach to protection of other information resources in such an entity (personal data, secrets of medical professions). The importance of ensuring information security in the context of privacy protection in organizations is mentioned by the ISO 29100 standard. Therefore, it seems reasonable to apply a comprehensive approach to information security in entities carrying out medical activities, taking into account legal requirements in

the scope of personal data protection, medical profession secrets and cybersecurity.

The special categories of personal data that include health data relate not only to the sphere of privacy but also to human intimacy. The effects of unauthorized interference in the latter sphere make it difficult to reverse effects on a person's personality and undermine his or her sense of security for a long time. In the case of electronic data processing, the risks associated with such an incident are significantly higher than in an analogue environment, due to the complicated structure of data sets and network protocols, which is a source of ever new vulnerabilities.

B. Security of eHealth, mHealth and Internet of Things solutions

ICT and the Internet network have already penetrated into many social and economic life areas. One of them is the healthcare sector, in which patients, medical personnel and public institutions use these tools to implement treatment or broadly understood objectives of healthcare policy and healthcare management [15]. In WHO documents, eHealth is presented as the use of ICT in healthcare [16]. Undoubtedly, the eHealth systems, which make it possible to implement health services in a new model, along with their provision at a distance, require special standards for patients' personal data processing, the specificity of these solutions taken into account, and being adapted to new categories of threats [17]. The eHealth notion is closely related to the issues of patient's remote care [18]. In healthcare environment such technologies are hallmarks of high quality services and for this reason they are naturally accepted by patients despite some interferences in their privacy [19].

Modern patient care processes currently use not only traditional computer systems and networks, but also Internet of Things and mHealth solutions. The WHO defines these technologies in its documents as medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices [20]. In European Commission documents Internet of Things solutions refer to "a dynamic global network infrastructure with self-configuring capabilities based on standard and interoperable communication protocols where physical and virtual things have identities, physical attributes, and virtual personalities and use intelligent interfaces, and are seamlessly integrated into the information network" [21]. Solutions of this kind are particularly useful for improving the situation of those with chronic diseases (diabetes and heart disease) and the elderly. As a result of using IoT technologies, they could be subject to constant health monitoring which, in the event of a sudden deterioration of medical indicators, would trigger an automatic call for help for the patient. Such tools allow for decentralization of treatment and diagnostics processes - through electronic personalized devices connected to the network, some activities in this area could be performed in relation to him at his home or workplace.

Besides obvious benefits of the widespread use of mHealth and IoT solutions in healthcare, one should not forget about the threats that may exploit the vulnerabilities of communication protocols, such as Bluetooth, to loss of confidentiality, data integrity and raise the possibility of identity theft [22]. Strongly symmetrical and asymmetric cryptographic algorithms with appropriately long and complex encryption keys are relatively rarely used in communication between such devices. It is also important to provide adequate technical and procedural guarantees that the user of such solutions will be properly informed of the extent to which are affect his information autonomy [23]. The mHealth and IoT technologies should be configured, even at the price of a certain reduction in their performance and universality, to ensure proper respect for values such as privacy.

C. Role of automated information processing and Big Data in healthcare

Automated processing and profiling derive not only information of the health situation of a particular patient, but also identifies a new and not fully recognized health threats to the population. In the latter case, whenever possible, anonymized or pseudonymized data should be used. The development of automated processing methods facilitates rational decision-making in healthcare. For example, such technical solutions in conjunction with artificial intelligence algorithms can be used to medical digital images recognition in diagnostic procedures [24]. Appropriate data analysis allows for greater resources usage optimization as a result of its better allocation, which eliminates waste, improves savings and the efficiency of business processes [25]. However, it is necessary to strike a proper balance between the values associated with the effective implementation of health policies or the management of an individualized treatment plan for a particular patient - in relation to the preservation of the essence of his or her privacy and information autonomy. It is of particular importance here to prevent patient discrimination based on health condition, which was determined on the basis of automated profiling methods. In accordance with Article 22(4) GDPR for special categories of personal data, making decisions based on automated processing, including profiling, is permissible if the patient's consent is obtained or is necessary for reasons of substantial public interest and on the sufficient legal basis. It seems that the last of these premises will apply, for example, in cases of epidemics and the need to respond quickly to its negative results on society in order to prevent irreversible serious effects on public health. Pursuant to Article 35(3)(a) GDPR, the data protection impact assessment is performed for the processes of systematic and extensive evaluation of personal aspects, which is based on automated processing, including profiling, and on which decisions are based that produce legal effects. The proper conduct of this procedure is of particular importance for maintaining satisfactory security level of special personal data categories, including data concerning health [26].

The concept of Big data refers to data sets characterized by the attributes of volume, velocity, variety and veracity, as well as values that are explored to discover such relationships between resources and new phenomena which have not yet been observed. It can therefore be concluded that these large databases are somewhat an "ocean of data", while the associated methods of analysis are "types of nets and fishing strategies" [27]. Big data is a source of new economic values as well as social and technical innovations

[28]. Analysis of large data sets not only makes it possible to effectively implement treatment processes using evidence-based medicine, which allows for deeper understanding of patient disease patterns. These technologies also support the development of a series of important strategic industries related to the national economy, people's livelihood, and national security [29]. The European Commission indicates that health records, gathered in a European perspective, can lead to better treatment of major chronic conditions and help to improve equal access to high quality health services for citizens [30]. Large data sets are not only a source of information about the phenomena observed in the present time, but are also the basis for precise forecasting of future trends and potential changes in environment. This allows us to react to possible negative situations in such a scenario that their effects are counteracted, in a way, ex ante, i.e. before they occur.

IV. Legal challenges related to the new paradigm of medical data processing

A. The importance of consistent and effective organizational and technical security of medical data

The use of personal data security measures relevant to privacy threats has a positive impact on the effectiveness of legal standards related to information processing. The ECHR jurisdiction has drawn attention to the importance of the adequate protection of medical records to respect the rights of data subjects. In case I v Finland [31], the Court found Finland to have violated its positive obligations to secure respect for private life pursuant to Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms, due to its failure to apply appropriate guarantees, through organizational and technological measures, the confidentiality of patient personal data in a public hospital. The Court of Justice of the European Union in its case law also takes into account these security solutions - as necessary elements of implementing the right to data protection that prevent accidental loss, alteration or unlawful erasure of these resources [32]. In order to guarantee an appropriate level of personal data protection, in addition to the implementation of purely legal obligations, it is also necessary to apply appropriate organizational, physical and technical security measures adequate to the risks associated with these assets.

Most of data handled in healthcare regardless of the form and place of processing should be subject to consistent rules guaranteeing their security. A holistic view of the issue of protecting information resources in healthcare, where gaps and the weakest links could be eliminated, not only allow for proper respect for patients' right to privacy, but will also help to improve the functioning of the processes of the entire sector by creating conditions for making decisions based on up-to-date data-driven reliable evidence.

As indicated in recital 35 of the GDPR, obligations arising from this act do not apply to identifiable data that relate to the deceased. Member States may provide for rules regarding the processing of personal data of such persons. Moreover, the medical professions secrecy guarantees the protection of patients' data not only during their lives but also after death. Protection in the medical sector of information concerning the deceased person is related to the specificity of the doctor-patient relationship based on trust, assuming respect for the patient's privacy and autonomy. This means that proper and lawful security of information resources processed in medical records cannot be based solely on the GDPR requirements.

B. Ensuring business continuity as a challenge for legal regulations

Secure information is characterized by its confidentiality, integrity and accessibility attributes. In the legal context, their behavior is guaranteed by the regulations contained in the GDPR and standards of the NIS Directive. The differences relate to the scope of protected information and the purposes of both regulations. The GDPR refers to the protection of filing systems and the processing of personal data due to the protection of the data subject's rights and freedoms. In the case of provisions from Article 21 (2)(c) of the NIS2 Directive, the business continuity of essential and important entities activities is protected by guaranteeing the proper functioning of information systems, which, in addition to personal data, may include resources such as know-how, technical data, financial and accounting data, etc., as well as software and hardware used to process them. Therefore, it seems that the basic assumption of these regulations is not competition, but their mutual complementation and the synergy effect that can thus be obtained.

Maintaining business continuity is of particular importance in healthcare units where it is necessary to guarantee stability and constant availability of treatment processes. Modern evidence-based medicine is based on the use of data processed via ICT, which should be characterized by reliability and high resistance to interference. With regard to personal data processing, both Article 32(1)(b) and (c) GDPR as well as the ISO 29100 standard indicate the importance of ensuring the availability of these resources in the event of the need for their use by an authorized user. This will allow patients to have constant access to quality-stable health services.

C. Role of legal guarantees ensuring the quality of medical data

As accurately pointed out by M. Safjan, in the contemporary information society, an individual can be simultaneously a beneficiary and a victim of modern information processing technologies [33]. From the point of view of challenges related to the use of ICT in healthcare, ensuring the quality of data is an important factor. Such resources processed in these solutions should be integral and accessible to authorized entities.

The principle of accuracy and quality enshrined in Article 5(1)(d) and ISO 29100 assumes that properly protected information throughout its entire life cycle should be accurate, complete, up-to-date, adequate and relevant for the intended use. The quality of information referring to the definition in ISO 9000 is the degree to which a set of inherent characteristics (confidentiality, integrity, availability, authenticity, accountability, non-repudiation, reliability) of these resources fulfils the requirements in healthcare business processes arising from the needs of its actors (patients, doctors and other healthcare professionals, scientists, government).

In the case of medical data, it is just as important as its confidentiality to preserve quality factors - the integrity and availability of such information resources. This means that it is equally important to safeguard against unauthorized access to guarantee the continuity of medical data processing. Unavailability of these resources may threaten more than just violation of the right to privacy of a person - this situation may pose a genuine threat to his or her health and life. An example confirming the role of ensuring the integrity and accuracy of medical data used in treatment processes may be an event that is the basis for the Regional Court (pol. Sąd Okręgowy) judgment in Katowice dated 12.12.2003 [34], resulting from accidental connection in the hospital database by the medical secretary of diagnoses of two different patients, which resulted in a medical error consisting in performing an unnecessary surgical procedure mutilating the patient. Ensuring adequate quality of data has a significant impact on the achievement of health protection objectives. These resources play a role not only in the field of medical therapy, but also in scientific research and management of the entire sector. Undoubtedly, ensuring a satisfactory level of data quality and maintaining its stability are stimulated by properly formulated legal requirements imposed on health care entities.

D. Legal challenges related to the interoperability of medical data

EU documents define interoperability as the ability of organisations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between these organisations, through the business processes they support, by means of the exchange of data between their ICT systems [35]. EU institutions support the process of eHealth systems interoperability improvement in cross-border healthcare. The e-health network was established in Article 14 of Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. A further proposed legislative step in the EU concerning a common area for the exchange of medical data between Member States is the proposal for a regulation on the European Health Data Space (COM(2022)197). Achieving a satisfactory level of interoperability between EHR systems in EU Member States will allow for the necessary access for doctors and patients to a comprehensive medical history, which should contribute to improving the quality of health services provided. Ensuring respect for the right to privacy and the protection of personal data also plays an important role in these information processing activities.

There are differences in regulations regarding medical confidentiality between countries, which is undoubtedly a significant barrier to the cross-border flow of medical data. Article 29 Working Party indicated that the legal requirements for the collection, access to and cross-border transmission of data concerning health might not be the same in all EU Member States, taking into account the possibly distinct conditions, including limitations, especially the conditions arising from Article 9(4) GDPR, concerning the processing of such resources [36]. EU Member States also take different approaches with regard to the protection of patients' rights. In some countries it is expressed in terms of the individual rights of patients. Nonetheless, in others it is primarily an obligation imposed on healthcare entities [37]. For example, in France, legal protection of medical confidentiality is considered stronger than in most of other European countries [38]. This obligation, however, is not absolute and is based on the theory of will assuming that it results from the content of a contract between the doctor and the patient. The French courts verbally adhere to the concept of strict protection of medical secrecy and, at the same time allow exceptions where the essential interests related to the disclosure of information covered by this restriction of its dissemination significantly outweigh the reasons for maintaining their confidentiality [39]. While the differences in the subjective (healthcare professionals obliged to protect the privacy of their patients) - and objective (concerning all information that comes to the knowledge of healthcare professionals when practicing the profession) scope of regulation of medical secrets in EU Member States are not significant, the list of exceptions to confidentiality results from the content of legal instruments and case law of individual countries. The regulation of these issues in EU applicable law is rather vague due to the fact that it goes beyond the shared EU competences in the areas of internal market and common safety concerns in public health matters, for the aspects defined in the Treaty on the Functioning of the European Union. However, this does not prevent EU institutions from using coordinating competences related to protection and improvement of human health to support and issue soft law, encouraging Member States to gradually harmonize the principles of observing medical confidentiality in the context of interoperable cross-border processing of medical data. The proposed solutions should naturally be compatible with the legal provisions adopted in the GDPR and NIS2 Directive.

V. WAYS TOWARDS EFFECTIVE REGULATIONS OF MEASURES ENSURING THE MEDICAL DATA SECURITY

A. Role of technical standards in increasing the effectiveness of legal requirements related to the protection of personal data and information security

The use of technical standards from the ISO 27000 family - in particular ISO 27001 containing requirements for the Information Security Management System, ISO 27002 covering guidelines for the implementation of such a system, ISO 27701 supplements the privacy policy management system contained in those standards and ISO 27799 indicating the use of this standard may undoubtedly be useful for ensuring a high level of information security in entities performing medical activities for specific information security problems in healthcare.

A properly managed risk management process is significant for ensuring the validity of organizational and technical measures in the entity performing medical activities. ISO 27005 or specific-industry standards may serve as a reference point in this regard. The codes of conduct provided for in the GDPR should allow for a comprehensive approach to risk management in information security for entities performing healthcare activities. A properly carried out risk management process makes it possible, with naturally limited resources (financial, material, personal and time), to ensure the

optimal condition of information security in current conditions by implementing measures to reduce risks exceeding the specified level of acceptance.

Technical standards also describe approaches to maintaining an adequate level of privacy in an organization environment where information processing operations prevail. The ISO 29100 document characterizes the privacy framework, while the ISO 29134 being developed contains guidelines for Privacy Impact Assessment. These standards comply with the GDPR and supplement the standards of the ISO 27000 family.

A structured and comprehensive approach to ensuring business continuity in the organization guarantees the adaptation of the ISO 22301 requirements. Implementation of this standard ensures capability of the organization to continue delivery of products or services at acceptable predefined levels following a disruptive incident. Business continuity management allows controllers access to data availability, and if they are also essential or important entities to apply such solutions, they are guaranteed to meet the requirements related to standards issued pursuant to Article 21 of the NIS2 Directive. Respecting standards of the sustainability of key information processes in an organization is particularly important in the healthcare sector, where disruption of access to data may result in a threat to such fundamental social values as patients' health and lives.

B. Vulnerability-driven approach to protection of information resources and related processes

It is good practice, not only from a medical point of view, but also in view of the management to prevent the occurrence of the causes of the problem, and not to focus solely on their effects. The earlier the vulnerability is reduced, the lower the potential likelihood and amount of costs of damage will be. In the field of technical protection of information processed in electronic form in healthcare, it is worth considering the development of industry-specific standards that will expand general approaches resulting from such knowledge bases as Open Web Application Security Project [40] sources or ENISA reports on the state of vulnerabilities [41]. Appropriate identification of vulnerabilities at the planning stage of specific personal data processing processes will allow for the most effective implementation of the privacy and data protection by design principles, as reflected in Article 25 GDPR.

The importance of vulnerability management in EU for effective information systems security has been highlighted in Article 12 of the NIS2 Directive. Information on these protection issues should be exchanged among competent entities in a secure communication environment both at a national and European level. Proper procedures in this area and efficient reaction to anomalies observed make it possible to reduce the risk of exploitation of identified vulnerabilities by threats, which may cause serious disruption to the activities of various entities. It should be mentioned that international technical standards, such as ISO/IEC 30111 and ISO/IEC 29417, refer to the management of this area of security and they should complement the legal regulations in this area. Sharing information about identified vulnerabilities in information systems enables a proactive and preventive approach to ensuring their security, unlike a reactive incident management model.

Pursuant to Article 25 of the GDPR, the vulnerability analysis should not be limited to the design phase of the processing activity, but should extend to its entire life cycle. Such a conclusion results from the content of Article 25 of the GDPR. The controller should identify and manage security vulnerabilities both at the time of the determination of the means for processing and at the time of the processing itself. This could be done not only on the basis of expert knowledge, but primarily on the basis of the analysis of empirical data from ISMS monitoring and related Intrusion Detection System/Intrusion Prevention System [42]. The latter tools are useful not only to detect the vulnerabilities of information resources, but mainly to identify threats directly affecting them.

It is necessary to manage technical safeguards life cycle. Cryptographic solutions (e.g. MD5 hash function [43]), they may no longer provide an adequate level of security. This is an effect of technological advances that allow these threats to bypass or break security, which undermines the effectiveness of information protection.

The hardware and software have vulnerabilities, the detection and disclosure of which is a continuous process. Some of them are identified by the threats (exploits) themselves for the first time (0-day vulnerabilities), which means that we would not find any advice or solutions in the available sources to counteract the effects of this incidents [44]. Working Party Art. 29 pointed to the growing importance for security of informing users as soon as possible about identified new vulnerabilities of IoT solutions [45]. The occurrence of such situations confirms the necessity for systematic hardening of the configuration of own information systems, which allows for self-improvement of the level of security of processed resources.

C. Role of documentation of the information security and personal data protection management system

Among the threats occurring on the Internet, the use of social engineering aimed at vulnerable users as well as Advanced Persistent Threat attacks is becoming increasingly significant. This means that the implementation of appropriate organizational solutions is no less important than ensuring physical and technical security of information systems. Properly maintained, structured and constantly updated documentation is their indispensable element. It contains not only the descriptions of procedures or current records but above all policies providing the framework for the information security management system functioning in the organization.

A consistent, comprehensive and current Information Security Policy, indicating the direction of management and support for security resources in accordance with business requirements and universally applicable regulations and other standards, should be the basic element of the information processing documentation in medical entities. The content of this document should make personnel, other cooperating entities and patients believe that their data are properly protected. It implements the principle of transparency resulting from Article 5(1)(a) of

the GDPR and expanded in Article 12. For this reason, the information security policy should be drafted in a language understandable to its recipients and should focus on the specific features of processing these resources in the organization. Such a document, which forms the basis for training in the organization and disseminating knowledge about data protection, increases awareness of threats and the methods of counteracting them in specific cases [46]. It leads to the creation and consolidation of a local proactive information security culture.

Information security policy may be uniform or may constitute a set of documents presenting selected thematic areas. Properly put in place and maintained policies are an important aspect of the implementation of the accountability principle laid down in Article 5(2) of the GDPR [47] and in ISO 29100 standard. This documentation describes the scope, limits, user roles and purposes of the personal data security management system functioning in a specific organization, which allows it to demonstrate the proper implementation of the other principles of protecting these resources [48].

According to the ISO 27003 standard, the content of the information security policy covers the following issues: scope, objectives, principles, key outcomes and related policies. The description of information security principles should refer to how they are implemented in a particular organization. The objectives of information security come not only from the content of legal instruments, but also relate to the nature of the organization's business activities and its mission. Moreover, this part of the policy confirms the implementation of the purpose limitation principle arising from Article 5 (1)(b) of the GDPR. Furthermore, the document should contain a list of key outcomes that will be achieved if the policy objectives are implemented, which will allow for evidence-based verification of its effectiveness. An appropriate adoption of the policies requires far more than the mere presence of these documents. In addition, the procedures for verifying the effectiveness of personal data security play an important role in the context of implementing the accountability principle. To sum up these considerations, the security policy may concern the information processes in the organization as a whole and create a comprehensive vision of their protection.

In a health care organization, not only information resources related to clinical, financial accounting and administrative processes should be properly protected, but also the documentation relating to the information security management system. Properly maintained documentation of information resources and personal data processing is a key organizational safeguard that allows the organization to demonstrate proper implementation of not only the accountability principle, but also the other principles arising from the GDPR. De lege ferenda EDPB and ENISA should consider formulating guidelines for maintaining documentation of information security management systems, with particular regard to the methods of its adequate protection. This proposal of the Art. 29 Working Party seems to be still justified for some controllers (such as individual medical practices) processing special categories of personal data, such as data concerning health. This will be a pattern for small entities to provide them with substantive support in implementing appropriate organizational security measures related to the specific characteristics of data processing.

D. Striking an appropriate balance between interoperability and information security

Ensuring interoperability is not only related to the effective implementation of business processes in the strict sense as a result of the effective exchange of data needed to implement such activities. At each of its levels, we should also consider taking into account the common standard of security-related exchange of health data. It is worth noting that in Directive 2008/57/EC of the European Parliament and of the Council of 17 June 2008 on the interoperability of the rail system within the Community, the concept of interoperability is understood as the ability of the rail system to ensure safe and uninterrupted train travel. In a similar vein, the interpretation of this term should promote the protection of health and privacy- to ensure the security of these information resources as well as the persons concerned is as important as achieving the purposes of effective processing health data.

The current level of ICT development allows for the implementation of the principle of time limitation by using cyclical tasks and triggers in databases containing electronic medical health records. These modern technical solutions have radically extended the data lifecycle, without its erosion as a result of forgetfulness and the promise of natural obliteration [49]. Therefore, a controller should have systematic procedures for data erasure embedded in the processing [50]. This effect can be achieved by using appropriately configured software, where the user's role is limited to exercising supervision over these processes.

There is a need for further debate on due measures for the observance of patients' rights with regard to interoperable cross-border exchanges of electronic health records, eHealth and telemedicine. Consideration should be given to further enhancing the role of trust services (eIDAS) in cross-border healthcare to guarantee authenticity and non-repudiation of health records exchanged between medical entities in the Member States. Such technical solutions allow patients to provide their medical consent in a secure way over the Internet. Particular emphasis should be placed on emergencies where information autonomy may be limited due to the risk to the health of the patient or other persons. In such a situation, however, full accountability of access to medical data should be ensured and the data subject would be guaranteed actual control rights over this process.

The confidentiality of information about patients in healthcare require effective protection. A controller is required to limit the number of those who can have access to personal data. This approach ensures the accountability of activities performed in EHR, thus facilitating the determination of responsibility for medical procedures confirmed by entries in this filing system. This is particularly important in the context of ensuring the patient-physician confidence not only in ICT itself, but primarily in medical data processing in which various health care entities are involved not only from a given EU Member State. Data exchange in cross-border healthcare requires

the accountability, authenticity and non-repudiation of related activities to preserve the confidentiality and integrity of these resources.

E. Principle of accountability and the privacy by design approach in relation to the concept of continuous improvement

Effective implementation by the controller of the accountability principle arising from Article 5(2) GDPR and the obligations related to appropriate protection of personal data processing activities pursuant to Articles 24 and 32 GDPR and the content of recital 78 should be based on the use of a continuous improvement approach. According to the ISO 27000 standard, it means recurring activity to enhance performance. This is an ongoing effort to improve key business processes in organization.

PDCA cycle is an iterative four-step (plan–do–check–act) model of process enhancement. It consists of the following phases:

- plan - includes formulation of goals, indication of actions leading to their achievement, together with the assigned resources;
- do - implementation of a previously established and approved plan;
- check - monitoring the changes introduced and observing deviations from the assumed indicators for the purposes;
- act - taking corrective actions in the event of significant deviations [51].

The application of this approach is characteristic of quality management systems based on the ISO 9001 standard and information security management implementing the requirements of the ISO 27001 standard. The issues of quality and ensuring information security are important aspects of the functioning of healthcare entities that use ICT solutions [52]. The PDCA continuous improvement cycle is compatible with the concepts of Privacy and Data Protection by Design due to compliance with their principles: Proactive not reactive approach, Privacy embedded into design and End-to-end security - full lifecycle protection [53]. These approaches to ensuring the security of information resources relate to the management of technical and organizational measures in healthcare and e-Health systems [54]. The DPIA process model has also iterative nature and review phase as PDCA. Moreover, the concept of continuous improvement is the foundation of IT system management methodologies such as COBIT, ITIL and ISO 20000-1. ISO 29100 also mentions the importance of periodic privacy compliance audits for information systems. This process enhancement model based on the PDCA cycle could also support the effective and proactive implementation of cybersecurity strategies in EU and its Member States [55].

As part of the process of continuous improvement of the ISMS based on the PDCA cycle, knowledge of data breaches and other cyber security incidents is used. It should be recalled that Article 33 GDPR (in relation to controllers) and Article 23 NIS2 (in relation to essential or important entities from the healthcare sector) introduce obligations to report incidents to supervisory authorities. Investigating the circumstances of an incident is part of the check phase, while action to counteract the effects of such incident is taken at the act phase. In the context of giving effect to the implementation of the approach based on continuous improvement and Privacy by Design, it is necessary to integrate the processes of improving the information security and personal data protection with the quality management processes in healthcare entities. Such a consistent approach to managing various areas of the organization's activities will allow for a synergy effect to be achieved.

F. Role of codes of conduct in healthcare

The use of soft law and self-regulation is a manifestation of a tendency to govern issues characterized by high dynamics of changes and relating to the boundaries between the legal system, business management and ICT. This approach is based on flexible instruments that do not manipulate the behavior of their addressees. Soft law solutions are not competitive but complimentary to universally applicable law.

Codes of conduct are used to organize the approach to ways and means of securing personal data in entities undertaking a similar type of activity. As recital 77 GDPR indicates, these are sets of guidelines for the implementation of appropriate measures and internal procedures to limit the risk of information processing and it is a way of demonstrating compliance with common legal and technical standards. Therefore, codes of conduct are voluntary accountability tools which set out detailed rules of data protection for controllers [56]. However, this should not be a one-size-fits-all approach where such seemingly versatile and universal document would apply to each and every healthcare entity. It seems reasonable to separate codes for hospitals, pharmaceutical companies, as well as individual and group medical practices due to the different nature and objectives of their activities.

A draft code of conduct on privacy for mobile health applications has been created for such ICT solutions in healthcare. It contains a description of the document (purpose, scope, governance model and enforcement), practical guidelines for software developers (consent management, application of data protection principles, fulfillment of the information obligation, data processing periods, security measures implementation, admissibility of advertisements, use of data for secondary purposes, transferring data to third parties, data breach response procedure, children's data handling) as well as a Privacy Impact Assessment template. The following principles of data processing are presented in the discussed code: purpose limitation, data minimisation, transparency, privacy by design and privacy by default as well as data subject’s rights. On 10 April 2017 Working Party 29 rejected this code due to its insufficient compliance with the GDPR requirements and the lack of added value in relation to Directive 95/46/EC [57]. Such a document should also guarantee appropriate enforcement mechanisms taking into account the data subject's rights [58]. However, this is an important attempt to develop a universal European standard for processing data concerning health via mobile devices.

Codes of conduct should not duplicate the content of legal instruments. Rather, they should complement them by

specifying the obligations imposed on the controller conducting specialized activities. It is also important to put in place effective mechanisms for independent verification of the requirements contained therein.

VI. Conclusion

In order to ensure an adequate level of information security in entities performing medical activities, it is not sufficient to merely comply with the scope of the GDPR and the national provisions of the NIS2 Directive. In this regard, national legal regulations relating to the protection of medical professional secrets also play a vital role. Undoubtedly, codes of conduct, industry-specific guidelines and technical standards can effectively supplement the requirements of the universally applicable law, thus ensuring an appropriate level of security of information resources processed in entities in the healthcare sector.

Maintaining information security, taking into account confidentiality, integrity and availability, requires the use of appropriate security measures tailored to the specificity of the activity of a given healthcare entity, which should be specifically oriented towards maintaining the continuity of the technical and organizational solutions used. It is possible to achieve this as a result of a properly implemented risk management process, taking into account the vulnerabilities of the resources processed by these entities to be protected.

In terms of the challenges related to the COVID-19 post-pandemic cyberspace environment and the aging of the societies of developed countries, the possibility of safe and effective exchange of medical data with due respect to patients' privacy is particularly important. Their interoperable uninterrupted data flow will enable the achievement of goals related to treatment processes and implementation of health policies.

Acknowledgment

The author acknowledges funding and support from the National Science Centre, Poland (NCN Preludium grant no. 2014/13/N/HS5/03523).

References

[1] Constitutional Tribunal judgement of 19 May 1998 (case K U 5/97), OTK of 1998, no. 4, item 46.

[2] ISO 29100:2011 Information technology — Security techniques — Privacy framework (ISO 2011) pp. 9.

[3] Article 29 Working Party (A29WP), Advice paper on special categories of data ("sensitive data")' (Ref. Ares(2011)444105, 20 April 2011) pp. 4.

[4] Article 29 Working Party (A29WP), 'Working Document on the processing of personal data relating to health in electronic health records (EHR)' (WP 131, 15 February 2007) .

[5] Explanatory memorandum to Recommendation CM/Rec(2019)2 of the Committee of Ministers to member States on the protection of health-related data (adopted by the Committee of Ministers on 27 March 2019 at the 1342nd meeting of the Ministers' Deputies) 4.

[6] Article 29 Working Party, Letter of 5 February 2015 to Paul Timmers (Director of Sustainable and Secure Society Directorate), Annex I – health data in apps and devices, Letter and Annex regarding health data in apps and devices, <https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf> accessed 19 January 2022, pp. 2-3.

[7] European Data Protection Board, Guidelines 3/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak (21 April 2020) pp. 4.

[8] CJEU Case C-101/01 Lindqvist (2003).

[9] ISO 27799:2016 Health informatics - Information security management in health using ISO/IEC 27002 (ISO 2016) pp. 3.

[10] This definition is based on Preamble to the Constitution of WHO as adopted by the International Health Conference, New York, 19 June - 22 July 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of WHO, no. 2, 100) and entered into force on 7 April 1948.

[11] Article 29 Working Party (A29WP),Working Document on Genetic Data (WP 91, 17 March 2004) pp. 5.

[12] W. Lowance, Privacy, Confidentality and Health Research, Cambridge University Press, 2012, pp. 111.

[13] D. Hallinan, M, Friedewald, P. De Hert, "Genetic Data and the Data Protection Regulation: Anonymity, multiple subjects, sensitivity and a prohibitionary logic regarding genetic data", Computer Law and Security Review, no. 29, 2013, pp. 322 - 323.

[14] E. Joh, "DNA Theft: Recognizing the Crime of Nonconsensual Genetic Collection and Testing", Boston Law Review, no. 91(2), 2011, pp. 672.

[15] J. Herveg, Y. Poullet, "Which Major Legal Concerns in future e-Health?" in P. Goujon, S. Lavelle, P. Duquenoy, K. Kimppa, V. Laurent (eds), The Information Society: Innovation, Legitimacy, Ethics and Democracy, New York: Springer, 2007, pp. 159-160.

[16] WHO, eHealth <https://www.emro.who.int/health-topics/ehealth/> accessed 5 July 2023.

[17] C. Dierks, "Legal and Social Responsibility in Health Service Chains" in B. Blobel, P. Pharow, M. Nerlich (eds), eHealth: Combining Health Telematics, Telemedicine, Biomedical Engineering and Bioinformatics to the Edge, Amsterdam: IOS Press, 2008, pp. 107.

[18] C. George, D. Whitehouse, P. Duquenoy, "Assessing Legal, Ethical and Governance Challenges in eHealth" in C. George, D. Whitehouse, P. Duquenoy (eds), eHealth: Legal, Ethical and Governance Challenges, Berlin Heidelberg: Springer, 2013, pp. 3.

[19] H. Nissenbaum, Privacy in Context: Technology, Policy, and the Integrity of Social Life, Stanford University Press, 2009, pp. 6.

[20] WHO, mHealth - New horizons for health through mobile technologies < https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1&isAllowed=y> accessed 10 March 2023, pp. 6.

[21] European Commission, Advancing the Internet of Things in Europe' accompanying the document COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Digitising European Industry - reaping the full benefits of a Digital Single Market (SWD (2016) 110, 19 April 2016), pp. 6.

[22] U. Pagallo, M. Durante, S. Monteleone, "What Is New with the Internet of Things in Privacy and Data Protection? Four Legal Challenges on Sharing and Control in IoT" in R. Leenes, R. van Brakel, S. Gutwirth, P. De Hert (eds), Data Protection and Privacy: (In)visibilities and Infrastructures, Dordrecht: Springer, 2017, pp. 70.

[23] L. Andrews, "A New Privacy Paradigm in the Age of Apps", Wake Forest Law Review, no 53, 2018, pp. 424.

[24] T. Davenport, R. Kalakota, "The potential for artificial intelligence in healthcare", Future Healthcare Journal, no. 5(2), 2019, pp. 97.

[25] Article 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (WP251, October 2017) pp. 5.

[26] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 (WP 248, 4 April 2017), pp. 8-9.

[27] B. Szafrański, „Realizacja zadań publicznych a big data" in Grażyna Szpor (ed.), Internet. Publiczne bazy danych i Big data, Warsaw: C.H Beck 2014, pp. 12.

[28] V. Mayer-Schönberger, K. Cukier, Big Data: A Revolution that Will Transform how We Live, Work, and Think, Boston: Harcourt 2013, pp. 12.

[29] M. Chen, S. Mao, Y. Zhang, V. Leung, Big Data Related Technologies, Challenges and Future Prospects, Cham: Springer 2014, pp. 22.

[30] European Commission, Shaping Europe's digital future (COM(2020) 67, 19 February 2020), pp. 12.

[31] ECHR Case 20511/03, I v Finland (2008).

[32] CJEU, Case C-293/12, Digital Rights Ireland, ECLI:EU:C:2014:238.

[33] M. Safjan, „Prawo do prywatności i ochrona danych osobowych w społeczeństwie informatycznym", Państwo i Prawo, no. 6, 2002, pp. 3.

[34] II C 911/01/5, Prawo i Medycyna 2005, No. 2 with the voice of M. Nesterowicz, pp. 122- 130.

[35] European Commission, New European Interoperability Framework. Promoting seamless services and data flows for European public administrations <https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf>, Luxembourg 2017, accessed 9 March 2023, pp. 5.

[36] Article 29 Working Party(A29WP), Letter on 11 April 2018 to Clemens-Martin Auer (e-Health Network Member State) regarding Agreement between National Authorities or National Organisations responsible for National Contact Points for eHealth on the Criteria required for the participation in Cross-Border eHealth Information Services, < https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=52057> accessed 1 March 2023, pp. 3.

[37] European Commission, Patients' Rights in the European Union - Mapping eXercise <https://health.ec.europa.eu/system/files/2018-01/2018_mapping_patientsrights_exe_en_0.pdf>, Luxembourg 2016, accessed 2 March 2023, pp. 18 - 19.

[38] M. Guedj, M. Munoz-Sastre, E. Mullet, P. Sorum, "Do French lay people and health professionals find it acceptable to breach confidentiality to protect a patient's wife from a sexually transmitted disease?", Journal of Medical Ethics, no. 32(7), 2006, pp. 414.

[39] S. Michalowski, "Medical Confidentiality and Medical Privilege - a Comparison of French and German Law", European Journal of Health Law, no. 5, 1998, pp. 112.

[40] OWASP, Top 10 - 2021 The Ten Most Critical Web Application Security Risks <https://owasp.org/www-project-top-ten/> accessed 20 January 2023.

[41] ENISA, State of vulnerabilities 2018/2019 - Analysis of Events in the life of Vulnerabilities <https://www.enisa.europa.eu/publications/technical-reports-on-cybersecurity-situation-the-state-of-cyber-security-vulnerabilities/at_download/fullReport> Heraklion 2019, accessed 3 March 2023, pp. 1 - 50.

[42] M. Stevens, A. Lenstra, B. de Weger, "Chosen-prefix collisions for MD5 and applications", International Journal of Applied Cryptography, no. 2(4), 2012, pp. 323.

[43] S. Sengan et. al., "Secured and Privacy-Based IDS for Healthcare Systems on E-Medical Data Using Machine Learning Approach", International Journal of Reliable and Quality E-Healthcare no. 11(3), 2022, pp. 1 – 2.

[44] M. Janiszewski, A. Felkner, P. Lewandowski, "A Novel Approach to National-level Cyber Risk Assessment Based on Vulnerability Management and Threat Intelligence", Journal of Telecommunications and Information Technology, no. 2, 2019, pp. 8.

[45] Article 29 Working Party (A29WP), Opinion 8/2014 on the on Recent Developments on the Internet of Things (16 September 2014), pp. 22.

[46] L. Stevens, C. Dobbs, K. Jones, G. Laurie, "Dangers from Within? Looking Inwards at the Role of Maladministration as the Leading Cause of Health Data Breaches in the UK" in R. Leenes, R. van Brakel, S. Gutwirth, Paul De Hert (eds), Data Protection and Privacy: (In)visibilities and Infrastructures, Dordrecht: Springer 2017, pp. 225.

[47] Article 29 Data Protection Working Party, Opinion 3/2010 on the principle of accountability (13 July 2010), pp. 4.

[48] R. Thomas, "Accountability - a modern approach to regulating the 21-century data environment" in H. Hijmans, H. Krankenborg (eds), Data Protection Anno 2014: How to Restore Trust?, Cambridge: Intersentia 2014, pp. 142.

[49] T. Gerety, "Redefining Privacy", Harvard Civil Rights-Civil Liberties Law Review, no. 2, 1977, pp. 288.

[50] European Data Protection Board, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default (13 November 2019), pp. 12.

[51] M. Taylor et al., "Systematic review of the application of the plan–do–study–act method to improve quality in healthcare", BMJ Quality & Safety, no. 23, 2014, pp. 293.

[52] A. Appari, M. Johnson, "Information security and privacy in healthcare: current state of research", International Journal of Internet and Enterprise Management, no. 6(4), 2010, pp. 280.

[53] A. Cavoukian, Privacy by Design. The 7 Foundational Principles <https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf> Toronto 2011, accessed 28 February 2023, pp. 2.

[54] G. Drosatos et. al., "Towards Privacy by Design in Personal e-Health Systems" in J. Gilbert et.al. (eds), Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies, Volume 5: HEALTHINF (ROME: SCITEPRESS 2016), pp. 473.

[55] A. Kańciak. "In search of EU law in the domain of cyberspace protection – the proposal based on the Cyber-PDCA model", Journal of Cyber Security Technology, no. 1(2), 2017, pp. 131.

[56] European Data Protection Board, Guidelines 1/2019 on Codes of Conduct and Monitoring Bodies under Regulation 2016/679 (12 February 2019), pp. 6.

[57] Reply to the letter of 7th December 2017 and a new draft code of conduct with the request of a positive opinion from the WP29 under the Data Protection Directive (11 April 2018).

[58] E. Mantovani, J. Antokol, M. Hoekstra, S. Nouwt, N. Schutte, P. Zilgalvis, J. Castro Gómez-Valadés, C. Prettner, "Towards a Code of Conduct on Privacy for mHealth to Foster Trust Amongst Users of Mobile Health Applications" in R. Leenes, R. van Brakel, S. Gutwirth, P. De Hert (eds), Data Protection and Privacy: (In)visibilities and Infrastructures, Dordrecht: Springer 2017, pp. 104.

Technical Solutions Supporting the Online RTBF in the CJEU and ECHR Jurisprudence

Nina Gumzej
University of Zagreb Faculty of Law, Chair of Information Technology Law and Informatics, Zagreb, Croatia
ngumzej@pravo.hr

***Abstract* - Legal solutions toward restricting online accessibility of content relevant to one's privacy are valueless without interdisciplinary cooperation and acknowledgment of technological developments. One example lies in the affirmed use of geo-filtering technology to support the scope of delisting, which ensures the more effective RTBF as a specific right related to the right to erasure, which is reserved for data subjects and implemented by search engines under EU law. Other two concern the measures of rearranging search results, and adding warnings on initiated proceedings in search engine results, which the CJEU acknowledged in certain cases of unsuccessful delisting requests, and which are according to their aim of providing currently accurate data/information traditionally directed toward original content producers. Fourth example lies in the recognized role of de-indexing technology that enables online publishers to restrict accessibility of their own content, and supports the right to private life under ECHR jurisprudence. The paper discusses and critically assesses those solutions. Delisting is shown to be less restrictive than de-indexing for the freedom of expression and at the same time less effective data protection-wise, taking into account also limited territorial scope, and is as such also consumed by de-indexing. Innovative CJEU measures may provide fair solutions for affected data subjects, but still require legal justification and proof of operation in the practice of search engines.**

Keywords - RTBF; geo-filtering; Google v. CNIL; delisting; de-indexing; GC and Others; TU and RE v Google; Biancardi v Italy; Hurbain v Belgium; blocking online content; digital archives; search engine

I. INTRODUCTION

Already in 2011, the European Network and Information Security (ENISA) reported that: „While it is impossible in general to remove data from the Internet once it was published, it might be possible to limit its accessibility (...) A natural way to "mostly forget" data is thus to prevent its appearance in the results of search engines, and to filter it from sharing services like Twitter. EU member states could require search engine operators and sharing services to filter references to forgotten data. As a result, forgotten data would be very difficult to find, even though copies may survive, for instance, outside the EU jurisdiction" [1]. The Court of Justice of the European Union (CJEU) acted on that foresight three years later in the *Google Spain* judgment (C-131/12), where it examined in detail the role of large internet search engines in the dissemination of information relating to individuals in the EU and interpreted the application of the EU data protection law requirements to search engines [2]. Thus, the data subjects may in certain cases request that certain information (personal data) are blocked from search results when users make a search query involving their name. With the exercise of delisting relevant information published on *third-party* websites is neither erased from the Internet nor removed from the index and cache of the Internet search engine. Search results retrieved upon a specific, very narrow name-based search query are merely *blocked*, e.g., they continue to provide links to delisted information in cases of search queries with other keywords [3]. Such a right is colloquially referred to as the "right to be forgotten" (RTBF), as defined and interpreted in the CJEU jurisprudence on the basis of applicable legislation, which is currently Article 17 of the GDPR [4]. Terminology used to denote this right, which data subjects normally invoke before leading search engine operators (Google and Bing being recently declared the very large online search engines [5] pursuant to the Digital Services Act) [6] and as used in this paper is mainly that of the "RTBF" or "delisting", which is also referred to in relevant Guidelines of the European Data Protection Board (EDPB) in [3] and in relevant procedures before Google [7].

Aim of this paper is to explore and discuss the evolution and role of different solutions (geo-filtering, rearranging search results, publishing warnings, de-indexing) impacting, directly or indirectly, the implementation of the online RTBF in the jurisprudence of the CJEU and that of the European Court of Human Rights (ECHR). Analysis in Part 2 sets off with an overview of the path towards the use of geo-filtering technology by Google, which it initially contested but eventually voluntarily implemented, and a discussion on its role in the furthering effect of the territorial scope of delisting, taking into account also relevant aspects of the *Google v. CNIL* judgment [8]. Further examined are innovative measures of rearranging search results and publishing warnings on initiated proceedings, which the CJEU acknowledged in *GC and Others* [9] and *TU and RE v Google* [10] judgments to apply in certain cases of unsuccessful delisting requests - as measures directed toward search engine operators. That analysis concludes with a discussion, on one hand, of the positive outlook for relevant data subjects who are able to and wish to make use of such solutions and on the currently unresolved issues relating to their implementation. Technical solutions relating to implementation of the RTBF have also been discussed in recent ECHR case law in terms of the balancing of the right to respect of private life (including the right to reputation) with the freedom of information by newspaper publishers maintaining online

archives. As I will explain in Part 3, the ECHR acknowledges concerns of the prolonged ease of access to published information via search engines and the related responsibility of noted publishers in implementing the technical measure referred to as "de-indexing" in certain cases, as a proper outcome of the balancing between conflicting fundamental rights. Mainly explored in Part 3 is the approach of the Court in *Biancardi v. Italy* judgment [11] and relevant developments in the current referral proceedings before the Grand Chamber in the *Hurbain v. Belgium* case [12]. While both cases certainly merit a detailed academic discussion [13] they are due to scope of paper not examined in substance and detail. Instead, focus is on approach of the Court toward delisting and de-indexing measures, delineation between the two measures and responsible parties to implement them, which further leads the discussion to the issue of the impact of de-indexing on relevant publishers and the exhaustion of the RTBF, from a data subject's perspective, in cases of successfully implemented de-indexing. Analysis in Part 3 is complemented by an overview of current technical documentation and policies, which support website owners toward implementing de-indexing (and other content-blocking measures) on the Google search engine. Part 4 concludes the paper.

II. MEASURES IN CJEU JURISPRUDENCE

A. Role and Evolution of Geo-filtering in Delisting

Google (which is in focus as the world's leading search engine in relation to whom the CJEU's RTBF jurisprudence evolved) currently implements successful delisting requests of data subjects by default across all of its EU domains, and on any other domain where it is established, on the basis of the IP address, that the search originates in the Member State of the data subject. However, this territorial scope issue was legally unresolved and thus left entirely at its discretion for a striking period of 5 years following Google Spain, i.e., up until the CJEU's second RTBF judgment in Google v. CNIL. As I will show, Google's decision to start using the geo-filtering technology resulted from the discussions and pressure of EU DPAs and particularly in light of court proceedings in France, which eventually led to the Google v. CNIL judgment.

Following Google Spain, the Article 29 Data Protection Working Party (Art. 29WP) [14] and Google with its Advisory Council [15] took opposing views on the delisting territorial scope. Art. 29WP considered that delisting should be implemented globally (i.e., on all domains, including .com.), in light of which it *never discussed* in [14] the possible technical measures, such as geo-filtering. Google's Advisory Council considered that the Court's judgment was not sufficiently precise and insisted on regional domain-based delisting. It supported this view with the practice of automatic referrals to local search engine versions when Internet users in Europe entered "google.com" in their browser, and with Google's claim that more than 95% of user searches from Europe are performed on their local versions. The Council also noted the concerns of opposing interests by both the Internet users outside of Europe and individual users within Europe. Unlike with the Art.29WP, the use of measures such as geo-filtering was here raised - though only by independent experts, who noted that there might be technical possibilities to prevent access to delisted content where searches are made in the EU territory. The Council rejected such option in [15] due to "concerns about the precedent set by such measures, particularly if repressive regimes point to such a precedent in an effort to "lock" their users into heavily censored versions of search results", and the lacking clarity on its effectiveness due to possible circumventions.

Google's domain-based delisting practice, which was limited only to its European versions, soon prompted DPAs reactions. In particular, the French data protection authority (CNIL) ordered a global delisting in 2015 [16]. In March 2016, amid the related court proceedings in France Google voluntarily changed its delisting practice to also include geolocation data pinpointing to users attempting to access relevant information from their EU State of residence, which corresponds to that of the data subject [17] ('glocal' implementation of the right to delisting [18]). According to [17], Google implemented these changes, i.e., the use of geo-filtering to such effect also retroactively, thus to all earlier delistings. Still unimpressed, the CNIL fined Google for failing to comply with its global delisting order and noted that while the new geo-filtering measure does constitute an improvement, it remains incomplete (delisted data can still be consulted by users located outside the territory affected by geo-filtering and it is still possible for users concerned to circumvent it, e.g., by use of the VPN) [19].

In Google v. CNIL the CJEU ruled *inter alia* that the "de-referencing" (English term used for delisting/RTBF in its jurisprudence following Google Spain) is in principle supposed to be carried out in respect of all the Member States, but not in cases of diverging public interests in Member States to access the relevant data. As regards global delisting, the EU law does not impose but it also does not prohibit it, whereby such delisting decisions are left to the national DPAs and the courts, i.e., to their balancing between data protection and freedom of information rights in light of their national fundamental rights standards [20]. A detailed analysis of the many complexities discussed and further raised in and as a consequence of this judgment, which attracted immense academic attention [21], falls outside the scope of this paper. To be pointed here is the Court's brief discussion on the role of "measures" such as geo-filtering (explicitly referred to as geo-blocking in the third preliminary question [22]). Specifically, it held that, if necessary, the search engine „may take sufficiently effective measures to ensure the effective protection of the data subject's fundamental rights", which must „meet all the legal requirements and have the effect of preventing or, at the very least, seriously discouraging internet users in the Member States from gaining access to the links in question using a search conducted on the basis of that data subject's name". While not specifying geo-filtering *per se* and criticized in literature for lack of detail on such an important issue [23], this note shows at least the Court's theoretical acknowledgment of technical measures that may be taken, though on pure discretion of the search

engine, upon balancing between the fundamental rights and according to particular circumstances of each case. However, specific technologies used, implementation details and related circumvention concerns have up to today never been discussed, at least under particular interpretative guidelines issued on those points specifically - most importantly by the EDPB. Additional technical measures to address the circumventions to geo-filtering have been pointed to in literature [24]. In any event, even despite its not being fool proof, the use of geo-filtering technology is considered a more acceptable alternative to global delisting orders, whereby the otherwise complex jurisdictional issues and in particular enforcement concerns are minimized [25].

B. Measures in Cases of Unsucccessful Delisting

In two of its other RTBF judgments the CJEU envisaged specific measures to be taken by search engines in cases of delisting requests, which are otherwise unsuccessful, upon balancing (such as due to prevailing public interest to easily access such data via a name-based search query) or due to lacking proof of the manifest inaccuracy of published content. Thus, in the *GC and Others* judgment (which I examined in detail elsewhere [26]) the CJEU considered the measure to be implemented by the search engine after establishing that the de-listing request should be rejected, even though relevant sensitive data are no longer accurate. More precisely, it concerns situations where it is established, upon balancing, that there is a prevailing public interest to easily access, via a name-based search query, published information relating to criminal proceedings brought against the data subject, concerning an earlier stage of the proceedings and no longer corresponding to the current situation. In such cases search engine operators should at the latest upon such examination of the delisting request adjust, i.e., rearrange search results so that the overall picture provided to internet users reflects the current legal position [27]. Such implementation of the "duty to relist" [28] means in particular that the higher ranking is implemented for search results linking to the webpages containing updated information, so that those links appear on top of the list. Measure therefore presupposes that the legal proceedings ended in favour of the data subject (whilst the currently inaccurate data on the earlier phase remain listed upon name-based search results), which issue should be brought to the attention of the search engine operator upon the delisting request.

Many issues remain unclear as regards this measure [29], such as whether this practice should extend to all of the outdated, thus currently inaccurate personal data, starting from the other sensitive data categories to even ordinary data, and to any other type of related legal proceedings. Furthermore, the CJEU did not provide any legal basis for such practice in its judgment (*obiter dictum*) and the issue was raised on its own volition. It may, however, be argued that such a practice might appropriately implement the GDPR data accuracy principle [30]. On the other hand, the resulting concern is mentioned broadening of its scope under that same principle. Otherwise, it is beyond doubt that the practice affects search engine's operations, which needs to be implemented into its automated ranking systems [31].

One of the noted issues that requires further clarification is that of the assumedly large scope of cases in which the updated information (reflecting current legal position) *was not published* [32]. An approach to resolving this lies in another measure, which the CJEU acknowledged in the subsequent *TU and RE v Google* judgment [10]. Specifically, where the operator does not accept the delisting request based on alleged inaccuracy of content (entire content referred to, or a part of it that is not minor in relation to whole content), which is not obvious (manifest), that operator should upon notice by the data subject add warnings concerning existing proceedings (administrative or judicial) disputing accuracy of such content, in relevant name-based search results [33]. Unlike in the previous case, at least according to circumstances of the case and the judgment itself it is possible to argue more clearly that this measure might have a broader scope of application, i.e., that it could cover the various types of situations and data, the current accuracy of which is contested. Furthermore, while the basic tenets *for asserting* the RTBF under Google Spain continue to apply (such as that neither proof of prejudice nor establishment of e.g., illegality of the relevant publication and removal from the source webpage are required) [34], described measure presupposes existence of at least some preliminary action taken in national proceedings toward disputing accuracy of published information (before DPAs or courts).

Both measures aim to provide a fairer outcome for affected data subjects (who wish to call on them) and appear to be a further application of the GDPR to search engines as *sui generis* data controllers, in the context of their own specific data processing activities, and related responsibilities, powers and capabilities [35]. However, the lacking clarity particularly on applicable legal bases is of concern. While as noted the CJEU did not provide explicit basis for rearrangement of search results, but which might be interpreted as implementation of the data accuracy principle, in TU and RE v Google it specified in [33] the purpose of publishing warnings to be ***inter alia*** that of providing the internet users „with information which continues to be relevant and up-to-date". While the latter concern is logical, allowing for a better informative Internet [36] and which could from that point of view justify both here described measures, the void of the crucial *inter alia* continues to potentiate the head-scratching line of "brushed-over" issues concerning the application of GDPR requirements (such as that of data accuracy, right to rectification and right to restriction of processing) [37] to search engine operators.

III. DE-INDEXING V. DELISTING, DIGITAL ARCHIVES AND TECHNICAL EXPERTISE IN ECHR CASE LAW

Not least on account of lacking EU harmonization in the complex balancing domain between the freedom of information and data protection rights [38] the CJEU's jurisprudence has so far focused strictly on RTBF as exercised by the data subjects before the search engine operators, and on corresponding duties of those operators, regulators and the courts, thereby leaving out the concerns of responsibility of online publishers to themselves take

action by restricting accessibility to disputed content. As I will explain, these issues gained prominence in recent cases before the ECHR. But first it is important to go back to the basics of the RTBF i.e., delisting according to Google Spain, which delineates it from the more traditional legal mechanisms that are directed toward the removal or modification of online content. One of the distinctive features of that right is that it can be asserted against search engine operators regardless of and independently of the data subject's exercise of the more traditional legal institutes against publishers directly (e.g., to have published information removed directly from the source). This is supported by several considerations, all of which are based on reasons of effective and complete protection of *data subjects* [39]. Specifically, they might not be able to succeed in removing their data online for reasons of technical ease of online data reproduction and for jurisdictional reasons where the publisher is not subject to EU law. Also, publishers may keep content online on the basis of their respective laws, including due to prevailing higher regard for the freedom of information [38]. Of relevance here is also the Court's acknowledgment of the fact that it is the publishers who can technically ensure that the data they publish are made unavailable to the search engines, e.g., via exclusion protocols or codes such as 'noindex', although it did not hold such (lack of) action by publishers a condition for establishing the search engine operator's responsibilities for data processing in the context of its own data processing activities, powers and capabilities [40]. Interestingly, in its analysis of the issue if search engine operators may be qualified as data controllers, the Advocate General argued that the only situations in which that operator could be the controller are where it ignored or disobeyed the publisher's (website administrator's) request not to index, i.e. to make the data unavailable for dissemination through the search engine, or where it ignored or disobeyed the publisher's request to update the cache memory (thereby, for example, continuing to show deleted content) [41]. More recently, the EDPB confirmed that search engines operators are themselves bound to carry out actual and full erasure of the URL to third-party content (i.e., de-index) in cases where they disregarded that party's (original publisher's) relevant request (de-indexing / exclusion protocol request) [42]. As a result, such measure has a broader scope than delisting, which is the mere blocking of links displayed in search results following a name-based search query and, as noted earlier, of limited territorial scope.

It is exactly the issue of publishers with online digital archives accessible to the public which is in focus of recent RTBF-related cases before the ECHR, specifically with respect to their role and liability concerning the implementation of noted de-indexing measures. Thus in *Biancardi v Italy* [11] the Court affirmed responsibility of online publishers i.e. owners of online newspaper websites (administrators of newspaper or journalistic archives accessible through the Internet) to deploy "de-indexing" in cases where content of information is itself not challenged nor the way information is published (thereby this case differing from *Węgrzynowski and Smolczewski v Poland* [43] and M.L. and *W.W. v Germany* [44]), but that of *prolonged ease of access to outdated published information via digital archives*. As such, that content is made easily accessible via search engines, thereby harming the individual's right to reputation in the wider context of the right to respect for private life.

In its judgment the Court was commendably attentive to terminological concerns and devoted an entire sub-section of the judgment to the issue of interchangeable uses of de-indexing, delisting and de-referencing concepts in the different sources of EU and international law [45]- which may certainly be confusing. For that reason, technical information on the relevant measure to be employed is vital in order to properly assess its intended aim and scope. It is therefore significant that the Court specifically referred to relevant technical information on de-indexing (as well as other tools enabling the blocking of accessibility of published content via search engines] [46] as a measure available to website owners and thus also to relevant online publishers (administrators of newspaper or journalistic archives accessible through the Internet in this case) [47]. Such technical aspects, most specifically on de-indexing, appear important towards supporting the Court's analysis of facts of the case and its decision that the relevant courts justifiably restricted the publisher's freedom of expression [48]. Specifically, domestic courts found publisher's liability i.a. due to their failure to de-index from the search engine the tags to published article, whereby accessibility to publisher's own content via search engines would be restricted, but would not in itself entail deletion of such content (article).

Discussions are currently developing in the same vein in the referral proceedings before the Grand Chamber in the case of *Hurbain v Belgium* [12]. The relevant court-ordered measure under consideration, which the ECHR Chamber earlier affirmed as proportionate, is that of anonymization of an outdated article available in the newspaper's digital archive. The article contained information on a fatal car accident and included the responsible driver's full name, who was in the meantime formally rehabilitated. In regard to the issue of the less restrictive measures (to anonymization) for the publisher, in the balancing between the right to respect for private life with the freedom of expression, the different content blocking measures such as de-indexing and delisting are discussed. As for delisting, taking into account its effect and key features (RTBF), there is no doubt that it constitutes the least restrictive measure, which only the data subjects may invoke. Interestingly, during the Grand Chamber hearing (09.3.2023) [12] the Belgian Government expressed the position (contrary to that of the publisher) that de-indexing (in the proper sense of adding a no-index tag by the publisher) might in fact be more restrictive for the freedom of the press, as it would lead to "virtual death" of an article (non-availability thereof in the search engines). In contrast, the measure ordered by the judges in national court proceedings (anonymization) would still keep the digital article easily accessible, but in an anonymized form. Technical realities of the online realm with endless redistribution

possibilities of once published content provides good validity (also) to that argument.

Appropriate cognizance of technological options in any case appears critical, particularly since "(evolving) technology can be turned into a virtue as well to better tailor measures and reach more granular verdicts that avoid that radical choices between publicity and deletion must be made" [49]. Taking into account the necessary technical evaluation of the different measures invoked, it is important to point to available documentation and support as provided by the leading search engines.

Thus, according to Google's *current* documentation and policy, website owners may block their article (webpage) from search results (which, as noted also logically include the owners of online newspaper websites) [50]. Currently also *quick removal*s are possible (within a day), but only for those website owners who verified their site in Google's Search Console [51]. These typically last to 6 months [52]. For *permanent removals* (but which do not take effect so quickly), website owners are advised to implement them with a noindex rule [53], which is set with a <meta> tag or HTTP response header, with the result that: "When Googlebot crawls that page and extracts the tag or header, Google will drop that page entirely from Google Search results, regardless of whether other sites link to it" [54]. Upon re-crawling the index, content would no longer be available in search results (options *are available* to speed up the process). It is important to also note the measures to ensure that not only Google but also any other search engine supporting the deindex rule does not index relevant content, and the disclaimer that some search engines *might interpret the noindex rule differently* (which means that content might still appear in their results). Various additional technical options are possible, such as for including (*and excluding*) the *sitelinks search box* in search engine's results, which is promoted as "a quick way for people to search your site or app immediately on the search results page." (…) "Google Search may automatically expose a search box scoped to your website when it appears as a search result, without you having to do anything additional to make this happen. This search box is powered by Google Search. However, you can explicitly provide information by adding WebSite structured data, which can help Google better understand your site (....) If Google Search already exposed a sitelink search box for your site, you can control certain aspects of the sitelink search box by adding WebSite structured data" [55]. Noted possibility of *automatic search-box exposure* by Google is important to take into account in cases of de-indexing.

As regards *non-site owners*, such as private citizens, they cannot use such tools but may submit the *personal data removal request form*s depending on the different types of content/data. They may notify Google also in cases of *outdated content*, if the webpage no longer exists or is significantly different from the current page version. Additionally, and more broadly, content removals are possible for various legal reasons, which include i.a. DMCA copyright violation reports, child sexual abuse imagery, and *requests for removals under the EU RTBF* [56]. Further relevant, but falling outside the scope of this paper are the duties arising from new EU acts such as the Digital Services Act [6] and from the self-regulatory standards, such as those to fight disinformation [57].

Going back to the pending Hurbain v. Belgium case, there appears to be validity in the argument that de-indexing might constitute an even more restrictive measure than article anonymization. However, not least due to entailed consequences of such potential precedent for publisher's operations and freedom of the press in general, the publisher would still prefer that measure to article anonymization. Without prejudice to proper determination of particular circumstances of the case in the current referral proceedings, in theory the revised approach on the balancing and thus on de-indexing, along the lines of that taken in Biancardi v. Italy, appears likely.

IV. Concluding Remarks

Complex relationship between the freedom of information and data protection online, where information once published eventually becomes outdated, irrelevant or incorrect, but still keeps on re-emerging with the simple name-based search queries, and thereby exercising continuing negative effects on affected data subjects, is and will continue to be a growing not only legal, but also general societal and economic problem. Nowadays there is a growing number of paid for services claiming improvement or taking control of one's online reputation, which include employment of various strategies to "push down" negative search results and achieve favorable content prioritizing. The responding legal protection from the point of view of EU data and privacy protection resulted in the RTBF, i.e., the right to delisting, which if successful results in the blocking of search results following a specific and very narrow name-based search queries, and which is implemented by a search engine operator upon request of the data subject. Examined evolution of Google's initially contested but eventually voluntarily and technically even retroactively implemented geo-filtering technology in delisting cases supports the effectiveness of the RTBF when combined with the regional (EU) domain-based delisting. However, the use of such technology and concerns discussed in the paper such as that on possible circumventions, as well as proposed solutions, combined with the related aspects of Google v. CNIL have still not been addressed - at least through EDPB's interpretative initiative. Thereby Google's currently standard delisting approach appears tolerated, possibly in light of its being a "difficult" compromise between the regional, and global domain-based delisting that the EDPB's predecessor insisted on. In any event, from the point of view of data subjects the limited territorial scope is one of the significant drawbacks of the RTBF, i.e., delisting.

If the narrow name-based search query and limited territorial scope of delisting are "sticks", then the CJEU offers the rearranged search results and added warnings on initiated proceedings in search results as "carrots". As such, these measures may lead toward a fairer outcome for those data subjects whose delisting requests were not successful under certain circumstances - and who wish to utilize such options in search engine results. Thereby the

CJEU broadened the scope and type of measures to be taken by very large search engine operators as an arguably further application of the GDPR to them as *sui generis* data controllers (and in the context of their own specific data processing activities, related responsibilities, powers and capabilities), but which still requires proper legal justification. This is so particularly since according to their aim of providing the currently accurate data, such measures are usually directed towards original content producers, also under the more traditional legal institutes.

The role and liability of online publishers in the maintenance (prolongation) of easy access to outdated published information in the digital archives, and their lack of action as regards requested de-indexing to make such content inaccessible to search engines, is an issue gaining prominence in ECHR case law. De-indexing as a technique defined and, in the end effectuated by the search engine operator, at website owner's request (*the functionalities of which may continue to evolve*), signifies the removal of the relevant page from search results regardless of the key word(s) used in the search query. Due to technology involved it also has as an effect the automatic removal of relevant data from search results of more search engines - should they all technically support the deindex rule. When comparing that measure to delisting, it is clear that the latter provides a less restrictive measure in terms of freedom of expression and is at the same much less effective in terms of individuals' data protection and privacy rights, also taking into account its limited territorial scope. In any case upon de-indexing the data subjects normally no longer require delisting from search engines, which is typically consumed by de-indexing and which in any event they might need to separately exercise against more search engines.

On a broader level it has for a while now become clear that any legal solution toward restricting online accessibility of content that is relevant to one's privacy and data protection rights is valueless without interdisciplinary cooperation to support at least the identification, and application of relevant technological developments. Very large internet search engines are certainly key players in that area. In April 2022 Google expanded its policy on removals of personal data that appear in Google Search [58], which it considers to be in line with its aim of helping people take more control of their online presence on its search engine. In fact, it proclaims even more broadly to be continuously "looking for new ways to ensure our policies and built-in safeguards reflect peoples' evolving needs and are easy to use" [59]. This in any case adds to ENISA's early recommendation in [1] on relevant multistakeholder cooperation: "Research communities, industry etc. should develop techniques and coordinate initiatives that aim at preventing the unwanted collection and dissemination of information (e.g., robot.txt, do not track, access control)".

Proper consideration of relevant technological developments and expertise is especially important for relevant regulators and courts in this area. Accordingly, this paper ends with an inspiring note by the German Constitutional Court: "Given that technical developments are ongoing and that they entail uncertainties regarding how and to what extent content providers can influence the dissemination of their articles on the Internet in interaction with search engines, it will fall to the ordinary courts to continue to shape effective and reasonable protective measures. Where reasonable, the courts, which have a considerable margin of appreciation in respect of all these measures, can also require the actors to develop new instruments" [60].

REFERENCES

[1] ENISA, "The right to be forgotten – between expectations and practice," 20.11.2012 (date of deliverable: 18.10.2011), p. 13 (point 4.4.), https://www.enisa.europa.eu/publications/the-right-to-be-forgotten/at_download/fullReport (accessed 06.4.2023).

[2] C-131/12, Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, EU:C:2014:317.

[3] EDPB, "Guidelines 5/2019 on the criteria of the Right to be Forgotten in the search engines cases under the GDPR (part 1)," version 2.0, 07.7.2020, pp. 5, 12 (points 8-9, 49).

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1, Corrigendum [2018] OJ L127/2.

[5] European Commission, "Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines," 25.4.2023, https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413 (accessed 25.4.2023).

[6] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277.

[7] Google, "Right to be Forgotten Overview," https://support.google.com/legal/answer/10769224#zippy=%2Cview-a-list-of-some-common-tools; "Personal Data Removal Request Form," https://reportcontent.google.com/forms/rtbf?visit_id=638185239021320897-1606126095&hl=en&rd=1 (accessed 06.4.2023).

[8] C-507/17 Google LLC v CNIL, ECLI:EU:C:2019:772.

[9] C-136/17 GC and Others v CNIL, ECLI:EU:C:2019:773.

[10] C-460/20, TU and RE v Google LLC, ECLI:EU:C:2022:962.

[11] Biancardi v. Italy (Application no. 77419/16), 25.11.2021.

[12] Hurbain v. Belgium (Application no. 57292/16), 22.6.2021 - referred to the Grand Chamber, hearing held on 09.3.2023: https://www.echr.coe.int/Pages/home.aspx?p=hearings&c= (accessed 20.4.2023).

[13] M. R. Allegri, "Dimenticare, rievocare, rappresentare: dove conduce la via dell'oblio," MediaLaws - Rivista di Diritto dei Media, vol. 2, 2022, pp. 81-123.

[14] Article 29 Data Protection Working Party, "Guidelines on the implementation of the Court of Justice of the European Union judgment on 'Google Spain and inc v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González' C-131/12," WP 225, 26.11.2014, p. 9.

[15] The Advisory Council to Google on the Right to be Forgotten, 06.2.2015, pp. 18-20 (point 5.4.), https://static.googleusercontent.com/media/archive.google.com/en//advisorycouncil/advisement/advisory-report.pdf (accessed 06.4.2023).

[16] CNIL, "La CNIL met en demeure Google de procéder aux déréférencements sur toutes les extensions du moteur de recherche," 12.6.2015, http://www.cnil.fr/linstitution/actualite/article/article/la-cnil-met-en-demeure-google-de-proceder-aux-dereferencements-sur-toutes-les-extensions-du-mote/ (accessed 12.6.2015).

[17] P. Fleischer, "Adapting our approach to the European right to be forgotten," 04.3.2016, https://blog.google/topics/google-

europe/adapting-our-approach-to-european-rig/ (accessed 06.4.2023).

[18] Y. Padova, "Is the right to be forgotten a universal, regional, or 'glocal' right?," International Data Privacy Law, vol. 9, pp. 15-29, 2019 at p. 18.

[19] CNIL, Délibération n° 2016-054, 10.3.2016.

[20] C-507/17 in [8] paras 53-73.

[21] J. Globocnik, "The Right to Be Forgotten is Taking Shape: CJEU Judgments in GC and Others (C-136/17) and Google v CNIL (C-507/17)," GRUR International, vol. 69, pp. 380–388, April 2020; J. Quinn, "*Google v CNIL*: Circumscribing the Extraterritorial Effect of EU Data Protection Law," in Data Protection Beyond Borders: Transatlantic Perspectives on Extraterritoriality and Sovereignty, F. Fabbrini, E. Celeste and J. Quinn, Eds. Oxford: Hart Publishing, 2020, pp. 47–62; G. Frosio, "Enforcement of European Rights on a Global Scale (July 13, 2020)", in Routledge Handbook of European Copyright Law, E. Rosati, Ed. Routledge, 2021, pp. 413-440, http://dx.doi.org/10.2139/ssrn.3650521; P.T.J. Wolters, "The territorial effect of the right to be forgotten after Google v CNIL," International Journal of Law and Information Technology, vol. 29, pp. 57–75, spring 2021; A. Klinefelter and S. Wrigley, "Google LLC v. CNIL: The Location-Based Limits of the EU Right to Erasure and Lessons for U.S. Privacy Law," North Carolina Journal of Law and Technology, vol. 22, pp. 681-734, 2021; I. Hadjiyianni, "The Global Reach of EU Law in the Digital Age: The Territorial Scope of Data Protection," in EU Internet Law in the Digital Single Market, T.-E. Synodinou, P. Jougleux, C. Markou and T. Prastitou-Merdi, Eds. Springer Nature Switzerland AG, 2021, pp. 311-335; E. P. Maat, "Google v. CNIL: A Commentary on the Territorial Scope of the Right to Be Forgotten," European Review of Private Law, vol. 30, pp. 241-262, 2022; M. Taylor, Transatlantic Jurisdictional Conflicts in Data Protection Law. Fundamental rights, privacy and extraterritoriality. Cambridge University Press, 2023, pp. 165-188.

[22] C-507/17 in [8] para 39.

[23] O. J. Gstrein, "Right to be forgotten: european data imperialism., national privilege, or universal human right?," Review of European Administrative Law, vol. 13, pp. 125-152, 2020 at p. 135; Wolters in [21]; J. Quinn, "Geo-location technology: restricting access to online content without illegitimate extraterritorial effects," International Data Privacy Law, vol. 11, pp. 294-306, 2021 at p. 303; Klinefelter and Wrigley in [21] at p. 717.

[24] D. Erdos, "Search Engines, Global Internet Publication and European Data Protection: a New *via Media*?," Cambridge Law Journal, vol. 79, pp. 24-27, 2020 at pp. 26-27; Wolters in [21] at pp. 70-71.

[25] B. V. Alsenoy and M. Koekkoek, "Internet and jurisdiction after Google Spain: the extraterritorial reach of the 'right to be delisted'," International Data Privacy Law, vol. 5, pp. 105-120, 2015 at p. 114; D. J. B. Svantesson, Solving the Internet Jurisdiction Puzzle. Oxford University Press, New York: NY, 2017 at p. 214; Quinn in [23] at pp. 305-306; G. E. M. Perez, "Global or Local? Freedom of Speech and Some Extraterritorial Court Decisions on the Internet," Queen Mary Law Journal, vol. 3, pp. 70-91, 2022 at p. 89; Klinefelter and Wrigley in [21] at pp. 717-718; Taylor in [21] at pp. 177-185.

[26] N. Gumzej, "'The Right to Be Forgotten' and the Sui Generis Controller in the Context of CJEU Jurisprudence and the GDPR," Croatian Yearbook of European Law and Policy, vol. 17, pp. 127–158, 2021 at pp. 142-156.

[27] C-136/17 in [9] para 78.

[28] F. Giovanella, "From the "right to delisting" to the "right to relisting," MediaLaws - Rivista di Diritto dei Media, vol. 2, pp. 124-144, 2022.

[29] Globocnik in [21]; Gumzej in [26] at pp. 154-155.

[30] GDPR in [4] Art. 5(1)(d); Globocnik in [21].

[31] "A guide to Google Search ranking systems," https://developers.google.com/search/docs/appearance/ranking-systems-guide?hl=en (accessed 06.4.2023).

[32] Globocnik in [21].

[33] C-460/20 in [10] para 76 (in connection with paras 68-75).

[34] C-131/12 in [2] paras 84-88, 96, 99.

[35] Gumzej in [26].

[36] Giovanella in [28] at p. 137.

[37] GDPR in [4] Art. 5(1)(d), Art. 16 and Art. 18.

[38] GDPR in [4] Art. 23(1); Art. 85; Art. 17(3)(a); Notifications under Art. 85(3) GDPR: "EU Member States notifications to the European Commission under the GDPR," <https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu/eu-countries-gdpr-specific-notifications_en> (accessed 06.4.2023); D. Erdos, "Special, Personal and Broad Expression: Exploring Freedom of Expression Norms under the General Data Protection Regulation," Yearbook of European Law, vol. 40, pp. 398–430, 2021.

[39] C-131/12 in [2] paras 84-86.

[40] C-131/12 in [2] para 39.

[41] "Opinion of Advocate General Jääskinen," 25.6.2013, ECLI:EU:C:2013:424, paras 91-93, 99.

[42] EDPB in [3] p. 5 (point 10).

[43] Węgrzynowski and Smolczewski v. Poland (Application no. 33846/07), 16.7.2013.

[44] M.L. and W.W. v Germany (Applications nos. 60798/10 and 65599/10), 28.6.2018.

[45] Biancardi v. Italy in [11] paras 53-56.

[46] Biancardi v. Italy in [11] para 50 incl. n. 2.

[47] Biancardi v. Italy in [11] para 51.

[48] Biancardi v. Italy in [11] paras 70-71.

[49] S. Verschaeve, "Going dark or living forever: the right to be forgotten, search engines and press archives," Master's thesis, KU Leuven Faculty of Law, p. 43, 25.5.2020, http://dx.doi.org/10.2139/ssrn.3669865 (accessed 06.4.2023).

[50] "Crawling and Indexing," "Removals," https://developers.google.com/search/docs (accessed 06.4.2023).

[51] "About Search Console," https://support.google.com/webmasters/answer/9128668?hl=en#zippy=%2Cproduct-requirements (accessed 06.4.2023).

[52] "Removals and SafeSearch reports Tool," https://support.google.com/webmasters/answer/9689846 (accessed 06.4.2023).

[53] "Remove a page hosted on your site from Google," https://developers.google.com/search/docs/crawling-indexing/remove-information (accessed 06.4.2023).

[54] "Block Search indexing with noindex," https://developers.google.com/search/docs/crawling-indexing/block-indexing (accessed 06.4.2023).

[55] "Sitelinks search box (WebSite) structured data," https://developers.google.com/search/docs/appearance/structured-data/sitelinks-searchbox (accessed 06.4.2023).

[56] "Remove your personal information from Google," https://support.google.com/websearch/troubleshooter/3111061; "Remove outdated content," https://support.google.com/websearch/answer/6349986; Remove "Outdated Content tool," https://support.google.com/webmasters/answer/7041154; "Personal Data Removal Request Form" in [7] (accessed 06.4.2023).

[57] "The Strengthened Code of Practice on Disinformation 2022," https://ec.europa.eu/newsroom/dae/redirection/document/87585 (accessed 06.4.2023).

[58] "Remove select personally identifiable info or doxxing content from Google Search," https://support.google.com/websearch/answer/9673730 (accessed 06.4.2023).

[59] M. Chang, "New options for removing your personally identifiable information from Search," 27.4.2022, https://blog.google/products/search/new-options-for-removing-your-personally-identifiable-information-from-search/ (accessed 06.4.2023).

[60] BVerfG, 1 BvR 16/13, 06.11.2019, http://www.bverfg.de/e/rs20191106_1bvr001613en.html at para 142 (related see also paras 133-135), accessed 06.4.2023.

Legal Assessment of the National Cybersecurity System in Poland in the Light of the New Developments in the NIS2 Directive

A. Besiekierska
Cardinal Stefan Wyszynski University, Department for Informatics Law, Warsaw, Poland
agnieszka.besiekierska@uksw.edu.pl

***Abstract* – The Polish Act on the National Cybersecurity System was adopted on 5 July 2018. Unlike with the NIS Directive, which it implemented, the scope of its regulation covered public administration. It turned out in the course of its application that the introduced regulations were rather ineffective. It was particularly visible in the case of local governments, which, as indicated by the reports of the Supreme Audit Office, showed numerous shortcomings in the implementation of relevant cybersecurity policies. For more than two years, work has been underway on a draft amendment to the Act on the National Cybersecurity System, which has aroused great controversy, and the ninth draft has now been published. Discussions are taking place primarily in the area of the supply chain cybersecurity and include its geopolitical aspects, which will undoubtedly be of great importance for the further development of the cybersecurity system in Poland and Europe.**

Keywords – cybersecurity; national cybersecurity system; NIS2 Directive

I. INTRODUCTION

In recent years, the number of ICT legislation has increased significantly in the EU, especially in the area of cybersecurity. It includes the NIS Directive, its successor – the NIS2 Directive, the Cybersecurity Act, sectoral acts, such as the DORA, as well as the proposed EU Cyber Resilience Act. A large number of legal acts cause problems with the absorption of the EU legislation at the national level. A negative example could be the protracted works in Poland on the implementation of the European Electronic Communications Code which had been planned to be completed in 2021 and are still pending (as of 19 March 2023). This intense legislator's activity is often referred to as the legislative tsunami [1].

In the maze of the existing and planned legal acts, the NIS Directive, and its successor, i.e. the NIS2 Directive, are the first cross-sectoral legal acts in the area of cybersecurity, as well as a starting point for establishing the National Cybersecurity System in Poland. The paper outlines the main principles of the NIS and NIS2 Directives and points out some pros and cons for the existing legal provisions in the context of the planned amendment of the Polish National Cybersecurity System. The research was based on the analysis of legal acts and literature, including reports on the implementation of obligations in the field of cyber security issued by the Supreme Audit Office, as well as the analysis of the position papers raised in the pre-legislative process for the draft amendment of the Act on the National Cybersecurity System.

II. NIS DIRECTIVE

Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive), adopted on 6 July 2016, imposed a number of obligations on the Member States concerning the establishment of specific institutions and cooperation mechanisms. Each Member State was obliged to set up the National Competent Authorities for network and information security, whereas the function could be performed by the already existing institution or institutions. The task of the National Competent Authority was to monitor the implementation of the provisions of the Directive at the national level throughout all sectors addressed by the regulation. Those Competent Authorities had the power to investigate cases of non-compliance in the field of network and information security, issue IT security guidelines as well as impose sanctions for non-compliance [2].

In order to strengthen the Member States' cooperation, each Member State had to establish a Single Point of Contact. Its task was to collect information on incidents on a national scale, and exchange information on significant international incidents with its foreign counterparts. The NIS Directive also provided for the creation of CSIRTs, i.e. a Computer Security Incident Response Teams. Member States could designate one country-wide CSIRT or build a network of sectoral CSIRTs covering market sectors [3].

The NIS Directive provided for two types of entities, their role being to fulfil cybersecurity obligations: operators of essential services from among those listed in Annex II to the Directive (energy, transport, banking, financial markets, health sector, drinking water supply and distribution, digital infrastructure) and digital service providers (online marketplace, online search engine, cloud computing service, as indicated in Annex III). The operators of essential services were required to assess the risk of cyber threats and to adopt appropriate measures to

ensure network and information security. They also had to report any incidents seriously threatening their networks and IT systems to the competent authorities. Incidents bearing a significant impact on the continuity of operations of operators were subject to mandatory reporting, which meant that the reporting thresholds were to be determined by the Member States in the process of implementing the provisions of the Directive. Digital service providers were subject to the so called "light touch" approach. It consisted in *ex post* supervisory activities, i.e. following an incident and only by the country where the service provider was based [4]. It is worth noting that the communications undertakings had not been included within the scope of the NIS Directive and remained subject to sector specific regulation, i.e. Directive 2002/21/EC (Framework Directive) [5]. The NIS Directive was to be transposed into the Member States' national laws by 9 May 2018.

III. NATIONAL CYBERSECURITY SYSTEM

In Poland, the NIS Directive was implemented by the Act of 5 July 2018 on the National Cybersecurity System, which created the National Cybersecurity System, and which is still the most important Polish legal act in the field of cybersecurity. The National Cybersecurity System provides for the competent authorities for network and information security, three CSIRTs, essential services operators, digital service providers and public administration entities [6]. Unlike the NIS Directive, which did not apply to public administration services, the Polish legislator decided to include the public sector. That was possible due to the fact that the NIS Directive served as a minimum harmonization framework which meant that it only set out certain minimum conditions that had to be met [7]. With their number amounting to approximately 3,000, public entities are currently the largest group within the National Cybersecurity System [8].

A significantly less numerous group are the essential services operators and digital service providers, with probably nearly 400 operators of essential services and about 50 providers of digital services [9]. The numbers are not exact, as the operators are designated and placed on the list of operators of essential services by the competent minister, and both the decision and the list of operators of essential services are confidential.

In practice, it turned out that the entities covered by the National Cybersecurity System do not fully fulfil the obligations arising from the Act. To a large extent this applies to public entities, which, despite the binding legal obligations to ensure the data security and integrity, often fail to comply with basic security principles and are exposed to cyberattacks. In its audit reports, the Polish Supreme Audit Office (NIK) negatively assessed the performance of tasks related to ensuring the security of the processed information, indicating that the public entities, especially local governments, lacked a systemic approach to ensuring information security. The public entities did not have information about their IT resources, did not perform risk assessments, as well as did not carry out an annual audit. 48% of them failed to make backups, improperly stored the backups or did not check the correctness of the copies made [10]. Failure to comply with basic security principles makes the public entities highly vulnerable to cyberattacks. As bad as it is, the situation may to a large extent result from the technical and organizational problems related to the high costs of professional IT support. Neither do the low financial fines which may be imposed for infringements on the basis of the Act support law enforcement.

IV. NIS2 DIRECTIVE

The Commission was obliged to periodically review the functioning of the NIS Directive and to report to the European Parliament (Article 23(2) of the NIS Directive), whereas the report on the results of the first review was to be presented by 9 May 2021. As part of the review activities, the Commission identified numerous issues relating to the implementation of the NIS Directive. First of all, it identified significant differences in the implementation and identification of essential services operators in the Member States. Since the Directive established a minimal legal framework for harmonization, each Member State could enjoy considerable freedom in implementing its provisions. As a result, there were different thresholds for identifying operators of essential services as well as different thresholds for reporting ICT incidents. That resulted in a highly fragmented market, with operators providing services in several Member States having to comply with different legal regimes. Further, the scope of the NIS Directive was found to be insufficient as it did not address certain sectors and operators, despite their key importance from the point of view of cybersecurity [11]

As a result of the review and the discussions which followed, the European legislator decided to adopt a new Directive, Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union (NIS2 Directive), which replaced the NIS Directive. The NIS2 Directive extends the subjective scope of its predecessor to cover new sectors and include, among others, public administration, food, chemical, sewage, industry (production of medical devices and medical devices for *in vitro* diagnostics, computer, electronic and optical products, electrical equipment, machinery and equipment, motor vehicles, trailers and semi-trailers, other transport equipment), waste management and space, and treats some sectors more broadly (e.g. expanding the scope of digital infrastructure). Unlike the NIS Directive, which distinguished between essential services operators and digital service providers, NIS2 divides entities from the above mentioned sectors into two categories: essential entities and important entities. Communications undertakings, i.e. providers of public electronic communications networks and/or providers of publicly available electronic communications service, fall within the scope of the NIS Directive as essential entities [12]. Such approach may potentially result in losing the experience with legal, technical and economic aspects of security which had been built in the telecommunications sector framework for over 10 years.

As regards the terms of management, operation and disclosure of security vulnerabilities, testing cybersecurity

levels and effective use of encryption, entities covered by the NIS2 Directive are subject to much stricter requirements than before [13]. The new Directive also includes more precise provisions in the field of incident reporting. In addition, Member States may require essential and important entities to obligatorily certify products, services and processes in accordance with the European certification schemes provided for under the Cybersecurity Act. Undoubtedly, mandatory certification would contribute to increasing cybersecurity. However, taking into account the long lasting works on European cybersecurity certification schemes and the controversies that those works raise, in particular in relation to cloud computing services, one may be concerned that such mandatory cybersecurity certification will not be introduced soon. Also the fact that the NIS2 Directive, similarly to the Cybersecurity Act, does not deal with the issue of cybersecurity certification in relation to digital resources can be considered a drawback [14].

Another new element is the introduction of a coordinated risk assessment of critical supply chains at the EU level. Following consultations with the Cooperation Group and ENISA, the European Commission may identify critical ICT services, systems or products that may be subject to a coordinated risk assessment which will correspond to the one that was carried out for the 5G network [15].

To ensure better law enforcement, the NIS2 Directive provides for high financial penalties for entities which fail to implement its provisions in a proper manner. Those penalties amount to a maximum of at least EUR 10.000.000 or up to 2% of a company's total annual worldwide turnover, whichever is higher. A novelty is the introduction of the company management's responsibility for compliance with cybersecurity risk management measures. Further, the NIS2 Directive also introduced provisions increasing the role of the European Commission, which will assess the implementation of the Directive in the Member States every 18 months, whereas the review should also address such issues as funds allocated to cybersecurity, resources and development of cybersecurity capabilities. Thus, unlike in the case of the NIS Directive, it will not only serve as an assessment of the transposition of the provisions of the Directive itself.

The NIS2 Directive has to be transposed into national laws by 17 October 2024.

V. PLANNED AMENDMENT TO THE NATIONAL SECURITY SYSTEM

The work on the amendment of the Polish Act on the National Cybersecurity System has been underway since 2020, with the first draft act published in September 2020 and the most recent, the ninth one, in January 2023. The amendment process has been carried out independently of the adoption of the NIS2 Directive and the planned provisions do not implement it. The release of the drafts was often preceded by public consultations, including hundreds of position papers, which makes the amendment process incomparably more intense than any other before. The non-completion of this long-lasting process hinders the announcement of the auction for 5G frequencies, as the planned amendment should introduce the security measures referred to in the EU 5G Toolbox. Due to the delayed assignment of the 5G frequencies, Poland is at the very end in Europe when it comes to the use of the 5G technology [16].

The most significant and, at the same time, the most controversial change introduced by the draft act amending the Act on the National Cybersecurity System relates to the development of the 5G network in Poland. It provides for a new status of communications undertakings, i.e. providers of communications services and networks, which are to be included in the National Cybersecurity System. The inclusion of communications undertakings has been combined with the planned provisions on the control of the supply chain. As part of the supply chain control, a procedure is to be introduced to identify a supplier as a high-risk one. The procedure would be initiated by the minister competent for computerization and might apply to a software or equipment supplier. When issuing a decision, the minister would take into account an analysis, prepared by a special collegial body, which would cover both technical as well as non-technical and political aspects, such as economic, intelligence and terrorist threats to national security or threats to the implementation of allied and European commitments, or the likelihood that the hardware or software supplier remains under the control of a country outside the EU or NATO territory. Should a decision to identify a supplier as a high-risk one be issued, the entities being part of the National Cybersecurity System would not be allowed to use ICT products and software within the scope covered by the decision and provided by the high-risk supplier, and would have to withdraw those used within a certain time limit (5 or 7 years). The decision would be immediately enforceable and courts would not have the capacity to stop its execution. In the intense public debates, which may be fuelled by different interest groups, strong opinions have been voiced that the planned changes raise controversies concerning their constitutionality, in particular regarding the freedom of economic activity or openness of proceedings, as well as those indicating the geopolitical situation in which the World is getting polarized along the West – Far East axis [17].

For the first time, the draft amendment creates a national cybersecurity certification system and defines rules and a procedure for cybersecurity certification of ICT products, ICT services or ICT processes, and thus implements the provisions of the EU Regulation Cybersecurity Act. By way of a regulation, the Council of Ministers may determine a national cybersecurity certification program for a given product, service or process, taking into account the need to develop requirements in accordance with current scientific and technical knowledge and with the aim of increasing cybersecurity in Poland. certification system corresponds with the provisions the Cybersecurity Act.

VI. CONCLUSION

The Polish legislator was ahead of the European one by including public entities within the national cybersecurity system. However, the Polish Supreme Audit Office (NIK) audits indicate a low level of cybersecurity

compliance by those entities. Currently, trying to control the maze of cybersecurity regulations, the Polish legislator has chosen a strategy that has not been successful so far. It is still dealing with 5G security requirements, creating a legal framework allowing the exclusion of hardware and software from high-risk vendors, while it is already time to start implementing the NS2 Directive. Taking into account the problems with the amendment of the Act on the National Cybersecurity System, the question arises how long it will take to implement the NIS Directive. Any delay in the transposition of the NIS2 Directive is harmful, as the Directive contains numerous legal solutions conducive to better law enforcement.

At the same time, the example of the works relating to the amendment shows how closely today's cybersecurity issues are related to global politics and how the political aspects may be even more important than the technical ones. This will probably result in works on any subsequent legal acts addressing cybersecurity issues, such as the aforementioned EU level coordinated risk assessment of critical supply chains even more complex and time-consuming.

REFERENCES

[1] A. Besiekierska, J. Mazur, A. Mednis, J. Mojsiejuk, W. Paluszyński et al., "The law is no substitute for common sense", Prawo nie zastąpi zdrowego rozsądku, Domena , 2022, vol. 2, pp. 4-9.

[2] S.Schmitz-Berndt, P. G. Chiara, "One step ahead: mapping the Italian and German cybersecurity laws against the proposal for a NIS2 directive", Int. Cybersecur. Law Rev., 2022, no. 3, pp.294-295.

[3] G. Szpor, "European regulation of network and information systems security and state sovereignty", Europejska regulacja bezpieczeństwa sieci i systemów informacyjnych a suwerenność państwa, in G. Szpor, A. Gryszczynska, Internet. Strategie Bezpieczeństwa, Warsaw, C.H. Beck, 2017, pp. 13-15.

[4] D. K. Kipker, "The EU NIS Directive compared to the IT Security Act- Germany is well position for the new European Cybersecurity Space, ZD-Aktuell 2016, 05363.

[5] S. Piątek, "Obligations of telecommunications undertakings related to cybersecurity", Obowiązki przedsiębiorców telekomunikacyjnych w zakresie cyberbezpieczeństwa, internetowy Kwartalnik Antymonopolowy i Regulacyjny, 2020, vol. 9 , no 2, pp. 28-41.

[6] P. Wajda, "Cybersecurity – sectoral aspects of the regulation", Cyberbezpieczeństwo – sektorowe aspekty regulacyjne, internetowy Kwartalnik Antymonopolowy i Regulacyjny, 2020, vol. 9 , no 2, pp. 14-26.

[7] A.Besiekierska, "Commentary to the Act on National Cybersecurity System", Ustawa o krajowym systemie cyberbezpieczństwa, C.H. Beck, Warsaw, 2019, p.103.

[8] Regulatory impact assessment (OSR) of 17 January 2023, https://mc.bip.gov.pl/projekty-aktow-prawnych-mc/630873_projekt-ustawy-o-zmianie-ustawy-o-krajowym-systemie-cyberbezpieczenstwa-oraz-ustawy-prawo-zamowien-publicznych.html., pp. 8-13, [access: 19.03.2023].

[9] Regulatory impact assessment (OSR) of 17 January 2023, pp.-8-13.

[10] Information on the results of the NIK au-dit, Implementation of public services for citizens using the ePUAP platform, Informacja o wynikach kontroli NIK, Realizacja usług publicznych dla obywateli zwykorzystaniem platformy ePUAP , 2021, https://www.nik.gov.pl/kontrole/P/20/004/ [access: 19.03.2023], Information on the results of the NIK audit, Information security management in local government units, Informacja o wynikach kontroli NIK, Zarządzanie bezpieczeństwem informacji w jednostkach samorządu terytorialnego, 2019 https://www.nik.gov.pl/kontrole/P/18/006/[[access: 19.03.2023], Information on the results of the NIK au-dit, Information security at remote work and mobile data processing (Informacja o wynikach kontroli NIK, Bezpieczeństwo informacji w pracy na odległość imobilnym przetwarzaniu danych, 2022, https://www.nik.gov.pl/kontrole/P/21/081/LOL/ [access: 19.03.2023].

[11] Z. Bederna, Z. Rajna, "Analysis of the cybersecurity ecosystem in the European Union," Int. Cybersecur. Law Rev., no. 3, 2022, pp. 45–46,

[12] A. Piechocki, K. Gorzkowska, "Planned changes in cybersecurity", Planowane zmiany w cyberbezpieczeństwie, PNT 2021, no 1,

[13] T. Sievers, "Proposal for a NIS directive 2.0: companies covered by the extended scope of application and their obligations", Int. Cybersecur. Law Rev., 2021, vol. 2, pp. 225–227.

[14] S.A. Salvaggio, S.A., N. González, "The European framework for cybersecurity: strong assets, intricate history". Int. Cybersecur. Law Rev., 2023, vol. 4, pp. 140–142; D. D. Stewart Ferguson, "European Cybersecurity Certification Schemes and cybersecurity in the EU internal market", Int. Cybersecur. Law Rev., vol. 3, 2022, pp. 54-55.

[15] EU Coordinated Risk Assessment published October 9, 2019 https://digital-strategy.ec.europa.eu/en/news/eu-wide-coordinated-risk-as-sessment-5g-networks-security, acces: 19.03.2023.

[16] Kearney 5G Readiness Index 2022: https://www.kearney.com/telecommunications/article/-/insights/time-is-tight-for-telcos-on-5g-strategies-even-as-the-european-rollout-lags , [access: 19.03.2023].

[17] A. Besiekierska, "Legal Aspects of the Supply Chain Cybersecurity in the Context of 5G Technology", Review of European and Comparative Law, 2022, vol. 51, no. 4, 129–147; R. Siudak, "Cybersecurity in Poland. From discourses to public policies", Cyberbezpieczeństwo w Polsce, Od dyskursów do polityk publicznych, Księgarnia akademicka, Cracow, 2022, pp. 165–170.

Online Audiovisual Content, Video Sharing Platforms and Regulation under DSA and AVMSD

H. Lisičar, T. Katulić i M. Jurić
Faculty of Law, University of Zagreb, Croatia
hrvoje.lisicar@pravo.hr, tihomir.katulic@pravo.hr, marko.juric@pravo.hr

Summary - In light of recent changes of regulatory framework for provision of digital services within the EU, this paper analyses interplay and overlapping of Digital Services Act and Audiovisual Media Services Directive with regard to the regulation of online audiovisual content. Implementation of DSA raises several questions which need to be clarified in order to enhance the legal certainty of the enforcement of EUs online content regulation. Both the Regulation and the Directive include provisions that pertain to video sharing platforms and their content. The paper examines their scope and the differences, especially in relation to protection of users and their rights.

Keywords – audiovisual media services, digital services, video sharing platforms, media, DSA, media content

I. INTRODUCTION

Over the past decade, the landscape of digital services and audiovisual media has changed dramatically due to technological advancements and changes in consumer behaviour. The European Union (EU) has recognized the need to ensure that the regulatory framework remains relevant and more efficient. As part of the two broader strategies, one for strengthening the digital single market[1] and the other for shaping the EU's digital future,[2] in recent years, there were several legislative initiatives pertaining to online audiovisual content and video sharing platforms. At the EU level, for several decades, audiovisual media content is regulated by the Audiovisual Media Services Directive 2010/13/EU (AVMSD).[3] Main goal of this Directive is to harmonize the rules and to ensure free reception and distribution of audiovisual services across borders while providing Member Stats with significant freedoms to regulate those services in line with their economic and cultural background. The AVMSD regulates main audiovisual players that are providing services under editorial responsibility, like television and video on-demand services and, since the last adaptation of AVMSD 2018/1808 in 2018,[4] the video-sharing platform (VSP) providers that do not fall under the scope of editorial responsibility but provide their services while organising the user generated audiovisual content that is distributed through their platforms.[5] With that, AVMSD is currently regulating different types of audiovisual content dissemination where some of its provisions refer only to certain types of providers, depending on the responsibilities and distribution method to the consumers.[6] More recently, another specific regulatory instrument, Digital Services Act Regulation (EU) 2022/2065 (DSA)[7] came into force, so AVMSD is no longer the only relevant instrument governing the dissemination of audiovisual content. DSA introduces more comprehensive platform regulation, which partly also includes the players (VSPs) already addressed by the AVMSD. In addition, DSA also regulates platforms as intermediaries which are of great importance for the distribution of audiovisual content. In this paper, we will discuss the interplay and overlapping of these two legislative frameworks, analysing their key provisions, objectives, and implementation in the EU member states. Furthermore, we will explore the new challenges that arise from these legal frameworks and their impact on the dissemination of online audiovisual content. The starting point is an in-depth analysis of relevant provisions of the DSA and AVMSD with regard to the scope of application for providers of online content.

II. THE SCOPE OF DIGITAL SERVICES ACT

As previously mentioned, DSA is a new legal instrument that was enacted in the EU as part of the "Digital Services Act package" together with the Digital Markets Act Regulation (EU) 2022/1925 (DMA).[8] Where DSA is concerned mostly with free speech and content focus aspects, DMA's focus is on economic aspects of digital services, mainly competition and antitrust issues.

The DSA came into force on 16 November 2022 and will be fully applicable for all entities in its scope across the EU from February 2024, except for the provisions that pertain to the obligation of online platforms that have 3 months (until February 2023) to report the number of active end users on their websites which will be taken into account when European Commission (EC) makes an assessment[9] as to whether a platform should be designated a very large online platform (VLOP) or search engine. Following such a designation decision by the Commission, the entity in question will have 4 months to comply with the obligations under the DSA, including

carrying out and providing the Commission with the first annual risk assessment exercise.[10]

The main objective of DSA is to enhance the responsibility of platforms in the dissemination of online content.[11] Since the introduction of the E-commerce Directive (ECD)[12] in 2000, online platforms have been mainly defined and regulated under the term information society services. ECD has also introduced the liability exemption regime for three types of information society services as intermediary services, where the service can be considered as a mere conduit, caching, or hosting service. Several Member States took a somewhat different approach when defining information society service and liability exemptions while transposing the ECD in their national regulatory frameworks, which then led to numerous cases where the Court of Justice of the EU (CJEU) had to interpret the definition of information society services[13] and provide guidelines on the application of the liability exemption.[14] Furthermore, in recent years, to answer the new challenges that are connected to the dissemination of audiovisual content in various sectors, some other specific definitions of different types of online platforms have emerged, like for example the notion of video sharing platforms in AVMSD. Following all those developments, in technology advancements, provisions of services and new regulations, where the digital landscape has changed dramatically, it was clear that the existing regulatory framework, mainly ECD, was not adequate and had to be revised and improved. Before the introduction of the new regulatory framework, European Commission (EC) conducted an impact assessment analysis[15] in which three possible options for revision of the existing regulatory framework were assessed. The first option included the introduction of limited measures against illegal activities, laying down the procedural obligations for online platforms to tackle illegal activities, in order to protect users' fundamental rights and ensure transparency. The second option considered the implementation of fully harmonised measures to incentivise actions from service providers, enhance transparency and address a wider set of emerging risks by empowering users. The third and preferred option was the implementation of asymmetric measures with stronger obligations for very large online platforms, further clarifications of the liability regime for online intermediaries and EU governance with reinforced oversight and enforcement. [16] Advantage of this approach is that platforms of different size, scope and influence can be subject to different regulation. In a final version of a new regulatory framework for digital services third option was implemented. ECD is kept in force (except Art. 12-15 of the ECD) with DSA rules modifying it, additional rules for online platforms that are not part of other regulations and directives are introduced, a new *ex ante* regime is established with DMA Regulation and sector-specific rules are still valid.[17]

Building upon existing provisions in ECD, DSA took a horizontal approach to regulate different platforms, meaning that, a targeted set of uniform, effective and proportionate mandatory rules[18] shall apply to all digital service providers operating in the EU, regardless of the type of service they offer or their specific business model. Expanding on previous liability exemptions for intermediaries in ECD, the novelty of DSA is that it introduces a layered approach to defining intermediary services. Art. 3(g) of the DSA defines three types of providers of intermediary service: a 'mere conduit' service, a 'caching' service, and a 'hosting' service. Additionally, hosting providers can be further categorised as online platforms according to Art. 3(i) or very large online platforms (VLOPs). This layered approach to intermediary services provides a different set of rules for hosting services, online platforms, and very large online platforms (VLOPs), where each category is subject to different obligations and responsibilities, depending on their role and level of influence in the online environment. "Simple" hosting services refer to intermediaries that provide storage services for user-generated content. They are typically not involved in the moderation or curation of the content and do not provide access to the public. Under the DSA, these hosting services would be subject to a set of general obligations that pertain to all hosting services[19] and are related to the designation of points of contact for Member States' authorities and recipients of the service, transparency, notice and action mechanisms, user complaints, and cooperation with law enforcement authorities.

The next step in this layered approach is the regulation of online platforms that are also hosting services, but they enable users to access and share content with others, where the platform stores and disseminates content to the public, like social media platforms, video-sharing platforms, and online marketplaces. Under the DSA, besides general obligations applicable to all hosting services, online platforms are subject to a broader set of obligations related to content moderation (Art 20, 23), dispute settlement (Art. 21), trusted flaggers (Art. 22), interface design and organisation (Art. 25), advertising (Art. 26.) and protection of minors (Art. 28.).

Finally, VLOPs are the largest digital intermediaries with a huge number of users (which have a number of average monthly active recipients of the service in the Union equal to or higher than 45 million - Art. 33 DSA) and most influence on society, such as Facebook, Google, YouTube etc. They are subject to an even stricter set of rules than regular online platforms and have more obligations under the DSA. These obligations pertain to identifying, analysing and assessing and mitigating any systemic risks (Art. 34, 35), they are subject to independent audit, providing at least one option for each of their recommender systems (Art. 38), additional online advertising transparency (Art. 39), they have to enable access to data that are necessary to monitor and assess compliance with DSA (Art. 40), they have to establish a compliance function (Art. 41) and have a transparency reporting obligation with regards to content moderation that they engaged in (Art. 42).

In summary, this layered regulation of intermediary services under the DSA aims to establish a different set of

rules and obligations that are appropriate for different types of intermediaries, depending on their role and level of influence in the society and digital landscape.

With regards to the territorial scope of DSA, Art. 2(1) states that it applies to intermediary services offered to recipients of the service that have their place of establishment or are located in the Union, irrespective of where the providers of those intermediary services have their place of establishment. This approach is similar as in some other instruments addressing the provision of digital services for recipients in the EU (most notably, GDPR).[20] Furthermore, this approach reflects the legislator's intention to protect EU citizens and their fundamental rights and is further specified in Art. 3(d) which provides a definition of meaning 'to offer services in the Union as enabling natural or legal persons in one or more Member States to use the services of a provider of intermediary services that has a substantial connection to the Union. A substantial connection to the Union, as per Art. 3(e), can result either from the provider's establishment in the EU or from specific factual criteria such as a significant number of European users or the targeting of activities towards one or more Member States.

With regards to the content, the main goal of the DSA is to harmonize the rules applicable to intermediary services with the objective of ensuring a safe, predictable, and trusted online environment addressing the dissemination of illegal content and dissemination of harmful content like disinformation,[21] while at the same time introducing safeguards for freedom of speech and other fundamental rights such as the right to an effective remedy, non-discrimination, rights of the child as well as the protection of personal data and privacy online. To achieve the above-mentioned goals, DSA introduces measures like 'voluntary initiative investigations and legal compliance' conducted by intermediaries, so called “Good Samaritan” clause, that do not preclude well-known established liability limitation regime i.e., "safe harbour".

III. ONLINE AUDIOVISUAL CONTENT AND THE SCOPE OF AVMSD

The AVMSD is the main regulatory instrument that governs the cross-border transmission of audiovisual media services in the EU. As a Directive, it establishes a minimum set of rules to which providers of audiovisual media services in all Member States must adhere. The original scope of the Directive (previously named Television without Frontiers Directive - TWFD) applied to the broadcasting of television, but in 2007[22] it was extended to on-demand audiovisual media services (VOD). Since providers of VOD services are, amongst others, online platforms, this was the first extension of the scope of the Directive to a certain type of services and content provided online, due to the developing media landscape. The second revision of AVMSD in 2018 further extended the scope to online services and online audiovisual content, introducing the provisions pertaining to video sharing platforms (VSP) and user-generated content (UGC). With this revision, for the first time, AVMSD is extended to online services and providers that do not have editorial responsibility.

AVMSD defines VSPs as a service whose principal purpose or of a dissociable section thereof or an essential functionality of the service is devoted to providing programmes, user-generated videos, or both, to the general public, for which the video-sharing platform provider does not have editorial responsibility, in order to inform, entertain or educate, using an electronic communications network.[23] It is clear that this definition also covers and pertains to hosting intermediaries.

With regards to the online audiovisual content provided by VSPs, Art. 28(b) of AVMSD 2018/1808 imposes the obligations on providers to take appropriate measures to protect minors from programmes, user-generated videos and audiovisual commercial communications which may impair their physical, mental, or moral development in accordance. Furthermore, they have to take appropriate measures to protect the general public from programmes, user-generated videos and audiovisual commercial communications containing incitement to violence or hatred directed against a group of persons or a member of a group based on any type of discrimination and from programmes, user-generated videos and audiovisual commercial communications containing content the dissemination of which constitutes an activity which is a criminal offence under Union law, like a provocation to commit a terrorist offence, offences concerning child pornography and offences concerning racism and xenophobia.[24] These measures are without prejudice to Articles 12 to 15 of ECD and they can't lead to any ex-ante control measures or upload filtering of content.[25] They comprise revised terms of service, the introduction of flagging systems and reporting of illegal and inappropriate content, a clear declaration of functionalities of VSPs, implementation of age verification or rating and control systems, establishing transparent and easy to use measures to resolve complaints and provision of effective media literacy tools.[26]

The territorial scope of AVMSD 2018/1808 extends the country of origin principle, a well established system for all audiovisual media content providers since the introduction of TWFD, to VSPs. This means that VSPs established on the territory of any Member State are under the jurisdiction of that Member State. If there is a case where VSP is not established on the territory of a Member State of the EU, it will still be under the jurisdiction of a certain Member State if it is part of a parent undertaking or subsidiary undertaking that is established on the territory of that Member State. It is the same case if it is part of a group and another undertaking of that group that is established on the territory of the Member State.[27]

Considering all the above-mentioned AVMSD provisions that govern on-demand audiovisual media services, video sharing platforms and online audiovisual media content, it is clear that AVMSD and DSA share a

lot of common grounds when online platforms and online audiovisual content regulation are concerned. Although the legislator clearly states that DSA is without prejudice to the rules laid down by other Union legal acts regulating other aspects of the provision of intermediary services, including AVMSD,[28] interplay between these two instruments and possible overlapping in their implementation must be considered.

IV. INTERPLAY AND OVERLAPPING OF DSA AND AVMSD

Laying out the scope of the DSA and AVMSD and provisions that pertain to the regulation of online platforms and online audiovisual content raises several questions which must be clarified to enhance the legal certainty of the enforcement of online content regulation.

As previously mentioned, and as stated in Recital (10) and Art. 2 of the DSA, it is not intended for DSA Regulation to affect sectoral legislation like AVMSD, rather it should complement them. It should address issues that are not addressed or not fully addressed in AVMSD. This makes a DSA *lex generalis* in relation to other specific sectoral regulations, including AVMSD. Although the relationship between these two documents is clearly established, there are still some unanswered questions and possible situations of legal uncertainty.

One of those questions can be linked to the regulation of VSPs since they should also be treated as online platforms as defined by DSA. AVMSD as a Directive imposes certain obligations on Member States to regulate VSPs regarding the provision of audiovisual content. The rules applicable to VSPs imposed by certain Member States can vary in their scope, detail and strictness. In cases where those imposed provisions are stricter and more detailed should they be reviewed, in order to comply with DSA or will the provisions of DSA as a Regulation prevail. Furthermore, VSPs can provide other ancillary services that do not constitute audiovisual media content but are closely linked to it and are not covered by AVMSD. It is clear that for these types of services DSA applies, but does this also means that jurisdiction and oversight of such VSPs will be divided between different regulatory bodies or only one regulatory body will be authorized to oversee all of the VSP's services, regardless of services that it provides. Also, when considering the provision of online audiovisual content, we should consider some specific forms of services, for example, streaming services. Those services are regulated by AVMSD. Since the goal of DSA's horizontal approach is to regulate all digital services, in which category of intermediaries should streaming services fall.[29] It is possible that streaming services in certain cases and depending on the means of provision of those services won't be covered by the DSA's definition of 'hosting service' or 'online platform'. This means that in certain cases all of the provisions pertaining to those types of services (notice and action mechanisms, internal complaint-handling system, trusted flaggers etc.) won't be applicable, although it is clear that they should be.

Regarding the provisions of AVMSD pertaining to the implementation of measures to protect minors and the general public from illegal and harmful content, AVMSD sets out general obligations for VSPs and Member States. Contrary to that, DSA as lex generalis, appears to have more precise demands when imposing measures to combat illegal and harmful content. When comparing provisions pertaining to VSPs and the publication of their terms and conditions, AVMSD requires that protection of minors and the general public from illegal and harmful content, and qualitative rules for audiovisual commercial communications should be included. On the other hand, Art. 14. of the DSA have more precise requirements, like the inclusion of information on any restrictions that they impose in relation to the use of their service in respect of information provided by the recipients of the service (policies, procedures, measures and tools used for the purpose of content moderation). In applying and enforcing these restrictions providers of services must act with due regard to the rights and legitimate interests of all parties involved, including the fundamental rights of the recipients of the service, such as the freedom of expression, freedom and pluralism of the media, and other fundamental rights and freedoms as enshrined in the Charter.[30]

When considering online advertising, provisions of AVMSD must be also supplemented with DSA's requirements in Art 26. for advertising on online platforms. Furthermore, with regards to provisions pertaining to mechanisms for reporting and flagging certain online content, Art. 28 (3)(d, e) of the AVMSD in general terms sets out obligations for VSPs to establish such systems and inform the users on their effects, while DSA in Art 16. includes more elaborate rules on the implementation and functioning of those systems. Furthermore, Art. 22. of the DSA imposes the obligation on providers to establish the systems to be able to receive notices submitted by trusted flaggers that are acting within their designated area of expertise.[31] Similar situation, where DSA's provisions are more concise and elaborate can be also found in the case of transparency obligations of platforms, users' complaint systems and out-of-court settlement of disputes.[32]This situation where *lex generalis* rules are more detailed and elaborate than rules set out in *lex specialis* can cause some practical problems in their implementation. Firstly, AVMSD as a Directive is already transposed by the Member States since it already came into force and implemented measures can be more detailed and stricter than the measures referred to in AVMSD. Secondly, AVMSD measures must be applied to all VSPs in a practicable and proportionate manner while taking into account the nature of the content, the harm it may cause, the characteristics of the category of persons to be protected, as well as the rights and legitimate interests at stake. On the other hand, DSA pertains to a specific category of intermediaries where certain provisions are applied according to their categorization and in the case of online platforms, their reach.

The provisions of the AVMSD relating to the establishment and territorial jurisdiction for VSPs that are not established in a Member State also significantly differ from the provisions in the DSA, which can also cause difficulties in their practical implementation. Art. 28a (2) of the AVMSD provides a criterion for determining the jurisdiction of VSP in cases where VSP is established in a third country but has a connection to a Member State via its parent or subsidiary undertaking or an undertaking member of the same group. At the same time Art. 56 (4) of the DSA stipulates that in case of a provider that does not have an establishment in the EU must appoint a legal representative and the Commission or Member State where that representative resides or is established shall have powers, to oversee and enforce the relevant obligations. This leads to a situation where the assessment of jurisdiction would have, in some cases, been conducted according to both legal instruments. It is also easy to imagine the situation where an online platform that also constitutes a VSP and does not have an establishment in the EU but provides services to the EU citizens has appointed a legal representative in one Member State, but also has a connection to another Member State via its parent or subsidiary undertaking.[33] Since AVMSD implements the country of origin principle, for that VSP jurisdiction must be determined under the AVMSD rules.

Finally, with regards to the online audiovisual content and liability exemption for providers, the interplay between these two instruments becomes once again apparent. Art. 28b of the AVMSD 2018/1808 stipulates that provisions pertaining to VSPs' obligations to implement measures to protect minors and the general public are without prejudice to articles 12 to 15 of the ECD (replaced by DSA Art. 4,5,6 and 8). This means that "safe harbour" still applies to VSPs.[34] We can take a look at YouTube as an example, which definitely qualifies as a VSP but also qualifies as VLOP under the DSA. For relevant online audiovisual content and audiovisual commercial communications that are hosted and provided by YouTube, the platform will have to apply all the measures from AVMSD to protect minors from content that may impair their physical, mental or moral development and the general public from illegal content, however, DSA will be applicable to the extent that the AVSMD or other EU legal acts, such as a Regulation on addressing the dissemination on terrorist content online,[35] do not contain more specific provisions.[36]

V. CONCLUDING REMARKS

The scope of the Digital Services Act (DSA) and Audiovisual Media Services Directive (AVMSD) raises several questions that must be clarified to improve the legal certainty of the enforcement of online content regulation. Although the relationship between the two documents is established, there are still uncertainties, particularly with regard to regulating video sharing platforms (VSPs). While AVMSD provides general obligations for VSPs, DSA imposes more precise requirements to combat illegal and harmful content, including the inclusion of information on restrictions that online platforms impose in relation to their service. Additionally, DSA's provisions on mechanisms for reporting and flagging content, transparency obligations of platforms, users' complaint systems, and out-of-court settlement of disputes are more elaborate than those set out in AVMSD. The differences in the establishment and territorial jurisdiction for VSPs between the two documents may also cause difficulties in their practical implementation. This creates the potential for confusion and legal uncertainty, particularly for smaller platforms that may struggle to comply with varying regulatory regimes across different jurisdictions.

Therefore, further harmonization of these provisions is necessary to ensure a consistent and effective online content regulation across the EU. This will be the task of competent authorities that will be responsible for the supervision of providers of intermediary services and enforcement of these instruments.

It is worth noting that the issues raised by the relationship between the DSA and AVMSD are not the only ones that require clarification. As online content regulation continues to evolve, new questions and challenges will arise that must be addressed in order to ensure legal certainty and effective enforcement. Especially with regard to media content. For that purpose, a legislator has already come out with a proposal of several new legislative instruments like for example European Media Freedom Act.[37]

In conclusion, while the DSA and AVMSD provide important frameworks for regulating online platforms and online audiovisual content, their relationship requires further clarification in order to enhance legal certainty and ensure effective enforcement. It is important that policymakers and regulators continue to work together to address these issues and ensure that online content regulation remains effective and up-to-date in the face of evolving challenges.

References

[1] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions *"A Digital Single Market Strategy for Europe"*, COM/2015/0192 final.

[2] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions *"Shaping Europe's digital future"*, COM/2020/67 final.

[3] Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive), OJ L 95, 15.4.2010, p. 1–24.

[4] Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities, OJ L 303, 28.11.2018, p. 69–92.

[5] Buiten, Miriam, The Digital Services Act: From Intermediary Liability to Platform Regulation (June 21, 2021). Available at SSRN: https://ssrn.com/abstract=3876328, p5.

[6] Etteldorf Christina Cole D. Mark, , Future Regulation Of Cross Border Audiovisual Content Dissemination, Media Authority of North Rhine-Westphalia, available at: https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Presse/Pressemitteilung/Gutachten_ExSum_lang_EN_Cole_2023_2.pdf [3.2.2023].

[7] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022, p. 1–102.

[8] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance), OJ L 265, 12.10.2022, p. 1–66.

[9] Art. 33. DSA

[10] Art. 34. DSA

[11] Cole D. Mark., Overview of the impact of the proposed EU Digital Services Act Package on broadcasting in Europe, Institute of European Media Law, p.3. available at: https://emr-sb.de/wp-content/uploads/2021/06/EMR_Legal-Issue-Paper-DSA-DMA-and-Broadcasting_Summary.pdf [5.2.2023].

[12] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), OJ L 178, 17.7.2000, p. 1–16.

[13] For example, in Case C-434/15 Press and Information Asociación Profesional Elite Taxi v Uber Systems Spain SL

[14] op.cit. 9, Cole D. Mark, p.7., for example in joined cases C-236/08 to C-238/08 against Google France SARL.

[15] European Commission, Impact Assessment, Digital Services Act package: deepening the Internal Market and clarifying responsibilities for digital services, available at: https://ec.europa.eu/newsroom/dae/redirection/document/72158 [7.2.2023.]

[16] Ibid.

[17] See more: Savin A., The EU Digital Services Act: Towards a More Responsible Internet, Copenhagen Business School. Aspen Publishers, March 2021, 15-25, available at: https://www.proquest.com/docview/2528510222

[18] Rec. (4) DSA

[19] Chapter 3, section 1 and 2 of the DSA.

[20] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), OJ L 119, 4.5.2016, p. 1–88

[21] Geiger C.,Frosio G. Taking Fundamental Rights Seriously in the Digital Services Act's Platform, Liability Regime, European Law Journal December 2022 available at: https://www.researchgate.net/profile/Giancarlo-Frosio/publication/350320059_Taking_Fundamental_Rights_Seriously_in_the_Digital_Services_Act's_Platform_Liability_Regime_European_Law_Journal_2022_forthcoming/links/60638e6e92851cd8ce7ad122/Taking-Fundamental-Rights-Seriously-in-the-Digital-Services-Acts-Platform-Liability-Regime-European-Law-Journal-2022-forthcoming.pdf

[22] Directive 2007/65/EC of the European Parliament and of the Council of 11 December 2007 amending Council Directive 89/552/EEC on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities (Text with EEA relevance) OJ L 332, 18.12.2007, p. 27–45

[23] Art. 1. (b)(aa) AVMSD 2018/1808

[24] Art 28b p.1 AVMSD 2018/1808

[25] Art 28b p.3 AVMSD 2018/1808

[26] ibid.

[27] Art 28a p.2 AVMSD 2018/1808

[28] Art. 2. p. 4 DSA

[29] About non-hosting intermediaries see more: Schwemer, Sebastian Felix and Mahler, Tobias and Styri, Håkon, Legal Analysis of the Intermediary Service Providers of Non-hosting Nature (July 1, 2020). Final report prepared for European Commission, University of Copenhagen Faculty of Law Research Paper No. 2021-108, Available at SSRN: https://ssrn.com/abstract=3798494

[30] Art 14.(4) DSA

[31] van Hoboken, Joris V. J. et.al. Putting the Digital Services Act Into Practice: Enforcement, Access to Justice, and Global Implications (March 10, 2023). Amsterdam Law School Research Paper No. 13, 2023, Institute for Information Law Research Paper No. 03, 2023, Verfassungsbooks. https://doi.org/10.17176/20230208-093135-0., available at: https://ssrn.com/abstract=4384266

[32] About out of court mechanisms see more: Husovec M, Will the DSA Work?, On Money and Effort, ibid.

[33] For conflict of laws with regards to territorial scope see more: Lutzi, T., The Scope of the Digital Services Act and Digital Markets Act: Thoughts on the Conflict of Laws February, 2023. Dalloz IP/IT, Forthcoming, Available at SSRN: https://ssrn.com/abstract=4367531

[34] For exemption of platforms see more: Genç-Gelgeç, B.Regulating Digital Platforms: Will the DSA Correct Its Predecessor's Deficiencies?, Croatian Yearbook of European Law and Policy, Faculty of Law – University of Zagreb, 2022, p 57

[35] Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online (Text with EEA relevance), OJ L 172, 17.5.2021, p. 79–109

[36] Similar situation can be found in relationship between DSA and CDSM Directive as explained in: Quintais J.P and Schwemer S.F., The Interplay between the Digital Services Act and Sector Regulation: How Special Is Copyright?, European Journal of Risk Regulation (2022), 13, 191–217

[37] Proposal for a Regulation of the European Parliament and of the Council establishing a common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU, COM/2022/457 final

Interplay of Gatekeepers' Obligations and Consumer Rights under the Digital Markets Act

I. Kanceljak*
* Faculty of Law, University of Zagreb, Department of Civil Law, Zagreb, Croatia
ivana.kanceljak@pravo.hr

Abstract – **Although the Digital Markets Act primarily aims to protect small undertakings in the digital sector in their relations with Big Tech undertakings by introducing contestable and fair markets, consumers as end users can also be considered important participants. The Digital Markets Act provides that large undertakings, if they meet the prescribed conditions, can be designated as gatekeepers with a special legal position. The main characteristic of this legal position concerns several obligations which they must adhere to while providing core platform services. Those obligations provide a better legal position for small undertakings but it is questionable if they also create benefits for consumers as end users.**

This article aims to explain the relation between gatekeepers` obligations and the legal position of consumers in the digital sector. It shows that it is possible for a normative regulation, originally aimed at providing contestable and fair markets, to bring many benefits to consumers. This can be seen through new special rights for consumers and even support of the rights that already exist for consumers through general consumer protection law.

Keywords – Digital Markets Act; gatekeeper; gatekeepers` obligations; consumer; consumer rights

I. INTRODUCTION

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (hereinafter: Digital Markets Act) [1] entered info force on 1 November 2022 and will apply from 2 May 2023. The Digital Markets Act and Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2] are two European Commissions` important legislative initiatives known under name "The Digital Services Act Package". [3].

The Digital Markets Act can be considered a legislative response to the imbalance between a small number of big undertakings and a vast number of small undertakings on the internal market. This imbalance is reflected through unfair practices and unfair competition imposed by large undertakings ("Big Tech" companies such as Google, Microsoft, YouTube, Meta, and Apple). [4] On the one hand, large undertakings which provide core platform services have gained considerable economic power, a better negotiating position, connections with a vast number of business users and end users (including consumers), which gives them access to large amounts of data. [5] On the other hand, small undertakings are faced with many barriers, no matter how innovative they are.

The main goal of the Digital Markets Act is to create rules aimed to ensure contestability and fairness for the markets in the digital sector in general. [6] To achieve this goal, large undertakings can be designated as "gatekeepers" pursuant to the Digital Markets Act, which means that they would have a new legal position determined by various obligations. Although the main goal of the Digital Markets Act is to interfere in relations between undertakings, it must not be forgotten that consumers are also important participants on the internal market. This article aims to investigate to what extent the Digital Markets Act takes into account consumers as end users, explain how their position has been changed, and to highlight actual benefits for consumers having in mind existing legal solutions designed for the protection of consumers.

II. LEGAL POSITION OF GATEKEEPERS AFTER THE DIGITAL MARKETS ACT

A. The term "gatekeeper"

The Digital Markets Act provides a set of rules applicable to undertakings which will be designated as gatekeepers by the European Commission. First of all, the Digital Markets Act defines the term undertaking "as an entity engaged in an economic activity, regardless of its legal status and the way in which it is financed, including all linked enterprises or connected undertakings that form a group through the direct or indirect control of an enterprise or undertaking by another" (Article 2 paragraph 27 of the Digital Markets Act).

Upon designation, such gatekeepers would be obliged to provide their services legally, aligned with *ex ante* regulation provided in the Digital Markets Act. [7] This means that the Digital Market Act provides special regulation which, by creating gatekeepers` obligations, from the point of view of EU competition law, tries to solve problems by preventing them [8].

Such regulation is applicable to undertakings with a place of residence both in the EU and outside of the EU if a core platform service is offered to business users or end users established or located in the Union (Article 1 paragraph 2 of the Digital Markets Act), designated as

gatekeepers. It is explicitly provided that electronic communications networks and services are excluded form application (Article 1 paragraph 3 of the Digital Markets Act).

The preamble to the Digital Markets Act emphasizes that "gatekeepers have a significant impact on the internal market, providing gateways for a large number of business users to reach end users everywhere in the Union and on different markets". [9] In addition, Article 2 of the Digital Markets Act defines a gatekeeper as an "undertaking providing core platform services". A core platform service means any of the following services: online intermediation services, online search engines, online social networking services, video-sharing platform services, number-independent interpersonal communications services, operating systems, web browsers, virtual assistants, cloud computing services along with online advertising services, including any advertising networks, advertising exchanges and any other advertising intermediation services, provided by an undertaking that provides any of the previously listed core platform services. [10] Some authors have highlighted that this list of services is a result of "the Commissions' deepest concerns about contestability and fairness". [11]

B. Designation of Gatekeepers

A new competence for the European Commission is to designate an undertaking as a gatekeeper. The Digital Markets Act introduces two different designation procedures. In the first one, the undertaking notifies the Commission (Article 3 paragraph 3 of the Digital Markets Act), and in the second one the Commission starts the procedure on its own initiative and in accordance with the powers provided in the Digital Markets Act (Articles 16 and 17). The goal of both procedures is to determine if an undertaking meets the thresholds set out in Article 3 paragraphs 1 and 2 of the Digital Markets Act.

Article 3 of the Digital Markets Act regulates three thresholds which must be met to designate an undertaking as a gatekeeper. First of all, an undertaking must have a significant impact on the internal market. It will be presumed that this prerequisite is met if an undertaking "achieves an annual Union turnover equal to or above EUR 7.5 billion in each of the last three financial years, or where its average market capitalisation or its equivalent fair market value amounted to at least EUR 75 billion in the last financial year, and it provides the same core platform service in at least three Member States" (Article 3 paragraph 2 (a) of the Digital Markets Act). Secondly, an undertaking must provide such core platform service "which is an important gateway for business users to reach end users". Again, it is presumed that this prescription is met if the undertaking "in the last financial year has at least 45 million monthly active end users established or located in the Union and at least 10 000 yearly active business users established in the Union, which will be identified and calculated in accordance with the methodology and indicators set out in the Annex to the Digital Market Act" (Article 3 paragraph 2 (b) of the Digital Markets Act). Thirdly, an undertaking must "enjoy an entrenched and durable position, in its operations, or it is foreseeable that it will enjoy such a position in the near future". It will be presumed that this last prerequisite is achieved if the above mentioned thresholds (from Article 3 paragraph 2 (b) of the Digital Markets Act) are met in each of the last three financial years.

C. Gatekeepers` Obligations

Upon designation, the undertaking`s legal position will significantly change. The rules that will apply to gatekeepers after designation can be found in Chapter III of the Digital Markets Act entitled "Practices of Gatekeepers that Limit Contestability or are Unfair". Although this chapter sets out numerous gatekeepers` obligations along with the obligation to comply with those obligations, it also regulates possible suspensions of certain obligations in exceptional situations (Article 9 of the Digital Markets Act). In order to create a functional system, the European legislator has envisaged special powers for the European Commission [12], among which of great importance are decisions on imposing fines [13].

For the purposes of this paper, gatekeepers` obligations are divided into three main categories. [14] Obligations in the first category mostly concern end users' personal data and processing of such data (Article 5 paragraph 2 and Article 6 paragraphs 2, 10 and 11 of the Digital Markets Act). The main goal to be achieved with this regulation for end users is to provide a free choice to opt in to data processing and also to change their mind concerning the consent given for data processing. [15]. Since the goal of this research is to discuss end users` legal position as consumers, those gatekeepers` obligations will not be further elaborated.

The second category of gatekeepers` obligations consists of obligations for gatekeepers concerning interoperability of number-independent interpersonal communications services (Article 7 of the Digital Markets Act). Those are specific obligations relating to only interpersonal communications services. [16] Although they do provide a better position for end users concerning their safety and choices to use a certain service, those rules do not take a general approach applicable to all services in the digital sector or have a connection to classical consumer rights therefore they will also be excluded from further analysis.

The third category of gatekeepers` obligations consists of numerous obligations principally regulating relations between gatekeepers and small companies as business users offering goods and services to other end users, both undertakings and consumers (Article 5 paragraphs 3-10 and Article 6 of the Digital Markets Act). Provisions regulating those obligations are taken as core gatekeepers` obligations relevant for this research. Hypothetically, they can directly or indirectly affect end users' position concerning their consumer rights. In the following research, the question on the relation of gatekeepers' obligations and the legal position of end users as consumers will be discussed from the point of view given through these obligations.

III. CONSUMERS AND THE DIGITAL MARKETS ACT

A. *The term "consumers"*

The Digital Markets Act does not define the term "consumer". Concerning the fact that the goal of this regulation is not primarily consumer protection, such an approach is hardly surprising. Some authors have emphasized that protection of consumers in the Digital Markets Act is only a reflex of the regulation aimed to provide a better position of business users. [17]

On the other hand, the Digital Markets Act does define the end user as "any natural or legal person using core platform services other than as a business user" (Article 2 paragraph 20 of the Digital Markets Act). Overall, this definition encompasses consumers who are, from the point of view of this paper, understood as consumers as defined in general legal sources of consumer protection law. In that sense, consumer means any natural person acting for the purposes which are outside his trade, business, craft or profession. Such general definition of the term "consumer" derives from Article 2 paragraph 1 of Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament [18], which has been amended by Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules [19].

B. *Legal solution for consumers under the Digital Markets Act*

Two legal solutions arising from the Digital Markets Act directly affect the legal position of consumers. One might say that they also indicate that the Digital Markets Act is actually quite oriented towards providing protection of consumer rights. Both legal solutions establish a legal framework for legal actions available to consumers in cases of breach of their rights arising from the Digital Markets Act.

The first significant provision amends Directive (EU) 2020/1828 of the European Parliament and of the Council of 25 November 2020 on representative actions for the protection of the collective interests of consumers and repealing Directive 2009/22/EC (hereinafter: Directive (EU) 2020/1828). [20] Article 42 of the Digital Markets Act provides that "Directive (EU) 2020/1828 shall apply to the representative actions brought against infringements by gatekeepers of provisions of this Regulation that harm or may harm the collective interests of consumers". [21] Without such amendment it would be impossible to even use representative action mechanisms for the protection of collective interests of consumers. That is apparent from the regulation of the scope of application of Directive (EU) 2020/1828 while representative actions brought against infringements by traders are possible only in explicitly regulated situations. [22]

Another legal solution is provided along with gatekeepers` obligations in Article 5 paragraph 6 of the Digital Markets Act. Under this provision "the gatekeeper shall not directly or indirectly prevent or restrict business users or end users from raising any issue of non-compliance with the relevant Union or national law by the gatekeeper with any relevant public authority, including national courts, related to any practice of the gatekeeper". In addition, the Digital Markets Act in Article 51 regulates the amendment to Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law (hereinafter: Directive (EU) 2019/1937). [23] The main goal of Directive (EU) 2019/1937 is to provide minimum standards for the protection of persons reporting breaches of Union law. Sources of law are, for the purposes of this specific regulation, divided into ten different categories. [24] Those categories must be interpreted by way of Part I of the Annex to Directive (EU) 2019/1937 which gives a list of legal sources divided into the said categories, providing the exact legal sources that can be breached and that are within the scope of the directive. In that sense, it should be noted that Article 51 of the Digital Markets Act amends only the category of legal sources concerning protection of privacy and personal data, and security of network and information systems (Part 1 of the Annex of Directive (EU) 2019/1937, part J.), but on the other hand, it does not amend the category concerning legal sources aimed to protect consumers (Part 1 of the Annex to Directive (EU) 2019/1937, part I.). Nevertheless, the amendment set out in Article 51 of the Digital Markets Act should be interpreted in favor of consumers, in a way that it is possible to report even breaches of provisions affecting the legal position of consumers as established by the Digital Markets Act (which will be further elaborated below). Also, it has to be emphasized that a person entitled to report a breach (Article 4 regulates the personal scope of Directive (EU) 2019/1937)[25], known as the "whistleblower", need not be a consumer. The legal position of "whistleblowers" is strictly connected with a work-based relationship and it is not necessary that he is also a consumer in a specific situation.

C. *Consumer rights in relation to gatekeepers` obligations*

It has been stated above that gatekeepers' obligations can be categorized into three different categories, while the third category refers to the core obligations that are general obligations for gatekeepers. [26] Those obligations form the basis for further research.

A special analysis of the provisions that are considered to regulate the core obligations has been carried out in order to gain more specific insight into the

relation between gatekeepers` obligations and the legal position of the consumer. To achieve this goal, two tests for the provisions were created.

The first test focused on the content of the norm. The first question in the test was: Does the provision create a subjective right for the consumer, likely to be followed by a claim or a request? If the answer was positive, it was necessary to describe such right, but if it was negative, a new question was asked. The second question was: Does the provision strengthen or give support to the consumer right or principle that already exists in consumer protection law? In the case of a positive answer, it was explained how it affects consumer rights. On the other hand, if the answer was negative, an additional question was asked: Does the provision in any way affect consumer rights? If so, this relation was explained having in mind that this effect can also be negative.

The second test concerned the relation of gatekeepers` obligations and consumers to find out how those provisions in general relate to consumer rights. The test consisted of only one question: To what extent do the provisions regulating gatekeepers` obligations contribute to a better legal position of consumers in the digital sector? It was possible to choose between three answers: a) the obligation contributes to a considerable extent, b) the obligation contributes to a limited extent, and c) the obligation does not contribute at all.

Teleological methodology was used in both tests, while they were based on the interpretation of the goals each provision achieves and having in mind the relation to consumers. The total number of obligations interpreted was 19 (n=19; Art. 5 paragraphs 3-10 (p. 9 and 10 were taken as one obligation concerning advertisement) created 7 gatekeepers` obligations while Art. 6 paragraphs 3-9, 12-13 (p. 3, 4, and 7 were taken as two obligations) created 12 gatekeepers` obligations).

The results of the first test led to several conclusions. Nine obligations (47.3 %) were found to create new specific rights for consumers (Article 5 paragraphs 4, 5 and 6; Article 6 paragraphs 3, 4, 7, 9 and 13 of the Digital Markets Act). This means that:

- consumer has a right to conclude contracts directly with traders or through a third-party online intermediation service regardless of the core platform that the consumer usually uses;
- the consumer may request access and use through any core platform service content, subscriptions, features or other items regardless of the core platform usually used or used to conclude a contract,
- consumer has a right to raise any issue of non-compliance with the relevant Union or national law without any negative consequences,
- consumer has a right to easily uninstall any software applications on the operating system of the gatekeeper, without prejudice to the possibility for that gatekeeper to restrict such uninstallation in relation to software applications that are essential for the functioning of the operating system or of the device and which cannot technically be offered on a standalone basis by third parties,
- consumer has a right to easily change default settings on the operating system, virtual assistant and web browser of the gatekeeper,
- consumer has a right to install and use third-party software applications or software application stores and set that downloaded software application or software application store as their default,
- consumer has a right to access the open internet and this right should not be limited by any technical means, all in connection with a right to switch between different undertakings providing internet access service,
- consumer has a right to achieve portability of data provided or generated through his activity by using the core platform service,
- consumer has a right to terminate a core platform service.

One obligation (5.2 %) supports consumer rights that are already known in consumer protection law (Article 6 paragraph 5 of the Digital Markets Act). It is important to emphasize that those core platform services that provide services which are the same as those provided by business users using the same platform core service (for example Amazon while it offers its own products along with other users products) are not allowed to give a better position in the ranking or indexing of their own products. This means that consumers will have a more transparent way of being informed on the offers according to the principle of transparency, they will be better informed and, finally, they will be able to make a reasonable decision for the conclusion of contracts. [27]

Eight obligations (42.1 %) do not create new rights for consumers and they do not support existing consumer rights but, in a way, do concern consumer rights (Article 5 paragraphs 3, 7, 8, and 9, Article 6 paragraphs 4, 6, 8, and 9 of the Digital markets Act). In that sense, several benefits for consumers arise from the Digital Markets Act. Overall, consumers will have the right to choose from different products and services in the digital sector and also to choose between different business users (traders) providing those products and services, regardless of their basic platform core service. They will be able to compare terms and conditions from even more traders and to choose those that they find the most suitable for their needs. For consumers, possibilities to choose will be more extensive while the market will become more open and the use of services more flexible, while some of them will not be bound directly to certain core platform services. From the technical point of view (i.e. concerning hardware and software), the possibility to freely choose will not be limited. Those conclusions support a very important and relatively new notion of "consumer choice", previously expressed by the ECJ in *Google and Alphabet v Commission (Google Shopping)*. [28] Some authors have already concluded that the Digital Markets Act creators had in mind this notion during the legislative procedure and explained that

"protecting the sovereign decision-making capacity of consumers becomes increasingly important in the digital world". [29] Also, beside consumer choice, some gatekeepers` obligations, such as those in case of more transparent marketing, might lead to a reduction of traders` costs and even a reduction of the final prices.

The first test showed that only one obligation (5.2 %) does not have any effect on consumer rights and consumers' legal position in general. It can also be concluded that gatekeepers` obligations do not bring into existence or undermine any existing consumer rights.

The second test also led to important conclusions on how gatekeepers` obligations contribute to a better legal position of consumers in the digital sector. The research has shown that nine gatekeepers` obligations (47.3 %) contribute to a considerable extent, nine obligations (47.3 %) contribute to a limited extent and only one (5.2 %) does not contribute at all to a better position of consumers.

IV. CONCLUSION

In general, the objective of regulation of the Digital Markets Act is the creation of a completely new legal order in the relations between Big Tech companies and small undertakings. It mainly consist of rules prohibiting abuse of rights and an overall better position of big undertakings toward small undertakings as business users. Along with that, it creates a framework for the functioning of this new system of rules by giving new competences to the European Commission. At first glance, the Digital Markets Act could lead to a conclusion that it does not bring much to consumer rights, but that conclusion would be wrong.

Two important amendments to the Digital Markets Act directly affect the position of consumers. Firstly, the amendments to Directive (EU) 2020/1828 open a new path for collective redress as an important instrument of consumer protection. Without such regulation, such protection would not even be possible. Secondly, any breach of gatekeepers` obligations may result in an issue of non-compliance. According to amendments to Directive (EU) 2019/1937, "whistleblowers" could report breaches of provisions provided in the Digital Markets Act including those provisions that aim to regulate the legal position of consumers as end users.

If the Digital Markets Act is seen only from the point of view of consumer, it can be said that consumers are not the main focus. In that sense, the Digital Markets Act does not regulate the term "consumer" but only the term "end user". End users can be both business users and consumers. Also, there are no provisions regulating consumers proper. The research on the legal position of consumers was based on an extensive interpretation of gatekeepers` obligations in order to determine how each provision interferes with the legal position of consumers. To obtain more concrete answers it was important to categorize gatekeepers` obligations and to separate those that have the most general scope and to classify them under the name "core gatekeepers` obligations". Those obligations were the basis for two tests applied to find exact answers. One test dealt with the content of the norm and the other concerned the overall position of the consumer after the Digital Markets Act.

The research has shown that nine obligations of the total of 19, create new rights for consumers. One gatekeepers` obligation supports the existing consumer right – the right to be informed along with the principle of transparency. Eight obligations affect the legal position of consumers with an emphasis on the free choice of consumers while only one obligation should not have any impact on consumers. Also, obligations should not have any negative effect on the existing framework for the protection of consumers.

The second test aimed to emphasize a possible positive impact on consumers' legal position in general. The results of the second test showed that nine gatekeepers` obligations contribute to a better legal position of consumers to a considerable extent, nine obligations contribute to a limited extent and only one does not contribute at all.

Both tests gave comparable data, which leads to some general conclusions. Although the Digital Markets Act initially cannot be seen as a legal source of consumer rights, it significantly changes their legal position. Consumers have also been faced with many barriers in the digital sector while using one platform service, which led to the impossibility to use services provided from other platform core services. New consumer rights will remove many obstacles, create a more flexible environment to gather information, conclude, terminate or even change contracts, to switch between core platform services or adjust digital content and services to their needs.

REFERENCES

[1] OJ L 265, 12.10.2022, p. 1–66.

[2] OJ L 277, 27.10.2022, p. 1–102; see more: Cauffman, C.; Goanta, C., A New Order: The Digital Services Act and Consumer Protection, European Journal of Risk Regulation, Volume 12 , Issue 4 , December 2021 , pp. 758 – 774; Busch, C.; Mak, V., Putting the Digital Services Act in Context: Brigding the Gap Between EU Consumer Law and Platform Regulation, EcCML, vol. 10, issue 3, 2021, pp. 109-114.

[3] See more: https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package (accessed: 12 January 2023).

[4] See: Recitals 2, 3 and 4 of Preamble of Digital Markets Act; R. Podszun, P. Bongartz, S. Langenstein, The Digital Markets Act: Moving from Competition Law to Regulation for Large Gatekeepers, EUCML, Vol. 10, Issue 2, 2021, pp. 60, 61.

[5] Recital 3 of Preamble of Digital Markets Act.

[6] Recital 7 of Preamble of Digital Markets Act. See also: Chiarella, M. L., Digital Markets Act (DMA) and Digital Services Act (DSA): New Rules for the EU Digital Environment, Athens Journal of Law, Volume 9, Issue 1, January 2023, pp. 37-39.

[7] On the concept of *ex ante* regulation see: H. Schweitzer, The Art to Make Gatekeeper Positions Contestable and the Challenge to Know What is Fair: A Discussion of the Digital Markets Act Proposal, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3837341 (03 January 2023), pp. 15, 24.

[8] Kreber, W., Taming tech giants with a per-se rules approach? The Digital Markets Act from the "rules vs. standard" perspective, Law & Economics, Concurrences No.3, 2021, p. 28, 29.

[9] Recital 6 of Preamble of Digital Markets Act. Schneider, H., The European Union`s Digital Markets Act Seeks to Regulate Competition with Little Regard to Impact on Consumers, available at: https://cei.org/studies/the-european-unions-digital-markets-act-seeks-to-regulate-competition-with-little-regard-to-impact-on-consumers/ (accessed: 12 January 2023).

[10] For definitions of the terms see article 2 paragraph 3-13 of Digital Markets Act.

[11] R. Podszun, P. Bongartz, S. Langenstein, op. cit., p. 63. See also: Recital 15 of Preamble of Digital Markets Act.

[12] See more in Chapter V "Investigative, Enforcement and Monitoring powers", articles 20 – 46 of Digital Markets Act.

[13] See more in article 30 of the Digital Markets Act.

[14] For other possible categorizations see: Monti, G., The Digital Markets Act: Improving Its Institutional Design, European Competition and Regulatory Law Review (CoRe), vol. 5, no. 2, 2021, p. 91.

[15] Recitals 36 and 37 of the Preamble of Digital Markets Act.

[16] Recital 64 of the Preamble of Digital Markets Act.

[17] Podszun, R., The Digital Markets Act: What's in it for Consumers?, Journal of European Consumer and Market Law, Volume 11, Issue 1, 2022, pp. 4.

[18] OJ L 304, 22.11.2011, p. 64–88.

[19] OJ L 328, 18.12.2019, p. 7–28.

[20] OJ L 409, 4.12.2020, p. 1–27. See more: Inchausti,F. G., A new European way to collective redress? Representative actions under Directive 2020/1828 of 25 November, Zeitschrift für das Privatrecht der Europäischen Union, Volume 18, Issue 2, 2021, pp. 61-80.

[21] See also article 52 of Digital Markets Act along with recital 104 of the Preamble of Digital Markets Act.

[22] See article 2 paragraph 1 along with Annex I of the Directive (EU) 2020/1828.

[23] OJ L 305, 26.11.2019, p. 17–56.

[24] See article 2 paragraph 1 (1) of Directive (EU) 2019/1937.

[25] See more: Stappers, J. T., EU Whistleblower Protection Directive: Europe on Whistleblowing, ERA Forum volume 22, 2021, pp. 87-100.

[26] See under: C. Gatekeepers` Obligations.

[27] On the importance of informing the consumer see more: Busch, C., The future of pre-contractual information duties: from behavioural insights to big data u: Twigg-Flesner, C. Research Handbook on EU Consumer and Contract Law, Edward ElgarPublishing, Cheltenham, UK + Northampton, MA, USA, 2016, p. 221-240.

[28] Judgement of 10 November 2021, Google and Alphabet v Commission (Google Shopping), T-612/17, ECLI:EU:T:2021:763, p. 588.

[29] Podszun, R., The Digital Markets Act: What's in it for Consumers?, *op. cit.*, p. 3.

Analysis of IoT Devices Security for Household Applications

M. Milenković, R. Nikolić, S. Čelan, P. Petrošanec
University of Zagreb, Faculty of Transport and Traffic Sciences/Chair of Transport Law and Economics, Zagreb, Croatia
mmilenkovic@fpz.unizg.hr

Abstract - According to predictions that the world consumption of Internet of Things (IoT) devices will exceed 772 billion dollars this year and considering the increasing spread of IoT devices and the number of investments in them, the paper will present the authors' view on the security of IoT devices. The paper is based on the conducted research through a quantitative analysis of the familiarity and safety of the use of IoT devices in the household. The paper intends to diagnose the growing potential risks that could threaten the security of IoT households across the EU. As one of the instruments for the protection of end users, the paper will instruct legislators on the minimum technical and security requirements IoT devices should use for end users to have more protection, and security in the rapidly growing category of household devices. The security of IoT devices will be analyzed based on the known attacks on IoT devices and the survey conducted among the population of citizens of the Republic of Croatia. *De lege ferenda*, legislators will be called upon to define the minimum "reasonable" security features that should become part of all IoT devices sold in the EU by 2030.

The key words: IoT devices, security, cybersecurity, data protection, household application.

I. INTRODUCTION

This paper continues the research on the security of IoT devices, which was conducted through a questionnaire among the adult population of the citizens of the Republic of Croatia (further: Croatia). All adults in Croatia who participated in the questionnaire had the opportunity to answer the questions regardless of their age and/or gender.

504 respondents participated in the research intending to analyze their awareness of IoT devices' possibilities. The majority of respondents (56.6%) believe that IoT devices and their use do not contribute (or contribute in a minor percentage) to improvements to the quality of their life. Furthermore, 2.9% of respondents replayed that their data is not exposed to potential threats on the Internet, while 56.4% answered that their data is somewhat exposed to threats, and 40.7% are aware of the dangers that can happen with their personal data on the Internet. [1] From the respondents' answers in the questionnaire, it can be concluded that citizens are not familiar with the operation and abilities of IoT devices and that they are not aware of the dangers to their personal data. It is important to note that the respondents are aware of the inadequate involvement of the system in terms of education and weak regulation, as well as insufficient information on the part of the manufacturers. From the previous can be concluded there is a need to improve the legal framework and education regarding awareness and protection of the personal data of end users.

Based on the mentioned research, the authors wanted to raise the level of awareness of the end users and propose the implementation of education in primary and secondary schools with an emphasis on higher education institutions in the field of technical professions. Previous also includes lifelong education of older citizens and also all the other citizens outside the schooling system.

In this paper, the authors review the security of IoT devices based on the research of recent literature, the previously mentioned questionnaire and the personal use of various small-scale devices.

Therefore, the aim of the paper is to present a realistic picture of the entire system and the potential and ubiquitous dangers to which this technology exposes end users, their households and their personal data.

II. ADVANTAGES AND DISADVANTAGES OF IOT DEVICES IN THE HOUSEHOLD

A. Literature review

Various authors [2] [3] [4] point out that IoT devices have more advantages than disadvantages, and some of the most important ones are listed below. Among the authors mentioned, *minimisation of human effort* stands out as the main advantage of IoT devices as they reduce the need for human intervention by communicating and connecting with each other and performing various tasks without human intervention.[3]

Some of the benefits also cited by the authors include *smart home systems* that are connected to the internet 24/7, *home automation systems and reliable energy management systems*. [5]

In addition, [6] states that the IoT allows users to *access information in real-time* from anywhere in the world. ‘’Users can connect to the app and collect data about their personal devices whenever they are able to connect to the Internet. One of the main benefits of IoT integration is that processes become more adaptive to what is actually happening in the real world. Inherently, this is based on events that are either detected directly or by real-time analysis of sensor data.’’

Also, *patient care* can be delivered more efficiently in real-time without the need for a doctor's visit. It gives them the ability to make decisions and provide evidence-based care, which is called Smart Health System [4]. It includes

smart devices and appliances that support human health, reduce medical costs and even shorten hospital stays. Such devices can be used to check and monitor various health problems and even fitness levels or the number of calories burned. They are also used to monitor critical medical conditions in hospitals and trauma centres.

The Internet of Things has changed the entire environment of the medical field by facilitating it with smart devices. [7] [8].

Another important aspect of peoples' lives is transportation, where the IoT is bringing significant improvements. Vehicles available in the market are equipped with pre-installed sensor devices that are interconnected and can predict and suggest routes taking into account the real-time traffic situation on the road (e.g., fleet tracking). The exciting things that the IoT offers today is the ability to pre-order a charging station for electric cars. [9] [6].

On the other hand, the disadvantages of IoT technology in homes, according to most authors, are as numerous:

First of all, users' data becomes accessible, and hackers can tap sensitive information. A data breach usually occurs when a third party (usually an attacker) can access the user's data without the user's knowledge and awareness. And this data can ultimately be used to manipulate the user, i.e., to manipulate their smart home devices.

Security, and thus privacy, are significantly threatened because IoT systems are connected to numerous devices that communicate within the network. This would mean that there is little to no privacy between the devices on the network and that data is exchanged to make the devices more efficient. IoT devices are more vulnerable to security threats and attacks due to weak technical quality standards.

Currently, there is a lack of adequate security solutions for IoT devices and applications that "lead the world of securely connected things towards the Internet of insecure things". [10] It is therefore necessary to pay attention to security threats and the resulting privacy issues.

B AUTHORS' VIEW

Based on the questionnaire mentioned, research of current literature and independent use of IoT technology in their households, the authors list the following main advantages and disadvantages of IoT devices, according to their own experience. Some of the main advantages of IoT devices that are relevant to the daily lives of smart home users are:

a.) the automation of activities, which reduces the time needed to perform certain actions in the household, usually via a smart home application. For example, the user of a smart home can remotely turn on the heating and cooling system remotely, in advance, and set the temperature in the smart home, also the oven in the smart home can be set in advance to have a ready meal when arriving at the household For example, there is also the possibility to unlock the house/flat via a mobile application to let the handyman in, without coming back home; it can also be done remotely;

b.) healthcare for patients can be significantly improved through the application of telemedicine. With the help of software and digital solutions, servers and IT equipment should be offered, staff trained and the basis for telehealth strengthened, which would enable disadvantaged groups in particular to access medical care (remote patient monitoring). This also includes elderly and frail citizens who can stay in their homes without having to go to old people's and nursing homes;

c.) simplification of traffic flows and smart traffic signals - the user of a smart household can check the fastest way home through the smart-home application and pre-set the opening time of the garage door;

d.) availability of information at any time - for example, a student living in a smart home can search the online library and study materials needed from home, and employees can also do their work remotely.

On the other hand, some of the disadvantages brought by connected IoT devices in the smart household are:

a.) loss of privacy and loss of data privacy, i.e., the availability of all data via the household's smart network via connected IoT devices, the user's personal and intimate data become available to attackers who would want to threaten such a household in any way;

b.) the danger of complete loss of control - the possibility of interruption for using the application for controlling the system of a smart (connected) household - blocking of IoT-connected devices by a third party, usually intruders, i.e., attackers, e.g., complete blocking of the system and unauthorized access to the household;

c.) misuse of connected devices in the household - a third party controls and downloads personal data via connected devices without the permission of the household owner and sells personal data, i.e., shares data with third parties, which can afterwards be used for marketing or any other not permitted purposes;

d.) low system features – end users have no control and often get a simplified interface, which limits the mitigation of vulnerable services. Accordingly, end users have almost no control or insight into the operation of the IoT household devices.[11]

On the one hand, we list the above benefits because the primary function of IoT devices is to improve the lives of the users of these devices and technologies that improve the lives of the general population, as the example of the use of IoT in the healthcare sector shows.

On the other hand, such devices are full of flaws that manifest themselves in the loss, disclosure and sale of personal data of end users without their knowledge to third parties where it is not specified who third parties are.

Therefore, one of the priorities is to propose a privacy standard that would decisively cover the obligations of

TABLE 1: Display of criteria that privacy policies of IoT device manufacturers meet *according to GDPR*

Criteria	Samsung	Google	AWAIR	Honeywell
Sharing of data with third parties without consent of end users	X	X	/X	/X
The possibility to change data by end users (editing, deleting, sharing, downloading)	/X		/X	/X
Right to erasure ('right to be forgotten')	X	X	X	X
Lack of regulation				

manufacturers and sellers of IoT devices. It will be further explained in chapter IV and chapter V.

All data from Table 1 are the results of the Privacy policies of the companies mentioned in this paper Samsung [12], Google [13], AWAIR [14], and Honeywell [15].

The companies in Table 1 are at the same time manufacturers and sellers of IoT devices and were therefore selected to be ranked in this comparison. The results from Table 1 show that the major manufacturers of IoT devices do not apply the GDPR regulations in full, which is evident from their privacy policies. Therefore, the companies that occupy a large part of the internal market of EU Member States are unintentionally putting end users personal data at risk. Their privacy policies are not transparent and easily readable to the average end user; on the other hand, they are complex and complicated to read even to experienced users.

Although Article 8 of Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on Consumer Rights [16] provides the consumer with such information in plain and intelligible language, in a manner appropriate to the means of distance communication, and where such information is provided on a durable medium, it must be legible.

Under Article 17 of the GDPR, the data subject should have the right to rectification of personal data and the "right to be forgotten" where the storage of such data infringes the Regulation or the law of the EU or of a Member State to which the controller is subject.

The data subject should have the right to have his or her personal data erased and no longer processed if the personal data are no longer necessary for the purposes for which they were collected or otherwise processed.

Likewise, where the data subject has withdrawn consent or objects to the processing of personal data concerning him or her, or where the processing of his or her personal data is otherwise not in compliance with the Regulation. [17]

The results of the privacy policies show that manufacturers avoid the provisions of the GDPR, i.e., they do not apply them entirely which is also visible from the Table 1 rationale.

Accordingly, it is clear that manufacturers do not inform end users about the sharing of their data nor about the capabilities of IoT devices, and for the mentioned reasons there are frequent breaches in IoT systems and the loss of personal data of end users, whose initial protection is mandated by the GDPR.

III. STATISTICAL DATA ON ATTACKS AND SECURITY OF IOT DEVICES

According to EUROSTAT statistics for 2022, the countries with the highest percentage of users who own IoT devices in their household (e.g., alarm systems, smoke detection systems, surveillance cameras, door locks or other internet-connected security-safety solutions for home) are Norway (24.37%), the Netherlands (22. 68%), Sweden (18.23%), Denmark (16.61%), Ireland (15.96%) and Estonia (15.72%). The Republic of Croatia is in 23rd place with 5.87% of users [18].

It is significant to point out that there is a certain lack of reliability among EU citizens that this technology brings with it. This is evident from statistical data, which shows that the percentage of users concerned about the security of IoT devices is higher than the percentage of users who own these devices.

Austria is the leading European country when it comes to the percentage of users who are concerned about security with 27.13% of users. It is followed by Finland with 23.36% of users, while Croatia is third with 20.39%. This claim can also be supported by the fact that in the countries mentioned, an even higher percentage of users are concerned about the privacy and security of the personal data generated by such devices. For Austria this percentage is 28.50%, for Finland 24.16% and for Croatia 21.10% [19].

This could indicate two possibilities:

a) citizens are not sufficiently familiar with technology, and/or

b) citizens are familiar with technology and are aware of the potential dangers such devices can cause. Based on the questionnaire conducted for the purpose of researching the awareness of the citizens of Croatia about the possibilities of IoT devices, it can be concluded that the citizens are not sufficiently familiar with IoT technology but are also aware of the shortcomings of the entire system. From above is evident that a large number of EU citizens do not trust the security of IoT devices.

It is important to note; when purchasing an IoT device, end users do not sign the consent for sharing of their data,

which is prescribed by the GDPR. [20] While when purchasing an IoT product, end users should always sign a consent form that lists all the potential risks the user is taking on by purchasing such a product and specifies the warranty period, i.e., the period until which the vendor commits to provide software support and security patches to each user. This is explained in more detail in the next chapter. It is also important to mention that in the last ten years, there have been a few significant attacks on IoT devices.

Such attacks have shown how vulnerable the IoT device network is and how easily breakable to steal the data stored by the device or even to block the entire system.

One of the largest IoT attacks in the history of IoT was the Mirai botnet, which occurred in October 2016 and infected over 600,000 IoT devices. [21]

The scale of the threat posed by IoT devices can also be seen in the example of the attack on Jeep and Tesla automated cars.

In 2015, there was a company-initiated "attack" on Jeep Cherokee vehicles that identified steering wheel vulnerabilities. In this planned presentation of a possible attack, it was found that it was possible to remotely control a car and 1.4 million cars of the said brand were taken out of circulation. [22]

This suggests how common attacks on IoT devices of all profiles have become.

In some cases, users cannot even tell that their devices are being monitored, i.e., that their personal data is being shared, and even if they find out, they still do not know with which third parties their data have been shared and to what extent.

IV. HOW TO REGULATE LEGAL FRAMEWORK AND PROTECT END USERS?

Nowadays, most governments consider data protection as a basic human right [23] but in practice it looks different.

However, existing privacy laws and regulations are not specific to IoT devices. We argue that they are insufficient to recognize important differences between general privacy scenarios and IoT-specific scenarios. [24]

The European general data protection framework includes the GDPR [20] but it does not explicitly mention IoT devices.

Neither Croatian national law; nor the Croatian Act on the Implementation of the General Data Protection Regulation [20] does not contain any statements and provisions applicable to the collection of data via smart home devices and the Internet of Things, and none of them specifies what happens to personal data and/or special types of personal data and with whom they are shared and for what purposes. IoT products and services enable the collection of large amounts of data, some of which as specific types of personal data can be potentially sensitive. Therefore, it is necessary to take protective measures without jeopardizing the users' privacy.

Recital 4 of the GDPR sets out the basic principles of the Regulation. It states that "the processing of personal data should be lawful, fair, and transparent to the data subject."

The GDPR [9] requires that the service provider ensures the safety and security of the processing. It considers this further in the introductory statement 39, elaborating the principles of data protection, especially the principles of purpose limitation, storage limitation, and data minimization: "Natural persons should be aware of the risks, rules, protective measures and rights related to the processing of personal data and how to exercise their rights related to such processing."

Manufacturers' privacy policies are usually not compatible with the obligation of the EU to provide for a high level of consumer protection, as mentioned in Article 169 [25] of the Treaty of the Functioning of the European Union.

For this reason, it would be necessary to oblige manufacturers and sellers to respect privacy standards. Privacy policies should include:

- a list of what kind of personal data manufacturers and vendors collect, maintain, or share,
- notifications for the end users about their data privacy,
- a reasonable consent consisting of risks IoT technology brings with all details not defined by default in the security standards,
- a guide for end users' access to download and edit their data,
- organizations and third parties the personal data are shared with, and their conditions. [26]

Specifically for Croatia, it would be necessary to adopt a regulatory privacy standard that obliges producers and vendors to have users' consent when selling IoT devices to end users.

Consent should be written in clear, comprehensible, and simple language; in terms of IoT device capabilities and by signing the consent the user accepts all given risks.

The authors consider these as minimal "reasonable" security features that should become part of all IoT devices sold in the EU in the next decade. This is currently not regulated by the GDPR nor any other Act of the *Acquis*. Updating IoT devices should not be left to chance. Privacy and security requirements will have to be implemented in the original design process of the device, network or storage system and GDPR will have to adapt accordingly.

Why is it necessary to adopt tougher criteria for producers?

Precisely because it is necessary to bring a higher level of product security with digital elements, which will ultimately increase user confidence in these products.

Vendors, on the other hand, should be responsible for keeping devices up to date.

IV. CONLUSION

The IoT spans a variety of industries, including healthcare, transportation, the consumer rights sector and

the private homes of end users. Even when the attack occurs, the average citizen might not be aware that any IoT device can be compromised. Cybersecurity is everyone's responsibility and an important issue for everyone, not just businesses.

The large number of devices connected to the internet produce enormous amounts of data. This is ultimately the biggest challenge in the IoT sector, as it is most difficult to secure IoT devices and the data exchanged over the network. As end users' private data is transferred between devices, various security challenges such as privacy, confidentiality, integrity and reliability need to be addressed.

It is also crucial for a large number of end users, as according to Statista [27] [28], the number of smart home users in Europe was estimated to reach 64.7 million people in 2022, and how much will the number of users grow in the next 5 years?

Therefore, data security is a key issue for all sectors, especially for households where a large amount of personal data is stored.

One of the aims of this paper was to highlight the current state of IoT device security, the transparency of IoT device manufacturers and vendors to end users, and give an insight into their privacy policies.

The authors proposed the key points of a data protection standard they believe should be mandatory for all manufacturers and vendors of IoT devices in EU member states.

They also pointed out that the privacy policies of individual IoT device manufacturers are not in line with the provisions of the GDPR and Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights [16], nor is the end user support system, which should be simple, transparent and easy to understand according to the legal framework mentioned in this paper.

As a result, the system of networked IoT devices in the smart home is exposed to dangers and risks that end users are not even aware of. This is evident from the statistical data in Chapter 3 for EU countries.

For this reason, their homes and their personal data are exposed to daily use by manufacturers, their partners and third parties who are not clear about who they are and under what conditions they process end users' data.

The General Data Protection Regulation does not oblige manufacturers as much as it should, given the number of IoT devices in households representing a large part of the market. On the one hand, the Regulation needs to be completed as it has not been amended or supplemented since its adoption, and on the other hand, the market is changing rapidly.

Therefore, the authors propose a more detailed check of compliance with the GDPR and, as far as the market for IoT services is concerned, to regulate this area more strictly in the manner written in this paper.

REFERENCES

[1] P. Petrošanec, S. Čelan, R. Nikolić, M. Milenković, Awareness of Croatian citizens about the advantages and disadvantages of IoT devices and their safety, MIPRO proceedings 2023, [Manuscript accepted and submitted for publication in MIPRO proceedings 2023]

[2] Aplus Topper, What is IoT? Advantages and Disadvantages of Internet of Things (IoT), 2022, https://www.aplustopper.com/advantages-and-disadvantages-of-iot/#Advantages_of_IoT, [Accessed 9th of March 2023]

[3] TechVidvan Advantages and Disadvantages of IoT, https://techvidvan.com/tutorials/advantages-and-disadvantages-of-iot/, [Accessed 9th of March 2023]

[4] Kumar, S., Tiwari. P., and Zymbler., M., ''Internet of Things is a revolutionary approach for future technology enhancement: a review'', Journal of Big Data, (2019) 6:111, https://doi.org/10.1186/s40537-019-0268-2, pp. 2-3., https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0268-2#Ack1, [Accessed 9th of March 2023]

[5] Geeks for Geeks, Advantages and Disadvantages of IoT, 2022, https://www.geeksforgeeks.org/advantages-and-disadvantages-of-iot/, [Accessed 8th of March 2023]

[6] Internet of Things From Research and Innovation to Market Deployment, 2014, River Publishers, p. 79

[7] Sfar AR, Natalizio E, Challal Y, Chtourou Z. A roadmap for security challenges in the internet of things. Digit Commun Netw. 2018;4(1):118–37.

[8] Minoli D, Sohraby K, Kouns J. IoT security (IoTSec) considerations, requirements, and architectures. In: Proc. 14th IEEE annual consumer communications & networking conference (CCNC), Las Vegas, NV, USA, 8–11 January 2017. https://doi.org/10.1109/ccnc.2017.7983271.

[9] Behrendt F. Cycling the smart and sustainable city: analyzing EC policy documents on internet of things, mobility and transport, and smart cities. Sustainability. 2019;11(3):763.

[10] Kholoud Y. Najmi, M. A. AlZain, Mehedi Masud, N.Z. Jhanjhi, Jehad Al-Amri, Mohammed Baz, A survey on security threats and countermeasures in IoT to achieve users confidentiality and reliability, Materials Today: Proceedings, 2021, Internet of Things 19 (2022) 100564, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.03.417., (https://www.sciencedirect.com/science/article/pii/S221478532102469X), Internet of Things 19 (2022) 100564

[11] O. Alrawi, C. Lever, M. Antonakakis and F. Monrose, "SoK: Security Evaluation of Home-Based IoT Deployments," 2019 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2019, pp. 1362-1380, doi: 10.1109/SP.2019.00013., p.5

[12] Samsung. *Samsung´s Privacy Policy*. Available: https://privacy.samsung.com/privacy/samsung [Accessed 3rd of April 2023]

[13] Google. *Privacy Policy*. Available: https://policies.google.com/privacy?hl=en-US [Accessed 3rd of April 2023]

[14] AWAIR. *Privacy Policy*. Available: https://store.getawair.com/pages/privacy-policy [Accessed 3rd of April 2023]

[15] Honeywell. *Honeywell Privacy Statement*. Available: https://www.honeywell.com/us/en/privacy-statement#english [Accessed 3rd of April 2023]

[16] Directive 2011/83/EU of the European parliament and of the Council on consumer rights, amending Council Directive 93/13/EEC and

Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32011L0083&from=EN [Accessed 29th of March 2023]

[17] The EU General Data Protection Regulation (GDPR), Oxford University Press, 2020

[18] Eurostat. *Internet of Things - use*. Available: https://ec.europa.eu/eurostat/databrowser/view/isoc_iiot_use/default/table?lang=en (Accessed 9th of March 2023)

[19] Eurostat. Internet of Things - barriers to use. Available: https://ec.europa.eu/eurostat/databrowser/view/isoc_iiot_bx/default/table?lang=en (Accessed 9th of March 2023)

[20] Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, preamble 39, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=HR, [Accessed: 20th of February 2023]

[21] V. Unterfingher, A Technical Analysis of the Mirai Botnet Phenomenon, 2021, Available: https://heimdalsecurity.com/blog/mirai-botnet-phenomenon/

[22] Geenberg A. *Wired, Hackers Remotely Kill a Jeep on the Highway—With Me in It.* Available: https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/ [Accessed 3th of April 2023

[23] Web Hosting Secret Revealed. *The Simple Privacy (and Cookie) Policy Guide for Website Owners.* Available: https://www.webhostingsecretrevealed.net/blog/blogging-tips/have-a-website-you-need-a-privacy-policy-heres-why/ [Accessed 3th of April 2023]

[24] A. Subahi and G. Theodorakopoulos, "Ensuring Compliance of IoT Devices with Their Privacy Policy Agreement," 2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud), Barcelona, Spain, 2018, pp. 100-107, doi: 10.1109/FiCloud.2018.00022.

[25] Consolidated version of the Treaty on the Functioning of the European Union OJ C 326. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012E/TXT&from=EN [Accessed 5th of April 2023]

[26] The Internet of Things, Legal Issues, Policy, and Practical Strategies, C., A., Suarez, C., H., Cwik, and L., Thomson, American Bar Association Book Publishing, 2019, pp. 25 -27

[27] Statista. *Number of smart home users in Europe in 2022.* Available: https://www.statista.com/forecasts/1283896/smart-home-users-europe-segment [Accessed 5th of April 2023]

[28] Jansen B. 60+ IoT Statistics, Facts & Trends. *vpn Alert.* Available: https://vpnalert.com/resources/iot-statistics/?gclid=EAIaIQobChMIre72h4nP_QIVQo9oCR1PAgJaEASAAEgK9xvD_BwE#IoT_Devices_Statistics [Accessed 9th of March 2023]

Privacy Notice Informativeness: in a Search for Benchmark

Marta Alić, Ph.D.
Zagreb University of Applied Sciences, Zagreb, Croatia
marta.alic@tzv.hr

***Abstract* - The term transparency in the disciplines of information management, business ethics and information ethics is commonly used for forms of information visibility and access to information, with the aim of reducing information asymmetry among stakeholders. The principle of transparency is also one of the main mechanisms in data protection and regarding regulations, such as General Data Protection Regulation. But requirements for assesing it's functionality are somewhat ambiguus and hard to evaluate. One of proposed metrics can be the amount of information the transparency mechanism is trying to convey. The basic (ex ante) tool of transparency is the publication of privacy notices, with the purpose of informing the individuals about the procedures related to the collection, sharing, use and storage of their personal data. By using lexical density as a metric to evaluate how much information there is in privacy statements of two major IT companies, Google and Microsoft, across five languages (Croatian, English, German, French, Italian) and comparison within set and generic sections of privacy notice, the aim of this paper is to identify reference scale for this, specific legal documents, between descriptive and explanatory narrative.**

Keywords - privacy notice, lexical density, informativeness, transparency, Google, Microsoft

I. INTRODUCTION

In modern times of information abundance, when the rights of the individual are strengthened, transparency is set as an integral part, a mechanism, in building a trust and a system in which the individuals take an increasing place as a stakeholders in decision making. Transparency is also the principle built-in the legislation of various areas of human activity. The individuals' right to their own choice is becoming stronger in the field of privacy protection, as a fundamental human right, and in today's world of digital transformation, it is increasingly focused on the protection of personal data in digital environments. In this context, privacy is considered to be the right and ability of persons to communicate through digital channels, while controlling the dissemination of information and being informed about who and how participates in the interaction, what information is exchanged and what the information is used for. Furthermore, individuals have broad rights to information transparency of the data collected and processed about them. The basic (ex ante) tool of transparency is the publication of privacy notice, a set of data aimed at familiarizing data subjects with the actions of the system or organization regarding the collection, sharing, use and storage of their personal data. Theoretically, an organization's privacy notice contains all the information that users need to be aware of in order to make informed decisions and control the processing of their personal data. The concept itself is defined as The Individual Participation Principle of the fair information practices. The principle refers to the right of individual „to have data relating to him communicated to him, within a reasonable time, at a charge, if any, that is not excessive; in a reasonable manner, and in a form that is readily intelligible to him“ [1]. Today, reporting on data practices is an important aspect of data protection frameworks and regulation, such as the General Data Protection Regulation (GDPR) [2].

II. RESEARCH CONTEXT

Although a lot of research is focused on examining the content of privacy policy notifications in relation to the regulatory requirements of individual data protection standards, there is a lack of studies that focus on the quality of the mechanism by which these contents are communicated to the target public.

Natural language processing and machine learning methods have also been used to examine the corpus of privacy notices from the standpoint of linguistics and philology. [3][4] through machine learning algorithms categorize different sections of priva5te policy statements, while [5][6][7] are aimed at examining the readability, ie. intelligibility of content. [8] suggests supervised learning in determining the categories of data processing covered by the privacy notice, and other approaches are directed towards comparing the parts, section of the text [9][10] of the privacy statement.

A. Lexical density

Lexical density determines how meaningful the text is, that is, how much information it conveys. Higher density texts are more descriptive and therefore contain more information. Essentially, communication in written form is denser in lexical meaning than that in the colloquial form [11] [12].

Lexical density, expressed by the proportion of lexical words (lexemes) in relation to the total number of words (occurrences), therefore affects the degree of ease with which the recipient understands the message in a particular communication process and can affect memory and sentence retention [13] .

Measuring lexical density is one of the methods used to describe discourse, and therefore depends on the language register and genre of the text. The higher the number of lexical words, the denser the lexical text, so in more

complex texts this value is closer to the value 1 [14]. Expository texts, such as news, informative and technical articles, have a higher lexical density than fiction. According to research by the website Analyze My Writing [15], articles on Wikipedia have a density between 55% and 58%, while a sample of articles from the BBC News and New York Times have a density between 56% and 58%. The same page also researched various forms of short fictional prose, the results of which indicate a lexical density between 48% and 51%. For comparison, interview transcripts were taken as a sample of colloquial language whose average lexical density was close to 45%.

The higher lexical density of non-fiction texts (40-65%) in relation to fiction texts (40-54%) is also pointed out by [16], while the upper limit of 40% is set for non-fiction texts [11]. Furthermore, [17] reveals how lexical density can vary drastically within a single text, which will be examined in the research conducted.

But lexical density can also be affected by the relationship between communication participants [11]. According to [16] and [18] specific guidelines, law enforcement orders and literature that authors expect to be available to the reader for re-reading tend to maximize lexical density. In principle, privacy policy statements can be added to this group as legal texts with some instructions for action, available for reading "on request". In studies of the lexical density of spoken and written material in different European countries and at different age groups [19], the results showed that the lexical density of the one population group is similar and depends on the morphological structure of the mother tongue and the sampled age group, given that lexical density for an individual grows with age and education, depending also on communication style, circumstances and health condition [20]. Therefore, the lexical density in European countries is also the highest for adults, while the variations assessed as lexical diversity were higher for teenagers as the same age group (13-year-olds, 17-year-olds).

III. METHODOLOGY

The research was conducted on the privacy policies of two technology companies, which have a large number of digital services in their portfolio through which they collect a large amount of data, including personal ones - Microsoft and Google.

Privacy notices are divided into sections taking into account the principles of Fair Information Practices – information / awareness, choice / consent, access / participation, integrity / security and implementation / legal protection, accepted by the Organization for Economic Co-operation and Development (OECD) [21] and in the works of [3] and [4]:

• collection: the section explains what information and how it is collected from data subjects;

• purpose: the section explains the purpose of data collection and its use;

• selection / control: the section explains choices and control options available to data subjects;

• sharing - the section states with whom (third party) and under what circumstances the information is shared;

• security - the section discusses standard practices for the data protection;

• retention - the section explains the procedure and deadline for retention of data;

• specific - the section lists practices that apply only to a specific group of users (eg. children, Europeans or residents of California);

• policy change - the section explains how users will be informed about changes in privacy practices.

In the first part of the research, the privacy notices were analyzed regarding the sections in 5 languages: Croatian, English, German, Italian and French. When determining the lexical density, the formula was used:

number of different words / number of total words * 100

and the amounts are shown as a percentage.

The analysis was performed with the Text Analyzer tool [22], and the texts were purified from the peripheral text, titles, bullets, domains and other elements that could affect the results.

In the second part of the research, notices in the Croatian language were analyzed in more detail concerning the relationship of full-meaning words, or lexems, (nouns, verbs, pronouns, adjectives, numbers, adverbs) and grammatical words (prepositions, conjunctions, exclamations, particles). The analysis was performed manually by the author, a professor of Croatian language and literature, using the formula:

lexical words / total number of words * 100

in order to obtain a percentage representation of the ratio.

IV. RESULTS ANALYSIS

A. First part

The first part of the research, related to the calculation of the ratio of the number of different words and the total number of words in the text of the privacy notice, shows comparative results in 5 languages (Croatian, English, German, French, Italian) according to the given sections.

The results of Google's privacy notice [23] in the five selected languages show a range of 43% to 57%, which is in line with the results for non-fiction texts. The highest average lexical density of privacy notice is in Croatian, while the lowest is in French. A comparative analysis of the lexical density of individual sections shows the highest lexical density (Table 1) of the section related to data security settings, except for German language. With the next section on higher density range, relating to policy change conditions, it is consisted of two short sections and their discourse implies an explanatory model of narrative, in contrast to the remaining parts of notices that focus mainly on specification or the listing of individual cases related to the topic and purpose of the section.

The lowest average lexical density (Table 1) in all languages has a section related to the collection, which in its content contains repetitive forms when listing all the necessary ways of collecting data.

Section / Language	Croatian	English	German	French	Italian	Average
Collection	42,25	28,73	34,59	28,66	31,36	33,12
Purpose	43,65	29,34	36,10	28,02	30,20	33,46
Select/Ctrl	39,42	25,27	31,99	26,55	29,92	30,64
Sharing	58,16	42,09	49,59	42,15	46,20	47,64
Security	85,1	59,23	79,48	66,02	72,63	72,51
Retention	64,00	47,47	53,25	49,28	47,67	52,33
Specific	34,20	25,20	33,25	25,25	26,13	28,81
Policy Change	75,00	63,76	71,54	58,89	69,33	67,70
Average	55,23	40,14	48,72	40,60	44,18	

The Microsoft Privacy Statement [24] in the five selected languages shows a range of 40% to 55% and the highest lexical density (Table 2) is shown in the Croatian language version, as in the case of Google's privacy notice.

Table 1 Google privacy policy lexical density (%)

Section / Language	Croatian	English	German	French	Italian	Average
Collection	45,80	31,81	36,64	31,78	38,07	36,82
Purpose	49,51	34,34	39,53	35,30	39,61	39,66
Select/Ctrl	49,32	35,40	41,61	35,22	40,15	40,34
Sharing	58,78	46,21	50,09	43,12	47,57	49,15
Security	73,57	61,25	57,94	56,96	65,92	63,13
Retention	62,50	51,12	60,71	51,60	58,22	56,83
Specific	55,88	41,54	44,26	41,83	45,59	45,82
Policy Change	62,57	57,05	59,03	51,67	55,20	57,10
Average	57,24	44,84	48,73	43,43	48,79	

The lowest lexical density (Table 2) is recorded in the English version of the Statement. Here, too, on average, the highest lexical density is in the security section, with the exception of the English version, which is also "followed" by the section on policy change conditions.

Table 2 Microsoft privacy notice lexical density (%)

The section on specific practices has the lowest lexical density in the Microsoft Statement, which is focused on aimed at determining the requirements under the California Consumer Privacy Act (CCPA), and they contain a distinct form of enumeration.

When comparing the average lexical density by sections in both organizations (Table 3), the results show low values (below 40%) precisely in the sections that are highly "enumerative", ie focus on specification rather than descriptive explanation, while the highest lexical density is recorded in sections with more explanatory expressions.

Table 3 Comparison of lexical density results of Google and Microsoft privacy notices (%)

Section	Google	Microsoft
Collection	36,822	33,120
Purpose	39,661	33,467
Select/Ctrl	40,344	30,640
Sharing	49,158	47,645
Security	63,134	72,510
Retention	56,833	52,336
Specific	45,822	26,139
Policy Change	57,109	67,708
Average	48,61038	45,44563

B. Second part

The results of a more detailed analysis of the lexical density of Google's privacy notice, conducted in the Croatian language version, shown in Table 4, confirm the highest density of the security section, followed by the section on policy change conditions, which is mainly descriptive. The lowest density was confirmed in the part related to the listing of data collection methods, followed by sections with the stated purposes and methods of selection, ie control, which focus on specifications, listing individual practices within the organization.

The average lexical density of the entire corpus (ratio of lexemes to total words) is 33.431%.

Table 4 Results of a detailed analysis of Google Privacy notice in the Croatian language

Section	Number of words	Different words	Grammatical words	Lexems	Density
Collection	786	360	159	201	25,57%
Purpose	721	357	149	208	28,84%
Select/Ctrl	590	291	125	166	28,13%
Sharing	495	291	101	190	38,38%

Security	193	142	33	109	56,47%
Retention	208	130	44	86	41,34%
Specific	569	318	109	209	36,73%
Policy Change	171	107	28	79	46,19%
Total	3733	1996	748	1248	

In the case of the Microsoft Privacy Statement, the results (Table 5) show the highest lexical density in the section thematically related to data security. Furthermore, the lowest lexical density is confirmed in the part that has the most enumerations (listings) - the conditions and settings of privacy under the California Consumer Privacy Act.

The average lexical density of the entire corpus (ratio of lexemes to total words) is 24.013%.

Table 5 Results of a detailed analysis of Microsoft Privacy Statement in the Croatian language

Section	Number of words	Different words	Grammatical words	Lexems	Density
Collection	1640	693	368	325	19,81%
Purpose	1560	681	332	349	22,37%
Select/Ctrl	1258	496	230	266	21,14%
Sharing	501	291	111	180	35,92%
Security	81	69	11	58	71,60%
Retention	374	239	81	158	42,24%
Specific	842	288	162	126	14,96%
Policy Change	132	99	27	72	54,54%
Total	6388	2856	1322	1534	

Comparison of both documents in Croatian language (Table 6), shows the difference for Google document by 10.898%, and for Microsoft 10.117%. The mutual difference between the results for the two statements is 3.164% for the analysis based on the formula for the ratio of the number of different words and the number of total words, and 2.383 based on the analysis of lexemes in relation to the total number of words, which shows consistent results that can serve as a reference values.

Table 6 Comparison of lexical density results of Google and Microsoft privacy notices in Croatian (%)

	1st part analysis		2nd part analysis	
Section	G	M	G	M
Collection	36,82	33,12	25,57	19,81
Purpose	39,66	33,46	28,84	22,37
Select/Ctrl	40,34	30,64	28,13	21,14
Sharing	49,15	47,64	38,38	35,92
Security	63,13	72,51	56,47	71,60
Retention	56,83	52,33	41,34	42,24
Specific	45,82	26,14	36,73	14,96
Policy Change	57,10	67,70	46,19	54,54
Average	48,61	45,44	37,71	35,327

G=Google; M=Microsoft

V. DISCUSSION

In the context of privacy protection, the requirement of inferability [25], the ability of respondents to act based on the information obtained, serves as a global practice in the protection of the rights of individuals based on the concept of consent.

Transparency, which ensures this requirement, is a complex concept that goes beyond simply ensuring the visibility and presentation of this information, and it also includes the quality of its mechanisms [26]. Furthermore, as a functional requirement of information systems, it is linked to the attributes of trust [5] [27], responsibility [28] [29], but also informed decision-making by stakeholders in a certain system [30] [31].

Privacy notices, therefore, should provide information in a concise, effective, and concise manner to avoid fatigue due to excessive information [32]. Furthermore, the recommendations of the Article 29 Data Protection Working Party in its Transparency Guidelines under Regulation 2016/679 refer to best practices for clear writing [33], which include providing information in the simplest possible way, while avoiding complex sentences and language structures.

Also, the information should be concrete and clear and should not be formulated in an abstract or ambiguous way. Also, when designing information in a digital environment, the Working Party calls for the application of a layered approach, which assumes a design and graphic arrangement that supports enumeration.

Therefore, privacy statements represent a specific form of linguistic expression in perceiving the results obtained. On the one hand, as somewhat legal documents, they have the specifics of exhibiting texts that tend to a higher lexical density, above 40%, while on the other hand, due to parts that rely on enumeration in their form, they retain the specifics of colloquial language, whose results show lexical density below 40%, and are marked by a lower representation of lexical words.

And with regard to this duality, the difference in the discourse, it is necessary to consider the results of this research.

VI. CONCLUSION

Results of this research show that when it comes to benchmarking informativeness of privacy notices there is a significant dependency to the characteristics of language used when shaping the content of these documents, as analysis on documents in Croatian language show higher lexical density than in four other compared languages. Respecitvely, there is a notable discrepancy between document sections within all inspected languages,

depending on the set target of information communicated within them. By determining the factors related to the characteristics of language when shaping the content of privacy notice and placing them in relation to the properties of transparency mechanisms, the research in question provides an insight into the importance of these aspects when shaping the strategy of information transparency, not only in the field of data protection and legal sciences, but the results can also be applied as a basis for the creation of guidelines for ensuring effective transparency tools, contributingt to the current literature in the field of privacy requirements engineering.

REFERENCES

[1] IAPP - International Association of Privacy Professionals, "Fair information practices," 2023. https://iapp.org/resources/article/fair-information-practices/

[2] European Parliament and of the Council, "Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," Off. J. Eur. Communities, vol. OJ L 119/1, pp. 1–88, 2016, [Online]. Available: http://data.europa.eu/eli/reg/2016/679/oj

[3] N. Guntamukkala, R. Dara, and G. Grewal, "A machine-learning based approach for measuring the completeness of online privacy policies," Proc. - 2015 IEEE 14th Int. Conf. Mach. Learn. Appl. ICMLA 2015, pp. 289–294, 2016, doi: 10.1109/ICMLA.2015.143.

[4] S. Wilson et al., "The Creation and Analysis of a Website Privacy Policy Corpus," in Proceedings ofthe 54th Annual Meeting of the Association for Computational Linguistics, 2016, pp. 1330–1340.

[5] T. Ermakova, H. Krasnova, and B. Fabian, "Exploring the impact of readability of privacy policies on users' trust," 24th Eur. Conf. Inf. Syst. ECIS 2016, no. April, 2016.

[6] A. K. Massey, J. Eisenstein, A. I. Anton, and P. P. Swire, "Automated text mining for requirements analysis of policy documents," 2013 21st IEEE Int. Requir. Eng. Conf. RE 2013 - Proc., pp. 4–13, 2013, doi: 10.1109/RE.2013.6636700.

[7] G. Meiselwitz, "Readability assessment of policies and procedures of social networking sites," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8029 LNCS, pp. 67–75, 2013, doi: 10.1007/978-3-642-39371-6-8.

[8] E. Costante, Y. Sun, M. Petkovic, and J. Den Hartog, "A machine learning solution to assess privacy policy completeness (short paper)," Proc. ACM Conf. Comput. Commun. Secur., no. November, pp. 91–96, 2012, doi: 10.1145/2381966.2381979.

[9] F. Liu, S. Wilson, F. Schaub, and N. Sadeh, "Analyzing vocabulary intersections of expert annotations and topic models for data practices in privacy policies," AAAI Fall Symp. - Tech. Rep., vol. FS-16-01-, pp. 264–269, 2016.

[10] R. Ramanath, F. Liu, N. Sadeh, and N. A. Smith, "Unsupervised alignment of privacy policies using hidden Markov models," 52nd Annu. Meet. Assoc. Comput. Linguist. ACL 2014 - Proc. Conf., vol. 2, no. June, pp. 605–610, 2014, doi: 10.3115/v1/p14-2099.

[11] J. Ure, "Lexical density and register differentiation," Appl. Linguist., vol. 443–452, 1971.

[12] M. A. K. Halliday, "Spoken and written language," 1989.

[13] C. A. Perfetti, "Lexical density and phrase structure depth as variables in sentence retention," J. Verbal Learning Verbal Behav., vol. 8, no. 6, pp. 719–724, 1969, doi: 10.1016/S0022-5371(69)80035-6.

[14] V. Johansson, "Lexical diversity and lexical density in speech and writing: a developmental perspective," Work. Pap. Linguist., vol. 53, no. 0, pp. 61–79, 2009.

[15] "Analyze My Writing." https://www.analyzemywriting.com/

[16] M. Stubbs, "Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture," 1996.

[17] T. Berber-Sardinha, "Comparing corpora with WordSmith Tools: How large must the reference corpus be ?," WCC '00 Proc. Work. Comp. corpora, pp. 7–13, 2000.

[18] D. Biber, Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure. John Benjamins Publishing, 2007.

[19] V. Johansson, "Toward a cross-linguistic comparison of lexical quanta in speech and writing," Writ. Lang. LiteracyWritten Lang. Lit., vol. 5, no. 1, pp. 45–67, 2002, doi: 10.1075/wll.5.1.03str.

[20] K. D. Bopp and P. Mirenda, "Prelinguistic predictors of language development in children with autism spectrum disorders over four-five years," J. Child Lang., vol. 38, no. 3, pp. 485–503, 2011, doi: 10.1017/S0305000910000140.

[21] OECD, The OECD Privacy Guidelines. 2013. doi: 10.1787/5kgf09z90c31-en.

[22] "Text Analyser," Online-Utility.org. https://www.online-utility.org/text/analyzer.jsp

[23] Google, "Privacy Policy," 15.10.2019, 2018, [Online]. Available: https://policies.google.com/privacy

[24] Microsoft, "Privacy Statement." https://privacy.microsoft.com/en-us/privacystatement

[25] G. Michener and K. Bersch, "Identifying transparency," Inf. Polity, vol. 18, no. 3, pp. 233–242, 2013, doi: 10.3233/IP-130299.

[26] M. Alić, "Model vrjednovanja informacijske transparentnosti politika privatnosti." [Online]. Available: https://dr.nsk.hr/en/islandora/object/ffzg%3A7203

[27] B. Zieni, "Software Requirements Engineering for Transparency," University of Leicester, 2021. [Online]. Available: https://s3-eu-west-1.amazonaws.com/pstorage-leicester-213265548798/31619780/2021ZIENIBPhD.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAILQQVTTAWSUNAFFA/20220423/eu-west-1/s3/aws4_request&X-Amz-Date=20220423T143746Z&X-Amz-Expires=10&X-Amz-Sign

[28] D. J. Weitzner, H. Abelson, T. Berners-Lee, J. Feigenbaum, J. Hendler, and G. J. Sussman, "Information accountability," Commun. ACM, vol. 51, no. 6, pp. 82–87, 2008, doi: 10.1145/1349026.1349043.

[29] D. Hess, "Social Reporting and New Governance Regulation," Bus. Ethics Q., vol. 17, no. 3, pp. 453–476, 2007, doi: 10.5840/beq200717348.

[30] C. Ball, "What Is Transparency?," Public Integr., vol. 11, no. 4, pp. 293–308, 2009, doi: 10.2753/PIN1099-9922110400.

[31] M. Turilli and L. Floridi, "The ethics of information transparency," Ethics Inf. Technol., vol. 11, no. 2, pp. 105–112, 2009, doi: 10.1007/s10676-009-9187-9.

[32] European Commission, "Guidelines on transparency under GDPR," pp. 1–40, 2018, [Online]. Available: http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227

[33] Europska komisija, "Pišimo jasno," pp. 1–16, 2011.

DeShame Croatia: Student Reactions and Consequences of Online Sexual Harassment in High Schools

L.Vejmelka*, R. Matković**, M. Rajter*** and T. Ramljak****
* University of Zagreb, Faculty of Law, Department of Social Work, Zagreb, Croatia
** Institute for Public Health of Split-Dalmatia County/Department for Mental Health, Split, Croatia
*** University of Zagreb, Office for Research
**** Center for Missing and Exploited Children, Višnjevac, Croatia
lvejmelka@pravo.hr

Abstract **- Online harassment and abuse present a contemporary phenomenon with potentially devastating effect on children's lives. The aim of this paper is to explore the reactions and consequences of online sexual harassment. The data were collected within the DeShame project in Croatia in 2020. The probabilistic stratified cluster sample consisted of 2016 pupils of high schools in Croatia (53,8%M; Mean age 15,74). The results show that 20,3% of children witnessed or experienced online sexual harassment in the last year. The most occurring feeling to online sexual harassment was feeling upset (34,2%) and worried (31,1%). However, research identified group of children who didn't want to go to school or see their friends and around every 6th child was worried that this incident will be mentioned again in the future that implies the shame they are feeling because of the involvement in such an event. Therefore, students choose independent problem-solving strategies. The most prevalent reaction after the incident was ignoring followed by blocking the perpetrators, seeking for the friend support or trying to speak to perpetrators or reporting to the platform. Seeking help from adults was significantly less prevalent with speaking to parents being the most salient. The results show that the prevalence of sexual harassment is high among adolescents and the future policies should build the activistic component among adolescents and cooperation to prevent the normalization of sexual harassment.**

Keywords – online sexual harassement; national sample; higih school students; reactions and consequences.

I. INTRODUCTION

Online harassment and abuse present a contemporary phenomenon with potentially devastating effect on children's lives. The studies on non-sexual and sexual abuse of children on the internet are often focused on the prevalence and the characteristics of victims and perpetrators. However, given the high socio-cultural complexity of the problem, there is a lack of data on the help seeking behaviors from victims. The aim of this presentation is to explore the reactions and consequences of online sexual harassment in the context of the creation of adequate policies and prevention and treatment programs.

II. CONCEPTUALIZATION OF PROBLEMATIC INTERNET USAGE, ONLINE VIOLENCE AND ONLINE SEXUAL HARASSMENT

The virtual environment offers everyone a platform to commit various risky as well as punishable behaviors, mostly aimed at harming children and young people but it is important to emphasize that the perpetrators can also be peers, i.e. children . Risky behaviors on the Internet are even more in focus due to the fact that in the past few years after the global pandemic number of the activities "moved" into the online environment.

Problematic internet use is an umbrella term that encompasses a variety of activities and all potentially problematic behaviors on the network, the use of which is advocated by numerous experts in the field of safe internet use [1]. What is particularly worrying is that most of children and young people do not perceive such risky behaviors, whether they come from known or unknown persons or their own risky behaviour, as dangerous or in most cases are not even aware of them, which greatly favors the further spread and development of such forms of harmful action. Cyberbullying is one of the most researched types of problematic internet use but there are various definition and conceptualization of this phenomena. For example, Hindjua and Patchin identify cyberbullying with classic bullying, describing it as repeatedly and intentionally harming others via a computer and a mobile or other electronic device [2]. Tokunaga considers electronic violence more broadly as *any behavior that interferes with modern technologies, in which communication between a group or individuals implies messages of an aggressive and hostile character intended to harm or cause discomfort to other people* [3]. Problematic Internet use also includes other types of online behaviors including internet addiction, internet gambling, online shopping addiction, excessive use of online pornography, and information dependence/ addiction and excessive information search. Internet gaming disorder is the first specific form of problematic internet usage included in the diagnostic manual [4,5] and it is treated as

a special form of addiction in the health system. What is specific to problematic internet use is that for a group of children and young people, this includes exhibiting other forms of risky or problematic behavior such as false identities (and misrepresenting age) and obtaining information about other people's bank cards (most often parents') for example when online gambling and compulsive shopping online.

Appearance of sexually risky behaviors of children online can vary from milder forms of sexting to involvement in criminal offenses against children, such as online sexual abuse and online exploitation of children and online sexual extortion and coercion or revenge pornography [6]. Today number of the materials with sexual content of underage children is produced by children themselves and it is defined as self-generated sexually explicit material [7].

The concept of online child sexual abuse on the Internet is most broadly defined as a form of child abuse in a digital environment that includes various behaviors and actions directed towards the child, with the aim of achieving a certain sexual service. Furthermore, online sexual exploitation of children occurs when an individual or group exploits an imbalance of power to force, manipulate or deceive a child and encourage him to engage in sexual activity through a virtual environment [8].

Sexting (word includes words sex and texting) for children is always a high risk behavior considering that illegal production of sexual material with children occurs. Sexting can lead to higher risk of online sexual coercion and extortion (so-called sextortion). In everyday life, sexting is one of the most common forms of sexually risky behavior in the online environment, defined as sending sexually explicit images, videos or messages over the Internet or using information and communication technologies [9]. Sexting is broadly use among various age groups (smart phone users with adequate level of digital competences) but very specific dimension is legality of sending or receiving sexual content. Consensual sexting is completely legal for adults but for children is part of criminal act considering production and dissemination of material with sexual content of the underage child. On the other hand, sexting is commonly used among adolescent population, and previous research (N= 1794) shows that high school students in Croatia do sext in high percentage. More, than half of the high school students in the sample received messages (51.6%), as well as sexually provocative photos or recordings of another person (52.4%) or an acquaintance (50.5%). Those high school students who received such contents, in a larger percentage, received them more than once, that is, on several occasions [10].

It should be emphasized that some of these acts are often committed by minors to the detriment of other minors and depending on the work and period available statistical data of criminal behavior in the online environment show 15-30% of perpetrators precisely from the group of minors [11].

III. CHILDRENS EXPOSURE TO ONLINE SEXUAL CONTENT –COMPARATIVE AND NATIONAL FINDINGS

According to EU Kids Online, a large comparative survey of children aged 9 to 16 from 19 European countries, 25% of children encountered inconvenience while using the internet. These inconveniences can range from cyberbullying, through hacking and profile theft to slow internet and technical issues, but they can also include online sexual harassment. Inconveniences on the Internet grow with age, where they are experienced by 20% of children aged 9-11 years to 34% of children aged 15-16 year [12].

National sample study HR Kids Online states that nearly 2/3 of children between the ages of 9 and 17 have seen sexual photos or a naked person's video content online in the past year without their intention of seeing them. This was encountered by 3/4 of children aged 9 to 11 years, more than 2/3 of children aged 12 to 14 years and almost 2/3 of children aged 15 to 17 years. The survey results show that in 2017, just over a quarter of children were exposed to sexual content. Children who were exposed to sexual photos are more likely to have access to the internet when they want or need to and spend more time online during the work week and weekends. Children who were exposed to sexual photos of naked people without intending to do so were more likely to close the site or app and clicked the button more often but also more often reporting cyberbullying than children who intended to see such content [13].

Over 2/3 of children aged 9 to 17 have seen sexual photos or a naked person's film online without intention to see them, while nearly a fifth of them have seen such content with intent. A quarter of children aged 9 to 17 have seen photos or videos depicting sexual acts in a violent way without their intention to see such content. A third of children aged 9 to 11 saw sexual photos related to violence without it being their intention and every 12 th child between the 9 and 17 years old received a message with sexual content and that message may include words, pictures or video of sexual content.
[13].

A. Prevalence rates of sexual harassment among children in different countires

Prevalence rates of online sexual harassment among children are diverse and not comparable due to methodological differences and different criterions for sexual harassment. Previous studies of online sexual harassment in different countries shows large scale of prevalence among youth below 10 % in northern Europe in developed countries from 15 to 40% in Spain and Chile and Germany shows as high as 68% percent of exposure to online sexual harassment [14, 15, 16, 17,18,19].

Prevalence of online sexual prevention is as low as 4 or 5 % in the north developed countries as Netherland or Sweden. In a Swedish-based sample of adolescents (N = 594, 12–20 years old), 9.26% of the girls and 5.73% of the boys had been called a fag or pussy against their will, or

someone had been commenting on their private life in a sexual manner in the last six month [14] Furthermore in a Dutch-based sample (N= 1,026, 18 – 88 years old), 5.6% of male adolescents and 19.1% of female adolescents had been sexually harassed online in the past 6 months [15] Online sexual harassment is common among youth in United States (from 13–18 years old), with prevalence rates for the past years ranging from 4 to 38% [20, 21]

Higher percentages were found in the large Chilean-based sample of adolescents (N= 18,872, 12 -17 years old) where 16.23% reported being online sexual harassed [16]. Study of Spanish-based sample of adolescents (N = 3,897, 12 - 17 years old) found online sexual victimization (sexual coercion, sexual pressure, online grooming by an adult, unwanted exposure to sexual content and violation of privacy for 39.5% of the participants [17] In German-based sample (N = 9,000, 10 – 50 years old) numbers are even higher with 68.3% participants identifies with experiences of online sexual harassment as a direct answer on one variable (receiving pornographic material or being asked intimate questions) [18, 19]

B. Gender differences in experiences of online sexual harassement

Previous research shows that gender makes a difference in the experience of exposure to sexual content. Boys were almost twice as likely to see sexual photos as girls and the experience of exposure to sexual content online also grows with age. Boys and older children talk more often about exposure to sexual photographs on pornographic websites. Thus, 12% of children aged 15 to 17 and 3% of children aged 12 to 14 saw photos on pornographic sites at least once a week. Messages with sexual content are received almost equally by girls and boys. Found based on age, with older children receiving such messages more often. Three-quarters of the children who received such messages received them via social networks, almost half as a message sent to a mobile phone, and a third via pop-ups and media platforms (YouTube, Instagram, Flickr) [14].

Some other research also points that exposure to sexual harassment is higher for girls and for other vulnerable groups as LGBT adolescents [21]. Study on 477 Croatian female adolescents indicates that female adolescents are more likely to experience online sexual harassment and find it more upsetting then their male peers. Their study explores the role of resilience in explaining the association between online sexual harassment and negative mood (i.e., depression and anxiety symptoms) among female adolescents and their results shows that overall 42.6% of female participants experienced online sexual harassment at some point [22]. Other study from Egypt also points the gender differences in the experiences of online sexual harassment. They suggest that cyber sexual harassment has become a problem of great concern since 79.8% of the female students (N=2350 female students mean age of 19.9±1.3 years) reported exposure to cyber sexual harassment at least once in past 6 months, and 69,9% students experienced cyber sexual harassment more than once during the same period [23].

C. DeShame research project

The national quantitative, research of DeShame Croatia continues the international comparative research conducted so far in 3 countries[1] and enables international comparative data in the field of sexually risky online behavior and exposure to sexual content and sexual harassment in the online environment.
Given that one of the project's goals was to encourage systematic research in other countries, the Center for a Safer Internet in Croatia[2] formed a national expert team with experts in the field of research safe use of the Internet. The purpose of the research with high school students DeShame Croatia is to deepen knowledge about the risky behaviors in high schools and to find out how much students are exposed to harmful content and sexual harassment in the online environment. This paper will present the results of the national data set for the variables of student reactions and consequences of online sexual harassment in high schools in Croatia.

The aim of the research sequence presented in this paper was to explore student reactions and consequences of online sexual harassment among high-school students on national sample and to address findings and practical implication for the practitioners and experts who works in direct everyday practice with children promoting evidence-based approach in detection, prevention and treatment of online sexual harassment.

IV. METHODS

A. Sample

The research was conducted among secondary school students in Croatia. In order to achieve a national sample, one high school was randomly selected from each county (total of 21) and three first and three third graders were randomly selected in these schools. The data was collected anonymously and using online questionnaires in classrooms. A total of 2,016 students participated, making 81.06% of the planned sample. The average age of the respondents is 15.74 years (min: 14; max19), and 53.77% (N=1084) are male and 43.9% (N=885) are female, and 2.33% (N=47) of the respondents did not want to answer the question about gender.

B. Instruments

This research used a questionnaire developed by the Childnet project team based on a review of existing literature, policy documents and the team's professional experience. With the consent obtained that the questionnaire may be used, the Questionnaire was

[1] https://www.childnet.com/what-we-do/our-projects/project-deshame/

[2] Safer Internet Center Croatia: making the internet a good and safe place https://csi.hr/, financially supported by the European Health and Digital Executive Agency - HaDEA of the European Commission, which made Croatia one of the countries that have an organized national center for a safer internet within the framework of EU funds with the support of several ministries, agencies and institutions.

translated into Croatian, adapted and piloted (Vejmelka et al., 2021).

Demographics contains questions such as age, gender, type of education, class, school success, sexual orientation, family environment, relationship status, etc. It also contains questions about whether the respondents have special educational needs, a diagnosed physical or psychological health problem or a disability.

Psychological experiences and impacts. In the introductory part of this question, sexual harassment is described in detail, and the question was: "*During the last year, have you seen or experienced sexual harassment on the Internet?*". Based on self-assessment, the participants could answer one of the three offered answers: 1- *Yes, I have seen or been the target of sexual harassment on the Internet; 2- No, I have not seen or experienced any of the above; 3- I do not want to answer.*

Participants who indicated that they had experienced online sexual harassment were then directed to a specific set of questions about the related psychological and social impacts. Participants who experienced sexual harassment on the Internet described their feelings through 12 statements (e.g. *I was scared, I was ashamed, I was worried about what my family would think of me, I didn't want to go to school*) and answered on a scale from 1-Strongly disagree to 5 - Strongly agree.

Exposure to online sexual harassment assessed how often, in the past year, other Internet users treated students in the ways described through 10 statements (e.g., "*My naked, half-naked, or inappropriate photos were shared with other people without my permission," "Someone constantly sent naked, half-naked or inappropriate photos or messages of sexual content, even though I did not want them*"). While the experiences of observers witnessing online sexual harassment are measured thorough experiences of how often respondents have seen other young people behaving on the Internet in the last year, across 22 different behaviors. Students answered both scales on a frequency scale from 1 = never to 5 = all the time.

Furthermore, participants had the opportunity to state what they did when they experienced sexual harassment on the Internet (e.g. *I didn't tell anyone, I ignored it, I told a parent, I told a teacher/school staff, I told friends, I reported it to a social network on where it happened (e.g. Facebook, Snapchat), I reported it to the police, etc*.). 11 statements were offered, and participants could choose multiple answers.

C. Etics

The research was conducted with the consent of the Ministry of Science and Education (CLASS: 602-01/21-01/00628; URBROJ: 533-05-21-0004) and the consent of the Ethics Committee of the Faculty of Law, University of Zagreb, and in compliance with the Code of Ethics for Research with Children [28]. After the principals of the schools gave their consent that the research could be carried out in their schools, the parents of the selected classes received detailed information about the aim and purpose of the research and could express their disapproval of their child participating with their signature. Before conducting the research, it was emphasized to the students who participated in the research that the research is anonymous, that their participation is voluntary, that they can withdraw at any time, and that the data can only be seen by members of the research team. For students who did not participate in the research (with the signature of their parents or who did not want to participate themselves), have been prepared content about sexual harassment, an online quiz and suggestions where they can seek help if they or someone close to them experiences sexual harassment., Professional training (before the research) and presentation of research results were organized for professional associates in schools.

V. RESULTS AND DISCUSSION

In DeShame Croatia study we asked high school students directly, during the last year, have they witnessed or experienced sexual harassment on the Internet? The results show that 20,3% of children witnessed or experienced online sexual harassment in the last year and quite high percentage of 7,2% of children didn't want to answer on this question. This leads to conclusion that a number is even higher.

Graph 1. Answers on the question have you been witnessed or experienced online sexual harassment in a last year

Table1. Consequences of online sexual harassment

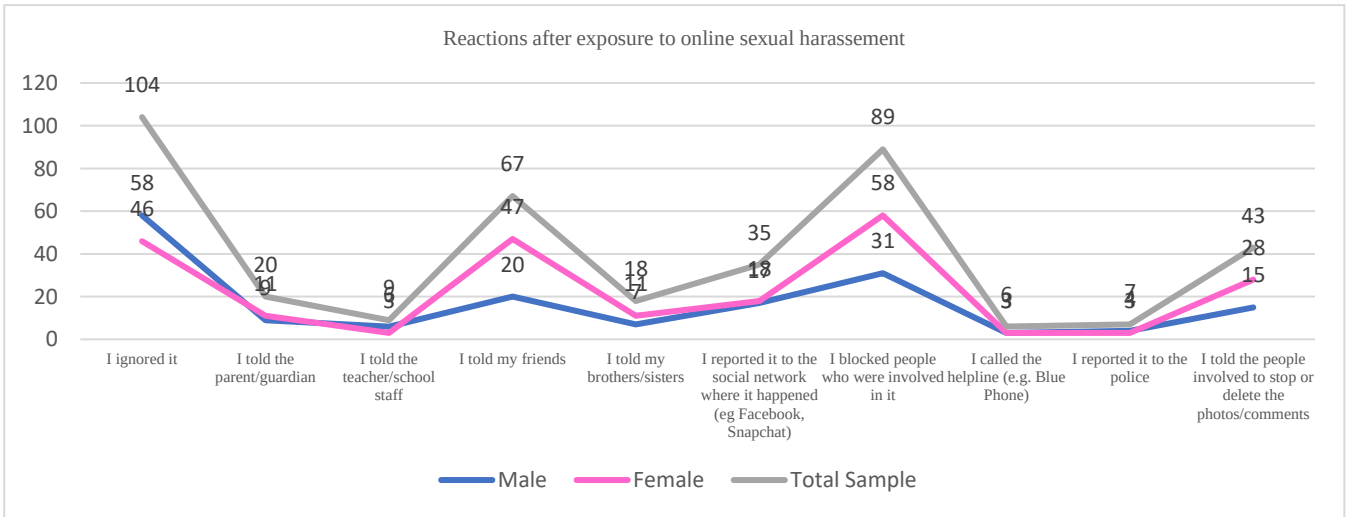
	N	M	SD	I completelydisagree	I mostlydo not agree	I neither agreenor disagree	I mostlyagree	I completelyagree
				%	%	%	%	%
I was scared	409	2,3	1,35	44,7	14,4	20,0	12,5	8,3
I was upset	409	2,6	1,50	36,2	13,7	15,9	19,3	14,9
I was worried	409	2,6	1,48	37,4	13,9	17,6	16,9	14,2
I blamed myself	409	1,9	1,29	60,6	13,2	13,2	4,6	8,3
I was ashamed	409	2,1	1,37	53,5	12,0	17,1	7,6	9,8
I felt lonely	409	1,9	1,29	60,4	12,0	12,7	7,8	7,1
I didn't know what to do	409	2,1	1,36	52,3	10,3	18,3	11,0	8,1
I was worried about what my family would think of me	409	1,9	1,36	63,3	6,1	14,2	8,1	8,3
I didn't want to go to school	409	1,6	1,06	73,3	6,8	13,0	3,4	3,4
I didn't want to see my friends	409	1,6	1,10	73,3	6,4	12,7	3,4	4,2
I was worried that it might be mentioned again in the future	409	2,0	1,36	59,2	8,3	15,6	8,6	8,3
I didn't mind	409	2,5	1,57	43,0	10,5	19,1	7,6	19,8

As we wanted a more reliable measure for determining witnessing and experiencing online sexual harassment, we calculated the scale results on the corresponding scales. The extent to which young people are exposed to sexual harassment was measured through 22 behaviors to which young people answered whether they had witnessed these behaviors in the past year. These results show that only 6.9% of young people in Croatia state that they have not seen such behaviors on the Internet, while 93.1% state that they have witnessed online sexual harassment behaviors. Exposure to sexual harassment in Croatia was investigated through measurement of 10 forms of experiencing online sexual harassment in the past year, and by this criterium 41.2% of high school students in a nationally representative sample experienced online sexual harassment at least once. One of the studies showed that the engagement in risky sexual online behavior was related to perceptions of peer involvement, risks, benefits, and vulnerability of this behavior. Adolescents who engage in sexual online behaviors perceived that more of their friends are also engaged in similar behaviors [24]. Opposite of this, at the DeShame Croatia research students were underestimated the involvement in the online sexual harassment. These results suggest that young people do not identify every behavior they experience under the online sexual harassment as such, pointing out to widespread character of the phenomenon and normalized behavior among high school population.

Graph 2. Reactions after exposure to online sexual harassement

Table 1 presents consequences of the online sexual harassment and most occurring feeling to online sexual harassment was feeling upset (34,2%) and worried (31,1%) or ignoring (27,4%). However around 7% of children didn't want to go to school or see their friends. Online violence involves a wide circle of observers, which means that the victim bears far more severe and long-lasting consequences. Electronic violence and traditional peer-to-peer violence, for example in the classroom, are often connected and there is a transfer of one form of violent behavior and setting where it occurs to another [26] .16,9% of children were worried that this incident will be mentioned again in the future. Results of the comparative European research states that half of the children felt somewhat guilty after unwanted exposure to sexual photos. A 1/3 of children who were exposed to sexual photos in a violent way online, and that was their intention, felt somewhat guilty [25]. Furthermore, in our research when asked what they did after the incident, the most prevalent reaction in total sample was ignoring (N= 104, 51,3%), blocking the perpetrators (N= 89, 47,9%), speaking to friends (N=67, 36,4%), trying to speak to perpetrators (N=43, 22,7%) or reporting to the provider (Graph 2.)

Comparative european research Eukidsonline points out that one in 10 children who have been exposed to sexual photos think their parents know nothing about their online activities. No child has tried to seek adult help or tried to report a problem online. From the results obtained, it is evident that a certain number of children who are disturbed by the experience remain silent about it and avoid seeking the help of adults [25].



Previous studies suggests that females are especially vulnerable group and pointing out in gender differences in experiencing and reacting on online sexual harassment events [12,22,23]. Gender differences were confirmed in DeShame Croatia research but in the direction that boys use more ignoring strategies (12% difference) and talks to friends less (21,7% difference) than their girl peers. Significant differences of reactions shows that boys will block the source of the harassment more often (27,2% difference) and are less likely to try to speak with perpetrators (13,2% difference) than girls when online harassment happens. Adolescents from the qualitative study were also concerned about gender dimension of online sexual harassment on social media sites [26].

According to study amongst a US adolescents (n = 1,152, 10-18 years old), nearly 15% reported online sexual harassment experiences [27]. The adolescents in the qualitative research indicated that were engaged in victim blaming by emphasizing that risky behavior (e.g., sexting, posting provocative pictures, or accepting too many friends) initigates sexual harassment pointing that a major role in preventing sexual harassment was allocated to the individual. Both boys and girls often referred to one's "own responsibility" for receiving sexual harassment, due to accepting friend requests from unknown people, posting "provocative" pictures in public, or sending intimate pictures [26]. It is precisely these attitudes of the people, peer groups, norms and cultural content of the environment that reinforce and confirm the feeling of personal responsibility and, ultimately, that victims of online sexual harassment choose strategies for solving problems independently instead of reporting criminal acts. This research as well as the previous ones suggesting that reporting rate of online sexual harassment is low and estimation of large dark number of non-reported cases still remains.

VI. Conclusion and Implications

Our results show that the prevalence of sexual harassment is high among adolescents and their reactions vary. Discrepancy in the results of the involvement in the online sexual behaviors is particularly worrying when comparing student answer on direct question with scale results. Prevalence rates of online sexual harassment among children are diverse and not comparable due to methodological differences and different criterions for sexual harassment.

Given that the prevalence rates of online sexual harassment among children are not comparable and differ greatly due to methodological differences and different criteria for sexual harassment, this research additionally confirmed the confirmation for standardization in this area. It is necessary to agree on a generally accepted definition and conceptualization of online sexual harassment, which will enable the operationalization of standardized instruments, which in turn will enable comparative results and monitoring of trends. These studies should become the basis for peer-based intervention aimed at reducing online sexual harassment.

When online harassment happens, they are most inclined to solve the situation by themselves or using the peer support, while they are less inclined to seek help from adults and support services. The most worrying reaction is ignoring the incident which in terms of the prevention of sexual harassment propagates the behaviour. The future policies should build the activistic component among adolescents and cooperation to prevent the normalization of sexual harassment. This result also points at the normalization of the online sexual harassment and widespread distribution of the phenomena among high school population. Important practical implications in this field are in the direction of building awareness, continuous education and internal rules on social media use and communication in the online environment. Simple, adequate, timely and well promoted ways of reporting are also crucial in fighting against online sexual harassment.

Acknowledgment

Special thanks to the principals of the schools for organizing the research in their institutions, to the professional associates in the schools who helped in the research conduction and to the children who participated in the research and thus contributed to the explanation of online sexual harassment of children and adolescents at the national sample. We also thanks to their parents who recognized the importance of this research and approved their participation.

References

[1] N. A. Fineberg, Z. Demetrovics, D. J. Stein, K. Ioannidis, M. N. Potenza , E. Grünblatt, M. Brand, J. Billieux, L. Carmi, D. L. King, J. E. Grant, M. Yücel, B. Dell'Osso, H. J. Rumpf, N. Hall, E. Hollander, A. Goudriaan, J. Menchon, J. Zohar, J. Burkauskas, G. Martinotti, M. Van Ameringen, O. Corazza, S. Pallanti, COST Action Network, S. R. Chamberlain, "Manifesto for a European research network into Problematic Usage of the Internet", European Neuropsychopharmacology, vol. 28(11), pp. 1232–1246, 2018. https://doi.org/10.1016/j.euroneuro.2018.08.004

[2] J. W. Patchin, and S. Hinduja, "Measuring cyberbullying: Implications for research", Aggression and Violent Behavior, vol 23, pp. 69-74, 2015.

[3] R. S. Tokunaga, "Following you home from school: A critical review and synthesis of research on cyberbullying victimization", Computers in human behavior, vol. 26, pp. 277-287, 2010.

[4] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders (DSM-5), American Psychiatric Association, 2013.

[5] World Health Organization,International Classification of Diseases 11th Revision (ICD-11). World Health Organization (WHO), 2018. https://icd.who.int/

[6] D. K. Citron, and M. A. Franks, "Criminalizing revenge porn", Wake Forest L. Rev., vol. 49, pp. 345, 2014.

[7] L. Vejmelka andJ. Jurinić, "Terminologija i konceptualizacija seksualnog iskorištavanja djece na internetu", Kriminologija & socijalna integracija, vol. 28(1), pp. 95-114, 2020. https://doi.org/10.31299/ksi.28.1.5

[8] Interagency Working Group on Sexual Exploitation of Children, Terminology guidelines for the protection of children from sexual exploitation and

sexual abuse. ECPAT International & ECPAT Luxembourg, 2016. http://luxembourgguidelines.org/

[9] M. Drouin, K. N. Vogel, A. Surbey and J. R. Stills, “Let’s talk about sexting, baby: Computer-mediated sexual behaviors among young adults”, Computers in human behavior, vol. 29, pp. A25-A30, 2013.

[10] G. Buljan Flander, K. Prijatelj, A. Raguž, M. Čagalj Farkas and E. Selak Bagarić, "Rasprostranjenost i navike sekstinga kod djece srednjoškolske dobi u Hrvatskoj", Napredak, vol. 162, br. 1 - 2, pp. 7-25, 2021. https://hrcak.srce.hr/259414.

[11] J. Jurinić, L. Vejmelka and M. Galiot, ”Seksualno iskorištavanje djece na internetu: od povijesnog prikaza do suvremenih trendova”, Policija i sigurnost, vol. 29, pp. 404-425, 2020.

[12] D. Smahel, H. Machackova,G. Mascheroni, L. Dedkova, E. Staksrud, K. Olafsson, S. Livingstone, and U. Hasebrink, “EU Kids Online 2020: Survey results from 19 countries”, EU Kids Online, 2020. https://doi.org/10.21953/lse.47fdeqj01ofo

[13] L. Ciboci, I. Ćosić Pregrad, I. Kanižaj, D. Potočnik, and D. Vinković, "National Survey on Child Safety on the Internet HR Kids Online", Society for Communication and Media Culture, 2020. http://hrkids.online/prez/EUKidsOnlineHRfinal.pdf

[14] S. Ståhl and I. Dennhag, “Online and offline sexual harassment associations of anxiety and depression in an adolescent sample”, Nord J Psychiatry, vol. 75(5), pp. 330-335, 2021. https://doi.org/10.1080/08039488.2020.1856924 .

[15] S. E. Baumgartner,P. M. Valkenburg, and J. Peter, “Unwanted online sexual solicitation and risky sexual online behavior across the lifespan,Journal of Applied Developmental Psychology, vol. 31, Issue 6, pp. 439-447, 2010.https://doi.org/10.1016/j.appdev.2010.07.005 .

[16] C. Guerra,C. Pinto-Cortez, E. Toro, E. Efthymiadou,and E. Quayle, “Online Sexual Harassment and Depression in Chilean Adolescents: Variations Based on Gender and Age of the Offenders”, Child Abus. Negl. , vol. 120, pp. 105219, 2021.

[17] I. Montiel, E. Carbonell, and N. Pereda, “Multiple online victimization of Spanish adolescents: Results from a community sample”, Child Abuse & Neglect,vol. 52, pp. 123-134, 2016. https://doi.org/10.1016/j.chiabu.2015.12.005.

[18] F. Staude-Mu¨ller, B. Hansen, and M. Voss, “How stressful is online victimisation? Effects of victim’s personality and properties of the incident”, European Journal of Developmental Psychology, vol. 9, pp. 260–274 2012.

[19] N. Henry, and A. Powell, “Technology-Facilitated Sexual Violence: A Literature Review of Empirical Research. Trauma”, Violence, & Abuse, vol. 19(2), pp. 195–208, 2018. https://doi.org/10.1177/1524838016650189

[20] G. Hill, and H. Kearl, Crossing the Line: Sexual Harassment at School. American Association of University Women, 2011. http://www.aauw.org/research/crossing-the-line/

[21] K. J. Mitchell, M. L. Ybarra, and J. D. Korchmaros, “Sexual harassment among adolescents of different sexual orientations and gender identities”, Child Abuse & Neglect, vol. 38(2), pp. 280–295, 2014. https://doi.org/10.1016/j.chiabu.2013.09.008

[22] K. Mitchell, and A. Štulhofer, “Online sexual harassment and negative mood in Croatian female adolescents”, European Child & Adolescent Psychiatry, vol 30(2), pp. 225-231, 2020. https://doi.org/10.1007/s00787-020-01506-7

[23] A. E. Arafa, R. S. Elbahrawe, N. M. Saber, S. S. Ahmed, and A. M. Abbas, “Cyber sexual harassment: a cross-sectional survey over female university students in Upper Egypt”, International Journal Of Community Medicine And Public Health, vol. 5(1), pp. 61–65, 2017. https://doi.org/10.18203/2394-6040.ijcmph20175763.

[24] S. E. Baumgartner, P. M. Valkenburg, and J. Peter, “Assessing causality in the relationship between adolescents' risky sexual online behavior and their perceptions of this behavior”, J Youth Adolesc., vol. 39(10), pp. 1226-1239, 2010. https://doi.org/10.1007/s10964-010-9512-y.

[25] Vejmelka, L., Matković Roberta & Rajter, M. (2022) Cyberbullying in COVID-19 Pandemic Decreases? Research of Internet Habits of Croatian Adolescents. Information, 13 (12), 586, 17 https://doi.org/10.3390/info13120586.

[26] K. Van Royen, H. Vandebosch and K. Poels, “Severe Sexual Harassment on Social Networking Sites: Belgian Adolescents’ Views”, Journal of Children and Media, vol. 9(4), pp. 472-491, 2015. https://doi.org/10.1080/17482798.2015.1089301

[27] J. E. Copp, E. A. Mumford, B. G. Taylor, “Online sexual harassment and cyberbullying in a nationally representative sample of teens: Prevalence, predictors, and consequences”, Journal of Adolescence, Vol. 93, pp. 202-211, 2021. https://doi.org/10.1016/j.adolescence.2021.10.003 .

[28] M. Ajduković, and V. Kolesarić, (Eds.), “Etički Kodeks Istraživanja s Djecom [Code of Ethics in Research with Children]. Etički Kodeks Istraživanja s Djecom. Vijeće za Djecu Vlade Republike Hrvatske; Državni Zavod za Zaštitu Obitelji, Materinstva Mladeži [Council for Children of the Government of the Republic of Croatia; State Institute for the Protection of the Family, Motherhood and Youth]. Zagreb. 2003. Available online: https://www.ufzg.unizg.hr/wp-content/uploads/2020/02/Etickikodeks-istrazivanja-s-djecom.pdf

Capacities of Western Balkan Economies (and Their Public Sectors) to Respond to Ransomware Attacks**

Djordje Krivokapić*, Andrea Nikolić* and Ivona Živković*
*Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia,
Masaryk University, Brno, Czech Republic
djordje.krivokapic@fon.bg.ac.rs
**This article is supported by the European Regional Development Fund (ERDF) project "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822).

***Abstract* - Ransomware became a global cybersecurity threat affecting not only individuals and private companies, but governments and other public bodies. The paper will firstly introduce the concept of ransomware and emphasize the implications of the global rise of ransomware. Secondly, the paper will provide an overview of the regulatory framework in relation to Ransomware in Western Balkan (WB) economies (Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, and Serbia), while an ethical and practical guidance will be proposed. Then, the paper will briefly present Western Balkan Case Studies, including recent key ransomware attacks occurred in the region. Lastly, the paper will provide recommendations regarding prevention and recovery from the ransom attack focusing on WB region.**

Keywords* - *Ransomware; Western Balkan Economies; Cybersecurity; Legal Regulation; Ethics

I. The concept of ransomware

Ransomware is a malicious software on computer servers that encrypts files and disables normal operation of the information systems, with the aim to get the payment in order to restore access to the system [1]. The payment is usually requested through digital currency in exchange for a decryption key, so anonymous crypto-currency payment systems encourage attackers to easier undertake the ransomware attack. The malicious actors do not discriminate. Ransomware affects both private and public sectors, small companies and big corporations, while covering various industries including health care, finance, government, construction, education, etc. The paper will briefly present the regulatory framework and recent key ransomware attacks that occurred in the Western Balkan region emphasizing the useful takeaways and proposing some recommendations regarding prevention and recovery from the ransom attack.

II. The implications of the global rise of ransomware

Cybersecurity attacks have been a growing concern for businesses and organizations around the world, as attackers are becoming increasingly sophisticated in their tactics. While analysts recorded a significant drop in ransomware attacks between 2016 and 2017, the first half of 2018 saw a 229 percent increase compared to the same interval in previous year [2]. In 2021, the FBI's Internet Crime Complaint Center received 3,729 complaints identified as ransomware costing victims over $49.2 million, and anticipated an increase in critical infrastructure victimization in 2022 [3].

The negative implications of ransomware attacks could be various, from paying the ransom itself to suffering some other significant losses. Furthermore, not only the primary target is affected, but the ransomware could cause harm to numerous stakeholders. The potential losses encompass both direct and indirect damages, including discontinuity of business operations, damage to health, loss of life, loss of income, loss of customers, reputational damage, etc. Therefore, the organization under ransomware attack could face legal suits by previous business partners and consumers asking for the compensation [4].

III. The regulatory framework in relation to Ransomware in Western Balkan

The ultimate question is how legal regulation could effectively regulate ransomware attacks and whether the specific regulation is required or the existing regulation is sufficient. Namely, ransomware attacks are already covered by existing regulation primarily in the areas of cybercrime, information security and data protection, but also in various other sectoral regulations, including telecommunication, energy sector, etc. Those regulations pose certain obligations aimed at prevention of the attack and some obligations, i.e. necessary steps that should be taken after the attack. Moreover, general criminal and civil liability standards could be applied to ransomware cases [5].

Cybercrime regulation is established by the Budapest Convention that has been ratified and implemented by 67 states, including Albania, Bosnia and Herzegovina, Croatia, North Macedonia, Montenegro, and Serbia [6]. It seems that Budapest Convention is sufficiently flexible to deal with the criminal law implications of ransomware. As a type of malware, ransomware is incriminated by "Data interference" (Article 4) and "System interference" (Article 5) [7]. National legal systems in practice complement these provisions with other criminal offenses such as extortion, ransom and coercion as well as national offenses related to cybercrime [8]. As the Budapest Convention turns 20, the Second Additional Protocol to the Convention on enhanced co-operation and the disclosure of electronic evidence is adopted and already signed by 35 states, including the Albania, Croatia, North Macedonia and Montenegro, while Serbia is the only one that did ratification [9].

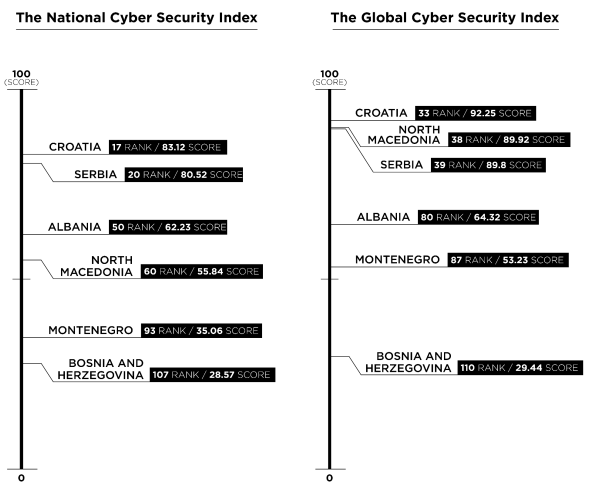
Further, an appropriate level of cyber resilience should be reached by developing various organizational, technical and legal measures regulated by information security, data protection and other sectoral regulation that impose preventative and reactive security measures on operators of information systems.

Directive on Security of Network and Information Systems ("NIS Directive") [10] established legal measures to improve cybersecurity in the EU. Recently, a review of the performance of the Directive has been performed, which led to the improvement by enacting the second version of the Directive in the last days of 2022 [11]. The review of the Directive catalyzed the institutional and regulatory approach to cybersecurity in the Union, paving the way for a significant change in mindset by expanding scope of public and private entities subject to the regulation. General Data Protection Regulation (GDPR) [12] on the contrary has a unified approach toward all EU members, and an extraterritorial reach, imposing the application of appropriate technical and organizational measures by all personal data controllers [13]. All these frameworks establish a risk based approach, meaning that the security measures are not set in stone but they require risk assessment, application of appropriate measures and constant monitoring and adaptation. The common tool to perform data security risk assessment is ISO 27001, the international standard for an information security management system [15].

Western Balkan economies are obliged to align their national legislations with the *acquis communautaire*, as a necessary step in the accession process to the EU. North Macedonia and Serbia have harmonized their national laws with GDPR, while Albania has partially harmonized its national law. Bosnia and Herzegovina and Montenegro have not harmonized their laws, but new draft laws harmonized with GDPR are in procedure. On the other hand, the harmonization with the NIS Directive of the Western Balkan economies has a different status. Albania mainly harmonized it by adopting the Law on Cyber security in 2017. In Bosnia there is no law on information security on a national level. However, there is a draft Law on Information Security and Security of Network and Information Systems. The similar situation is in North Macedonia that has no overarching cyber security law, but there have been three Draft Laws, with the latest being published in July 2021 and is in line with the NIS directive. Finally, Montenegro and Serbia are mainly harmonized with the NIS Directive. Namely, Montenegro has adopted Law on Information Security, Law on Determining and Protecting the Critical Infrastructure and Law on Data Secrecy, while Serbia has adopted the Law on Information Security in 2016 which was amended in 2019 to further harmonize with the NIS directive.

Preparedness of WB economies to prevent cyber threats and to manage cyber incidents is assessed by two indexes: National Cyber Security Index [15] and Global Cybersecurity Index [16], which are presented in the table I below.

TABLE I. National and Global Cyber Security Index



The threat of ransomware attacks became an imminent and foreseeable threat due to the rapid increase of ransomware attacks in recent years. Thus, ransomware should be treated with special care, including mandatory conduct of the risk analysis and applying a set of appropriate measures on a case to case basis.

IV. Ethical and other dilemmas regarding paying the ransom

Even though ransom payments are not criminalized per se, it can put the victim in a problematic position [17],

but the question is: is paying a ransom morally right or wrong? A second important question that arises from this ethical dilemma is whether payment of ransom encourages this type of criminal behavior?

On one side by refusing to pay the ransom to the abductors we have potentially discredited their attempts to make an illegal financial gain [18]. However, as it is always the case with moral decision-making, an additional question arises, how to justify a situation in which we refuse to pay for the ransom and lose data that are of enormous material and logistical importance to the organization?

The situation is further complicated in cases where human lives are, directly or indirectly, affected. Medical institutions are, for example, one of the most frequent victims of ransomware attacks. In those cases potential loss of data on medical history or diagnoses could lead to irreversible damage to patients' lives [19]. Also, in an organizational sense, the damage that would be caused to the company could lead to job loss and consequently a series of social and economic problems that would be borne by the employees [20].

In a process of trying to make an ethical decision, of whether to pay for ransom or not we can engage two main ethical approaches that can help us in the decision making process. The first one values the common good rather than putting individuals interests at the forefront. That approach is better known as a consequentialist or utilitarian approach. The main argument that stands behind utilitarian viewpoint in a case of not paying ransom is that when we calculate the consequences of having to pay the ransom it will potentially lead to an increase in the number of such cases in the future [21]. This could be explained by the fact that if we follow the utilitarian logic, we could say that they value the quantity of the aggregate people to be saved from future attacks rather than risking all of them for the sake of one person's interest [22].

Another crucial element that is valuable to mention is related to the problem of reputation (company, stakeholders or a country) in the case of ransom payments. In this case, the calculation is clear: the entity that once pays the ransom is targeted as a favorable victim to which the attackers will return in the future.

On the other hand, if we consider the second most represented school of ethical thought, the decision of whether to pay the ransom or not seems quite simple. Deontological approach is built around a list of morally prohibited acts- simpler acts that are prohibited no matter what [23]. Therefore, a deontological approach can unequivocally lead us to the conclusion that under no circumstances we can pay a ransom. Main principle of deontology says that the action we conduct must be fundamentally morally good as it is. However, this dilemma is not as simple as it seems, especially when it comes to paying a ransom in cases where human lives are involved.

In any case the decision whether to pay the ransom or not depends on many factors. So, we can conclude that there is no unified answer to the primary set question. But there are several factors that we need to consider. The most important ethical consideration is what consequences would be on people’s lives, their personal data, and critical infrastructure. Beside that we need to consider the outcomes of our action on a stakeholder. Also, we need to be prepared for damage mitigation, since whatever decision we choose, it will cause some damage. At the end we need to consider what are the effects of our decision on our reputation. As mentioned, if we choose not to pay, we will probably build a good reputation and potentially save ourselves from future attacks. On the other hand, we need to deal with outcomes that this decision brings and the ability to deal with potential losses.

V. Recent key ransomware attacks in Western Balkan

In recent years, and especially during the second half of 2023., there have been several high-profile incidents that have made headlines in the Western Balkans.

A. Serbia

In March 2020, a ransomware attack on critical infrastructure occurred at the public utility company "Informatika" in Novi Sad, Serbia [24]. This incident was the first publicly known case of a ransomware attack on critical infrastructure in Serbia. The attack resulted in the inability to access Informatika servers, compromising around 2000 accounts, preventing employees from accessing email servers, compromising backup copies, causing city cameras to stop working, and making it impossible to pay through the company's counter or process wages, sick leave records, and other personnel records. The city administration of Novi Sad worked with external IT companies and Informatika's own expert team to create a new hardware-software architecture for the company's information system to prevent any future attacks. Two years later, in June 2022., the Republic

Geodetic Institution (RGI) announced that their recent computer sabotage was carried out from abroad using the Ransomware Phobos virus. The attack has been qualified as an attack on sovereignty of the Republic of Serbia, 28 services of RGI were unavailable for at least a week which caused numerous consequences for citizens, businesses and public administration [25]. No breach of personal data occurred and no ransom note was found in the attack messages according to the RGI [26]. The institution was able to restore the full functionality of the system within a one month, through phased implementation, with the support of external experts from Serbia, EU, USA and Russia [27]. *The compliance of the security teams was closely monitored to ensure the recovery process was not hindered. More than 3000 computers, several hundred servers and more than 3 petabytes of data were thoroughly checked, using three different security systems, to ensure all viruses were cleaned and not left behind, which prevented any subsequent remote activation or reactivation system blockages.* Access to the web services from outside of Serbia is blocked and still active as a preventive measure which reduced daily cyber attacks on RGI on average from 26.000 to only 3 [28].

B. Montenegro

On August 20th, 2022, the well-known hacker group Cuba Ransomware targeted and successfully breached the server infrastructure of the state institutions of Montenegro [29]. The group claims to have obtained a significant amount of data, including financial documents, source code, and correspondence with bank employees. The attack caused widespread disruption, including the inability to use official emails, the unavailability of e-governance services, and the shutdown of the government's website and other web services. The attack also affected critical infrastructure systems, such as the power supply and water supply systems. In response to the attack, the FBI Cyber Action Team, NATO, US, France and the Great Britain embassy have all become involved in the investigation. The consequences of the attack have been severe, with services unavailable for at least 20 days, and at least 2471 workstations and 150 workstations within 11 institutions compromised and locked [30]. The government of Montenegro has also announced that data has been compromised, and at least 500 eGovernment services and numerous email servers are unavailable. The Agency for National Security claimed that coordinated Russian security services are behind the cyber attack, qualifying this attack as politically motivated [31].

C. Bosnia and Herzegovina

Bosnia and Herzegovina has seen an increase in ransomware incidents in recent years, particularly in 2022. In 2020, numerous Bosnian municipalities were under ransomware attacks [32]. In 2022 the Parliament of Bosnia and Herzegovina was attacked by a crypto-locker virus which caused the following consequences: personal workstations of employees were out of operation, mail servers were offline, web pages were unavailable, and certain databases were locked [33]. According to the officials' statements, the Parliament was not the only target: "Bosnia and Herzegovina's Presidency and the Council of Ministers have been targeted too." [34]. These incidents demonstrate the growing threat of cyberattacks in Bosnia and Herzegovina and the Intelligence and Security Agency of Bosnia and Herzegovina has also warned of the increasing risk of cyber attacks in the country and urged all state institutions and private companies to take immediate protective measures. The Federal Police Administration has confirmed that four reports of hacker attacks were received дуринг Аугуст 2023, which is consistent with the recent trend [35].

D. North Macedonia

The website of the Ministry of Education was recently targeted in a hacker attack in September 2022 [36]. The Ministry claimed that the attack did not cause any damage as there are no background databases on the website. Another attack at the same time targeted the Agriculture Ministry by the ransomware group BlackByte. The Ministry confirmed that some documents were compromised during the attack and its work blocked for some time, but it denied having lost any significant documents and data in the process [37].

E. Albania

On July 18, 2022, the Albanian government announced that it had to temporarily close access to online public services and other government websites due to disruptive cyber activity [38]. The incident was investigated by Microsoft Detection and Response Team (DART), which were engaged by the Albanian government. Microsoft stated that they had high confidence that the Iranian government sponsored the cyber-attack, and that the attack was conducted on July 15, 2022, and it caused disruption to government websites and public services [39]. After a wave of cyber-attacks targeting state institutions during 2022, the private data of thousands of

individuals has been published online without any remedy for affected individuals [40].

F. Conclusion

Overall, these case studies, which are only the most prominent cases amongst many, demonstrate the significant impact that cybersecurity attacks can have on critical infrastructure in the region, causing widespread disruption and sometimes even forcing organizations to pay significant ransoms to regain access to their systems. All the attacks were carried out from abroad, highlighting the need for better security measures and international collaboration to prevent such attacks in the future.

Regarding the attacks on the public sector, it is evident that incident response included cybercrime police and prosecution, national security agencies, mechanisms of international cooperation as well as many private actors from country and abroad. Recovery has been performed with more or less success but minimal progress on investigation and prosecution of attackers has been made (or publicly available). Considering that no perpetrator has been prosecuted and that it is unlikely that the cyberattackers will face any legal consequences for their actions, it seems that similar cyber attacks can be expected in the future.

Transparency of the incidents to the public has been very low and only in situations when citizens were already being affected by the attacks. Public reports on the incidents were not made available except in Serbia, where the Commissioner on data protection and National CERT made short statements on their website. The authorities usually did not inform public or expert community on specific consequences of the attacks and their resolutions. Finally, while there are occasional reports from private sector on necessity of ransom payment, there is no such confirmation from public sector. In general, it seems that low and unclear transparency with very limited proactivity further undermined public trust in the capacities of states to counter cyber attacks.

Case studies are packed with examples of the impact of cyber attacks on public sector targets, business entities and citizens. Potential losses are a consequence of business interruptions, incident response costs and third party damages (material and non-material), especially in situations when public sector operations have been interrupted. However, there have been no studies on the assessment of the impact of the attacks. Finally, there is no process available to affected parties to report or claim damages nor any accountability by the governments.

The motivation behind the attacks seems to be double: financial gain but more frequently in the last period of time, we can identify ideological motivation in the context of international relations.

VI. Recommendations regarding prevention and recovery from the ransom attack

The Western Balkan economies should further undertake the regulatory reform and build capacities to respond to the cyber threats. Firstly, all countries should develop their cybersecurity capacities by aligning with the revised NIS Directive. Additionally, they should adopt GDPR standards in their data protection legislation. Finally, they should improve their capacities for international collaboration in the area of cybercrime by conducting ratification of the Second protocol to Budapest convention. However, regulatory changes are just the precondition for building resilience of Western Balkan economies from ransomware and other cyber threats.

National Cyber Security Index and Global Cybersecurity Index represent some sort of a guidelines indicating a need to improve current prevention measures and procedures during a cyber incident. However, the index status did not correspond with real risks and damages suffered by WB economies. For example, Albania was well positioned on the Index scale, but has suffered the greatest damage.

There are several best practices that organizations can follow to help prevent and mitigate the impact of ransomware and other cybersecurity attacks. All organizations need to have preventive cybersecurity measures in place to protect against attacks, and to be prepared to respond if an attack does occur. This includes regular backups, software updates, and staff training. Additionally, having incident response plans in place can help organizations quickly and effectively respond to a ransomware attack if one were to occur. This plan should include steps for identifying, containing, and eradicating the malware, as well as steps for restoring the affected systems and data. It's also important to have a plan for communication and coordination with internal stakeholders as well as external parties such as law enforcement, legal team, and cyber insurance providers.

References

[1] ENISA Threat Landscape an annual report on the status of the cybersecurity threat landscape. 2021 available at: www.enisa.europa.eu/publications/enisa-threat-landscape-2021

[2] Help Net Security, “Ransomware back in big way”, 2018, available at:

www.helpnetsecurity.com/2018/07/11/2018-sonicwall-cyber-threat-report/

[3] FBI, Internet Crime Report 2021, available at: www.ic3.gov/Media/PDF/AnnualReport/2021_IC3Rep- ort.pdf

[4] Krivokapić, Đ., Nikolić, A., Stefanović, A. and Milosavljević, M., "Financial, Accounting and Tax Implications of Ransomware Attack", Studia Iuridica Lublinensia, vol. 32, no. 1, p. 197. 2023. available at: https://journals.umcs.pl/sil/article/view/14957

[5] Krivokapić, Đ., and Nikolić, A., "Legal Obligations and Liability in a Ransomware Attack", Zbornik radova Kopaoničke škole prirodnog prava – Slobodan Perović , pp. 185-192, 2023.

[6] Parties and Observers to the Budapest Convention. available at: www.coe.int/en/web/cybercrime/parties-observers

[7] T-CY Guidance Note #7 of the Explanatory Report of the Budapest Convention, available at: https:// rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTM Content?documentId=09000016802e 70b4

[8] Putnik, N., Milošević, M., and Cvetković, V., "Rensomver kao pretnja bezbednosti - društveni i krivičnopravni aspekti (Ransomware as a Security Threat – Social and Criminal Legislation Aspects)", Sociološki pregled, No. 1, 328-353, 2022.

[9] Signatures and ratifications of Treaty 224. available at: www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=224

[10] Directive 2016/1148/EU of the European Parliament and of the Council concerning meas- ures for a high common level of security of Network and Information Systems across the Union, OJ C218/1, 2016.

[11] Directive EU 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148.

[12] Regulation on the protection of natural persons with regard to the processing of person- al data and on the free movement of such data, and repealing Directive 95/46/EC - Data Protection Directive.

[13] Šarkić, N., Prlja, D., Damnjanović K., Marić V., Živković, V., Vodinelić, V., and Mrvić-Petrović, N., Pravo informacionih tehnologija. Pravni fakultet Univerziteta Union and Službeni glasnik, 2009, p. 156.

[14] ISO/IEC 27001 - Information security management, available at: www.iso.org/isoiec-27001-infor- mation-security.html

[15] National Cyber Security Index. available at: https://ncsi.ega.ee/methodology/

[16] Global Cybersecurity Index. available at: https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E/

[17] Krivokapić, Đ., Nikolić, A., Stefanović, A. and Milosavljević, M., "Financial, Accounting and Tax Implications of Ransomware Attack", Studia Iuridica Lublinensia, vol. 32, no. 1, pp.191-211. 2023. available at: https://journals.umcs.pl/sil/article/view/14957

[18] Dutton, Y.M., "Funding Terrorism: The Problem of Ransom Payments", San Diego L. Rev., vol. 53, p.335, 2016. available at: https://digital.sandiego.edu/cgi/viewcontent.cgi?article=1079&context=sdlr

[19] Flashpoint, How Ransomware has become an 'Ethical' Dilemma in the Eastern European Underground, 2017. available at: https://flashpoint.io/blog/ransomware-ethical-dilemma-eastern-european-underground/

[20] Hofmann, T., !How organisations can ethically negotiate ransomware payments!, Network Security, vol. 10, pp.13-17, 2020. available at: https://www.sciencedirect.com/science/article/pii/S1353485820301185

[21] Howard, J.W., "Kidnapped: The Ethics of Paying Ransoms", Journal of Applied Philosophy, vol. 35, pp.675-688, 2018. available at: https://www.sciencedirect.com/science/article/pii/S1353485820301185

[22] Kumar, Y., Baktybayev, B. and Anuarbek, Y., Ransom Dilemma: An ethical problem for the government?, 2021. available at: https://repository.apa.kz/xmlui/bitstream/handle/123456789/587/RANSOM%20DILEMMA%20-%20AN%20ETHICAL%20PROBLEM%20FOR%20THE%20GOVERNMENT.%ea%88%90%e0%b3%a6.pdf?sequence=1&isAllowed=y

[23] Loi, M. and Christen, M., Ethical frameworks for cybersecurity. Springer International Publishing, 2020, pp. 73-95. available at: https://library.oapen.org/bitstream/handle/20.500.12657/47324/9783030290535.pdf?sequenc#page=88

[24] Share Foundation, Kako je Novi Sad otet i zaključan, 2021. available at: www.sharefoundation.info/sr/kako-je-novi-sad-otet-i-zakljucan/

[25] Politika, Hakerski napad na katastar, 2022. available at: www.politika.rs/scc/clanak/509880/Hakerski-napad-na-katastar

[26] The record of the performed inspection supervision. available at: www.sharefoundation.info/wp-content/uploads/Nadzor-RGZ-Zapisnik.pdf

[27] The Interview with the Director of the Republic Geodetic Authority of the Republic of Serbia. available at: www.rgz.gov.rs/content/Vesti/2022/07/Intervju.png

[28] The Director of the Republic Geodetic Authority of the Republic of Serbia at the Kopaonik Business Forum. available at: www.rgz.gov.rs/vesti/5514/vest/direktor-rgz-a-na-kopaonik-biznis-forumu

[29] Bloomberg, Ransomware Attack Sends Montenegro Reaching Out to NATO Partners, 2022. available at: www.bloomberg.com/news/articles/2022-09-01/ransomware-attack-sends-montenegro-reaching-out-to-nato-partners?leadSource=uverify%20wall

[30] Antena M, Camaj: Sajber napad direktno uticao na 150 računara u 10 institucija, kriptovano 17 informacionih sistema, 2022. available at: www.antenam.net/politika/264774-camaj-sajber-napad-direktno-uticao-na-150-racunara-u-10-institucija-kriptovano-17-informacionih-sistema

[31] Security Week, Montenegro Reports Massive Russian Cyberattack Against Govt, 2022. available at: www.securityweek.com/montenegro-reports-massive-russian-cyberattack-against-govt/

[32] Klix, Napadi na općinske baze podataka u FBiH sve učestaliji: Uoči izbora ugroženi matični registri, 2020. available at: https://archive.vn/lyoCU

[33] The record, Bosnia and Herzegovina investigating alleged ransomware attack on parliament, 2022. available at: https://therecord.media/bosnia-and-herzegovina-investigating-alleged-ransomware-attack-on-parliament

[34] N1, Bosnia's state IT systems disabled for two weeks now due to cyber attack, 2022. available at: https://n1info.ba/english/news/hina-bosnias-state-it-systems-disabled-for-two-weeks-now-due-to-cyber-attack/

[35] BHRT, Postoje indicije da bi CIK BiH mogao biti hakovan na dan izbora, 2022. available at: https://bhrt.ba/postoje-indicije-da-bi-cik-bih-mogao-biti-hakovan-na-dan-izbora

[36] Radio Slobodna Evropa, Hakovana web stranica Ministarstva prosvjete i nauke Sjeverne Makedonije, 2022. available at: www.slobodnaevropa.org/a/severna-makedonija-hakeri-ministarstvo-obrazovanja/32027577.html

[37] Balkan Insight, North Macedonia Ministry Denies Covering up Ransomware Attack, 2022. available at: https://balkaninsight.com/2022/09/26/north-macedonia-ministry-denies-covering-up-ransomware-attack/

[38] Mandiant, Likely Iranian Threat Actor Conducts Politically Motivated Disruptive Activity Against Albanian Government Organizations, 2022. available at: www.mandiant.com/resources/blog/likely-iranian-threat-actor-conducts-politically-motivated-disruptive-activity-against

[39] Microsoft, Microsoft investigates Iranian attacks against the Albanian government, 2022. available at:

www.microsoft.com/en-us/security/blog/2022/09/08/microsoft-investigates-iranian-attacks-against-the-albanian-government/

[40] Balkan Insight, Albanians Mull Options as Data Security Takes New Hit, 2023. available at: https://balkaninsight.com/2023/01/25/albanians-mull-options-as-data-security-takes-new-hit

Status of Open Data (Sub)Ecosystem in Croatia: National Open Data Portal

A. Musa, P. Đurman
Faculty of Law, University of Zagreb, Zagreb, Croatia
pdurman@pravo.hr

***Abstract* - The main purpose of this paper is to assess the state of open data (sub)ecosystem in Croatia via application of open data assessment framework developed within the Horizon 2020 TODO (Twinning Open Data Operational) project. TODO assessment framework consists of four main areas - governance, availability, portal and impacts. We have applied assessment framework to Open Data Portal of the Republic of Croatia and institutional data as a selected category of open data. The results of analysis have shown that there is a lot of potential for improvements in all analyzed aspects, although most important formal requirements and technical features are in place.**

Keywords – open data (sub)ecosystem; institutional data; TODO project; assessment framework; national open data portal; Croatia

I. INTRODUCTION

Open data are one of the most salient developments in the area of e-government and e-participation. The availability of data to everyone in an open and machine-readable form, free of charge represents a specific mechanism for achieving government transparency, which goes much further in accomplishing the values of open government than the traditional transparency (access to information, whatever the form). The reuse of open data for commercial or noncommercial purposes also promotes participatory government [1], since the users (i.e. the public – individuals, civil society organizations (CSO), private businesses, media etc.) constitute a critical element in generating the final outcome of the data (re)use - different applications, sophisticated business product based on open data (such as legal information portals, business portals), scientific research and analyses, etc.

Although the literature advocates several potential advantages of open data, which can basically be divided into democratic and economic benefits, their implementation in practice has been difficult to measure. While the impacts of open data may be challenging to evaluate, different assessment frameworks capturing one or more aspects of open data initiatives (such as data quality, availability, legal context etc.) have been developed and widely applied worldwide during the past decade. While the results of such assessments can be useful in formulating and modifying national open data policies, as well as in knowledge transfer and good practice exchange between the countries, they are characterized by some serious shortcomings, both conceptual and methodological.

The main purpose of this paper is to assess the state of open data in Croatia via application of the assessment framework developed during the Online Training Program (OTP) of Horizon 2020 TODO (Twinning Open Data Operational) project, and to discuss the usefulness of such evaluations based on the interpretation of the assessment results. The TODO assessment framework is applied to Croatian open data sectoral subsystem – institutional data, which is provided within central open data Portal.

II. OPEN DATA EVALUATIONS

A. Existing models for assessment

During the last two decades, open data has strongly been advocated, especially at the international level by different actors (CSO, IT experts, businesses) and organizations (OECD, UN, WB). While the main goal behind the open data initiatives lies in increasing democratic potential of government and acquiring economic benefits, up until today it is not clear to what extent have those aims been achieved in specific countries [2]. While such wide-range effects of open data may be very difficult to capture, there is a number of evaluation frameworks encompassing different aspects of open data initiatives. Since open data is still relatively young phenomenon within e-government development, evaluation is needed as a benchmarking tool in comparisons of different countries or organizations, as well as a means for monitoring the progress of open data initiatives and formulating new open data policies [4]. As stressed by Susha et al. [5], they can "give credible insight into the strengths and weaknesses of a particular country and thus serve as powerful incentives for further improvements."

Different assessment frameworks that have been developed during the past decade can be differentiated according to open data area they asses or the type of metrics methodology they use. Open data assessments can be based on the area (i.e. aspect) of assessment - open data readiness assessments, open data implementation assessments and open data impact assessments [4]. Ref. [6] elaborates the difference between the evaluation models starting from the type of research method. While subjective models focus on users' opinion and perceptions of open data aspect (such as benefits, usefulness and other), objective models use „predefined metrics and values of them towards the assessment of specific benchmarks regarding the evaluated aspect (e.g. impact and readiness assessment)". Therefore, objective evaluation models are based on quantitative research

methods, while subjective models can use both qualitative (open questions) and quantitative (Likert scale) methods, or a combination of both types.

Based on the existing information systems research, Charalabidis et al. [6] identify eight relevant subjective streams, four of which pertain to the category of quantitative models (information systems evaluation, information systems acceptance, information systems success and e-services evaluation) and four of them characterized by either qualitative or quantitative research methods (maturity assessment, readiness assessment, post adoption and impact assessment). Subjective quantitative evaluation aims at obtaining users' opinion about different aspects of their user experience, such as perceived usefulness, user-friendliness, future behavior, etc.), assessed on the five or seven-point Likert-scale. Within the second category of subjective evaluation methods, which can be both qualitative or quantitative, maturity models are preoccupied with the ability of an organization or a country for an improvement, which is commonly conceptualized through different stages or phases. Readiness assessments focus on organizational factors acting as an enabler or barrier to opening the data. Post adoption assessments deal with issues of acceptance, routinization, and assimilation of open data, while impact assessments are aimed at measuring the impact of open data at different levels (macro – effect on the society and economy, meso – impact on the specific sector, and micro – effects of specific datasets usage). Objective Evaluation Models rely on specific quantitative measures according to which different aspects of open data can be assessed, for example data quality, popularity of datasets and similar. Examples include indexes such as Open Knowledge Foundation, Open Data Barometer, and many others [6].

Alongside potential benefits of such assessment frameworks, the literature has also identified several limitations, both conceptual and methodological nature. With open data still being young research area, some concepts may be differently understood and defined, with varying indicators (for example, data quality). Assessment frameworks, as well as open data initiatives themselves, have been developed by different actors, with different points of departure and purposes, which is why they are very context specific and most useful when applied in specific situation. Since particular assessment frameworks differ in their scope and focus, as well as methodology, they can result with varying country ranks [5]. Evaluation of open data usually takes form of benchmarks, which in general suffer from the problem of shallowness, often with scarce benchmarks and inability to measure the progress [5]. They rather capture an immediate snapshot of the phenomenon instead of the real, wider picture. Therefore, future research should encompass „more in-depth and focused look at what else can be learnt from open data benchmarks“ [5]. Finally, the existing evaluation frameworks largely focus on data supply and data environment, neglecting the user perspective [2], which is also an important shortcoming.

One of the most relevant open data benchmarks in Europe is *Open Data Maturity Report* [3], published annually by the European Commission and assessing open data maturity in four dimensions – policy, portals, impact and quality. Countries are placed within four categories according to their score: ‘trend-setters’, ‘fast-trackers’, ‘followers’ and ‘beginners’. In latest, 2022 Report, eight countries scored best as ‘trend-setters’: France, Ukraine, Poland, Ireland, Cyprus, Estonia, Spain and Italy. In 2021, the category of best performing countries included France, Ireland, Spain, Poland, Estonia and Ukraine. Malta and Slovakia are examples of countries remaining at the bottom of the list for the past several years, while some countries have shown a remarkable growth between reporting period, such as Hungary in 2022 and Ukraine in 2021. In general, countries score better in policy and portal dimension, in comparison to quality and, especially, impact dimension.

Croatia has been placed slightly above the EU average in the last few cycles of the Report, with a continuous increase of open data maturity until 2021. In 2019, Croatia scored 69% (EU average 66%), in 2020 - 82% (EU average 78%), and in 2021 – 84% (EU average 81%). In 2022, Croatian score decreased to 72%, compared to EU average of 79%. However, Croatia remained within the category of ‘followers’ (alongside Austria, Belgium, Bulgaria, Croatia, Finland, Germany, Hungary, Luxembourg, the Netherlands, Portugal, Romania, Sweden, Switzerland and Serbia) as it was placed in 2021, together with Finland, Sweden, Greece, Bulgaria, Latvia, Romania and Czech Republic. In 2020, Croatia was among the ‘fast-trackers’, alongside the Netherlands, Greece, Finland, Germany and others [3].

B. TODO OTP Assessment framework

Within the project "TODO – Twinning Open Data Operational" [7], a Horizon2020 Twinning project, an online training program has been developed and implemented covering the fundamental and more advanced areas of open data research and practice. Online training program consists of three modules. First module involves an introduction to open data, focusing on individual lectures and quizzes on basic concepts of open data and open data ecosystem. Second module - assessing open data – includes the development of conceptual-methodological framework for the analysis of open data on Croatia, which was performed by interdisciplinary groups from different faculties of the University of Zagreb and their supervisors from Delft University of Technology (the Netherlands) and University of the Aegean Greece). Final, third module - Open Data in Croatia – focuses on practical application of the previously developed framework on the Croatian open data ecosystem or its part (subsystem).

Open data assessment framework developed as a result of interdisciplinary team work within the second module is a subjective, combined (mostly quantitative) assessment framework, aimed at assessing national open data ecosystem or its part, such as specific sector or level of government. The questionnaire mostly comprises multiple choice questions or ranking option (close-ended questions), but it also includes few open-ended questions reflecting the estimation or interpretation of the evaluator. Therefore, the assessment framework can be categorized

as subjective assessment framework mostly relying on quantitative research measures, with some quantitative measures. The questionnaire encompasses 53 questions in total, which are divided into four thematic areas. *Open Data Governance* is covering issues of open data policy (questions 2-7), which corresponds the characteristics of open data maturity and readiness assessments. Second part, *Open Data Availability*, comprises questions related to legal aspects of open data (questions 8-23), while technical aspects are covered in third part of the assessment - *Open Data Portals* (questions 24-41). Final part of the framework, *Open Data Impact*, contains questions on initial impacts of open data at meso and macro level as well as some questions related to the user perspective (questions 42-53).

In this paper, the TODO OTP Open data assessment framework is applied on the Croatian National Open Data Portal (*Portal otvorenih podataka*) [8], which contains different institutional data. Since Open Data Portal is one of the thematic areas within the assessment framework, including set of questions referring to the Open Data Portal itself, we opted to analyse institutional data as a specific (sectoral) open data subsystem. Therefore, within the sections on Open Data Governance and Open Data Availability we refer to the domain of institutional data as an Open Data ecosystem, while in the section Open Data Portals, the Croatian Open Data Portal was assessed specifically. Each area of assessment – governance, availability, portals, impacts - is briefly described according to the indicators of TODO assessment framework.

Institutional data represents a separate category of open data, together with political and public administration data, including data on organisational and functional aspects of political and administrative organisations and other public sector bodies at different levels of government (state, local, regional). Examples of such data would include data on election results, electoral constituencies, campaign financing, public officials, list of public sector bodies, register of national minorities' councils, schedule of government meetings and similar [8].

TABLE 1. Areas of the TODO OTP Open Data Assessment framework

OPEN DATA GOVERNANCE	***OPEN DATA AVAILABILITY***	***OPEN DATA PORTALS***	***OPEN DATA IMPACTS***
Policies *Strategies*	*Legal aspects – licenses, fees, registration*	*Technical aspects – formats, functionality, metadata, web statistics*	*Initial socio-economic effects* *User perspective*

III. APPLICATION OF THE TODO OTP ASSESSMENT FRAMEWORK ON CROATIAN OPEN DATA (SUB)ECOSYSTEM

A. *The state of open data in Croatia*

Open data policy was formally lately adopted in Croatia – in 2018, although the open data have been introduced in 2013 by the Law on the Right to Access Information (LRAI) [10] as part of the systemic regulation of the right to access information. In some sectoral areas (for example, geodata) open data existed earlier. Legal obligation to publish open data and accompanying requirements are stipulated by the Law on the Right to Access Information, which transposes the PSI Directive. The Law requires that each public body has to ensure that data is made public on the internet whenever it is possible, in machine readable formats (CSV, XLS, XML, JSON, HTM). Open data is available at the Open Data Portal of the Republic of Croatia, which represents central Portal for assembling and (re)use of public sector open data. The Portal was established in 2015 by the Ministry of Public Administration, and with the support of Croatian Regulatory Authority for Network Industries and Omega software firm. In technical terms, a combination of two systems was used (Drupal and CKAN), modelled after the British Open Data Portal (Data.gov.uk To Go). The new and upgraded Open Data Portal is now run by the Central Office for the Development of Digital Society. It contains around 2.143 datasets. Alongside Open data portal, open data is available on public bodies' official websites and some specialized portals (such as environment, geodata, statistics, etc.) [10]. The Open Data Portal is linked to the European Open Data Portal thus adding to the European wide pool of data open for the reuse.

Local governments have been largely lagging behind in making available their open data when compared to national. Although subject of the requirements of the LRAI, majority of local units are failing to publish data and information in machine-readable formats. Currently, four local units have established open data portals (Zagreb, Rijeka, Virovitica, Varaždin), with several other cities and municipalities offering some registers and other information in open formats. However, the number of such datasets is low, and most commonly accessible in excel file. Front-runners are City of Zagreb, with Portal established in 2015 and 129 datasets, and the City of Rijeka (Portal introduced in 2016, with 182 datasets currently). The quality of the portals datasets is around the level 2 or 3 (out of 5), encompassing areas of education and health, business and tourism, environment, energy, transport, public finances, infrastructure, institutional and statistical data, etc.

B. *Open Data Governance: Policies and strategies*

There is a formal Open data policy in Croatia (*Politika otvorenih podataka*) [12]. It was adopted on July 19th 2018 by the Croatian Government as a strategic direction for further development of the transparency and openness of public administration. The policy applies to all public authorities at national, regional and local level, i.e. government organizations, local and regional government, agencies and other bodies with public authorities. Although it is not specified to which bodies it applies, it can be assumed that Open Data Policy applies to all public authorities, as it appears from the provisions of the Law on the Right to Access Information (LRAI). The adoption of the policy was stimulated by the European PSI Directive, as well as the EU benchmarking report on the Open Data Maturity, showing no formal policy nor strategy was adopted during the 2017. With regard to the

scope of government datasets that Open data policy aims to make available, it comprises all public bodies (national as well as local), which are according to the provisions of the LRAI obliged to publicise, in principle, all data except those exempted by the law (personal data, classified information, and similar).

Open data policy envisages that the Action Plan has to be adopted (by the Council for State Information Infrastructure, a Government body). However, no formal strategy on open data nor Action plan has been adopted yet (Action plan is in the process of drafting). However, there are initiatives and institutions promoting open data policy. Open Government Partnership Action Plans in Croatia (both for the period 2017-2020 and 2022-2023) envisaged continuous progress in opening data as one of the measures of openness policy with main purpose to increase the amount of open data and promote the reuse of data [13].

Government organization responsible for the coordination of Open data policy in Croatia is Central government office for the development of digital society. Government Office provides support to the Croatian Government in the development of digital infrastructure and digital public services, and it promotes further development of digital society in different aspects (economy, public sector, citizens) in accordance with the EU policy on digital economy and society [14].

In addition, the Government has established in 2018 Coordination for ODP actions implementation, which is comprised of 8 members - two representatives of the Central government office, two representatives of Information Commissioner, two representatives of Ministry of Administration and two representatives of Central government office for civil society. The responsibilities of the Coordination include drafting of an Action Plan and monitoring its implementation. However, in formulation and coordination of open data policy are involved other government/public authorities, such as Information Commissioner who had very important role in fostering the adoption of the policy as well as Government Office for Civil Society.

Therefore, open data policy is largely fragmented and lacks strong institutional steering mechanism, as well as vivid political will, since the majority of policy achievements have been accomplished so far due to personal efforts of individual leaders of involved institutions and not due to unambiguous and continuous government (institutional) support.

C. Open Data Availability

For the purpose of assessing the aspect of open data availability, we analysed two datasets. First one is a register of public authorities, which is a database provided by Information Commissioner and is available at its official website (http://tjv.pristupinfo.hr/). Database is also available at the Open Data Portal, as well as by using generic search terms, placed as the first results of the search. In addition, it is available at some other, non-government platforms [15].

Dataset is accessible without prior registration and entirely free of charge. Dataset is published in national open licence (*Otvorena dozvola RH*), which is by a content equivalent to CC-BY licence, i.e. the user has to indicate the source or link where the information is available [16].

The second analysed dataset was National census – population by territorial units (last census from 2021) [17]. The dataset is findable using generic search terms, and also available at the website of the Croatian Bureau of Statistics which is the provider of the dataset [18], alongside the Open Data Portal. Registration is not required, i.e. anyone can access the dataset without prior registration. Dataset is also available free of charge. Dataset is published under national Open data licence which requires the indication of the source of the information by the user. In comparison to first dataset (database of public authorities), second dataset (National census) was somewhat harder to find, due to large number of similar datasets. In addition, the newest dataset with population census of 2021 has been published as open data only recently, due to data processing after conducting a census.

The analysis of two selected datasets (*Database of public authorities* and *National census*) showed high level of open data availability. Both datasets are in general easy findable, free of charge, without registration and published under national open licence. However, this does not have to be the case with some other datasets available at the Portal, and the analysis results of two datasets cannot be generalized to other institutional data.

D. Open Data Portal

Within the third part of the assessment, Open data portal of the Republic of Croatia (*Portal otvorenih podataka RH*) was analysed in general. The national open data portal is envisaged to serve as a data hub for the collection, categorisation and distribution of open public sector data. The portal represents a type of metadata catalogue that enables easier searching of open data. The Portal harvests datasets from local open data portals as well.

With regard to Berners-Lee categorization, the level of datasets mostly includes a mix of star levels, mostly 2 to 3 stars. Out of 2.148 datasets, formats used comprise: XLSX (492), CSV (384), XLS (372), HTML (238), XML (113), aspx (66), PDF (406), WMS (22), JSON (45), KML (45) SHP (45). With regard to search functionality, it is possible to alter option for topic, publisher, format, stars and the frequency of data update, and to sort the results by the importance, title, last updated and openness of the dataset. Option for advanced search offers opting for a preferred period of data publishing and the period of last update. Open data are mostly available via download service, although other services are available as well (a discovery service, a viewing service, e.g. web mapping service, an API, a request form). A preview function before downloading is also available. There are no download options for available datasets, only a single button with which dataset can only be downloaded as one complete file.

Metadata is in principle documented adhering to a metadata standard (e.g. ISO 19115, DCAT). Up until now, metadata is available in national language only. With regard to its completeness, there are missing metadata fields, although most important data are listed. Data provenance / data source(s) are clearly listed in metadata. Actuality of the datasets is clear, but only actual version according to update date is listed in metadata (with no specification of last update). Static datasets are updated, but not most frequently, while dynamic datasets are in near real-time (slight delay) although dataset is still useful for (near) real-time applications (being refreshed).

At the Open Data Portal there is an area to showcase most popular datasets and their applications, but only older applications are shown (not updated for over one year). Feedback options have been improved with introducing a new Portal, with separate section 'Your suggestions' including options to make a proposal for new dataset, advancement of an existing dataset or suggestion for a correction. There is also a section for comments and questions. There is a possibility to subscribe to RSS feed re update notifications and an option to send message to data provider directly (each dataset lists a contact). Other media for interaction (e.g. special social media for the Portal) are not in place, and Portal does not include an option to upload datasets.

The new Portal also introduced web statistics at the Portal, including the number of visits, downloads, published datasets per months, number of suggestions for advancements of datasets and for new datasets, and whether they have been addressed. Some web statistics has also been contained in the Reports of the Information Commissioner which has closely been cooperating with the Government Office on the open data issues. Having in mind all the mentioned technical characteristics of the Portal, its overall user-friendliness can be assessed as medium to high, since some options are lacking or could be improved – metadata language, updating, feedback options.

E. Open Data Impact

The final part of the assessment framework included section on the impact of open data. For the open data ecosystem, we have analysed – institutional data, scientific studies or reports published showing its (potential) economic or societal benefits have not been published yet. For example, the applications that are using open data are not reported on the Portal (last application reported in 2017). However, there are some activities aimed at promoting open data (re)use, such as hackathons (both national and local) and workshops and conferences on open data organized by universities (mostly University of Zagreb). The role of the academia is in that respect most relevant. Events promoting open data are also organized by government organizations, mostly by Central government office for the development of digital society and Information Commissioner, but are primarily focused on civil/public servants' education on open data (webinars, workshops, trainings). Surveys among (potential) open data users within the institutional data ecosystem have not been carried out yet, to our knowledge.

Hence, with respect to open data impact, it can be argued that different types of impact – societal, economic, or other – have not been extensively assessed yet. However, some effects and possible impacts are addressed in scientific studies and papers, largely produced by TODO consortium members. This research derives from different scientific disciplines and areas of open data usage (geospatial, transport, agricultural, legal), including interdisciplinary studies [19]. Level of user inclusion (as well as the assessment of open data re-use and users) remains rather low, with activities still primarily aimed at open data promotion and raising awareness.

IV. CONCLUSION

The application of the TODO assessment framework on institutional data as part of Croatian open data ecosystem has shown some interesting insights but has primarily pointed at wide area for improvements. First, following its formal introduction, Open Data Policy has remained in its beginnings, with important regulations still missing (strategy, action plan). Policy itself is fragmented within different strategic documents and acts referring to related issues. The coordination of open data policy is also rather fragmented and could be improved by systemic steering as well political commitment. Second, availability of analyzed datasets was assessed rather high, with most of technical requirements for open data in place (findability, charges, registration, licenses). Third, the assessment of the Open Data Portal in whole showed medium to high level of user-friendliness, with some deficiencies. Evaluation of open data impact is probably the most problematic aspect of the assessment, since comprehensive and in-depth impact assessments are still lacking. There are few scientific and other types of studies and reports focusing on different aspects of open data effects, which have emerged as a result of HORIZON2020 TODO project. Finally, the assessment framework pointed out some peculiarities of the national open data availability features, but the usefulness of such an approach could be significantly improved by a comparative approach which could explain some of the differences and similarities between specific countries' status of open data. Nevertheless, the project has significant role in informing and promoting open data (re)use and its benefits, primarily within the academia and public sector organizations in Croatia. What is still needed is raising awareness on the user side as well (different categories of potential users, general public).

ACKNOWLEDGMENT

This paper is a result of the TODO project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857592.

REFERENCES

[1] A. Musa, I. Koprić, Ivan, D. Dobrić Jambrović, and P. Đurman, "ICT supported participatory governance at the local level in Croatia," in M. Doucy, M. Dreyfus, N. Noupadya, S. Guerard, B. Kingwesi, A. Tudor, Eds. Democratic and Electronic Changes in Local Public Action In Europe: REvolution or E-volution? Paris: Institut Francophone pour la Justice et la Démocratie - Louis Joinet, 2022, pp. 315-336.

[2] F. Welle Donker, B. van Loenen, "How to assess the success of the open data ecosystem?", International Journal of Digital Earth, 10(3), pp. 284–306, 2017.

[3] Open Data Maturity Report 2022, available at https://data.europa.eu/en/publications/open-data-maturity/2022. Full versions of the reports, as well as overviews, are available annually from 2015 onwards.

[4] G. Vancauwenberghe, "Assessing Open Data" in B. van Loenen, G. Vancauwenberghe, J. Crompvoets, Eds, Open Data Exposed. The Hague: Asser Press, pp. 149-172, 2018.

[5] I. Susha, A. Zuiderwijk, M. Janssen, A. Grönlund, "Benchmarks for evaluating the progress of open data adoption usage, limitations, and lessons learned", Social Science Computer Review, 33(5), pp. 613-630, 2015.

[6] Y. Charalabidis, A. Zuiderwijk, C. Alexopoulos, M. Janssen, T. Lampoltshammer, E. Ferro, "The World of Open Data: Concepts, Methods, Tools and Experiences", Cham: Springer, 2018.

[7] https://todo-project.eu/en TODO project is aimed at improving the interdisciplinary scientific excellence and innovation capacity of the University of Zagreb in the field of open data to boost the supply and use of open government data in Croatia and beyond. Project partners from the University of Zagreb include Faculty of Geodesy as the coordinating institution, Faculty of Electrical Engineering and Computing, Faculty of Organization and Informatics, Faculty of Law, Faculty of Transport and Traffic Sciences, Faculty of Agriculture, Delft University of Technology and University of the Aegean. Project has started with its activities in October 2019 and will be completed by March 2023.

[8] https://data.gov.hr/

[9] P. Đurman, D. Nikić Čakar, D. Boban, "Open election data: evidence from Croatia in a comparative perspective", Interdisciplinary description of complex systems, 20(2), pp. 149-161, 2022.

[10] Law on the Right to Access Information („Official Gazette", 25/2013, 85/2015, 69/2022)

[11] For instance, geospatial data is available on several platforms - https://www.geohrvatska.hr/, https://geoportal.nipp.hr/, https://geoportal.zagreb.hr/ (local portal), etc.

[12] https://rdd.gov.hr/politika-otvorenih-podataka/281

[13] https://udruge.gov.hr/partnerstvo-za-otvorenu-vlast-271/271

[14] https://rdd.gov.hr/o-sredisnjem-drzavnom-uredu/9

[15] https://imamopravoznati.org/body/list/all

[16] https://data.gov.hr/otvorena-dozvola, https://data.gov.hr/library/pravilnik-o-vrstama-i-sadr%C5%BEaju-dozvola-kojima-se-utvr%C4%91uju-uvjeti-ponovne-uporabe-informacija

[17] https://data.gov.hr/ckan/dataset/statistika-u-nizu-popis-stanovnistva-kucanstava-i-stanova-2021

[18] https://dzs.gov.hr/naslovna-blokovi/u-fokusu/popis-2021/88

[19] For instance: L. Hrustek, M. Tomičić Furjan, F. Varga, A. Džidić, B. van Loenen, D. Šalamon, "Towards Digital Innovation: Stakeholder Interactions in Agricultural Data Ecosystem in Croatia", Interdisciplinary description of complex systems, 20(2), pp. 190-209, 2022.; A. Miletić, K. Kević, A. Kuveždić Divjak, "Otvoreni podaci u službi održivog upravljanja vodama i morskim okolišem Hrvatske", Geodezija i vode, Zbornik radova 15. simpozija ovlaštenih inženjera geodezije. Opatija, pp. 57-62, 2022; M. Vujić, L. Dedić, M. Tomičić Furjan, I. Pihir, "The benefits of open data in urban traffic network", in L. Knapčíková, D. Peraković, A. Behúnová, M. Periša, Eds., 5th EAI International Conference on Management of Manufacturing Systems, Cham: Springer, pp. 267-282, 2021.

Information Security among SMEs in Hungary - An Overview

N. Mike*, E. Krén** and T. Kecskeméti**
* Corvinus University of Budapest / Librantis, Budapest, Hungary
** Librantis, Budapest, Hungary
all authors: firstname.lastname@librantis.hu

***Abstract* - Information and cyber security are important for SMEs. The level of security contributes greatly to the competitiveness of companies, but it is greatly underrepresented among SMEs in Hungary. The study aims to answer the question of whether accelerated digitization harms information security in Hungary. The analysis mainly focuses on companies actively involved in e-commerce during and after the Covid-19 era. Further, the trends in information security are compared within Hungary and the European Union, highlighting a local-specific deficit. The study presents the results of 2020, 2021, and 2022 quantitative research conducted by Digimeter company, as well as publicly available data from the European Union's DESI index (Digital Economy and Society Index) and NCSI (National Cybersecurity Index). The expected results of the research confirm that the lack of information security is visible in Hungary.**

Keywords* - *information security, data protection, data clustering.

I. INTRODUCTION

Readers may know the story of the three little pigs. There are several lessons to be learned from this story. The first is that consolidating an infrastructure is best done by trial and error. After being chased from the haystack and destroyed woodshed, the piglets finally reach the safe place: a brick house. The second lesson is that danger often takes the form of a similar threat. The aim of the wolf has always been to catch the piglets, only the manner has not changed. The third lesson is that strong infrastructure can be a strategic advantage. This means that improved security requires proactive planning.

In this study, the authors set out to investigate the information security (IS) level in small and medium-sized enterprises (SMEs) in Hungary. The need for the study is motivated by the increasing contribution of IS to competitiveness.

The authors argue that IS remains underrepresented in the SME sector. Even among European Union (EU) Member States, the level of IS in Hungary is considered immature, but there are already signs of its need. This study also aims to identify the sectors where SMEs have enhanced IS.

II. RESEARCH METHODOLOGY

The study attempts to answer the question of whether accelerated digitalization has a negative impact on the level of IS in the lives of SMEs in Hungary. The analysis focuses on companies actively involved in e-commerce and does not identify them individually. The research is based on a case study where manipulation of participants' behavior is neither the goal nor the option [1]. The desired effect is not intervention, but observation and reporting.

Among the available research frameworks, a pragmatic approach is chosen [2], which allows for effective focusing of the research question [3, 4] and its examination through multiple lenses [5,6].

The study begins with a review of the literature to provide the theoretical background and briefly introduces the indicators used in the study. Then the reader is led to the research findings and their evaluation.

To answer the proposed research question data provided by the Digimeter Index (DI) was used. DI is operating as a questionnaire-based empirical research tool to assess the digital presence of companies in Hungary.

The sample size varies from year to year: 777 responses submitted in 2020; 757 in 2021 and 674 in 2022. No different level of significance is observed: The most represented group consists of responses from SMEs with 5-9 employees, established or working in Pest county. The results are presented using cluster analysis [7]. The paper concludes with a description of the limitations of the research and a conclusion.

III. LITERATURE REVIEW

IS stands for the protection of information during its creation, processing, storage, and transmission. The disposal can be achieved through logical, technical, physical, and organizational measures that compensate for the potential loss of confidentiality, integrity, and availability [8].

IS should be managed globally according to the ISO /IEC 27001 standard [21]. Organizations are expected to comply with the standard, which covers security management, security of corporate assets, and IT security expectations.

Understanding of IS as a concept is context dependent as it carries some level of subjectivity [9]. The following two definitions summarize what the authors consider appropriate for IS:

"An activity or process, a capability or skill, a condition by which information and communication systems and the information they contain are protected from damage, unauthorized use, modification, or exploitation." [10].

"Cybersecurity is a collection of tools, policies, security concepts, safeguards, guidelines, risk management approaches, measures, training, best practices, and technologies that can be used to protect the cyber environment and the assets of the organization and the user." [11].

Companies, but especially SMEs, are forced to deal with IS. Their business operations depend on the use of information technologies and network systems which are essential to support their decision-making processes. This dependence can make them particularly vulnerable to threats from IS as they have limited human and technical resources and limited ability to address such vulnerabilities [12].

In 2012, there were almost 30,000 enterprises in Hungary, not counting sole proprietorships and micro-enterprises. Both international and Hungarian surveys show that companies do not pay enough attention to IS. The situation is particularly depressing for SMEs. The expectations placed on them are changing rapidly and dramatically, as is their business environment.

Although SME IS activities or lack thereof, pose less risk, they must be constantly monitored and renewed. There are many ways to IS. To raise awareness the European Union Agency for Cyber Security (ENISA) has produced a specific publication to help businesses develop IS [13]. There is a need to invest in IS systems that in turn protect against security incidents [14].

SMEs need an IS system that is affordable, easy to implement, easy to use, and prevents damage from security incidents. An unused IS system is like a lock on an unlocked door. A security system that does not prevent theft does not protect users or SMEs against security incidents [15]. Assessing the IS level of SMEs in Hungary the authors used the following indexes:

A. Digital Economy and Society Index (DESI)

DESI is a composite index measuring the progress of EU countries in the digital economy and society. The index is calculated by the European Commission (EC) and covers four key dimensions of the digital economy and society: (a) human capital, (b) connectivity, (c) digital technology integration, and (d) digital public services [16].

B. National Cybersecurity Index (NCSI)

NCSI is a measure of a country's overall cybersecurity posture that takes into account several factors, including the country's legal framework, technical infrastructure, and public awareness and understanding of cybersecurity issues. The index provides a comprehensive picture of a country's cybersecurity environment and efforts. It is used to assist in assessment and improvement.

C. Global Cybersecurity Index (GCI)

GCI is a composite index developed by the International Telecommunication Union (ITU) to measure the cybersecurity readiness of countries around the world. The GCI aims to provide a comprehensive overview of countries' cybersecurity capabilities and help identify areas for improvement. The GCI measures cybersecurity readiness through the following main pillars: (a) legal measures, (b) technical measures, (c) organizational measures, (d) capacity building, and (e) cooperation. These pillars are then broken down into sub-indicators that are used to assign a score to each country [17].

D. Digimeter Index (DI)

DI is a digital presence measurement tool developed by Smartcommerce Consulting, Reacty Digital, Virgo, and eNET to measure the digital readiness and capabilities of organizations in Hungary [19]. The tool provides a comprehensive assessment of an organization's digital maturity and covers several areas. The DI consists of six sub-indices: (a) digital presence, (b) digital daily life, (c) corporate governance, (d) sales and marketing, (e) digital finance, and (f) IT security [20]. The assessment is usually conducted online, and respondents are answering a series of questions about their digital capabilities in each of the areas covered above.

IV. RESEARCH RESULTS

Based on the data received from DI, the companies were divided into four different categories using cluster analysis. The differences between the groups in terms of IS can be seen in Figure 1. The distribution of scores for each group differs significantly. A simple interpretation is that the clusters formed as a group of companies all have different IS systems and strategies.

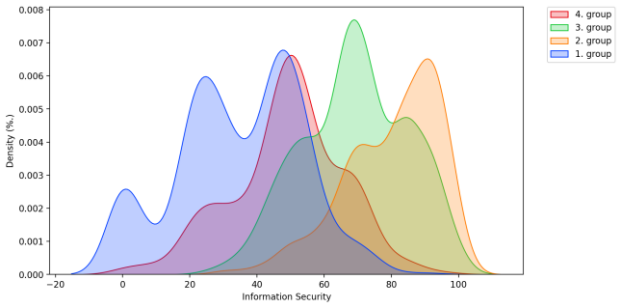


Figure 1. The density of cluster groups.

Staying with Figure 1, it is worth noting that the peaks of the distribution curves showing the mode of each category cluster around these values: (a) 45 points for 1. group; (b) 92 points for 2. group; (c) 67 points for 3. group; (d) 50 points for 4. group. The average scores were 34, 80, 69, and 49 points, respectively. This shows that each of the clusters formed has, on average, different IS systems or measures, if any, in the companies. An obvious question is if the overall score of the IS index separates the clusters in this way, how much variation is to be expected in the other aspects and characteristics?

It is also important to note that the most frequent scores, i.e. the mode of each group, are invariably 10-20 points lower than the mode of the next category. This shows that moving up to the next group is only possible at the price of

implementing significant additional security measures, even at the strategic level.

This is confirmed by the greater divergence of the 2. group from the other clusters. It should be noted, however, that the average IS score for this group is 80 points, which shows that, as with the other groups, it is not a completely homogeneous group, so there is still room for improvement even among the companies that score below average.

The assumption that companies with a lower level of digitalization care less about IS finds further support. Figure 2 describes the relationship between IS as a subset of DI and DI in isolation. SMEs with a lower index also either protect information to a lesser extent or have not yet focused on this aspect of their business. Taking into account the trends of recent years, one can even conclude that this level of preparedness is likely to move to a higher level shortly.

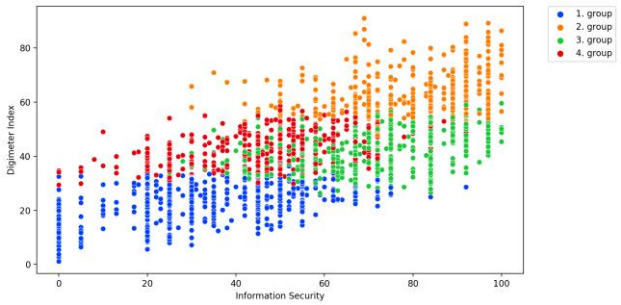


Figure 2. Clustered relationship between IS and DI.

Further analysis shows that one of the most meaningful links is between IS and sales and marketing, as shown in Figure 3. Companies that are leaders in digitalization are thus placed in this category not only because of their different digital presence but also because the IS level within these companies is higher. The sales and marketing activities confirm exactly this, as companies that frequently advertise and sell online must have up-to-date knowledge in the area of IS and data protection.

One of the main reasons for this is likely to be that compliance with legislation and the ever-changing demands of a digital presence requires constant adaptation. The need to adapt and stay up to date can only be achieved through a strong focus on this area. Without this, it would be exceedingly difficult for companies to continue their digital expansion.

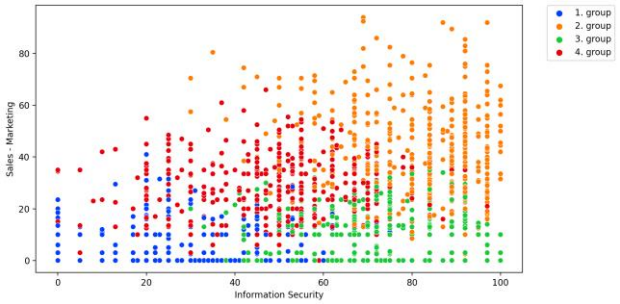


Figure 3. Clustered relationship between IS and Sales - Marketing.

Meanwhile, DESI has been measured every year in the European Union since 2014, so that changes in Hungary's situation can be tracked continuously. According to DESI, Hungary ranks 22nd out of 27 member states in 2022 [16].

Hungary's overall score of 43.8 shows that the country has developed in line with the EU average in recent years, but still lags significantly behind in the areas assessed in this study. The most disappointing result is in the integration of digital technologies. In this area, Hungary ranks 25th, with 14.5 points behind the EU average [16].

In terms of sub-scores, we see improvements in several areas, but the results show that there are still many Hungarian companies that do not sufficiently exploit the potential of digital technologies.

Resource planning systems that allow information to be exchanged electronically are used by 21% of companies [16]. 13% of companies have some form of social media presence, which shows that Hungary is well below the EU average in these areas [16]. The use of different systems and platforms does not mean that all companies use them properly and safely. The use of advanced technology would help protect the information, but Hungarian companies perform worst in this area. By advanced technology, the authors mean artificial intelligence, which is used by 3% of companies, Big Data, which is used by 7%, and cloud technology, which is the most widespread and is used by 21% of companies [16].

The DESI surveys SMEs in three areas: (a) online retailers (18%); (b) e-commerce sales (12%); (c) cross-border online sales (9%). There is a slight increase in all three areas. The results show that Hungarian SMEs are lagging in digitalization. One-third of the companies have a basic level of digital presence, which is extremely low compared to the EU average of 55%. In the coming years. It is crucial that SMEs move as close as possible to the EU, which is closely linked to the focus on data protection and IS.

According to the NCSI, Hungary ranks 35th in the world based on the last review, which took place on 13 October 2022, with a score of 65.53 out of a maximum of 100 [17]. The data before the 2022 measurements, taken from 2018 and 2019, clearly show a stagnant situation. This means

that Hungary's information and cyber security have neither increased nor decreased over the last four years.

The above-mentioned score of 65.53 is the result of four components, of which the ICT development index is the least acceptable. If the reader takes only this one indicator as a yardstick, Hungary slips back to 48th place [17]. This observation is consistent with the statements in the DESI.

Also in the ITU 2020 report, Hungary ranks 35th among the countries studied by researchers and 22nd among European countries [18]. ITU research produces a GCI that assesses the commitments of countries participating in the research, by answering to an 82-question questionnaire, divided to five pillars. Figure 4 illustrates Hungary's strengths and potential for improvement.

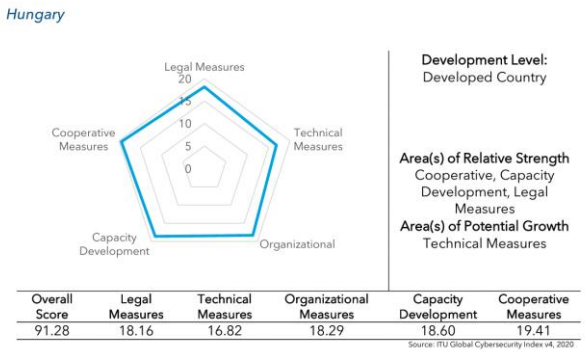


Figure 4. Global Cybersecurity Index (GCI) 2020. [4]

Overall, measures in four of the five pillars are considered sufficient, with room for further improvement in technical measures.

The areas of relative strength are due to successful national transposition of Directive no. 2016/1148 of the European Parliament and of the Council (the NIS Directive) [22], as well as Directive 2022/2555 of the European Parliament and of the Council (the NIS 2 Directive) [23]. Hungary had modified 40 of its national laws as a result of these legal instruments. The most national legislation is the Act CXII of 2011 on informational self-determination and freedom of information. This Hungarian law governs the protection of personal data and the right to access information of public interest. The Act CXII of 2011 is also affected by the applicability of GDPR [24]. This presents the *de lege lata* state concerning IS in Hungary.

The paper is not focused on providing *de lege ferenda* recommendations for changes or improvements to the existing legal framework. However, some potential and generally acceptable recommendations could include: (a) enhanced cybersecurity measures, that are proposing new laws or amendments to existing ones that focus on promoting stronger cybersecurity measures, including incident response planning, risk assessment, and the adoption of best practices across all sectors.

(b) explicit support for SMEs, as they lack the resources and expertise to adequately address IS issues, by inserting specific provisions for support programs, training, and resources specifically aimed at these organizations.

Of course, the question arises as to what would motivate businesses to spend on technical improvements if the performance on these is not outstanding at the national level. The answer lies in the global nature of cyber threats, as any business, regardless of its geographical location, can be the target of a cyber-attack.

V. LIMITATIONS

This section of the paper acknowledges several limitations of the study. Firstly, the use of data for only three years (2020-2022) provides a narrow overview of the situation, although these years were critical for digitization due to the pandemic.

Second, the number of respondents to the questionnaires varied each year, and this influenced the extent to which companies' attitudes towards IS have changed.

Third, there is a general lack of data on digitalization in Hungary, which made it challenging to conduct a qualitative survey. IS and cyber defense is underdeveloped and under-researched in Hungary, which limits the availability of information to companies and affects their perception of the seriousness of the issue.

VI. CONCLUSION

The analysis concludes that forced digitalization has harmed the level of IS in the life of SMEs in Hungary. The research shows that a critical mass of SMEs that have started digitization in the last two years have low levels of IS.

The results suggest that SMEs in Hungary are not able to cover a wide range of IS due to a lack of time, facilities, and expertise. The authors emphasize that all businesses are vulnerable to cyber-attacks and that SMES shouldn't excuse themselves by saying that they are small and not priority targets. Moreover, Hungary is lagging in this area, not only in e-commerce but also among SMEs as a whole.

The authors recommend repeating the survey, as the rapid development of IS can lead to major changes in one or two years.

REFERENCES

[1] Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study Design and Implementation for Novice Researchers. The Qualitative Report, 13(4): 544-559.

[2] Mertens, D.M. (2005). Research methods in education and psychology: integrating diversity with quantitative and qualitative approaches (2nd ed.) Thousand Oaks: Sage.

[3] Mackenzie, N., & Knipe, S. (2006). Research dilemmas: paradigms, methods and methodology Issues in Educational Research. 16. 193-205.

[4] Creswell, J.W. (2003). Research design: qualitative, quantitative, and mixed methods approaches (2nd ed.) Thousand Oaks: Sage.

[5] Edmondson, A. & McManus, S. (2007). methodological fit in management field research. Academy of management review. 32(4): 1155-1179.

[6] Mullarkey, M. T. and Hevner, A. R. (2018). An elaborated action design research process model. European journal of information systems. 28(1): 1-15.

[7] Simon, J. (2006). Applications of cluster analysis in marketing research. Statistical Review. 84(7): 627-649.

[8] Ključnikov, A. & Mura, L. & Sklenár, D. (2019). Information security management in SMEs: factors of success. Entrepreneurship and Sustainability Issues. 6(4): 2081-2094.

[9] Craigen, D., Diakun-Thibault, N., & Purse, R. (2014). Defining Cybersecurity. Technology Innovation Management Review, 4(10): 13-21.

[10] DHS. (2014). A Glossary of Common Cybersecurity Terminology. National Initiative for Cybersecurity Careers and Studies: Department of Homeland Security. 1 October.

[11] ITU (2009). Overview of Cybersecurity. Recommendation ITU-T X.1205. Geneva: International Telecommunication Union (ITU). Link: https://rb.gy/xszzxm (accessed on 02.02.2023).

[12] Sadok, M., Alter, S. & Bednar, P. (2020). It is not my job: exploring the disconnect between corporate security policies and actual security practices in SMEs, Information and Computer Security, 28(3): 467-483.

[13] Dr. Michelberger Pál & Lábodi Csaba (2012). Vállalati információbiztonság szervezése, Tanulmánykötet - Vállalkozásfejlesztés a XXI. században II. Óbuda University, Keleti Faculty of Business and Management.

[14] Simmonds, M. (2017). How businesses can navigate the rising tide of ransomware attacks. Computer Fraud St Security, 2017(3): 9-12.

[15] Bryan, L. L. (2020). Effective Information Security Strategies for Small Business. International Journal of Cyber Criminology, 14(1): 341-360.

[16] Digital Economy and Society Development Index, 2022 Hungary. Link: https://rb.gy/oq2ddk (accessed on 28.01.2023).

[17] National Cyber Security Index (2023). Link: https://rb.gy/vlp67s (accessed on 29.01.2023).

[18] Global Cybersecurity Index 2020. (2021). International Telecommunication Union. Link: https://rb.gy/f2wove (accessed on 29.01.2023).

[19] Bak, G., & Reicher, R. (2022). A vállalkozások és a digitális fejlődés, *in*: Baráth N.E. & Mezei J. (eds): Rendészet-Tudomány-Aktualitások 2022. Doktoranduszok Országos Szövetsége, Budapest, ISBN: 978-615-6457-06-6.

[20] Smartcommerce Consulting, Reacty Digital, Virgo & Enet. (2020). Digimeter. Link: https://digimeter.hu/ (accessed on 15.01.2023).

[21] ISO standards. Link: https://rb.gy/yxvakx (accessed on 03.02.2023).

[22] OJ L 194, 19.7.2016, pp. 1–30.

[23] OJ L 333, 27.12.2022, pp. 80–152.

[24] OJ L 119, 4.5.2016, pp. 1–88.

Anti-corruption Measures in the New Croatian Public Administration Office Management Regulation

Goran Vojković, Tihomir Katulić
Faculty of Transport and Traffic Sciences, University of Zagreb, Croatia
Faculty of Law, University of Zagreb, Croatia
goran.vojkovic@fpz.hr, tihomir.katulic@pravo.hr

Abstract – **The beginning of 2023 marked the started of the application of the new Office Management Regulation in the Republic of Croatia. The Office Management Regulation updates the rules and measures of office management in the performance of state administration, other state bodies, local and regional self-government units and legal entities with public authority with the requirements of digitalization and electronic communications. The new legal regulation introduces digital first, electronic office operations, also including provisions on certain anti-corruption measures. These measures are intended to prevent abuses in the office management system of the public administration system. This paper describes how the introduction of modern electronic business can simultaneously introduce anti-corruption measures and increase the legality and transparency of the work of the state and public administration. The model introduced by Croatia can certainly be an interesting example for other countries in the process of office business reform. The paper also explores the potential application of industry self-regulation anti-bribery standards.**

Keywords: office management; public administration; paperless work process; anti-corruption; anti-bribery;

I. INTRODUCTION

On July 1st 2021, the Government of the Republic of Croatia adopted a new Regulation on Office Management (Croatian: *Uredba o uredskom poslovanju*), a basic document regulating the management of state and public administration in the Republic of Croatia (further: Regulation). [1] The provisions of this Regulation are aimed at the functioning of public sector and related bodies, and shall apply to the office management of state administration bodies, other state bodies, local and regional self-government units and legal persons with public authority. Along with the Regulation, an Instruction was adopted on the numerical markings of letters and the content of office business records (hereinafter: Instruction). [2]

According to the Croatian Public Administration System Act (Croatian: *Zakon o sustavu državne uprave*), the tasks of the state administration are the implementation of state policy, direct implementation of the law, inspection supervision, administrative supervision and other administrative and professional tasks. Also, the affairs of state administration include any work entrusted to a certain legal entity by the law governing a particular administrative area and determined as work performed as a public authority. [3]

Article 3. of the Regulation explicitly provides: "Public law bodies conduct official correspondence with each other and with natural and legal persons primarily electronically."

This is undoubtedly the most crucial change in a seemingly mundane legal requirements pertaining to public sector office management legal framework in Croatia, prioritizing electronic communication and data interchange between public sector bodies as well as with citizens and commercial entitites and continuing on a path of development of internet based services that in the earnest started with the e-Građani (e-Citizen) project in 2014.[4] The Regulation provides for full application of this provision from January 1, 2023. At the moment of preparing this paper, while there is no current research into the status of the application of the Regulation, anecdotally the public sector bodies are struggling with implementation of requirements, ranging from planning and acquiring adequate IT systems and support to process planning and staff training.

The Regulation includes a few interesting and novel anti-corruption measures. The legislator, i.e. the proponents of the Regulation, long concluded that non-transparent operations, abuses and even corruption and other punishable procedures in the public sector start precisely from the system of office operations, as its archaic paper based processes, combined with other factors influencing the behaviour and performance of public sector employees foster an environment prone to misuse. As documentation in the system of state and public administration is managed through the system of office operations, clear, precise and transparent office operations would present an efficient barrier to irregularities.

In the legislative system of the Republic of Croatia, secondary, subordinate legislative acts (bylaws) are usually not adopted with extensive explanations of the acts themselves, as in the case of primary legislation such as

laws. Therefore, the reasons for the adoption of proposed provisions as were adopted have not been officially explained by the legislators, nor were subjected to public scrutiny through a formal public consultations.

However, the intention and the desire to introduce effective anti-corruption measures can be read from the new procedures. We consider these measures important, both for Croatia as well as for other countries in the broader Western Balkans region with the same or similar public administration tradition and background, undergoing the various transitions converging on the process of EU Member State candidacy. The transition in digitalization of public administration processes is an important step in this regard, and it is likewise important to detect and explain these measures and their application and purpose.

II. PUBLIC ADMINISTRATION OFFICE MANAGEMENT PRINCIPLES AND STANDARDS

In general, public administration office management operations are based on broad principles that should help guide the development of such systems, their behaviour and efficacy. These principles include:

- Principle of lawfulness and accuracy
- Principle of expediency
- Principle of simplicity, uniformity and transparency
- Principle of frugality
- Principle of civility

The principle of lawfulness and accuracy requires that office work and acts be based on real facts and other data and circumstances determined clearly, precisely, impartially and in a lawful manner. Principle of expediency mandates that office work be performed as quickly as possible within prescribed deadlines. Of course, efficiency must not come at the expense of legality and accuracy. The principle of simplicity, transparency and uniformity is understood to require that civil servants, during office work, should avoid any action, procedure or communication that could lead to unnecessary complications in official proceedings, redundant files and records. The principle of frugality means that in office operations, the greatest efficiency should be achieved with the least amount of money and material resources, energy and time of employees and clients being spent. Finally, the civil servants need to perform their function in a civil and respectful way, in handling documents and communicating with citizens and parties, as well as among themselves, avoiding quarrelsome and conflicting behaviour, in a calm and polite manner. [5]

In addition to the principles, general documentation management standards are also applied to office operations, for example standards such as HRN ISO 15489-1: 2016 Information and documentation - Records management - Part 1: Concepts and principles (ISO 15489-1:2016). [6] The mark HRN ISO means accepting a foreign standard in the original. This part of ISO 15489 establishes the core concepts and principles for the creation, capture and management of records. [7]

III. TOWARDS FULLY DIGITAL PUBLIC ADMINISTRATION

Article 3 of the Regulation states: "*Public bodies conduct official correspondence with each other and with natural and legal persons primarily electronically*." This provision means that electronic communication with clients is the rule, and falling back to paper based (including file folders and classic paper based documentation handling) correspondence becomes (an increasingly rare) exception. At the same time, even when dealing with paper-based communication, the data on the physically opened (paper based) case file is also kept in the electronic system, which as we will state in more detail later in the paper, can represent an anti-corruption measure in itself.

In Croatian administrative practice, there have been cases (some were the subject of criminal proceedings) of unlawful and/or fraudulent document changes in existing public administration files. The obligation to keep records of the file in electronic form (even when the file itself is physical) cannot prevent the replacement of an existing document in the file with another one, but it does prevent the prohibited retroactive insertion of documents (eg forgotten consents to a decision already made). Therefore, we consider the obligation to keep electronic records and the movement of documentation through the system, even when the actual case file is on paper, a practical anti-corruption measure.

Another important anti-corruption measure introduced by the new Regulation is the introduction of a new "unique identifier". The regulation states in its glossary: "*a unique identifier is a written identifier that enables verification of originality, completeness and traceability when exchanging data between office business information systems, and is created within the office business information system by machine, by random selection*".

As this is a completely new concept in Croatian office management legislation, it is additionally explained in a separate document: "*The unique identifier is used for the exchange of data between information systems.*" It is assigned to every communication (automatically within the information system of office operations), which is created and received in the body, but is not expressed in the communication itself. The unique identifier of a communication is conceived as a UID ("unique identifier") which is guaranteed to be unique among all identifiers used for a certain type of object and for a specific purpose. GUID or HASH are usually used in systems today. [8]

At this point, those familiar with current technological trends, may react at the very term and description of a unique identifier as a possibility of introducing blockchain technology into the work of state and public administration.

To be fair, blockchain based technology as a foundation for a system of office management in public administration is not referenced in the Regulation itself. However, interestingly, it is mentioned in the National Plan for the Development of Public Administration until 2027: "*At the global level, blockchain technology has been in the center*

of attention for some time, due to the advantages it can offer to public administration, for example in simplifying administrative procedures or managing public registers. Also, artificial intelligence has significant potential for application in public administration." [9] We believe that blockchain would represent another revolutionary change, but for that, as a further step, a wider acceptance of electronic office management in the workspace of public administration is required.

IV. ELECTRONIC RECORDS OF ACTIVITES

According to Article 7 of the Regulation, documents are recorded in the information system of the office operations of the public law body. This is a mandatory provision for which there are no exceptions, for example such as that small organizations could continue to conduct paper-based office operations. This is also an important anti-corruption measure. According to Art. 9 of the Regulation, documents received through the information system of office operations or other information systems are processed electronically, recorded in the information system of office operations and, as a rule, are not printed. According to Art. 10 of the Regulation, on documents received in physical form, the name of the public legal body that received the document, the date of receipt, the numerical code and the unique code of the document are printed. The instruction in Article 12 regulates the content to be written on documents received in physical form. These are: the name of the issuing public body, the date of the receipt, the numerical code of the submission (classification code and serial number of documents within the case) and the unique code of the document (linear or 2D barcode). The unique code of the item is not the same as the unique identifier, it is the numerical code of the item in a machine-readable form.

The situation explained above means that the document submission first has to be entered into the information system, in order to print the specified data (usually a label will be used for practical reasons), which is a significantly better solution than the previous reception stamp, which was used to stamp the submitted document and then data was entered manually. With this kind of manual process it was possible to enter data later and, for example, "adjust" the reception time. This is how it was entered in the information system. Among the anti-corruption measures, we can also mention the provision of Art. 11 of the Regulation according to which the receipt of a document received in physical form is confirmed at the request of the party after recording in the information system of office operations. The afore-mentioned certificate is printed with the prescribed content from the information system. This completes a precisely arranged business process - the document needs first to be recorded in the information system, and then a confirmation of receipt is printed. [10] This new model prevents abuses and omissions possible with paperwork and manual entry of data on document application. It was recorded exactly when the document was received, it received its identifiers, and the confirmation of receipt is in accordance with this information.

As cases are now mandatorily kept in electronic records, art. 18 of the Regulation states that in the information system of office operations, records of first-degree administrative proceedings, records of second-degree administrative proceedings and records of non-administrative proceedings are kept separately in the manner prescribed by the instruction. This is one of the key anti-corruption measures in the new Regulation, which significantly reduces the possibility of non-transparent and punishable behaviour and procedures. Namely, until now it was possible to keep paper registers and this was in fact the prevailing and dominant form of record keeping in public administration. Paper registers are individual papers of a larger format in which information about subjects was entered, and then after the end of the calendar year, these individual papers would be permanently bound. In practice, this was a generator of potential irregularities. It was, for example, trivial to leave blank columns where written ones would be entered later. Also, it was possible to reprint the entire page with different content (they are bound as we said at the beginning of the year for the next year). While sometimes used as a measure to correct incorrect or fill missing information, this possibility fostered behaviour that could and did lead to committing serious criminal offenses, but at the same time in practice it was very difficult to prove with paper trail. There were (and are) many examples of such irregularities if not outright misuses. For example, approvals for carrying out various activities on the sea coast, maritime domain, are granted on request. [11] It is therefore important who submitted the request first - with paper bookkeeping, it was illegal, but in practice it was easy to "reserve" an empty number for a favored bidder. Such behavior was difficult to detect or prove in a criminal investigation or subsequent court proceedings.

Among further anti-corruption provisions in the new Regulation, we can also mention that of Art. 15 which provides that by entering the first document or entry in the information system of office operations in the case records of administrative or non-administrative proceedings, a case is established. Namely, it is not possible to create an "empty" case, without going through a letter and then subsequently "adding" a written (either input - submission, or output - document) that empty case, which was the practice in some places, especially with output documents.

V. OTHER ANTICORRUPTION MEASURES

Among the provided anti-corruption measures (even if this provision is not exactly new), we can also include the provision of Art. 21 of the Regulation according to which the uniqueness of a document within a public legal body is indicated by a unique document designation.

The designation is data that forms an integral part of the document and is unique within the information system of office operations and contains the numerical designation of the body in whose information system the document is entered, the classification designation of the subject to which it belongs and the serial number in the subject, and is expressed in the form of a linear or 2D bar code. Therefore, every document in the public system has its own unique number.

Among the anti-corruption measures provided we can also mention the arrangement of powers on the information system according to Art. 31 of the Regulation, all officials of a public legal body must have access to the information system of office operations in accordance with the assigned

authorization decided by the management of the body. The Regulation does not elaborate this provision in detail, so in its application it is best to apply the official information security standard HRN EN ISO/IEC 27001:2017. [12]

Among the anti-corruption measures, we can certainly include the use of a qualified electronic signature and a qualified electronic seal in addition to a handwritten signature and a classic seal. Electronic seal and signature are defined by Regulation (EU) no. 910/2014 [13]

Shortly after the adoption of the Regulation 910/2014, the use of electronic seals and signatures was also defined by the basic regulation governing administrative procedure - the Law on General Administrative Procedure [14], where Article 97 stipulates: "*The decision shall be signed by an authorized official by hand or with a qualified electronic signature. A decision issued from the information system of a public legal body can only be certified with a qualified electronic seal.*"

Although the use of electronic signatures in state and public administration is still more an exception than a rule and is just beginning (the latest introduction project is ongoing from 2022)[15], many certificates and records have been issued with an electronic seal over last decade and a half, for example those from the e-Citizens system [16]. The introduction of e-signatures and e-seals is a milestone anti-corruption measure. In the last thirty years of Croatian legal practice, there have been numerous cases of falsified documents, from simple travel orders, where expert signatures were also carried out in the proceedings to forgeries of medical records required for the pension system or documents required by the court company register or land registry. The use of an electronic signature represents significantly greater security.

Finally, according to the Regulation, Art. 62. the information system of office operations ought to have the ability to connect and exchange data with the information system for monitoring the implementation of the law regulating the general administrative procedure in which the application of legal institutes in administrative procedures is recorded. The Law on General Administrative Procedure has a similar provision, facilitating remote monitoring by the competent ministry (currently Ministry of Justice and Administration).

The Republic of Croatia has an extremely large number of public law bodies, so their supervision on the ground was complex, expensive and deficient (for example, Croatia has 555 local cities and municipalities as self-government units in addition to 20 regional units - counties). Then there are hundreds of public institutions such as schools and numerous other organizations Regulation is applicable to. These provisions facilitate the possibility for the officials of the ministry in charge of supervision to carry out the same remotely, without the need to go out into the field. The very possibility of such supervision will have an anti-corruption chilling effect.

VI. UNDERSTANDING, IMPLEMENTING, AND AUDITING ISO 37001 ANTI-BRIBERY MANAGEMENT SYSTEM (ABMS) IN PREVENTING BRIBERY IN PUBLIC OFFICE ADMINISTRATION

Bribery and corruption in general are global problems that pose significant threats to economies, societies, and organizations, degrading the rule of law and introducing an aspect of legal uncertainty discouraging to societal development and economic investment. It is of no wonder that apart from government activities - more stringent legislation, better white collar crime detection rates and improving prosecution capabilities by public authorities, a variety of initiatives to understand and prevent these issues have appeared to augment these efforts in a form of industry self-regulation standards.[16]

The International Organization for Standardization (ISO) has developed the ISO 37001 Anti-Bribery Management System (ABMS) to help organizations combat bribery and corruption as a framework for organizations to prevent, detect, and respond to bribery (International Organization for Standardization, 2016). It aims to help organizations establish an ethical culture and maintain compliance with anti-bribery laws. [17]

The standard was released in October 2016. It can be used as a certification standard because it is a so called Type A standard. [18] The "high-level structure" and architecture of ISO 37001 are the same as those of all other contemporary management systems standards. The first of the standard's two parts outlines the conditions for an anti-bribery management system. A thorough annex explains broad features of anti-bribery management and provides information on use. The integrated management system may also use it as a module for quality management, environmental management, and other management functions. [19]

Understanding the ISO 37001 standards is beneficial for several reasons:

- Corruption is a global concern
- Legal compliance
- Reputation and trust

As bribery and corruption are on the rise worldwide, with devastating economic and social consequences, industry self-regulation standards can augment existing and future regulation efforts and enforcement.[20] Understanding the ISO 37000 family of standards may help organizations address concerns by developing effective preventative internal anti-bribery policies and procedures that are easy to follow, audit and act upon. [21]

Adhering to these standards can also help organizations demonstrate compliance with various international and national anti-bribery laws. This understanding can facilitate better risk management and reduce potential legal penalties. Organizations that understand and implement antibribery controls can signal their commitment to ethical

business practices, enhancing their reputation and trust among stakeholders. [22] As with most self-regulation standards, while the process of implementation requires an investment in time and resources, its successful completion requires several key elements, such as the genuine commitment of the highest level of the organization management. This includes establishing an anti-bribery policy, assigning responsibility for anti-bribery efforts, and providing necessary resources. Further, an organization looking to implement needs to undertake an honest and thorough risk assessment activity, investigating and evaluating potential bribery risks, which will in turn result in applying adequate controls and measures to mitigate recognized risks. Finally, adequate training and awareness program and measures for continuing monitoring and improvement are also required. [23]

Finally, self-regulation standards such as ISO standards often contain provisions ensuring their continuing application and validity of the certification. As with all similar standards, there are various process phases involved in auditing conformity. [24] These procedures guarantee that an organization's Anti-Bribery Management System (ABMS) is successfully maintained and improved over time. To independently ascertain the status of the certification, the organization chooses an auditor or audit team with the necessary credentials and anti-bribery management system experience before the audit can start. To ensure a fair evaluation, the auditor should be impartial and independent. To comprehend the design and use of the ABMS, the auditor examines the organization's anti-bribery policy, practices, and other pertinent material. During the audit, this review aids the auditor in locating probable non-compliance hotspots and areas that need more research. Often, an external auditor visits the organization's location to collect proof of the use and efficacy of the ABMS, conducts interviews with staff members, management, and pertinent outside parties. The auditor evaluates how well the organization's ABMS complies with ISO 37001 criteria, relevant anti-bribery regulations, and the anti-bribery policy of the organization. Any non-conformities, observations, and areas for improvement noted during the audit are recorded by the auditor. To achieve a thorough grasp of the issues and their ramifications, these results are then reviewed with the organization's management. The auditor writes an official audit report outlining the findings, conclusions, and any suggestions for improvements or remedial actions. The organization then carries out a corrective action plan based on the audit's findings to deal with any non-conformities or potential areas for improvement. To confirm the efficacy of the corrective actions and guarantee continuous adherence to ISO 37001 criteria, the auditor may undertake follow-up audits or reviews. [25]

The auditing procedure should be seen as a chance for organizations to continuously enhance their ABMS. The identification of improvement areas, the promotion of a compliance culture, and the reduction of bribery risks are all aided by routine internal and external audits. Regular audits contribute to the development of an ethical culture, transparency, and compliant business processes in addition to helping to maintain compliance. [26]

VII. CONCLUSION

As we stated in the introduction, by-laws in the Republic of Croatia are adopted without extensive explanations of individual provisions, but in the new Regulation on office management in public administration we see several very important changes that can be classified as anti-corruption related provisions, and roughly divided into three categories.

The first is the mandatory keeping of electronic records in the information system. At first glance, this measure is contrary to the principle of frugality, especially in the case of smaller organizations, but it is precisely the small organizations with only a few employees where abuses and pressure on officials (often a single clerk in charge of the office) to do something contrary to the regulations can be expected the most. Data entered into the information system is significantly more difficult to change ex post facto, either in terms of content or in terms of changing the time when it was received.

The second is the increasingly widespread use of advanced electronic signatures and seals, which significantly increases the reliability of signatures - as we mentioned, with handwritten signatures, especially when we are talking about a large number of documents, officials will usually use some abbreviated forms of signature that are very difficult to forensically check and establish their originality.

Ultimately, even if not explicitly stated, the new model of electronic office business represents the basis for the introduction of further advanced information technologies such as blockchain based services and is also the subject of interest of the European Commission. As elaborated in the EC document Blockchain for Digital Government: "*In the Public Administration field, blockchain technology can potentially be leveraged in a variety of contexts, and strengthen different use cases in terms of security, interoperability and transparency*."[28]

While there is always room for further development, the regime established by the new Regulation is definitively an improvement especially as it respects the reality that not everything can be managed electronically immediately, but leading in the direction of fully digitized work processes, as the technical and organizational capabilities of those subject to the Regulation develop and new technologies are introduced, taking into account the need to strongly and actively work in terms of preventing abuses, especially corruption and bribery.

In this sense, the Croatian example can be interesting for other countries, especially in the European East and Southeast, and especially for countries that are candidates for future accession to the European Union.

REFERENCES

[1] Regulation on office management - Uredba o uredskom poslovanju, Narodne novine,* 75/21

[2] Instruction on numerical signatures of documents and records of office management activities - Naputak o brojčanim oznakama pismena te sadržaju evidencija uredskog poslovanja, Narodne novine, 132/21

[3] Zakon o sustavu državne uprave, Narodne novine, 66/19

[4] G. Vojković, (2022) "Electronic Office Management of Public Administration in Croatia".Skala, K. (ed.) MIPRO 2022 45th Jubilee International Convention Proceedings.

[5] R. Odbaša, Uredsko poslovanje tijela državne uprave Republike Hrvatske (skripte), Osijek, 2007.

[6] HRN ISO 15489-1:2016 Information and documentation -- Records management -- Part 1: Concepts and principles (ISO 15489-1:2016), Hrvatski zavod za norme, Zagreb, 2016

[7] ISO 15489-1:2016(en), ISO, Available: https://www.iso.org/obp/ui/#iso:std:iso:15489:-1:ed-2:v1:en. [Accessed 7 April 2023]

[8] Ministry of Justice and Public Administration (Ministarstvo pravosuđa i uprave), Frequently Asked Questions on the new Regulation on Office Management (Često postavljana pitanja u vezi primjene Uredbe o uredskom poslovanju,) https://rdd.gov.hr/UserDocsImages//dokumenti//%C4%8Cesto%20postavljana%20pitanja%20-%20Uredba%20o%20uredskom%20poslovanju.docx [Accessed 7 April 2023]

[9] Ministry of Justice and Public Administration, National Public Administration Development Plan 2022-2027 (Nacionalni plan razvoja javne uprave za razdoblje od 2022.) do 2027. godine. March 2022. https://mpu.gov.hr/UserDocsImages/dokumenti/Strategije,%20planovi,%20izvje%C5%A1%C4%87a/Nacionalni%20plan%20razvoja%20javne%20uprave%20za%20razdoblje%20od%202022_2027.pdf [Accessed 7 April 2023]

[10] G. Vojković: Uredsko poslovanje, spisovodstvo i upravljanje pismohranom, Narodne novine, Zagreb, 2018.

[11] D. Bolanča, "Koncesijska odobrenja na pomorskom dobru", Pomorski zbornik, vol.38, br. 1, str. 159-170, 2000. [Online]. Dostupno na: https://hrcak.srce.hr/53009. [Citirano: 22.03.2023.]

[12] HRN EN ISO/IEC 27001:2017, Information technology -- Security techniques -- Information security management systems -- Requirements (ISO/IEC 27001:2013 including Cor 1:2014 and Cor 2:2015; EN ISO/IEC 27001:2017)

[13] Uredba (EU) br. 910/2014 Europskog parlamenta i Vijeća od 23. srpnja 2014. o elektroničkoj identifikaciji i uslugama povjerenja za elektroničke transakcije na unutarnjem tržištu i stavljanju izvan snage Direktive 1999/93/EZ (SL L 257/73 28. 8. 2014. – u daljnjem tekstu: Uredba (EU) br. 910/2014)

[14] General Administrative Process Act of the Republic of Croatia - Zakon o općem upravnom postupku, Narodne novine, 47/09, 110/21.

[15] Introduction of electronic signature and electronic seal, https://vlada.gov.hr/vijesti/uvodjenje-e-m-potpisa-i-e-m-pecata-digitalizirat-ce-poslovanje-drzavne-i-javne-uprave/33866 [Accessed 7 April 2023]

[16] Public e- government service of the Republic of Croatia - E-Građani, Informacije i usluge, gov.hr.

[17] Kohl, H. (2020). Standards for Management Systems - Standards for Compliance, Anti-bribery and Corporate Social Responsibility, p. 358, Springer Nature Switzerland 2020. doi:10.1007/978-3-030-35832-7

[18] *Ibid.*

[19] T.P. Kartika, N.R. Pratama, M. Dachyar: The Implementation of Anti-Bribery Management System Designed to Mitigate the Bribery Risk in State-Owned Enterprises: A Case Study of the Indonesia Agribusiness Industry, Proceedings of the First Australian International Conference on Industrial Engineering and Operations, Management, Sydney, Australia, December 20-21, 2022

[20] D.S. Handinita, L. Julaeha, D.L. Soewito: Strengthening Anti-Bribery Governance through Integration of GCG with SNI ISO 37001: 2016 ABMS, Asia Pacific Fraud Journal, vol 6.1. June 2021

[21] Transparency International, Corruption Perception Index 2022, available at: https://www.transparency.org/en/cpi/2022, [Accessed 7 April 2023]

[22] T. Mostovei: Perspectives for Implementation of the Anti-bribery Management Systems in the Republic of Moldova, Eurolimes, pp. 393, 2018

[23] H. Gehring, JP Mean: Implementing ISO 37001 to Manage Your Bribery Risks, Global Trade and Customs Journal Vol.3/5 p. 191 (2018).

[24] ISO 37001 – A Complete Guide: Practical Tools for Self-Assessment, The Art of Service, Edition 2020 and 2021.

[25] *Ibid.*

[26] *Ibid.*

[27] L. Veselovska, J. Zavadsky, Z. Zavadska: Mitigating bribery risks to strengthen the corporate social responsibility in accordance with the ISO 37001, Corporate Social Responsibility and Environmental Management, vol 27/4 July 2020

[28] European Commission, Joint Research Centre, Digital Economy Unit (JRC/B6), Blockchain for digital government, available at: https://joinup.ec.europa.eu/sites/default/files/document/2019-04/JRC115049%20blockchain%20for%20digital%20government.pdf [Accessed 7 April 2023]

* Narodne novine is the offical gazzete of Republic of Croatia.

Privacy Paradox and Generation Z

Marta Alić*, Lucija Sopić **
* Zagreb University of Applied Sciences, Zagreb, Croatia
** LeverUP Consulting Adria, Vukovarska 269f, 10000 Zagreb, Croatia
malic@tvz.hr
lucijasopic@hotmail.com

Abstract - **As technology advances daily, so are the challenges in preserving one's privacy. Being the generation that has been born in such a highly technological environment, members of Generation Z, born between mid-to-late 1990s and the early-to-mid 2000s, have been engaging in privacy related transactions more than any generation before. The issue of privacy is becoming more pronounced, along the possibilities of individuals controlling their data, that can lead to the discrepancy between attitudes about privacy preservation and actual behavior, that has become known as the "privacy paradox". By looking at this paradox through generational attitudes toward privacy, organizational practices and related legal frameworks, as well as the contemporary context of the sharing economy, the research aims to give insight whether members of the affected generation can be classified as fundamentalists, pragmatists, or unconcerned about privacy as a classification used by Alan Westin.**

Keywords – privacy paradox; Generation Z; Alan Westin; privacy index; digital economy

I. INTRODUCTION

In today's world, where information abounds, privacy is becoming a luxury. Activities that until recently were private are now becoming a source for analyzing individuals' interests, characteristics, beliefs, worldviews, and intentions for profit. By using the numerous Internet services, individuals (un)knowingly share a great deal of data with various actors: with each other, with organizations, and with government agencies.

Although the systems of the modern, digital economy is based on the exchange of data for the benefit of all stakeholders and society as a whole, the possibilities of data misuse, such as discrimination [1][2] and manipulation, are alarming.

The issue of privacy is becoming increasingly apparent, as is the ability of individuals to control their data. Advances in information technology have made it possible to collect and use personal data invisibly. As a result, individuals rarely know exactly what information others have about them and how that information is used and with what consequences. And if these individuals do not have such information, they are likely to be uncertain about how much information to share.

As Generation Z is first generation born into technology, its members are aware of the benefits, but also the risks it brings. Level of concern about privacy can indicate if their attitudes can classify them as fundamentalists, pragmatists and unconcerned.

This paper presents the quantitative research of the Generation Z attitudes regarding the existence of privacy paradox. The following sections provide an overview about the privacy paradox and main characteristics of Generation Z. The Alana Westin's privacy indexes methodology is also emphasized.

II. PRIVACY PARADOX

Calo [1] conceptualizes harm, or the negative consequences of privacy violations, through the categories of subjectivity and objectivity. While the subjective category of privacy damage is the perception of unwanted observation, which can cause undesirable mental states "within the individual", such as anxiety and discomfort, the objective privacy damage can occur when personal data is used to justify an adverse effect on a person, i.e. abuse data without the knowledge or consent of the individual. In both categories there is a certain insecurity of the individual regarding privacy, which can affect the behavior of individuals, as well as concern for its preservation.

Since the late 1970s, Alan Westin [2] has conducted over 30 privacy studies. For each he created a measurement instrument with one or more privacy indices to summarize his results and show privacy trends. By using broad, not context-specific questions, he segmented individuals into three groups: privacy fundamentalists, pragmatists, and privacy unconcerned. When asked directly, many individuals appear to be fundamentalists; they claim to care a lot about privacy and express particular concern about losing control over their personal data or having unauthorized access to it [3][4]. The unconcerned individuals do not care if their privacy is being violated because they look primarily at the benefits they can achieve by it, while, finally, the pragmatists, who are aware of the risks, calculate the risks and service benefits [5].

Anyway, the doubts about respondents' attitudes in predicting actual privacy behaviors have prompted numerous studies [4] and the discrepancy between individuals' attitudes and their behaviors has become known as the "privacy paradox".

The reasons for such inconsistency may be complex. A prudent concern for privacy may be a longer-term commitment that conflicts with the immediate gratification and convenience of Internet use as privacy policies are long and complicated [6]. Also, many individuals may simply be misinformed and believe that

their privacy is much better protected than it actually is [7].

There is now ample evidence of irrational cognitive biases in decision making and they can be predictable. In the privacy context, research has shown that willingness to disclose private information increases as perceived control over the disclosure of that information increases, even though it makes people more vulnerable to others accessing the information, whereas a decrease in perceived control has the opposite effect, even though more protection exists [8]. This is another paradox related to control and risk-taking behavior: the greater the perceived control, the more likely one is to take a risk. And further research [9] shows that perceived relative risk, i.e. the risk of one alternative compared to another, can have a greater impact on individual behavior in actual circumstances than differences in absolute risk. Thus, if a privacy policy is changed to lower the level of protection, it may have a greater impact on individual behavior than a policy that does not provide privacy protection in the first place.

Another illustration of such irrationality is the status quo bias and default settings, as consumers are much less likely to change default privacy settings, even when the settings do not reflect their actual preferences. That is specially emphasized when consumers encounter new technologies, such as the Internet of things smart devices [10].

A. Privacy as an economic calculation

Thus, the decision about privacy is the result of a rational "calculation" of costs and benefits [11], but also social norms, feelings, and heuristics. And each of these factors can influence behavior differently.

In the age of the digital economy, individuals are constantly engaged in private transactions, even if the privacy tradeoffs are intangible or the exchange of personal information is not visible or a major component of the transaction. For example, a search query on an Internet search engine has the same value as selling personal information (preferences, interests) in exchange for a service (search results). Although individuals sometimes give up personal information for various benefits or discounts, in other cases they voluntarily incur significant costs to protect their privacy. On the other hand, the desire for interaction, socialization, discovery, and recognition (or fame) is a powerful human motive, no less fundamental than the need for privacy.

Today's electronic media offer unprecedented opportunities to satisfy these motives. Through social media, disclosures can build social capital, increase self-esteem [12] and fulfill ego needs [13]. The report of the European Union Horizon 2020 research project [14] on privacy protection on the so-called sharing economy platforms shows that, despite relatively high privacy concerns, users on both sides (providers and users) believe that the benefits they derive from participating in the platforms outweigh the privacy risks. In addition, the majority of service providers express some concern about misuse of their data and loss of control over their online presence due to negative reviews or comments from other users. The whole concept of so-called impression management, the strategic sharing of personal data in order to create a more advantageous online presence, causes anxiety among users. Moreover, privacy concerns are contextual and can take forms ranging from extreme concern to apathy. That is, everyone is a stakeholder in one of Westin's defined groups, depending on time and place [7].

III. GENERATION Z

The date and age range of Generation Z are defined differently by different researchers, but it consists of people born between the mid-1990s and mid-2000s who are the first generation to be unaware of the world without electronic devices.

Digital is thus the most important characteristic that is almost always associated with Generation Z [15].

Members of this generation are very open-minded, accepting of diversity and promoting it [16]. And with the availability of information, proactivity in learning has developed, but at the expense of less developed social skills and a sense of isolation [17]. Some of the characteristics that distinguish them from other generations are realism and pragmatism-they want business stability and look to the future, and they want to know what is happening around them and be in control [18].

Social networks play a big role in the lives of Generation Z. Various trends are constantly evolving on social networks that are spreading very rapidly globally. This generation uses social networks to express themselves and their opinions, exchange information, communicate, comment and rate various products and brands, share content [19]. This is also reflected in the generation's consumption habits, as its members use the information available on the Internet to make the best possible purchase - by researching different products and then making a choice [17]. In addition, they place a high value on convenience in their purchasing decisions - saving time, ease of purchase, delivery methods, ease of use, and timely information about products [20]. As a result, members of this generation are aware of threats online - they recognize identity theft, virtual abuse and scams where the attacker impersonates someone else [21].

IV. RESEARCH METHODOLOGY

A. Research design

This analysis draws on Westin's research measuring attitudes and concerns about privacy and provides data on how those attitudes and concerns change over time. It is designed to examine the original issues, but in the context of today's age, audience and topics.

In regards to scaling the results, research methodology focuses primarily on Westin's earliest privacy index, the General Privacy Concern Index, which was developed as part of the 1990 study of 1,255 adult respondents from the U.S. public conducted for Equifax Inc. [22]. Designed as a Likert scale, the study questions target concerns about

current threats to privacy, claims that companies are asking for too much personal information, claims that the government is monitoring citizens, and claims that consumers have lost control over the sharing of their information. Responses to these questions were used to assign each respondent to a group concerned about privacy, resulting in a kind of nominal scale: a) high concern - three or four responses that leaned toward concerned, b) medium concern - two responses concerned about privacy, and c) low concern - one or no concerned response. This led to a division into fundamentalists, pragmatists, and unconcerned. Index is used as a basis for this study to establish continuity in monitoring privacy paradox throughout time. Questions in these research were adapted to reflect the contemporary context, such as the digital economy, as Westin also examined the public's general level of concern about privacy and also examined attitudes toward specific privacy-related issues, such as trust in organizations that handle personal information, acceptance of a national identification system, and even the use of medical records for research [23].

B. *Research sample*

The research was conducted in first half of September 2021. with a survey questionnaire that was disseminated through social networks Linkedin and Facebook. On the Linkedin network, the link to the anonymous survey was published on the personal profile of second author, a member of Z generation, while on Facebook the link to survey was published in various groups where there is a younger population, and these were mostly different groups intended for students. Ethical standards of anonymity and confidentiality were respected. The total number of respondents is 474 and survey competition time was approximately 5 minutes.

The distribution of respondents by gender is unequal, i.e. there are almost 3 times more female respondents than men. From total number of respondents, 75.6% are female, while 23.8% of them are male. Also, the significant age prevalence is within respondents born in 2002., that corresponds with education level, as 62,9% of respondents have declared high school as a completed level of education. Quarter of respondents (25,5%) have completed undergraduate studies and rest of the respondents, born in mid-1990s have graduated from university.

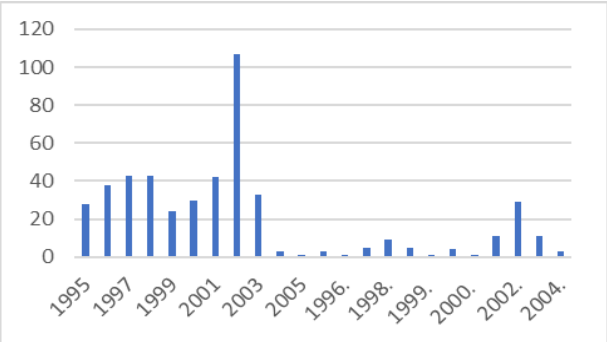


Figure 1 Respondents' distribution in age

The majority of respondents is born in between years 1995 and 2003, with most of them born in 2002, shown in Figure 1.

V. ANALYSIS AND RESULTS

A. *Privacy paradox trends*

According to the results, shown in Table 1, almost all respondents claim that protecting the privacy of personal data is important to them (85% believe that it is completely important, and 14% that it is mostly important to them). Also, a large proportion of respondents are concerned about threats to their privacy (30% are completely concerned and 42% are mostly concerned).

In relation to organizational practices majority of respondents agree that today's companies ask for too much personal information (42% strongly agree and 39% mostly agree) and that consumers have lost all control over the circulation and use of personal information by organizations (31% strongly agree and 35% mostly agree), as they show more disagreement (28% mostly disagree) with the statement that organizations handle data collected from users in an appropriate and discreet manner than 25% of respondents that mostly agree to this statement.

TABLE 1 ATTITUDES TOWARDS PRIVACY

	Strongly agree	Mostly agree	Mostly disagree	Strongly disagree	Not sure
Protection of the personal data is important	85%	14%	0	1%	0
I am concerned about threats to the privacy of my data	30%	42%	20%	4%	4%

Respondents find tracking practice resulting with targeted ads too invasive as 48% respondents strongly agree with this statement, and 37% mostly agree, as one third (33%) of respondents finds that privacy policies do not necessarily fully reflect the real data management practices within organizations. One fourth (25%) of respondent is not sure of their attitude towards these transparency mechanisms.

Following the results on organizational practices, prevalent number of respondents (35% mostly disagree and 13% strongly disagree) that their privacy rights are adequately protected even by the law as they disagree (39% mostly disagree and 17% strongly disagree) that existing laws and organizational practices provide a sufficient level of consumer privacy protection.

Results in Table 2 show respondents' opinions in relation to organizational practices.

As research was conducted in echoes of Covid-19 pandemy respondents were asked to assess privacy and security measures regarding respected medical records. The prevalent majority of them (27 strongy agree and 25% mostly agree) found that measures in such circumstances were decreased.

TABLE 2 OPINIONS IN RELATION TO ORGANIZATIONAL PRIVACY PRACTICES

	Strongly agree	Mostly agree	Mostly disagree	Strongly disagree	Not sure
Modern organizations ask for too much personal information	42%	39%	12%	2%	6%
Consumers have lost all control over the circulation and use of personal information by organizations	31%	35%	20%	3%	11%
Organizations handle data collected from users in an appropriate and discreet manner	15%	25%	28%	12%	19%
I believe that privacy policy documents do not necessarily fully reflect the real data management practices within organizations	17%	33%	19%	5%	25%
Tracking my movements on the Internet and displaying targeted ads based on the places I visit is too invasive	48%	37%	11%	2%	2%

Table 3 shows respondents' opinions in relation to privacy regulations.

However, when the question related to algorithms was put in the context of transparency, the respondents expressed an equal difference in their attitudes - 40% of the respondents feel discomfort due to the non-transparent segmentation of users, and 40% of the respondents do not feel this discomfort, as shown in Table 4.

TABLE 4 OPINIONS IN RELATION TO ORGANIZATIONAL PRIVACY PRACTICES

	Strongly agree	Mostly agree	Mostly disagree	Strongly disagree	Not sure
I feel uncomfortable knowing that the platforms collect my data and process it using algorithms whose principle of operation is not explained	38%	32%	2%	8%	2%
I believe that sharing economy platforms do more harm than good	9%	13%	32%	27%	20%
I feel uneasy because of non-transparent algorithms of user segmentation of the sharing economy platforms	15%	25%	28%	12%	19%

Privacy paradox in sharing economy context

As privacy paradox can be more conspicuous in the environment of sharing economy services, where the functioning of the algorithms of such platforms is mostly untransparent to end users, consequent statements were focused to examine respondents' attitudes towards it, as well as the "calculation" of costs and benefits in such cases.

The majority of respondents (38% fully agree, and 32% mostly agree) expressed that they feel uncomfortable knowing that platforms collect their data and process it using algorithms whose working principles are unknown to them. Also, most respondents (32% mostly disagree and 27% strongly disagree) believe that sharing economy platforms do not bring more harm than good.

TABLE 3 OPINIONS IN RELATION TO PRIVACY REGULATIONS

	Strongly agree	Mostly agree	Mostly disagree	Strongly disagree	Not sure
My privacy rights as a consumer are adequately protected by law and business practices	**7%**	**31%**	**35%**	**13%**	**14%**
I believe that regulatory bodies should control the work of organizations and the use of collected data	**56%**	**32%**	**6%**	**2%**	**4%**
Existing laws and organizational practices provide a sufficient level of consumer privacy protection	**6%**	**20%**	**39%**	**17%**	**18%**
Institutional bodies should be able to monitor citizens' private messages in justified cases of crime prevention	**12%**	**24%**	**17%**	**29%**	**18%**

VI. DISCUSSION

Overall, looking at the answers to all the questions, it can be concluded that sampled members of Generation Z find that privacy protection important and that they are concerned about the collection of personal data. Their concern can be classified as high in terms of organizational practices (extent of data collection, lack of control over data, targeted advertisement and trust in privacy policies), although they agree and disagree in equal proportion weather organizations handle data collected from users in an appropriate and discreet manner. Although they advocate privacy rights regulation, they distrust the existing system, showing concern in law and business practices. Also, they do not support threats to privacy by various authorities for the purpose of "greater good", as they also find circumstances within Covid-19 crisis unmanaged and intrusive of their privacy.

Finally, in context of digital economy members of subject generation show strong consensus about algorithmic collection and use of data, while they show no concern about the (un)transparency of such practices. They are also trustful of sharing economy platforms, finding them useful.

Discussing results in relation to Westin's groups, the prevalent results suggest high concern about privacy that can classify members of survey generation as fundamentalist about organizational and regulatory practices as Westin classified fundamentalists as generally distrustful of organizations asking for their personal information and worried about the accuracy of computerized information and additional uses made of it. They are also in favor of laws and regulatory actions in forcing their privacy rights.

But paradox can be stipulated in relation to modern economy and platforms as respondents show coherent pragmatism towards contemporary practices of organizations, finding them discrete, appropriate and useful as they calculate the benefits to them. By Westin this group estimate various consumer opportunities and services, protections of public safety or enforcement of personal morality against the degree of intrusiveness of personal information sought and the increase in government power involved, ie. risks [23].

Moreover, from another perspective, results show high rates of "not sure" answers that can relate to unconcerned group classification. These results can be indicatory and should be taken in consideration, as they are significant in relation to organizational practices, as well as regulatory ones, but also in terms of digital economy environment.

VII. CONCLUSION AND FUTURE WORK

Considering the brevity of the study, as one of its limitations, it is not possible to determine the dominant group into which the members of Generation Z, i.e. the sample under study, could be classified.

However, the results of the study can serve as a guide for further development of the privacy indexes in the modern environment through future research.

A contribution of research at the application level can be recognized as a useful insight in generation Z practices for experts in the field of privacy protection.

REFERENCES

[1] R. Calo, "The Boundaries of Privacy Harm," Indiana Law J., vol. 86, pp. 1132–1161, 2011.

[2] A. F. Westin, Privacy and Freedom. New York: Atheneum, 1967. doi: 10.2307/3479272.

[3] M. J. Culnan and P. K. Armstrong, "Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation," Organ. Sci., vol. 10, no. 1, pp. 104–115, 1999, doi: 10.1287/orsc.10.1.104.

[4] H. J. Smith, S. J. Milberg, S. J. Burke, and O. N. Hall, "Information Privacy: Measuring Individuals' Concerns About Organizational Practices," MIS Q., vol. 20, no. 2, pp. 167–196, 1996.

[5] A. F. Westin, "Social and Political Dimensions of Privacy," J. ofSocial Issues, vol. 59, no. 2, pp. 431–453, 2003, [Online]. Available: www. forstbuch. de

[6] A. M. Mcdonald and L. F. Cranor, "The Cost of Reading Privacy Policies," A J. Law Policy Inf. Soc., pp. 1–22, 2008.

[7] C. J. Hoofnagle and J. M. Urban, "Alan Westin's Privacy Homo Economicus," Inf. Priv. Law eJournal, 2014.

[8] L. Brandimarte, A. Acquisti, and G. Loewenstein, "Misplaced Confidences: Privacy and the Control Paradox," Soc. Psychol. Personal. Sci., vol. 4, no. 3, pp. 340–347, 2013, doi: 10.1177/1948550612455931.

[9] I. Adjerid, E. Peer, and A. Acquisti, "Beyond the Privacy Paradox: Objective versus Relative Risk in Privacy Decision Making," SSRN Electron. J., pp. 1–43, 2017, doi: 10.2139/ssrn.2765097.

[10] M. W. Bailey, Seduction by technology: Why consumers opt out of privacy by buying into the internet of things, vol. 94, no. 5. 2016.

[11] M. Laufer, Robert and Wolfe, "Privacy as a Social Issue and Behavioural Concept," J. Soc. Issues, vol. 33, no. 3, pp. 22–41, 1977.

[12] C. Steinfield, N. B. Ellison, and C. Lampe, "Social capital, self-esteem, and use of online social network sites: A longitudinal analysis," J. Appl. Dev. Psychol., vol. 29, no. 6, pp. 434–445, 2008, doi: 10.1016/j.appdev.2008.07.002.

[13] C. L. Toma and J. T. Hancock, "Self-Affirmation Underlies Facebook Use," Personal. Soc. Psychol. Bull., vol. 39, no. 3, pp. 321–331, 2013, doi: 10.1177/0146167212474694.

[14] G. Ranzini, M. Etter, and I. E. Vermeulen, "Privacy in the Sharing Economy: European Perspectives," SSRN Electron. J., 2017, doi: 10.2139/ssrn.3048152.

[15] N. Negroponte, Being Digital. Alfred A. Knopf, Inc., 1995.

[16] K. Parker, N. Graf, and R. Igielnik, "Generation Z Looks a Lot Like Millennials on Key Social and Political Issues," 2019. [Online]. Available: https://www.pewsocialtrends.org/2019/01/17/generation-z-looks-a-lot-like-millennials-on-key-social-and-political-issues/

[17] The Annie E. Casey Foundation, "What Are the Core Characteristics of Generation Z?," 2021. https://www.aecf.org/blog/what-are-the-core-characteristics-of-generation-z

[18] T. Francis and F. Hoefel, "'True Gen': Generation Z and its implications for companies," McKinsey & Company, 2018. https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/true-gen-generation-z-and-its-implications-for-companies

[19] O. Dolčić and M. Alić, "Generation Z Buyer's Journey," in Proceedings of Sciencefora international conference Dubrovnik, 2020, pp. 60–64.

[20] C.-V. Priporas, N. Stylos, and A. K. Fotiadis, "Generation Z consumers' expectations of interactions in smart retailing: A future agenda," Comput. Human Behav., vol. 77, pp. 374–381, 2017, doi: https://doi.org/10.1016/j.chb.2017.01.058.

[21] C. Seemiller and M. Grace, “Generation Z: Educating and Engaging the Next Generation of Students,” About Campus, vol. 22, no. 3, pp. 21–26, 2017, doi: 10.1002/abc.21293.

[22] Stephen T. Margulis, Jennifer A. Pope, and Aaron Lowen, “The Harris-Westin’s Index of General Concern About Privacy: An Attempted Conceptual Replication. ,” GPD Int. Surv. Res. Work., pp. 1–18, 2006, [Online]. Available: https://www.sscqueens.org/sites/sscqueens.org/files/Margulis_paper.pdf

[23] P. Kumaraguru and L. F. Cranor, “Privacy Indexes: A Survey of Westin’s Studies,” 2005.

Zračna luka kao osnova za uvođenje inovativnih tehnoloških rješenja u postupku kreiranja pametnog grada zračne luke

dr. sc. M. Milenković, Ph.D.*
* Sveučilište u Zagrebu, Fakultet prometnih znanosti/Katedra za pravo i ekonomiku prometa, Zagreb, Hrvatska
mmilenkovic@fpz.hr

Sažetak - Kako zračna luka može postati središnji dio regije? Ovaj rad pokušat će objasniti preduvjete za stvaranje pametnog grada zračne luke. Današnje zračne luke postale su ne samo mjesta polijetanja i slijetanja zrakoplova, već i generatori gospodarskog razvoja, dajući otvoren pristup državi ili gradu, s namjerom povezivanja gradova i zračnih luka u jedinstvenu cjelinu. U svijetu sklonom brzim promjenama, zračne luke su se usredotočile na širenje i povezivanje svojih zajednica, ne samo u smislu turizma, već i u smislu implementacije IoT tehnologije kroz integrirani transport, prilagodbu novim uvjetima, usmjerenost na kupca i korisnička iskustva. Ovaj rad nastoji odrediti mogućnosti implementacije inovativnih tehnoloških rješenja kroz postupke javne nabave koje je predložila Europska komisija, a koja se mogu koristiti za oporavak zračnih luka u EU nakon pandemije covid-19.

Ključne riječi: pametni grad, pametni grad zračne luke, IoT, inovativna tehnološka rješenja.

I. UVOD

Izazovno je definirati definiciju pametnog grada zračne luke te je stoga potrebno prvotno razjasniti točno ono što se podrazumijeva pod pojmom „pametnog grada", „pametne zračne luke" te pod izrazom „pametnog grada zračne luke". Definicija „pametne" zračne luke povezana je s definicijom „pametnog grada". Pametni gradovi primjenjuju tehnološke mogućnosti na inovativan način kako bi se stvorilo pogodnije za život i održivo urbano okruženje u kojem su gradovi dužni zadovoljiti potrebe svojih građana kroz različite tehnološke sustave i različite metode prijevoza, a u središte dolazi zračna luka. Značajan dio pametnog grada čini prometna infrastruktura, zbog toga i zračna luka mora biti „pametna" kako bi imala određene koristi od kojih ne samo grad oko zračne luke, već i čitava regija može napredovati.

Ovaj rad će pokušati razjasniti točno ono što se podrazumijeva pod pojmom pametnog grada zračne luke, uključujući njegove prednosti i nedostatke te kako ih se može jednostavnije prevladati koristeći dostupna inovativna tehnološka rješenja kroz postupke javne nabave.

Digitalizirane zračne luke mogu pomoći pametnom gradu zračne luke kroz ekonomske prihode te povećati gospodarski razvoj regije. Kroz digitalizaciju i uz podršku javno-privatnih partnerstava takav bi se model mogao primijeniti u svim većim zračnim lukama u EU.

Naime, posljednje desetljeće započelo je najnevjerojatnijim promjenama u zračnoj industriji. Tome koristan primjer je napuštanje letenja preko čvorišta i uporabe direktnih letova što je značajno izmijenilo zračni promet. Tome su također pridonijele promjene u zrakoplovnoj industriji, novi tipovi zrakoplova i očekivanja putnika. Ono što je također značajna stavka je da zrakoplovi danas imaju znatno manju potrošnju goriva i lakše održavanje, a s novim gorivima doći će i do velikih promjena u infrastrukturi samih zračnih luka. Štoviše, dolaskom niskotarifnih zračnih prijevoznika te nestankom navedenih vrsta zrakoplova u zračnom prometu u posljednjem desetljeću došlo je do značajnih promjena.

Novi zrakoplovi s učinkovitom potrošnjom goriva smanjili su operativne troškove, čineći letenje rutama od točke do točke (poin-to-point) prihvatljivijim. Drugo, niskotarifni zračni prijevoznici su preuzeli nebo. Odlazak putničkog Boeing -a 747 i još nekih drugih zrakoplova u povijest doveo je do promjene u zračnom prometu te do modernizacije zračnog prijevoza. Na primjer, kako bi putnici stigli iz Zagreba u London, morali su letjeti u Frankfurt ili iz nekog drugog europskog hub odredišta (čvorišta), a pojavom niskotarifnih zračnih prijevoznika tzv., „low-cost'' avioprijevoznika te su se linije promijenile. Putnici su počeli putovati direktnim letovima što je uvelike smanjilo potrošnju goriva, troškove spuštanja i ponovnog polijetanja aviona.

Pametne zračne luke/digitalni aerodromski gradovi današnjice koriste različite sustave te pri tome uključuju informatizaciju javne usluge kao što su pametno upravljanje rasvjetom, prometom i organizacijom prijevoza, upravljanje otpadom i vodama, administrativne politike, sigurnost dijeljenja podataka, energetsku održivost i moderne informacijske usluge, koje su pridonijele razvoju pametnog grada zračne luke.

S jedne strane, gradovima kakve do sada poznajemo upravlja se i nadzire svakim sustavom kao zasebnom jedinicom, što zauzvrat proizvodi više troškova za porezne obveznike uz mala ili nikakva poboljšanja u kvaliteti života, dok s druge strane pametne zračne luke, odnosno pametni gradovi zračne luke IoT tehnologijom upravljaju sustavom kao jednom kompleksnom cjelinom. Na taj način se svi troškovi smanjuju brzom prilagodbom i prihvaćanjem tehnološkog napretka i automatizacijom

raznih sustava koji iskorištavaju potencijale IoT tehnologije do maksimuma.

Dakle, 'pametni gradovi' koriste informacijske i komunikacijske tehnologije (ICT), kao što su pametne mreže i uređaje Interneta Stvari (dalje: IoT) za stvaranje veza i interakcija između sustava, na taj način smanjujući troškove i poboljšavajući kvalitetu života građana tijekom procesa.[1] Stoga će ovaj rad posebnu pozornost posvetiti razvoju pametnih gradova zračne luke koji mogu značajno utjecati i unaprijediti gradove i regije. Zračne luke su važan dio regije i pametnog grada. Istodobno, imaju stratešku važnost za područje koje opslužuju, a istovremeno su povezani s nizom zemalja u kojima su se sve više integrirali u cjelokupni prometni sustav. [2] Dakle, mogu osigurati veću dobrobit države pridonoseći mogućnostima zapošljavanja i gospodarskom razvoju.

Ovaj će rad stoga pokušati detaljno potvrditi da su prednosti digitaliziranih zračnih luka daleko veće od potencijalno štetnih učinaka. Također, kako se klimatski uvjeti mijenjaju, bit će potrebno prilagođavati se i prihvatiti promjene koje će uslijediti, no neki od prijedloga niti neće biti izvedivi i primjenjivi.

Covid-19 kriza u kojoj se živjelo proteklo gotovo tri (3) godine je dodatno potvrdila da su zračne luke spremne za veliku transformaciju i da je vrijeme da se njihove najvažnije usluge i procesi promptno digitaliziraju. Razvidno je da zračni promet ima najviše promjena od svih prometnih grana.

II. RAZVIJANJE PAMETNOG GRADA ZRAČNE LUKE

Za stvaranje pametnog digitalnog grada potrebno je pripremiti uvjete koji zadovoljavaju povezivu prometnu infrastrukturu unutar grada te poboljšati pravni okvir za implementaciju novih infrastrukturnih rješenja. Širenje komercijalnog razvoja prema zračnim lukama danas otvara vrata proizvođačima urbanog rasta jer moderne digitalne zračne luke postaju važne destinacije, a ne služe samo za zaposlenje nego i za kupovinu, trgovinu, poslovne sastanke, zabavu i slobodno vrijeme.

Stoga zračne luke preuzimaju mnoga obilježja središnjih gradskih poslovnih četvrti i sve više djeluju kao regionalne i nacionalne točke multimodalnog komercijalnog razvoja. To je dovelo do evolucije u funkciji i obliku te transformacije urbanih zračnih luka u aerodromske gradove. Također, današnje zračne luke se više ne mogu oslanjati samo na broj putnika koji prolaze kroz njih, potrebno ih je modernizirati i tehnološki i konceptualno, uz usklađivanje sa zakonskim okvirom te vremenom i brzinom tehnološkog napretka. [3]

Promatrajući u globalnom kontekstu, izvan EU, zračne luke su prepoznate kao vrlo slične pametnim gradovima jer su usmjerene na poboljšanje poslovanja, održivosti, iskustva i poslovnu vitalnost. [4] Na primjer, šangajska zračna luka „Jewel Changi" često se prikazuje kao najpametnija zračna luka na svijetu. Riječ je o spoju održive arhitekture i tehnologije.

Oni koriste IoT senzore, analitiku podataka i AI kako bi osigurali pozitivno korisničko iskustvo za putnike koji putuju kroz njihovu zračnu luku.

Može se spomenuti i primjer dobre prakse kroz europsku zračnu luku München, koja je u suradnji s „Massachusetts Institute of Technology" (MIT) osmislila koncept pametnog grada zračne luke.

Cilj je izgraditi radno okruženje za organizacije koje će dizajnirati i razvijati proizvode i usluge pametnog grada te ih potom testirati u zračnoj luci München.

LabCampus GmbH surađuje s MIT-om u podršci tehnološke infrastrukture za najsuvremenije koncepte energije, mobilnosti i urbanog planiranja.

Iako još nije pametna zračna luka u potpunosti, zračna luka München pokazuje kako zračne luke mogu koristiti svoje objekte na tlu i zatim testirati nove koncepte unutar svojih terminala na odgovarajući način.[4]

Slika 1 ilustrira koja su rješenja moguća korištenjem videoanalitike, budući da mogu dati trenutne rezultate koji su mjerljivi u ključnim područjima zračne luke.

Implementacijom ovih rješenja i razmatranjem načina na koji zračne luke mogu koristiti video-inteligenciju, umjetnu inteligenciju i internet stvari velike količine kapitala mogu se prenijeti u zračne luke kako bi se postigao cilj da postanu pametne zračne luke i u nekom trenutku u bliskoj budućnosti dobiju priliku zamijeniti, odnosno upotpuniti pametne gradove.

Ono što je jedna od novina pri korištenju inovativnih IoT tehnologija je način na koji se podaci prikupljaju putem video inteligencije i kako se mogu interpretirati i koristiti za utjecaj na iskustvo putnika unutar zračne luke.

Kao primjer, moguće je uzeti tako prikupljene informacije i učiniti ih inteligentnijima i učinkovitijima pomoću računalnog vida i strojnog učenja, poznatijeg kao videoanalitika.



Slika 1: Kako osmisliti i dizajnirati pametnu zračnu luku? [4]

Također, infracrvena tehnologija može se koristiti noću za povećanje protoka video podataka. Nadalje, nedavni napredak koji pruža mnoge nove mogućnosti je i uvođenje 3D lidara unutar same zračne luke, a lidari se mogu koristiti za praćenje različitih vrsta korištenja prostora.

Kao što je spomenuto u izvješću „Jesu li zračne luke sljedeći pametni gradovi?'': „Današnja inovativna rješenja također imaju mogućnost automatske anonimizacije

podataka prikupljenih zaštitnim mjerama poput maskiranja privatnosti, transparentnosti i 3D lidara. Nakon toga slijedi prikupljanje svih informacija i primjena umjetne inteligencije za analizu ulaza i stvaranje trodimenzionalnog prikaza zračne luke u stvarnom vremenu koji ne bilježi nikakve osobne podatke." Što je najvažnije, takav pristup omogućuje točno praćenje kretanja unutar zračne luke bez poznavanja identiteta pojedinaca. [4]

Zračne luke su slične gradovima koji trebaju ogroman izvor energije za rad i pružanje usluga. Stoga je s jedne strane od velike važnosti sigurna, ekonomična i, što je najvažnije, pouzdana opskrba električnom energijom u zračnim lukama. Također se razmatraju i potencijalni utjecaji klimatskih promjena jer se većina izvora energije još uvijek uvelike oslanja na fosilna goriva. Većina zračnih luka kupuje električnu energiju od dobavljača električne energije, zbog čega je potrošnja energije veliki operativni trošak za zračnu luku, a umjesto da osigura vlastitu energetsku učinkovitost kroz postavljanje vlastitih izvora napajanja električnom energijom. Osim toga, instalacija tehnologija obnovljive energije na licu mjesta u zračnoj luci može smanjiti ugljični otisak s minimalnim utjecajem na rad zračne luke, kao što su npr. krovni solarni paneli. [5] Sredstva za instalaciju takvih panela bi se mogla ostvariti te provoditi kroz postupke javne nabave inovativnih i digitalnih rješenja, što je između ostalih bio i jedan od ciljeva Europske komisije za 2021. godinu, a kako je navedeno u Uputama obavijesti za javnu nabavu inovativnih rješenja [6].

III. INFRASTRUKTURNI UTJECAJ

A. Razvoj grada zračne luke

Širenjem zračnih luka na hotele, nekretnine, objekte za iznajmljivanje, aktivnosti u slobodno vrijeme, razvoj uredskih četvrti i ostalih dodatnih sadržaja dovodi do koncentracije urbanih klastera koji se razvijaju u gradove zračnih luka. Ovakav tip grada usmjeren na urbano poduzetništvo predstavlja gospodarsku snagu koja izravno i neizravno utječe na razvoj regije. [7]

Koncept digitalnog pametnog grada zračne luke se pak logično pojavio kao proširenje prometnog čvorišta zračne luke i karakterizira ga vrlo visoka razina složenosti.

Za takvo funkcioniranje potrebno je infrastrukturno posložiti te razraditi prometne tokove unutar grada, pametnu mrežu povezanu sa svim drugim čvorištima unutar zračne luke te također i umreženi, odnosno povezani IoT sustav.

Zanimljivo je da su u 2017. godini svjetski zračni prijevoznici prevezli oko 4,1 milijardu putnika dok su prevezli 56 milijuna tona tereta na 37 milijuna komercijalnih letova. Zrakoplovi svaki dan prevezu preko 10 milijuna putnika i robu u vrijednosti od oko 18 milijardi dolara. Ovo samo ukazuje na značajan ekonomski utjecaj zračnog prometa na svjetsko gospodarstvo, što dokazuje da zračni promet čini 3,5 posto svjetskog bruto domaćeg proizvoda (BDP), odnosno 2,7 trilijuna dolara i stvorilo je 65 milijuna radnih mjesta diljem svijeta. Zračni sektor se umnožava i nastavio je rasti i nakon krize izazvane covidom-19. Najnovije procjene pokazuju da će potražnja za zračnim prijevozom rasti za 4,3% godišnje u sljedećih 20 godina. [8] Time se razvija i značajna potreba za sve brže rastućom tehnologijom, no sredstva za izgradnju tako povezanih pametnih gradova zračne luke jesu teško dobavljiva.

Usporedbe radi, globalna statistika Organizacije međunarodnog civilnog zrakoplovstva (ICAO-a) za 2019. godinu, a ujedno i najproduktivniju godinu prije krize izazvane virusom covid-19 pokazuje da je ukupan broj putnika prevezenih u redovnim linijama u 2019. godini porastao na 4,5 milijardi, što je 3,6 posto više nego prethodne godine, dok je broj letova dosegao 38,3 milijuna u 2019. godini, što je povećanje od 1,7 posto. [9]

Stoga EU i druge zemlje izvan šireg europskog područja pokušavaju pronaći rješenje za smanjenje ugljičnog otiska, no i uz daljnje godišnje povećanje zračnog prometa.

B. Održiva budućnost

Poučena sve težom gospodarskom situacijom Europska Komisija je donijela novi financijski instrument za potporu ekoloških i klimatskih akcijskih projekata iz kojih gradovi mogu povući sredstva. Oni uključuju: „Financijski instrument za okoliš i klimatske akcije" (LIFE), Program „Obzor 2020" i „Inteligentna energija za Europu" (IEE). Te bi navedeni programi trebali omogućiti jačanje provedbe digitalizacije kroz održivost kao dio pametnog grada, odnosno pametnog grada zračne luke. Cilj je holistički pristup u kojem se svi procesi odvijaju primjenom IoT tehnologije i inteligentnih transportnih sustava. [9]

Europska komisija navodi da se u 'pametnim gradovima', digitalne tehnologije pretvaraju u bolje javne usluge za građane, bolje korištenje resursa i manji utjecaj na okoliš. [13]

Stoga, neke promjene mogu uštedjeti ogromne količine energije, a iste se mogu postići primjenom aktivnih mjera kao što su implementacija sustava automatizacije. Pametan grad pa tako i pametan grad zračne luke iskorištava digitalne tehnologije za bolje korištenje resursa i manje emisija. Takva unapređenja se mogu postići inovativnom javnom nabavom razvojem inovativnije mreže gradskog prijevoza, poboljšanom opskrbom vode i objektima za odlaganje otpada te učinkovitijom rasvjetom i grijanjem zgrada. [13]

Sve navedeno može omogućiti i stvaranje pametnih i digitaliziranih zračnih luka.

A novim mehanizmom za oporavak i otpornost pokušava se potaknuti javna ulaganja, posebno nakon krize izazvane covid-19 virusom, koja je usporila zračni promet za 90%. [14] [15] Ovim mehanizmom Komisija je 27. svibnja 2020. predložila glavni pravni okvir instrumenta „Sljedeće generacije" EU-a.

Takav privremeni instrument oporavka omogućuje Komisiji prikupljanje sredstava za rješavanje neposredne gospodarske i društvene štete uzrokovane pandemijom koronavirusa.

Cilj Mehanizma je podržati javne naručitelje kako bi mogli na bolji način doprinijeti gospodarskom oporavku, također digitalnoj tranziciji, kao i otpornosti Europske unije na nastalu gospodarsku situaciju. [17]

Predviđa se da će veliki dio ovih ulaganja biti kanaliziran upravo kroz postupke javne nabave.

Procjena EU-a također pokazuje da Europa iskorištava samo polovicu potencijala javne nabave inovativnih rješenja za poticanje gospodarskog oporavka.

Dakle, postoji značajan nedostatak ulaganja, posebno u digitalnu nabavu te u usluge istraživanja i razvoja. Međutim, oba mehanizma su jednako važna za jačanje strateške autonomije i konkurentnosti Europske Unije. [15]

IV. PROVEDBA JAVNE NABAVE INOVATIVNIH RJEŠENJA

Javna nabava inovativnih rješenja može postati obećavajući instrument za promicanje uporabe naprednih inovativnih tehnologija i informacijsko-komunikacijskih usluga pružajući javnim tijelima poticaje za suradnju s inovativnim tvrtkama kao prvim kupcima. Navedeno može uključivati zaštitu kritične infrastrukture, pametne gradove i sve moderne digitalne tehnologije koje zahtijevaju rješenja za smanjenje troškova za zračni, pomorski i željeznički promet kojim upravljaju ili koji nadziru javna tijela. [15]

EU države članice ne smiju izravno pomagati privredi i prema tome idu k modelu inovativnih rješenja, za razliku od npr. azijskih modela gdje nije neizvedivo proračunskim sredstvima pomoći razvoju državne zračne luke i/ili zračnog prijevoznika. Naime, azijskim zemljama inovativne javne nabave nisu neophodna jer iz proračuna npr. Kine, Tajvana ili Koreje se mogu povući određena sredstva za pomoć zračnoj luci ili brodogradilištu, dok proračunski novac država članica EU nije moguće tako koristiti jer se sredstva moraju davati transparentno Stoga javna nabava inovativnih rješenja dolazi kao pravovremeni korak Europske komisije, posebno u trenucima gdje su mnoga važna poduzeća u situaciji krize.

Procjenjuje se da ukupna vrijednost javne nabave u EU pokriva oko 14% bruto domaćeg proizvoda EU [20], što je relevantno za gradove zračne luke jer to iznosi više od 1,9 trilijuna eura, stoga je javna nabava važna za konkurentnost EU zbog svoje veličine, posebno u sektorima kao što su energetika, promet, obrana, IT i zdravstvene usluge.

Očekuje se da će promicanje javne nabave za inovativna rješenja utjecati na nacionalna gospodarstva država članica kroz različite kanale poput gospodarskog rasta koji proizlazi iz priljeva privatnog kapitala i smanjenja fiskalnog opterećenja države kroz bolju vrijednost za novac. Također, takva rješenja mogu pridonijeti gospodarskom oporavku EU-a nakon krize izazvane koronavirusom kroz bolja javna ulaganja. [9]

Kako bi dodatno pomogla zrakoplovnoj industriji i drugim sektorima prometa i poduzetništva u prevladavanju krize, Europska komisija izdala je smjernice u kontekstu Komunikacije, pod nazivom „Obnavljanje istraživanja i inovacija – prilika za Europu da oblikuje svoju budućnost". [10]

Objasnimo stoga što je javna nabava inovativnih rješenja i kako ona može transformirati zračne luke u pametne digitalne gradove.

Što se tiče, javne nabave inovativnih rješenja ona se odnosi se na svaku nabavu koja uključuje barem jedan od sljedećih aspekata:

- predmet nabave je proces inoviranja - usluge istraživanja i razvoja - s (djelomičnim) rezultatima,
- predmet nabave su rezultati inovacija.[11]

Riječ je o digitalnim rješenjima koja do sada nisu bila korištena u praksi, među ostalim namijenjena su i za bržu prilagodbu tržištu usluga zračnog prometa.

Sam pojam „inovacija“, prema članku 2. (22.) Direktive 2014/24/EU Europskog parlamenta i Vijeća od 26. veljače 2014. o javnoj nabavi [17] naglašava visoku vrijednost inovacije kao jednog od glavnih pokretača budućeg rasta europskih pametnih gradova te glasi:

„inovacija" znači provedba novog ili značajno poboljšanog proizvoda, usluge ili postupka, uključujući, ali ne ograničavajući se na postupke proizvodnje, građenja ili izgradnje, nova metoda stavljanja na tržište ili nova metoda organizacije u poslovnoj praksi, organizacije radnog mjesta ili vanjskih odnosa između ostalog s ciljem pomaganja rješavanja društvenih izazova ili kao potpora strategiji Europa 2020. za pametan, održiv i uključiv rast".

Alat koji je dostupan svim privatnim i javnim poduzećima u EU je javna nabava inovativnih rješenja koja nije dostupna u Aziji, a upravo tamo se nalazi većina svih digitalnih i pametnih gradova te pametnih gradova zračne luke. Primjerice, očekuje se da će godišnji globalni prihod Smart Cityja dosegnuti 88,7 milijardi dolara do 2025. godine, a većina će zauzimati najbrže rastuće azijsko područje. Prema Milenković et. al. u azijskim zemljama financiranje 'pametnih gradova' provodi se putem državnih poticaja, lokalnih subvencija, javno-privatnog partnerstva i projekata privatnog poduzetništva. [11]

Stoga kupnja inovativnih proizvoda i usluga može pridonijeti kvaliteti javnih usluga i ostvariti ekonomske, ekološke i društvene koristi stvaranjem novih ideja kroz inovativne usluge.

Samo na taj način može pomoći u promicanju održivog gospodarskog rasta.

Stoga, direktive EU:

1. Direktiva 2014/24/EU o javnoj nabavi [17],
2. Direktiva 2014/23/EU o dodjeli ugovora o koncesiji i
3. Direktiva 2014/25/EU o nabavi subjekata koji djeluju u sektoru vodnog gospodarstva, energetskom i prometnom sektoru te sektoru poštanskih usluga

su otvorile nekoliko prilika za ulaganja u današnje pametne gradove, kao sutrašnje pametne gradove zračne luke. Na način da su zadržale osnovne zahtjeve tržišnog natjecanja, transparentnost, jednaki tretman i usklađenost s pravilima EU-a o državnim potporama, a s ciljem poticanja zapošljavanja i inovacija uvodi se novi postupak nabave pod nazivom partnerstvo za inovacije (čl. 31. Direktive 2014/24/EU).

Trenutno zračne luke polako prolaze kroz unutarnju transformaciju, a digitalni alati su ti koji im mogu pružiti novi koncept vezan uz oglašavanje, digitalne interaktivne trgovine, programe vjernosti i mobilne aplikacije za putnike.

Digitalna transformacija zračnih luka nije isključivo prihvaćanje tehnologije. Također se radi o poslovnoj transformaciji u digitalnom svijetu. Radi se o razvoju procesa i usluga kako bi se putnicima i kupcima pružilo bolje iskustvo.

Kao što je spomenuto u članku „Utjecaj digitalne transformacije na zračne luke": „Takva transformacija također omogućuje besprijekorno korisničko iskustvo u zračnoj luci i besprijekornu integraciju sustava i usluga, uključujući sve dionike od sigurnosnih zračnih prijevoznika, carine, koncesija i zemaljskih operatera.

Digitalna transformacija u zračnim lukama koristi tehnološka rješenja kao što su unutarnje geolokacije, upravljanje identitetom, upravljanje protokom putnika, rudarenje podataka, umjetnu inteligenciju (AI) i Internet stvari za poboljšanje sigurnosti i pojednostavljenje njihovog poslovanja." [12]

Putem aplikacija za svoje putnike, zračne luke mogu dodatno poboljšati lojalnost kupaca i cjelokupno iskustvo.

Prijave bi trebale uključivati ažuriranja statusa leta i promjene izlaza u stvarnom vremenu, pristup širokom izboru maloprodajnih i ugostiteljskih objekata u zračnoj luci, mogućnost zarađivanja i iskorištavanja nagradnih bodova te dobivanje ažuriranih informacija o promocijama i događajima u zračnoj luci.

Pametne zračne luke putnicima također žele pružiti osjećaj dobrodošlice putem mogućnosti geoprostorne inteligencije koja pomažu putnicima da se brzo kreću kroz zračnu luku s interaktivnim kartama i traže najprikladnije načine da dođu do određenih točaka u zračnoj luci. [12]

Zračne luke moraju svakodnevno surađivati s vladinim agencijama. U tom kontekstu, inovacija može pomoći procesima za poboljšanje iskustva i sigurnosti putnika. A implementacija procesa automatizacije može poboljšati operativnu učinkovitost. [18] No, budući da biti pružatelj tehnologije nije primarna misija zračne luke, zračne luke obično moraju surađivati s drugim organizacijama kako bi pružile inovativne proizvode ili usluge. [19]

S druge strane, kao i mnoge druge industrije, zračne luke se suočavaju i s učincima sve većeg pritiska na okoliš. A kao rezultat toga, zračne luke digitaliziraju svoje procese; zrakoplovne kompanije moderniziraju flote zrakoplova kako bi ih učinile ekološki prihvatljivijima, smanjujući tako svoj ugljični otisak. Dakako, lokacija zračne luke ima značajnu ulogu, a ključnu ulogu imaju i raspoloživi sadržaji kojima zračna luka raspolaže.

Kao rezultat toga, zračne luke diljem svijeta prisiljene su usvojiti planove upravljanja okolišem te politike i strategije koje su ekološki prihvatljivije. [19]

Nadalje, istraživanje Digitalne transformacije zračne luke [20] ispitalo je osam ključnih aktivnosti, s namjerom identificiranja specifičnih digitalnih tehnologija za koje se očekuje da će se primjenjivati u budućnosti.

To su: Big Data i napredna analitika, Tehnologije u oblaku, Internet stvari, Pametni strojevi i roboti, Integrirani ekosustavi/Mobilna dodana vrijednost, Pametni energetski sustavi i autonomni sustavi, koji su pokazali da su zračnim lukama potrebne hitne promjene u digitalizaciji i ubrzanju procesa. [21]

Znači, cilj zračnih luka je digitalizacija i poboljšanje iskustva putnika, slično kao i u pametnim gradovima koji svojim građanima nastoje donijeti znatno bolje uvjete za život. Trenutno važeći hrvatski Zakon o zračnim lukama [22] i Zakon o zračnom prometu [23] RH ne pružaju mogućnost korištenja inovativnih IoT tehnologija te na taj način otežavaju primjenu istih u području digitalizacije zračnih luka, no sam Zakon o javnoj nabavi RH [24] u čl. 86., st. 2. definira mogućnost korištenja inovacija u postupcima javne nabave te navodi da: „javni naručitelj može koristiti partnerstvo za inovacije ako ima potrebu za inovativnom robom, uslugama ili radovima koju ne može zadovoljiti nabavom robe, usluga ili radova već dostupnih na tržištu.", stoga je u tom segmentu usklađen s Direktivom 2014/24/EU o javnoj nabavi te nudi mogućnost korištenja inovativnih usluga i u području unapređenja zračnih luka.

V. ZAKLJUČAK

Javna nabava inovativnih dobara i usluga ključna je za pametne gradove zračne luke jer pridonosi ne samo učinkovitosti javnih usluga i rješavanju njihovih potreba, kao što su mobilnost, okoliš i sigurnost, nego također pomaže u stvaranju radnih mjesta i jačanju poslovne konkurentnosti na europskom tržištu. [25]

Kriza uzrokovana koronavirusom s kojom je svjetska populacija živjela gotovo 3 godine značajno je oštetila zračni promet, stoga su promjene nužne posebice u vidu digitalizacije i korištenja inovativnih i digitalnih rješenja.

Također, situacija s pandemijom pokazala je da EU i ostatak svijeta nisu spremni za ovakve ili slične događaje te da su hitno potrebne pripreme za buduće izvanredne situacije.

Predlaže se; *de lege ferenda,* da se dopuni pravna regulativa [22] i [23] i to sukladno novorastućim tehnologijama te da se navedeni zakoni dopune uz mogućnost primjene inovativnih tehnoloških rješenja kroz postupke javne nabave, a sukladno ranije navedenim smjernicama Europske komisije, Direktivi 2014/24/EU o javnoj nabavi te hrvatskom Zakonu o javnoj nabavi.

Štoviše, u nadolazećim klimatskim promjenama gdje većina narativa poziva na smanjenje ugljičnog otiska tehnološki moderni hubovi prvi će kreirati koncept pametnog aerodromskog grada i prvi već koriste metode javne nabave inovativnih rješenja, kako je prikazano u radu, jer će samo na taj način moći biti održivi i konkurentni.

Između ostalog, neki znanstveni izvještaji predviđaju i preporučuju značajno smanjenje unutar sektora zračnog i pomorskog prometa gašenjem velikog broja zračnih i morskih luka, poput recimo spomenute studije pod nazivom „Absolute Zero" [26], no autor ne dijeli njihovo oduševljenje zatvaranjem globalne logističke mreže

zračnih luka i pomorskih luka jer takav princip neće biti moguće ostvariti zbog potrebe prekooceanskih letova i dostave velikih količina tereta.

Bez zračnog i pomorskog prometa transport sirovina, ljudi i roba ne bi bio moguć, a život bi postao sasvim nezamisliv, uz dužno poštovanje prema klimatskim promjenama.

Trenutačno mnogi izazovi s kojima se suočavaju pametne zračne luke premašuju kapacitete, sposobnosti i postignuća dosadašnjih institucija i tradicionalnih procesa upravljanja. Stoga su potrebni novi i inovativni oblici upravljanja i hitna implementacija novih tehnoloških rješenja navedenih u ovom radu.

Navedena tehnološka i digitalna rješenja zahtijevaju korištenje javne nabave inovativnih rješenja te modela javno-privatnog partnerstva i ugovora o koncesiji za prevladavanje ovih poteškoća.

Međutim, potrebna su daljnja dubinska istraživanja za većinu sadašnjih europskih čvorišta kako bi se utvrdilo mogu li se zračne luke nositi sa svim digitalnim napretkom i pronaći načine da zaštite zračni promet od izumiranja te razviti uspješne i međusobno povezane gradove digitalnih zračnih luka kao središta regije.

LITERATURA

[1] R. Riva Sanseverino (ed.), Competitive Urban Models, in Smart Rules for Smart Cities, vol. 12, Springer, 2014, pp.1-14.

[2] A. Graham, Managing Airports, An international perspective, (Fifth edition),Routledge, 2018, p. 1

[3] M. Milenković, The Applicative Croatian Model of Concessions for Managing Airports with Comparative Analysis, University of Ljubljana, Faculty of Law, Zagreb 2021, p. 128

[4] Airports International, Jesu li zračne luke sljedeći pametni gradovi?, Hitachi, Key Publishing Ltd, 2020, pp. 2-7, https://www.airinternational.com/article/hitachi-are-airports-the-next-smart-cities, Preuzeto 10/02/2023

[5] V. H. Nam, GREEN SUSTAINABLE AIRPORTS: THE DEPLOYMENT OF RENEWABLE ENERGY AT VIETNAM AIRPORTS. IS THAT FEASIBLE?, Journal of Mechanical Engineering Research and Developments (JMERD), 2019, p. 1, https://www.researchgate.net/publication/335496704, Preuzeto 14/02/2023

[6] Europska Komisija, Pregled financiranja iz EU-a, https://eufundingoverview.be/funding/procurement-of-ict-innovation-ppi-pcp, Preuzeto 16/02/2023

[7] M. Krylova, Potentials for the development of "Airport City " concept in Russia, 2016, DOI: 10.13140/RG.2.2.35239.75688, , p.7, https://www.researchgate.net/publication/313793778_Potentials_for_the_development_of_Airport_City_concept_in_Russia, Preuzeto 14/02/2023

[8] ICAO, https://www.icao.int/Meetings/FutureOfAviation/Pages/default.aspx, Preuzeto 14/02/2023

[9] ICAO, https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx, Preuzeto 24/02/2023

[10] Priopćenje Komisije. Nova industrijska strategija za Europu, COM (2020) 102 final, 10.3.2020, https://eur-lex.europa.eu/legal-content/HR/TXT/PDF/?uri=CELEX:52020DC0102&from=HR, Retrieved 10/02/2023

[11] M. Milenković, M. Rašić, G. Vojković, Using Public Private Partnership models in smart cities– proposal for Croatia, MIPRO 2017/DE-GLGPS, pp. 1656- 1661

[12] ISG, The Impact of Digital Transformation on Airports, https://isg-one.com/articles/the-impact-of-digital-transformation-on-airports, Preuzeto 12/02/2023

[13] Europska komisija, Smart cities, https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en, Preuzeto 22/02/2023

[14] IATA, Passenger air traffic dropped 90% at COVID-19 peak, https://www.airlines.iata.org/analysis/passenger-air-traffic-dropped-90-at-covid-19-peak, Retrieved 20/02/2023

[15] ICAO, Učinci novog koronavirusa (COVID ‑ 19) na civilno zrakoplovstvo: analiza ekonomskog učinka, Kanada 2022

[16] Europska Komisija, Obavijest Komisije, Smjernice za javnu nabavu inovativnih rješenja, Bruxelles 2021, C (2021) 4320 final, OJ C 267, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021XC0706(03)&from=HR, Preuzeto 14/02/2023

[17] Direktiva 2014/24/EU Europskog parlamenta i Vijeća od 26. veljače 2014. o javnoj nabavi, OJ L 94, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0024, (in force), Preuzeto 1/3/2023

[18] Airports Council International, Inovacije u zračnim lukama u Sjevernoj Americi, SAD 2020, https://airportscouncil.org/wp-content/uploads/Airport-Innovation/Innovation-at-Airports-in-North-America.pdf, Preuzeto 24/02/2023

[19] G. Baxter, P. Srisaeng, G. Wild, An Assessment of Airport Sustainability, Part 2—Energy Management at Copenhagen Airport, Resources 2018, 7, 32; doi:10.3390/resources7020032, https://www.researchgate.net/publication/325357230_An_Assessment_of_Airport_Sustainability_Part_2-Energy_Management_at_Copenhagen_Airport, Preuzeto 25/02/2023

[20] Airport digital transformation From operational performance to strategic opportunity, Arthur D Little, 2018, https://amadeus.com/documents/en/airports/research-report/airports-digital-transformation.pdf, Preuzeto 25/02/2023

[21] Single Market Scoreboard Report by the European Commission (2018). Available online: https://ec.europa.eu/internal_market/scoreboard/performance_per_policy_area/public_procurement/index_en.htm, Retrieved 1/3/2022

[22] Zakon o zračnim lukama, NN 19/98, 14/11, 78/15, https://www.zakon.hr/z/439/Zakon-o-zra%C4%8Dnim-lukama, Preuzeto 20/04/2023

[23] Zakon o zračnom prometu, NN 69/09, 84/11, 54/13, 127/13, 92/14, https://www.zakon.hr/z/177/Zakon-o-zra%C4%8Dnom-prometu, Preuzeto 20/04/2023

[24] Zakon o javnoj nabavi NN 120/2016, https://narodne-novine.nn.hr/clanci/sluzbeni/2016_12_120_2607.html, Preuzeto 20/04/2023

[25] S. Manika, Mechanisms for Innovative-Driven Solutions in European Smart Cities, Smart Cities 2020, Switzerland, p. 535, 3, 527–540; doi:10.3390/smartcities3020028, Preuzeto 1/3/2023

[26] Absolute Zero, UK FIRES, 2019, UK, https://www.repository.cam.ac.uk/bitstream/handle/1810/299414/REP_Absolute_Zero_V3_20200505.pdf?sequence=9&isAllowed=y, Preuzeto 3/03/2023

EE

International Conference on

ENGINEERING EDUCATION

Steering Committee

Chairs: Predrag Pale, University of Zagreb, Croatia
Branimir Pejčinović, Portland State University, United States
Juraj Petrović, University of Zagreb, Croatia

Members: Leslie Martinich, Competitive Focus, Austin, TX, United States

Program Committee

Members: Tomislav Jagušt, University of Zagreb, Croatia
Branimir Pejčinović, Portland State University, United States
Juraj Petrović, University of Zagreb, Croatia
Ana Sović Kržić, University of Zagreb, Croatia

PAPERS

Development of Renewable Energy Course for Electrical Engineering Program

Fadi R. Shahroury, Ibrahim Abuishmais, and Hani H. Ahmad
Princess Sumaya University for Technology Amman, Jordan
fadi@psut.edu.jo

***Abstract*—With the increased penetration of renewable energy in Jordan, higher education institutes are required to equip students with the knowledge and skills to thrive in the market. In response to that, this paper presents a newly developed course in renewable energy systems (RES). The designed course contains a practical part in addition to theoretical knowledge. The laboratory embedded in this course includes state-of-the-art equipment in the field of photovoltaic modules, an off-grid system with storage and a wind turbine. The course offers a unique teaching method that bridges theoretical and practical knowledge gaps. As this course is part of an ABET-accredited program, the course learning outcomes are designed carefully to align with ABET outcomes. Upon completing this course, students show a better ability to design and implement a renewable system. An example of a small-scale off-grid photovoltaic (PV) system is demonstrated in this work. At the end of the course, all students are asked to fill out a feedback questionnaire to assess the teaching method's effectiveness and evaluate if all learning outcomes are met. The results of this pilot project are auspicious and students' overall perception was very positive.**

Keywords*—Renewable energy awareness, education, ABET, off-grid, renewable energy systems.

I. Introduction

Jordan's total energy consumption in 2020 is over 18,500 gigawatt-hours (GWh). The distribution of electric energy consumption is illustrated in Fig. 1. The Figure shows that residential and industrial sectors dominate energy consumption with 45.8% and 22.1%, respectively. Most energy is generated by non-renewable and nonenvironmentally friendly sources such as crude oil and natural gas [1], [2]. The Jordanian government recognizes the harmful effects of these resources and sets up plans to use renewable energy sources, such as wind and solar. The Jordanian government encourages and incentivizes investors to build utility-scale wind and solar farms to achieve this goal, creating a high demand for specialized and skilled personnel in the market [3].

As a response, higher education institutes such as Princess Sumaya University for Technology (PSUT) started to develop curricula in this field to support the growing demand. In order to provide the electrical engineering students at PSUT with state-of-the-art knowledge in the renewable energy field, the university actively develops renewable energy courses, laboratories, and on-campus off-grid solar stations. This allows students to gain theoretical and hands-on experience with photovoltaic systems components such as solar cells, inverters, and batteries. In addition, the students learn about wind energy and how renewable energy is integrated into residential and commercial enterprises, as well as gain an understanding of the adverse economic and health impacts of coal-based energy generation.

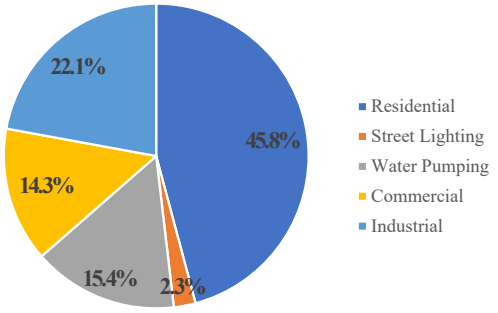


Fig. 1: The distribution of electric energy consumption in Jordan.

Program-level accreditation is increasingly becoming popular worldwide as it ensures a good quality of degree programs and graduates [4]. The Accreditation Board of Engineering and Technology (ABET) is the main accreditation body for engineering programs. More than 3709 programs in the United States and worldwide have gained ABET accreditation [5]. This paper presents the experience of the engineering school at Princess Sumaya University for Technology in developing a renewable energy laboratory that complements the theoretical part of the course. Thus, the overall course provides a unique experience to students as it combines theoretical and real-world environmental aspects of the subject matter.

The rest of the paper is organized as follows; the advanced renewable energy course is described in Section II. The course learning outcomes are discussed in Section III. Section IV presents the main results and discussion. Finally, the conclusions are drawn in Section V.

II. The Advanced Renewable Energy Course

The advanced renewable energy course is a three-credit-hour course for bachelor's students in their fifth year. The prerequisite of this course is an introductory course to renewable energy systems. The theoretical part of this course covers the following topics :

- Off-grid PV system design.

- Energy storage systems; types and design.
- Fuel cell technologies.
- Smart grid: architecture, communication and measurement technologies, including power management units (PMUs).
- Power electronics for smart grid and renewable energy systems.
- Modern distribution networks.

In addition to regular class lectures, the course contains laboratory work that accounts for 9 contact hours [6]. The Lab consists of four experiments listed in Table I. The first experiment explores the different semiconductor materials used for PV cells. Here, students identify three different types, namely, Monocrystalline (c-Si), Amorphous Silicon (a-Si) and Cadmium Telluride (CdTe), as shown in Fig. 2. This experiment aims to expose students to different types of semiconductor materials and their characteristics.

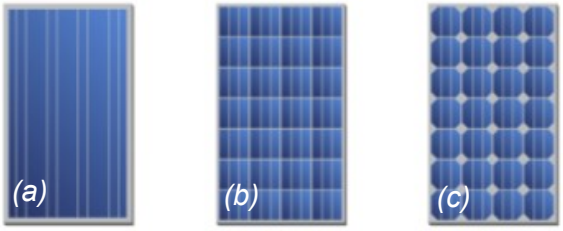


Fig. 2: PV module types : (a) Thin-film CdTe, (b) Amorphous Silicon a-Si and (c) Mono Crystalline Silicon c-Si.

The second experiment explores PV panel performance under different tilt azimuth angles. Here a compass and a protractor are used to determine the module tilt and azimuth angle. The objective here is to enable the student to understand the impact of these angles on the power produced by the module. Related topics like sun path diagram, development of shadow diagram and tracking systems are also clarified. Fig. 3 depicts the used PV modules with a movable custom-made structure. The PV module tilt angle Σ is set using a built-in protractor and the azimuth angle ϕ_C is determined using a generic compass.

In the third experiment, students are introduced to off-grid PV systems. This part is complementary to lectures dealing with off-grid systems and storage sizing. The experiment comprises two parts: the first part deals with the off-grid PV system components, where students study an existing large-scale system and examine its main parameters. On the other hand, the second part teaches students to build a small-scale off-grid PV system. Fig. 4 shows the large-scale off-grid system components. The system comprises a string of PV panels with a peak rating of 1250W mounted on the roof, a maximum power point tracking (MPPT) charge controller, a battery energy storage system (BESS), an inverter and a distribution board containing fuses and circuit breakers. The BESS is constructed using 4 lead-acid absorbent glass mat (AGM) batteries; each with a 300Ah capacity [7]. The storage system has two strings, each containing two 12V batteries

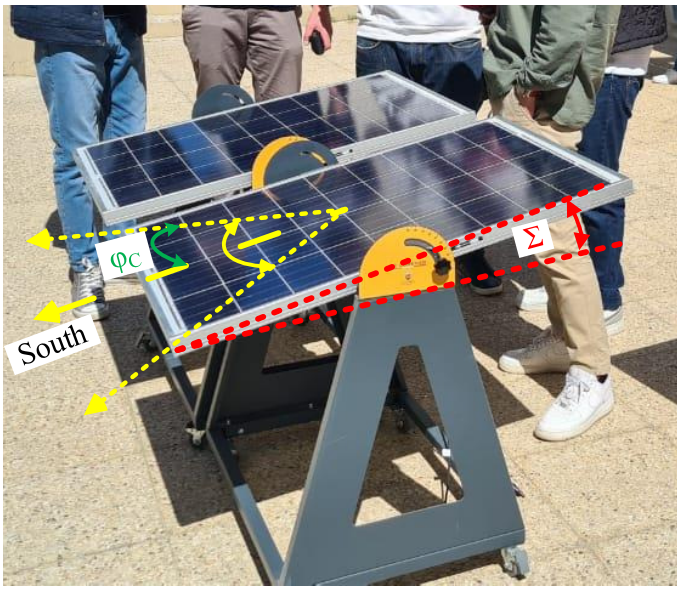


Fig. 3: PV panel performance under different tilt azimuth angles setup.

connected in series resulting in a 24V DC bus. The total system capacity is 600Ah. With an average insolation level of 5.5kWh/day in Amman [8], the system is expected to generate 2000kWh/year. This annual energy production is calculated using the peak-hours approach estimation method [9]. Fig. 5 shows a photograph of the system except for the PV panels.

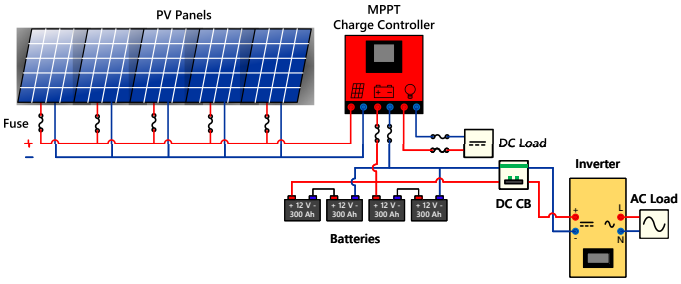


Fig. 4: The schematic diagram for the off-grid system.

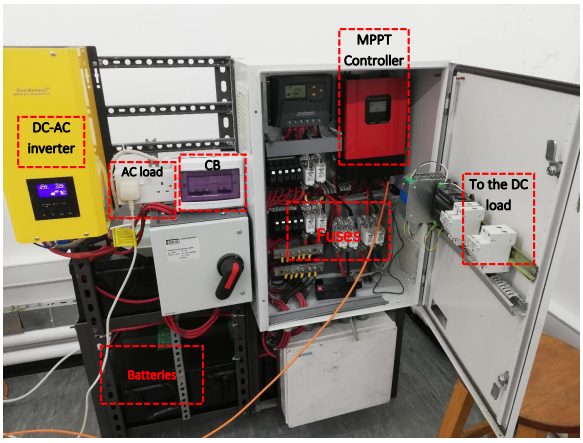


Fig. 5: A photograph of the off-grid PV system.

The last experiment deals with wind energy, particularly the characteristics of a wind turbine. The experiment introduces students to the effect of the number of blades on the turbine performance. Students can evaluate the wind turbine performance at different speeds by using two, three, and four blades turbines and measuring the power delivered to a load connected to the generator. This part of the experiment clarifies the theoretical analysis done during the course that relates the tip speed ratio to the rotor's efficiency. By increasing the number of blades, the turbine reaches the highest efficiency at a lower wind speed. The students can develop the wind turbine's power curve in the second part of this experiment. The objective here is to identify the key wind speeds: cut-in, rated and cut-out speeds and differentiate them from the idealized characteristics of the turbine. The power curve is developed by changing the wind speed and measuring the power delivered by the turbine. The power versus wind speed curve of the used turbine is depicted in Fig. 6. This setup is delivered by the De Lorenzo Group [10].

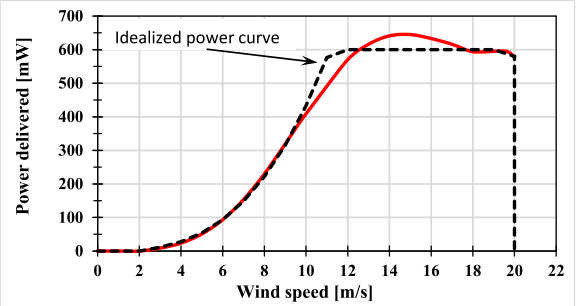


Fig. 6: Measured (solid) and idealized (dashed) power curves of the used wind turbine.

TABLE I: List of experiments.

Experiment No.	Experiment Title
1	Types of PV modules material and I-V characteristics
2	PV panel performance
3	Off-grid PV systems
4	Wind turbine characteristics

III. The Course Learning Outcomes

Laboratories and practical classes are vital parts of the engineering learning process. Specific course learning outcomes (CLOs) are achieved by performing practical experiments. This course includes the renewable Lab to enhance the teaching process and improves students' understanding of the main topics in the field of solar and wind energy. Furthermore, skills in solving engineering problems, the ability to function effectively in a team and communication skills are also fostered during the Lab. these CLOs are achieved by dividing the class into teams and assigning a design task for each. The key learning outcomes upon completion of this course are:

CLO1 for types PV modules material and I-V characteristics:

- Identify different semiconductor materials used in the PV cells.
- Understand the impact of semiconductor material on the PV module parameters such as power rating, efficiency and temperature derate factor.

CLO2 for PV panel performance:

- Compute and measure the power delivered from a module at different tilt angles.
- Describe the impact of collector azimuth angle on the PV module performance.
- Describe the performance and applications of connecting modules in series or parallel.

CLO3 for Off-grid PV systems:

- Identify the main components used for off-grid PV systems and describe their functionality.
- Compute the required storage size considering the battery type, load demand, power availability and the site temperature.

CLO4 for wind turbine characteristics :

- Describe the design of the wind turbine and the effect of the number of blades on the turbine performance.
- Construct the wind turbine's power curve and identify the cut-in, rated and cut-out wind speeds on the constructed curve.

Since this course is part of an ABET-accredited program, the instructor is also required to assess ABET students' outcomes (2, 3, 5 and 6) [5].

IV. Results and Discussion

The learning process in this course comprises two phases. The first phase targets enhancing theoretical knowledge via lectures and homework assignments. The second phase offers an opportunity to gain hands-on experience in the Lab. To demonstrate the effectiveness of this strategy, students were asked to design an off-grid PV system for a household before conducting an experiment. In this off-grid PV system, the calculation of the load demand, the design of the PV array and battery sizing using lead-acid batteries are covered. An example off-grid system built by a group of students is shown in Fig. 7. This learning process enables the students to understand the practical experiment better and compare measurements to theoretical calculations.

The aim of this work is to enable higher education institutes to develop high-tech courses in the renewable energy field. For continuous improvement purposes, a questionnaire was distributed seeking students' feedback about the course. The survey instrument is designed based on 18 items in four dimensions (The effectiveness of the learning pedagogy, Lab status equipment, educational learning outcomes and their relevance to ABET, and students' overall perception). The Likert-type scale, a 5-point scale ranging from "Strongly disagree" to "Strongly agree," was used to collect the data from the 14 undergraduate students enrolled in the course. Table II lists the items

TABLE II: The students' feedback survey questionnaire items.

Item No.	Effectiveness of the learning pedagogy	Strongly agree (%)
1	The Lab was organized in a manner that helped me understand the underlying concepts.	92.9
2	The Lab gave me the confidence to do more advanced work in the subject.	85.7
3	I believe that what I'm being asked to learn in this Lab is important for me.	85.7
4	The Lab associated with the RES course increased my knowledge in the PV and wind energy resources field.	64.3
5	The Lab complemented the material covered in the lectures.	92.9
Item No.	**Lab equipment status**	**Strongly agree (%)**
1	The equipment used in the renewable energy Lab is suitable for the learning proposes.	92.9
2	The Lab experiments were up-to-date (modern).	64.3
3	The RES Lab experiments helped me understand the practical design of PV systems.	85.7
4	After conducting the RES Lab experiments, I will be able to design a simple off-grid PV system.	64.3
Item No.	**CLOs versus ABET's student outcome**	**Strongly agree (%)**
1	The RES Lab helped me acquire and apply new knowledge in the PV and wind energy fields.	78.6
2	The RES lab enhanced my ability to function effectively on a team formed by my classmates and the instructor.	78.6
3	The RES Lab enhanced my ability to develop and conduct appropriate experiments to characterize different PV module types.	85.7
4	The RES Lab enhanced my ability to develop and conduct appropriate experiments to better understand the wind turbine characteristics.	64.3
5	The skills I learned while writing the RES Lab report helped me write an effective and concise report to communicate experimental results and conclusions.	85.7
6	After taking the RES Lab, I can follow the design procedure for designing an off-grid system introduced in the course to create an off-grid system that meets specific design requirements.	71.4
Item No.	**Students' overall perception**	**Strongly agree (%)**
1	I would highly recommend this course to other students.	92.9
2	Overall, this course met my expectations for the quality of the course.	85.7
3	The course was helpful in my progress toward my degree in Electrical Power and Energy Engineering.	92.9

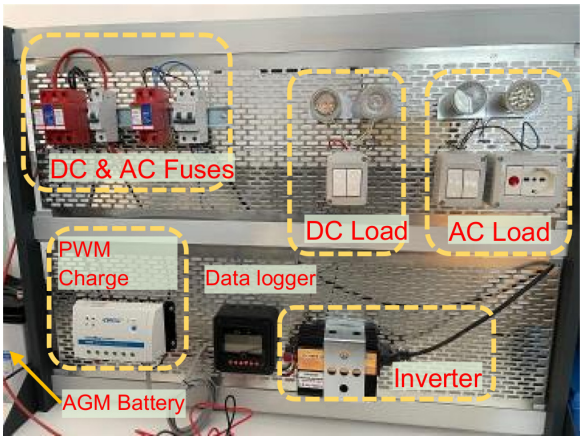


Fig. 7: A simple off-grid PV system built by a group of students.

included in the survey and the percentage of strongly-agree responses.

In the "Effectiveness of the learning pedagogy" dimension, the overwhelming majority of students find the teaching method very effective; 93% assert that the Lab complements the lectures and increases their knowledge. The overall average of the "Strongly agree" responses in this dimension is 84%.

As for the "Lab equipment status," most students believe the Lab has up-to-date apparatus suitable for learning. A 64.3% of students responded with a "Strongly agree" that they are confident replicating a simple off-grid PV system after taking this course. The overall average of this dimension's "Strongly agree" responses is 76.8%.

In the "CLOs versus ABET's student outcome" dimension, the survey reveals that most of the learning outcomes of this course are met. 98.7% of students "Strongly agree" that the Lab helps them acquire and apply new knowledge in the field of RES. A 93% of students assert that the course enhances their ability to function effectively in a team and enhances their skills in technical writing. All students either agree or strongly agree that their ability to develop and conduct experiments in the RES field has improved. The overall average of the "Strongly agree" responses in this dimension is 77.4%.

In the last dimension, "Students' overall perception," 85% of students think the course has met their expectations, and 92.9% would highly recommend this course to other students. The overall average of this dimension's "Strongly agree" responses is 90.5%.

Soliciting other input not covered by the survey, students are invited to write their comments on their course experience and potential future enhancement. Here are samples of students' written comments:

"This course was of great benefit and helped me further understand the renewable energy world and was most enjoyable because it had an integrated lab within the course, which greatly helped me understand the topics discussed during this course and I hope more courses implement the same strategy in the following years."

"I really enjoyed this course; it was accompanied by practical implementation, which our courses lack the most; it also sets you up for real work, where a lot of research and reports are written."

"The course increased my knowledge in battery types

and PV systems." "It was an amazing course, honestly."

V. Conclusion

This course was held purely as an extension course in renewable energy for undergraduate students. The majority of students really enjoyed it. The development of an advanced RES course that combines theory with laboratory came as a response to the increasing awareness of clean energy needs. The teaching style used in this course, combining theoretical knowledge and practical experience, has proven to be very effective. This course's quality of teaching and learning is enhanced as students acquire various skills through observation and practical experimentation. This course achieves different learning outcomes in the field of renewable energy technology in general, PV and wind in particular. The designed course successfully conveyed the necessary knowledge of renewable energy to students. For example, the students demonstrate the ability to design and build a small-scale off-grid PV system. All students are asked to complete a feedback survey questionnaire at the end of the course to evaluate the teaching strategy's effectiveness and determine whether all learning objectives have been accomplished. The outcomes of this initiative are promising, and students generally had a very high opinion of it. The aim of this work is to enable higher education institutes develop high-tech courses in the renewable energy field. Future work will be intended to improve the teaching methodology by implementing a flipped-classroom approach [11] and upgrading the renewable energy Lab so students can access it remotely [12].

References

[1] S. S. Alrwashdeh, "Energy sources assessment in jordan," *Results in Engineering*, vol. 13, p. 100329, 2022.

[2] S. Sandri, H. Hussein, and N. Alshyab, "Sustainability of the energy sector in jordan: Challenges and opportunities," *Sustainability*, vol. 12, no. 24, p. 10465, 2020.

[3] G. Abu-Rumman, A. I. Khdair, and S. I. Khdair, "Current status and future investment potential in renewable energy in jordan: An overview," *Heliyon*, vol. 6, no. 2, p. e03346, 2020.

[4] H. Lucas, S. Pinnington, and L. F. Cabeza, "Education and training gaps in the renewable energy sector," *Solar Energy*, vol. 173, pp. 449–455, 2018.

[5] ABET Student Outcomes. [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2021-2022

[6] EE 24582 Renewable Energy. [Online]. Available: https://sites.google.com/psut.edu.jo/ee-24582-renewable-energy/home

[7] E. A. Batter. Batters. [Online]. Available: {https://www.everexceed.com/ev-agm-battery_p98.html}

[8] A. Qasaimeh, "Solar energy optimization through seasons: case study in jordan," 2012.

[9] G. M. Masters, *Renewable and efficient electric power systems*. John Wiley & Sons, 2013.

[10] D. Lorenzo. Wind power plant. [Online]. Available: {https://edquip.co/en/de-lorenzo/wind-power-plant}

[11] F. Shahroury, "E-learning during covid-19 epidemic: Experience of a university from jordan," *Academy of Strategic Management Journal*, vol. 21, no. 4S, pp. 1–6, 2022.

[12] F. R. Shahroury, L. Al-Tarawneh, and A. Al-Zoubi, "In-the-online-class remote lab in post covid-19 pandemic," in *2022 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 2022, pp. 677–682.

Hybrid Agile Approach in Software Engineering Education – A Case Study

M. Kaluza, S. Candrlic and M. Asenbrener Katic
Faculty of Informatics and Digital Technologies, Rijeka, Croatia
marina.kalu5a@gmail.com, sanjac@inf.uniri.hr, masenbrener@inf.uniri.hr

***Abstract* - By definition, software engineering is a systematic approach to software development. However, to succeed in today's dynamic business environment, development teams must adapt and respond to change quickly. To prepare students for their role in real-world software development teams of which they will soon be a part, teachers attempt to create as realistic a project environment as possible. The software development education model presented in this paper is based on the hybrid agile approach. It combines both the agile approach, which is suitable for dealing with changes in requirements, and the planning-oriented, systematic approach, which is traditionally used in software engineering. In addition, the students' attitude towards this hybrid agile model is investigated.**

Key words – hybrid agile, agile, software engineering, education

I. INTRODUCTION

Higher education courses are prone to many changes over time as the industries and jobs, for which universities prepare students evolve and change. Software engineering is one of the fundamental courses of study programs in the field of informatics, computer science and engineering. Software engineering is a systematic approach to software development in which the software development process is viewed as a series of phases and steps [1], [2].

Traditional development approaches are focused on predefined phases. Requirements are fully specified before coding begins. This approach gives large companies a structure for monitoring, and simplifies the understanding and scheduling of tasks. Managing large teams of developers is easier because the end goal and all requirements are specified before the development phase. Some of the drawbacks include lack of adaptability, delayed testing, and delayed product release. The lack of customer involvement in development can lead to dissatisfied customers who have waited too long for a product that ends up not providing everything the customer wants. One of the most commonly cited traditional models is the waterfall model [1], [2], [3].

The Waterfall model consists of seven development phases: requirements analysis, design, coding and implementation, testing, operation and deployment, and maintenance. Each of these phases must be fully completed before moving to the next phase. In the first three phases, all planning is done, and only when all coding and testing is completed, can the project be deployed. The last phase does not end, because perfecting the product and improving of its functionalities will continue as long as the product is used [3].

With the publication of the agile manifesto, a rapid development of many different agile methods began. Agile approaches were created to overcome the shortcomings of traditional approaches. The agile approach focuses on tasks, users and team members, rather than documentation, and it focuses more on responding to change rather than following a fixed plan. Development in agile approaches is carried out through iterations. Unlike traditional approaches, user involvement and feedback are crucial in the development process. Some advantages of agile approaches are easier adaptability to changes, quick user feedback, and earlier product implementation. On the other hand, agile approaches lack documentation, which can be a problem when new members join the project. It is more difficult to measure progress because the end result is not completely known. Short iterations can also lead to avoiding the development of some large features due to their complexity and not leaving enough time to design ideas, which can lead to later changes due to user dissatisfaction [4], [5], [6].

Scrum is the best-known agile methodology. In Scrum there are three different roles: Scrum team, Scrum master and product owner. The Scrum team is a team of different experts who work together to develop the product. The Scrum master is a member of the team who makes decisions when there are disagreements within the team. The product owner provides feedback to the user. The Scrum methodology is based on development in iterations, sprints. At the end of each sprint, a new product increment is achieved. A Scrum sprint lasts from at least a week to a month and consists of brief planning, development of selected tasks from the backlog, testing, and at the end a sprint review is conducted. Progress is presented to the project owner in the review meeting. Development is monitored through daily scrums when developers discuss what smaller tasks should be completed that day. Scrum promotes good communication between team members and ensures regular communication with the product owner, which is why it is possible to reduce costs, since in each development step only what the project owner needs is

done, and all members are well acquainted with the requirements and needs of the project [7].

In addition to traditional methodologies and the agile approach, companies today also use a hybrid agile approach. This approach combines the advantages of traditional and agile methodologies to produce a methodology which can be used with minimal customization [8], [9]. Hybrid approach is usually based on waterfall model and some agile methods. The focus is on the need for an initial plan, which is not as extensive as in the waterfall model, but consists of initial documentation, time and cost estimates and main goals. In the next step, development is carried out according to an agile model. For example, if the Scrum methodology is used, the sprint iteration begins in this step [9], [10]. Since the main development is done using an agile approach, this approach to software development can also be called a hybrid agile approach, which is the focus of this paper.

The paper describes the implementation of the hybrid agile approach in a software engineering course and the evaluation by the students. After the introduction, the second chapter presents related work on the implementation of the hybrid agile approach in industry and software engineering courses. The methodology that covers course structure and the evaluation of the implemented approach is described in the third chapter. The fourth chapter presents the survey results and their discussion. Finally, the conclusions and plans for future work are presented.

II. RELATED WORK

Many different studies have been conducted on the implementation of traditional and agile methods in work and educational environments. The main focus of this paper is on the hybrid agile approach.

A. Development methodologies used in work environments

Previous research shows that companies mostly use different hybrid approaches for software development [11], [12]. The authors in [11] explored which methodologies are used in different organizations. Out of 6 respondents, 5 use a hybrid development model, consisting of the waterfall model and the Scrum methodology, while only one respondent uses a pure agile approach with the XP methodology. Based on this, it can be concluded that in the organizational environment, when developing complex projects, it is usually not possible to exclude documentation, which is emphasized in traditional methodologies but neglected in agile methodologies. The authors in [12] collected survey data on global software development (GSD), which is characterized by distributed and large-scale development and concluded that GSD needed to move away from the use of traditional approaches, resulting in companies primarily using some type of hybrid approach to development. They also noted that pure agile approaches are somewhat rare in GSD, which is consistent with [11].

The authors in [13] discuss adoption of a hybrid development model in companies outside the IT industries. The implementation of the hybrid methodology was conducted under the supervision and guidance of three agile methodology coaches who participated in this study. Companies that moved to a hybrid model experienced several benefits of implementing some agile practices in development, such as creative contributions from individuals, as well as better communication between different teams and more transparency within the team.

When companies hire new employees, they want applicants to have extensive knowledge that universities have difficulty providing [14]. The authors in [14] noted that because expectations are too high, companies are forced to hire individuals who lack soft skills, which proves to be a greater disadvantage than less knowledge. Software engineering education needs to focus more on computer science and the current tools used by the companies that will employ the students upon their completion of higher education. On the other hand, companies need to hire people who are ready for teamwork even if they have less knowledge than other applicants.

B. Implementation of agile approaches in education

The authors in [15] conducted a comprehensive study of software engineering education papers that address software engineering trends. The most common trend used in education was agile software development. The authors of the paper found that, while the trend of implementing agile approaches in education is the most popular, it comes with its own challenges. This paper provides guidelines for practitioners, researchers, and educators.

The following papers describe how hybrid agile practices have been implemented in the higher education environment. The most commonly cited agile methodology in software engineering education is Scrum [16-20], which is not surprising since Scrum is often used in industry. The following papers show the implementation of agile approaches in higher education courses.

The authors in [16] presented a course in which a project for a real user was carried out. A team of 6 students participated in this course. Students were assigned roles and each was responsible for a different aspect of the project. The main problem was that the students had to learn about the technologies and procedures and apply them at the same time applying them. In addition, the students had to be encouraged to communicate with the users, so that the final product met user expectations. The authors concluded that not all agile practices are appropriate for the educational environment, but that they should still be used. Another conclusion is that the project needs to be designed simply enough that it can be done in a single course and students have enough time for other courses during the semester.

The authors in [17] did not face the problems mentioned in [16] that the students did not know fundamentals and technologies because their course consisted of a 3-week fast-track theoretical course and a 10-week product development course. Each team had the same assignment and was given a user with whom they

had to analyze and specify product requirements. Students could choose which agile methodology and practices they wanted to use. The course was deemed successful as all teams were able to successfully produce a final product. It was found that weekly meetings and user involvement played the biggest role in project success as students were warned if they were going in the wrong direction or taught how to adopt more agile practices.

The authors in [18] had also introduced agile development in the form of 1-week and in later years 2-week assignments for teams. The teams worked on all assignments separately, so the final results were different. Students confirmed that this type of project in a course gave them a better understanding of software development.

In contrast to the before mentioned sources, the following initially implemented traditional methodologies and then gradually implemented Scrum [19] or completely switched to Scrum [20]. Both papers found that the introduction of Scrum increased student interest. The authors in [19] initially taught students a traditional approach to software analysis. The agile approach was introduced gradually so that students could become more familiar with agile practices. Daily meetings encouraged students to actively participate. With this approach, student satisfaction in learning and developing software was observed as they developed a project that had a value to the end user. The authors in [20] faced the problem of students' working at the last minute when using traditional methodology which led them to switch to agile methodology. The introduction of Scrum in education was done in a controlled environment where students were given detailed instructions for the goals of each sprint. The authors also wanted to prevent non-participation in the team, so team work was eliminated and assessments were introduced at the end of each sprint. To promote the team aspect of agile methodologies, a platform was introduced where students could communicate, ask questions, and help each other. The conclusion is that independent projects like this, with the introduction of some aspects of agile approaches, can teach students about agile paradigms and development.

III. METHODOLOGY

The Software engineering course is one of the fundamental courses of the Graduate University Study Program Informatics at the Faculty of Informatics and Digital Technologies, University of Rijeka. The course was offered in the winter semester with 2 hours of lecture and 2 hours of practical work per week. The workload for the course is 6 ECTS credits. The course is taught by two teachers: a professor and a teaching assistant.

The aim of the course is to acquire knowledge in the field of software engineering and covers software development phases: requirements analysis, project development, team software development and software testing. It covers the application of both agile and traditional methods, techniques, and approaches that help with planning, team organization, and task management during software development within a specific timeframes and resources.

The course includes the following topics: Models of software development; Traditional, agile, and hybrid approaches to software development; Methods and techniques used in different phases of software development; Team management; Analysis and management of user requirements; Estimation of resources for software development; Risk management; Software design and architecture; Implementation; Construction of program code in collaboration; Refactoring; Testing; Version management; Software documentation; Professional responsibilities of software engineers; Software re-engineering.

Course outcomes are listed below. It is expected that upon completion of all course assignments, students will be able to:

- O1. Distinguish basic concepts, methods, techniques, and approaches in the field of software engineering, particularly as they relate to traditional and agile approaches.
- O2. Develop models of a system based on an analysis of user requirements and market needs in a given domain.
- O3. Estimate the resources required to build the software.
- O4. Plan software development considering the various roles of development team members and users in a software development team project.
- O5. Based on the analysis performed and the project created, create the software in the chosen development environment and prepare its documentation.
- O6. Perform tests based on the planned test cases and document the test results.

Course content is delivered through project-based learning, so that assignments simulate real-life situations.

The activities used to assess the acquisition of learning outcomes are a written exam (max. 30 points), project meetings (0-20 points for active participation, preparation, and proposed solutions) and software development using a hybrid agile approach (0-50 points).

At the beginning of the semester, students were divided into five teams of four. Each team member was assigned a primary role (project manager, designer, programmer, and tester), with responsibilities for the work performed. The teachers presented two main topics/themes that were used as global user requirements. Staring from these, each team developed its project idea. Each topic was selected by at least two project teams competing to develop better software for that topic.

The project began with a software requirements analysis to derive the key requirements for software development. Each team chose its own development framework. For each project meeting, the teams were given several tasks and deliverables to complete. Each meeting focused on a different deliverable, e.g., data model, mockup, test cases and scenarios, program logic and flow, etc. During the meetings, students presented

what they had done so far, commented on their solutions to the tasks, and set new tasks and plans for the next project meeting. There was a total of 4 project meetings and a final presentation of the software created by the end of the semester. A combination of traditional face-to-face classes and online classes supported by a learning management system (Moodle LMS) was used. Figure 1 shows the main activities in the course.

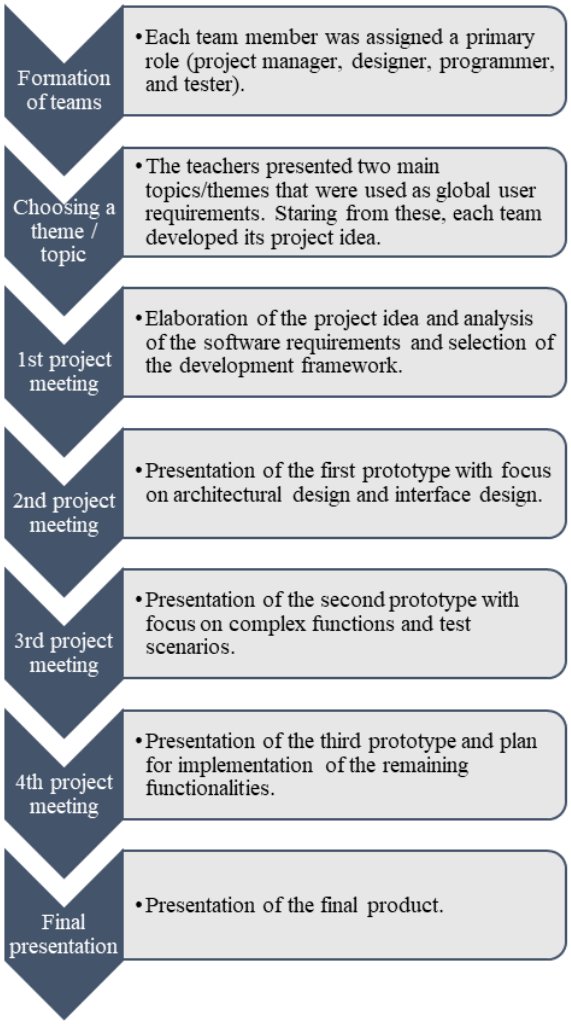


Figure 1. Activities in the course

After completing all teaching activities, students were asked to participate in a survey. The questionnaire was created using Google Forms. It consisted of questions with predetermined answers (yes/no), questions with a Likert scale of 1 to 5, where 1 was "strongly disagree" and 5 was "strongly agree", and open-ended questions. The survey was sent to all 20 students taking the "Software Engineering" course in 2020/2021, and all completed the survey. The Likert scale questions are shown in Table I, and the responses to the remaining questions are discussed in the text, both of which are listed in the next chapter.

IV. Survey Results and Discussion

Students on the teams were assigned roles (project manager, architecture designer, programmer, and tester), and most (95%) were satisfied with the role they chose. When asked if they would prefer to perform all activities equally, without being responsible for only one group of activities (e.g., as a tester for all testing activities), they had mixed opinions. 60% of students thought independent, separate roles were a better option, while 40% would prefer to participate equally in all activities, without being a separate person responsible for each activity. The latter group explained that with independent roles, team members are not equally motivated to participate and do not gain the same knowledge during the project. When asked how students should choose a role on the team, most agreed that everyone should try to expand their knowledge. However, when asked how this could be achieved, they had different ideas. Some felt that everyone should choose a role where they know more (60%) and then help the rest of the team learn, a few students felt it was better to choose a role where they are less experienced (15%) so that they can learn through the project, and others (25%) could not decide between the two because they felt that all students should participate in every aspect of the project.

Most students (75% strongly agreed and 10% agreed) felt that it was easier to work in a team with friends, i.e., they favoured independent division into teams rather than random assignment to a team. Opinions were divided on the question of whether the choice of the main project topic should be predetermined by the teacher: 35% of the students disagreed with this statement, 35% agreed, and 30% did not know.

Almost all agree (45% strongly agree, 45% agree) that the project meetings are a good way to evaluate progress. They think that it was good that the tasks for the project meetings were set 3 weeks in advance. One of the questions on which there was no unanimous opinion was: is it better to set the content and required documents of the meeting in advance, or should the meeting be flexible without full planning in advance. The general opinion is that some level of meeting organization is necessary to promote progress, but that too much documentation should be avoided to be consistent with the spirit of agile development. Some students suggested writing tasks on a Kanban board each week.

When asked if everyone participated equally, they were divided. Consistent with the previous question, students were asked how they would describe their contribution (Figure 2). Almost half of the students (45%) felt that their contribution was equal to that of the rest of the team, 35% and 20% felt that they had contributed less and more, respectively, to the final solution than the rest of the team.

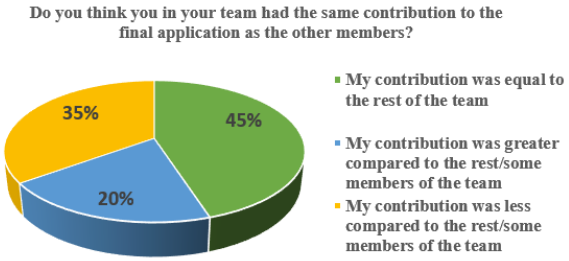


Figure 2. Contribution in team

TABLE I. QUESTIONS WITH LIKERT SCALE ANSWERS

Question	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)
Having roles in the team organization is a good way for all members to acquire equal knowledge	5	20	50	25	0
I am satisfied that we could chose our team	5	0	10	10	75
It is easier to work in a team with friends	0	10	15	10	65
Teacher should assing students into teams	25	50	10	5	10
In our team, the responsibilities were shared - everyone was responsible for the task they undertook (or were assigned)	0	20	20	35	25
The teacher's instructions for a main project topic (general user request) were sufficient to develop the project idea	0	0	15	20	65
The fact that we were competing with another team on the same topic motivated me to work harder and better	5	25	15	30	25
I would prefer if each team had a different topic as a main project topic, instead of having team competition	25	45	15	10	5
I would prefer each team to define on their own main project topic and idea	10	25	30	10	25
Project meetings are a good way to evaluate the progress of the project	0	0	10	45	45
Project meetings are a good way to evaluate the contribution of each individual team member	0	10	25	40	25
For a successful participation in a team (such as during this project), all team members should have the same level of knowledge	15	40	25	15	5
For a successful participation in a team (such as during this project), team members should have complementary levels of knowledge	0	10	5	50	35
It is good that tasks for each project meeting are set in advance (3 weeks before the meeting)	0	0	5	35	60
I would prefer that all tasks for all the meetings are known at the beginning of the semester	15	35	10	15	25
I would prefer that instead of having pre-defined content and required documents for each meeting, problems faced by the team and the functionalities to be included in the next version are discussed freely during the meeting, without planning it in advance	10	35	15	30	10
The assessment done during meetings encouraged me to actively participate in all project meetings and prepare the required deliverables of the project activities (documentation).	0	0	15	55	30
All members of my team participated equally actively in the project	5	25	20	40	10
Having a prototype required for one of the project meetings motivated me to work on the project.	0	5	0	75	20
I would prefer teachers to evaluate only the final product, without the checkpoints that were introduced during the project meetings	35	45	10	5	5
I believe that this project gave mean understanding of the problems and insight to real team software development.	5	5	10	55	25
In general, I think the experience of working on this project was useful.	0	10	0	30	60

It is interesting to note that only the students who held the roles of architecture designers and testers felt that their contribution to the final solution was less than the contribution of the other team members. The role of programmer was considered the most challenging by more than half of the students (55%), 25% considered the role of project manager the most challenging, and 20% considered all roles equally challenging, with no focus on the roles of architecture designer and tester (Figure 3).

When asked which approach, they thought should be the first choice for the software engineering course project, half of the students answered that it was a hybrid agile approach, the other half chose the pure agile approach, while no one chose the traditional approach. Even though they do not have much practical experience to judge this issue, the students' opinion was based on the theoretical ground they acquired in this this course.

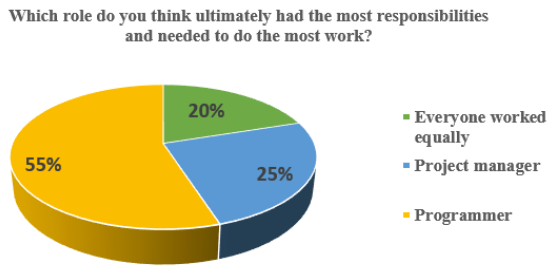


Figure 3. Hardest role

The students' opinions on the questions about the course organization and the project are shown in Table I. A Likert scale of 1 to 5 was used, with 1 representing "strongly disagree" and 5 representing "strongly agree".

Students generally found this project to be a useful experience, as all topics were adequately supported by theoretical classes and they were satisfied that it was possible to choose the development framework in which to develop the application. Suggestions for improvement included the introduction of "user" feedback to bring the agile approach closer. It was also suggested to introduce the Scrum methodology and define the objectives until the next meeting at the end of the previous meeting. In addition, it was suggested to introduce individual evaluation in the teams, as some students felt that not everyone contributed equally to the final product.

V. CONCLUSION

The survey results confirm some of the conclusions drawn in previous research on agile and hybrid agile methods in teaching. It was found that continuous monitoring of project progress is necessary, while evaluation of only the final product is rejected. Students believe it would be useful to introduce as many features of agile development as possible, less structured meetings and the introduction of some form of feedback immediately after the meeting.

However, as shown in the survey results, many students think that it is beneficial to have roles in the

team organization, which was also the case in [16]. Another opinion, as well as the conclusions from [17-19], is that roles should not be used to avoid the possibility of only 1 or 2 students can complete the whole project alone. However, it would be useful to improve the organization of roles or find another way to evaluate individual work, since half of the students felt that not everyone contributed equally to the project, and that some roles had more tasks and responsibilities at the end. The authors suggest dividing work by tasks rather than roles, which is more in line with the spirit of the agile approach. This would prepare the ground for all students to contribute equally to the development process through different types of tasks.

The competitive nature of the course organization was very motivating for the students. As in [16-20], a predefined project topic and implementation method (using specific tools) was considered a good way to go. However, unlike [16], [18-20], in the implementation of this project it was possible to choose a development tool that the students were satisfied with.

In our future work, we plan to improve the presented hybrid agile model. To make the agile approach more prevalent, this model will introduce a work management tool. This will allow teachers to better monitor the work done by students, as well as their individual contributions. Our future research will also focus on methods to assess individual contribution in software development team projects.

ACKNOWLEDGEMENTS

This work has been fully supported by the University of Rijeka under projects uniri-drustv-18-73 and uniri-drustv-18-140.

REFERENCES

[1] KPI Partners, “Traditional vs. Agile Software Development Methodologies”, 2023. Available: https://www.kpipartners.com/blog/traditional-vs-agile-software-development-methodologies [Accessed:15-Jan-2023]

[2] Geeksforgeeks, “Difference between Traditional and Agile Software Development”, 2023. Available: https://www.geeksforgeeks.org/difference-between-traditional-and-agile-software-development/ [Accessed:15-Jan-2023]

[3] B. Lutkevich, “Waterfall Model”, Available: https://www.techtarget.com/searchsoftwarequality/definition/waterfall-model [Accessed:18-March-2023]

[4] S. Al-Saqqa, S. Sawalha, & H. AbdelNabi, Agile software development: Methodologies and trends. International Journal of Interactive Mobile Technologies, 14(11), 2020.

[5] A. M. Gheorghe, I. D. Gheorghe, I. L. & Iatan, Agile Software Development. Informatica Economica, 24(2), 2020.

[6] Active Collab, “Advatages and Disadvatages of Agile Project Management”, 2017, Available: https://activecollab.com/blog/project-management/agile-project-management-advantages-disadvantages [Accessed:18-March-2023]

[7] A. Srivastava, S. Bhardwaj and S. Saraswat, "SCRUM model for agile methodology," 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 2017, pp. 864-869, doi: 10.1109/CCAA.2017.8229928.

[8] J. Noll, & S. Beecham, How agile is hybrid agile? an analysis of the helena data. In Product-Focused Software Process Improvement: 20th International Conference, PROFES 2019, Barcelona, Spain, November 27–29, 2019, Proceedings 20 (pp. 341-349).

[9] M. A. Jabar, S. Abdullah, Y. Y. Jusoh, S. Mohanarajah and N. M. Ali, "Adaptive and Dynamic Characteristics in Hybrid Agile Management Model for Software Development Project Success," 2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS), Johor Bahru, Malaysia, 2019, pp. 1-5, doi: 10.1109/ICRIIS48246.2019.9073337.

[10] N. Smits, “A Hybrid Software Development Method”, 2022 Available: https://www.devfacto.com/blog/a-hybrid-software-development-method [Accessed:18-March-2023]

[11] N. Yahya, & S. S. Maidin, The Waterfall Model with Agile Scrum as the Hybrid Agile Model for the Software Engineering Team. In *10th International Conference on Cyber and IT Service Management (CITSM)*, 2022, pp. 1-5.

[12] M. Marinho, J. Noll, I. Richardson & S. Beecham, Plan-driven approaches are alive and kicking in agile global software development. In *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2019, pp. 1-11.

[13] F. P. Zasa, A. Patrucco & E. Pellizzoni, Managing the hybrid organization: How can agile and traditional project management coexist?. *Research-Technology Management*, *64*(1), 2020, pp. 54-63.

[14] V. Garousi, G. Giray, E. Tuzun, C. Catal and M. Felderer, "Closing the Gap Between Software Engineering Education and Industrial Needs," in *IEEE Software*, vol. 37, no. 2, pp. 68-77, March-April 2020, doi: 10.1109/MS.2018.2880823.

[15] O. Cico, L. Jaccheri, A. Nguyen-Duc & H. Zhang, Exploring the intersection between software industry and Software Engineering education-A systematic mapping of Software Engineering Trends. *Journal of Systems and Software*, *172*, 110736, 2021.

[16] M. Olszewska, S. Ostroumov and M. Olszewski, "To Agile or not to Agile Students (With a Twist): Experience Report from a Student Project Course, *2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2017, pp. 83-87, doi: 10.1109/SEAA.2017.54.

[17] D. F. Rico and H. H. Sayani, "Use of Agile Methods in Software Engineering Education," *2009 Agile Conference*, 2009, pp. 174-179, doi: 10.1109/AGILE.2009.13.

[18] K. Fertalj, B. Milasinovic, B., & I. Nizetic, Problems and experiences with student projects based on real-world problems: a case study. *Technics Technologies Education Management*, *8*(1), 2013., pp.176-186.

[19] B. Bruegge, M. Reiss and J. Schiller, "Agile Principles in Academic Education: A Case Study," *2009 Sixth International Conference on Information Technology: New Generations*, 2009, pp. 1684-1686, doi: 10.1109/ITNG.2009.76.

[20] M. Madeja and M. Biňas, "Implementation of SCRUM Methodology in Programming Courses," *2018 16th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2018, pp. 333-340, doi: 10.1109/ICETA.2018.8572161.

Computer-supported Education of the Reserve Officers of the Armed Forces

Miro Čolić
Croatian Military Academy "Dr. Franjo Tuđman", Zagreb, Croatia
miro.colic@morh.hr, miro.colic@fer.hr

***Abstract* - The modern environment of warfare and the development of technology, which are changing faster than ever in the history of mankind, require of members of the armed forces to acquire completely new and constantly develop and reshape the existing competencies (knowledge, skills, independence and responsibility). Current models of competence acquisition, especially for reserve officers, which are carried out according to the old model in which "ex catedra" is almost the exclusive form of teaching, do not ensure the achievement of competences for the implementation of tasks in a modern environment. The paper will propose a new concept of education for reserve officers - computer-supported education. This new model implies a significantly greater application of modern education methodologies and can be adapted to different levels of education, lengths of time as well as forms of education, not only for reserve but also for active officers. In the paper, components of the education process are specified and described, with the fact that the preparation and implementation of the two-week final Computer Assisted Exercise (CAX), which is essential for achieving the desired learning outcomes, is specially elaborated.**

Keywords - planning, military operations, computer-supported education, learning outcomes, Computer Assisted Exercises - CAX, e-learning, teaching strategies.

I. INTRODUCTION

The fourth industrial revolution, the so-called digital revolution, which began at the beginning of the third millennium, will, according to everything, change society, the way that people live, and also the way of warfare more than the past three industrial revolutions. The cause of these changes is not technology per se, but the principles of modern warfare, some of which are: multidimensionality, unpredictability, surprise and comprehensiveness [1].

Although the end of the digital revolution and all its consequences on the conduct of military operations and warfare is not certain, it is already clear that the current competencies of individual members of the armed forces, as well as the military capabilities of units (Functional components - DOTMLPFI: Doctrine, Organization, Training, Materiel, Leadership Development, Personnel, Facilities, Interoperability) that they achieve through current education and training processes, should be changed to a large extent so that the military organization, as well as the security services, can carry out their tasks in accordance with the strategic defence and security documents of an individual state. Accelerating the digitization process of education and training of members of the armed forces, as well as members of other components of the defence and security system, is one of the possible answers to this problem. Drafting of the officers' competences of the armed forces of the countries: NATO 2011 [2], and the current drafting of the competences of military officers in the EU and the first five standards of military occupations in the Republic of Croatia published on December 21, 2021 showed once again that the digitization process is increasingly important and actually unavoidable in the process of reshaping education and training for the needs of the armed forces. The adoption of the National Security Strategy of the Republic of Croatia (*Official Gazette No.* 73/2017) [3] and several other documents from the field of defence in the Republic of Croatia have partially created the prerequisites for the formation of a high-quality and sustainable training system for the reserve component of the Croatian Armed Forces (CAF). Computer-supported education, unlike the current one, can use a whole series of "tools", for example: distance learning, e-learning, different simulation models (virtual and constructive simulations) [4], Geographic Information Systems - GIS [5], virtual and augmented reality [6] etc.

II. COMPUTER SUPPORTED EDUCATION CONCEPTS AND LEARNING OUTCOMES

The new model of education will be explained through the description of a two-month course for reserve officers of the armed forces - "Military Operations Planning Course" with the aim of acquiring competencies in the field of staff work and/or command. The course is organized in two parts (Figure 1): A - theoretical part (lasts six weeks) and B - practical part (lasts two weeks), totalling eight weeks. A - the theoretical part consists of a part that is carried out in the classroom (twice a week - the first and fourth week) and the second part that is carried out by e-learning (twice a fortnight - the second and third and the fifth and sixth week).

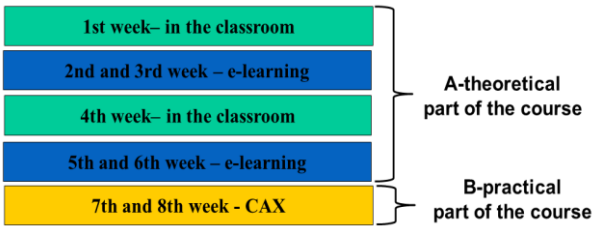


Figure 1. Course schedule by week

The practical part of the course is a Computer-Assisted Simulation Exercise - CAX and lasts two weeks (the seventh and eighth week).

The most significant learning outcomes that each student should achieve after completing the training are the following:

• state and analyse the reasons for new approaches to operations planning from the aspect of finding answers to poorly structured non-linear problems

• evaluate the interrelationships of the elements of national power in the approach to creating a response to a crisis situation

• plan a military operation with application of the prescribed methodology, current documents and procedures and the use of theoretical knowledge from operational skills

A. *Teaching strategies and methods*

The last listed learning outcome is the most comprehensive in terms of theory and, according to Bloom's taxonomy, is the highest level of the course itself.

Of the teaching strategies for achieving the given learning outcome, the teaching strategy, cooperative learning strategy and situational learning strategy would be applied [7]. Situational learning (situational cognition) [8] would be carried out through the context of problem solving activities. Collaborative learning [9], based on constructivist theory, implies the joint work of military trainees in a group working on the joint task of planning a military operation. At the same time, the overall success of the group depends on individual contribution of each member of the group (and vice versa). The course would be conducted by alternately engaging the participants through lessons in the classroom (campus-focused), online teaching (online focus) [10] and through preparation and implementation of the CAX exercise. Of special importance in this phase is the so-called - Brainstorming. During the last two weeks (CAX), collaborative [9] and situational learning [8] would prevail. In the seventh week, it is advisable to use the "Jigsaw" method of teaching (teaching that uses the puzzle technique), because all participants as a whole in that week (and the last week) would be in the role of command/headquarters of a battalion/brigade level unit. Therefore, the task that would be given to all together as battalion/brigade headquarters should be divided into components or functional units (battalion/brigade headquarters). When implementing a war game - during the creation of an operation concept and deciding which course of action (COA) to choose for a particular operation, artificial intelligence (AI) and Big Data technology can be used. Before that, participants will conduct a tour of the terrain where the planned operation will be carried out ("Command Reconnaissance") using real and virtual means (GIS, virtual reality, maps...).

III. COMPUTER ASSISTED EXERCISE – CAX

A. *Stages of CAX preparation and implementation*

1. Concept & Specification Development
2. Planning & Product Development
3. Execution
4. Post Exercise Analysis & Reporting

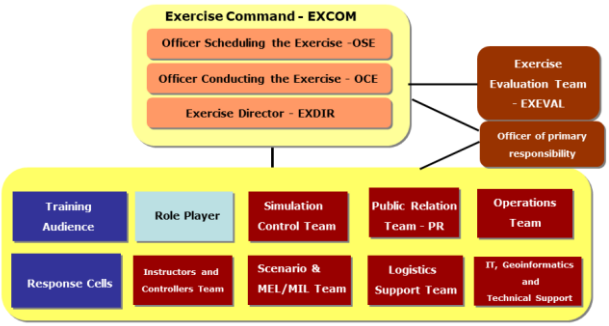


Figure 2. CAX organizational structure – concept

B. *Management of information systems in CAX*

In the case of preparation and implementation of the CAX system components from Figure 3 [11], they consist of:

Management

In the exercise (Figure 2), management is divided into "senior" and "operational" management. The management group (EXCOM) of the exercise belongs to the senior management (Manager, Commander and Director of the exercise, who are supported by the Officer of primary responsibility in communication with other teams), and the members of the individual teams from the exercise belong to the operational management. In order to understand the functioning of Management, it is good to distinguish the phases of exercise implementation: development of the specification and concept (Concept & Specification Development), planning and preparation of the exercise (Planning & Product Development), implementation (Execution) and analysis and reporting after the exercise (Post Exercise Analysis & Reporting).

Organization

Shown in Figure 2, it consists of the teachers and employees of the Simulation Centre and officers undergoing training. The components of the exercise are:

- **The Exercise Command** is the most responsible management or "senior" management in the exercise. It manages all teams and other components. The command of the exercise consists of: the Manager, the Commander, and the Director of the exercise, who are supported in communication with other teams by the Officer of primary responsibility.
- **Scenario and MEL/MIL** (Main Events List / Main Incidents List) **Team:** creates a scenario, i.e. incidents and events for the exercise, the implementation of which will ensure the achievement of the goals of the exercise, i.e. achievement of the learning outcomes of the education/training. In our case, it is a constructive simulation model JCATS. TABLE I shows the data that are components of the exercise scenario.

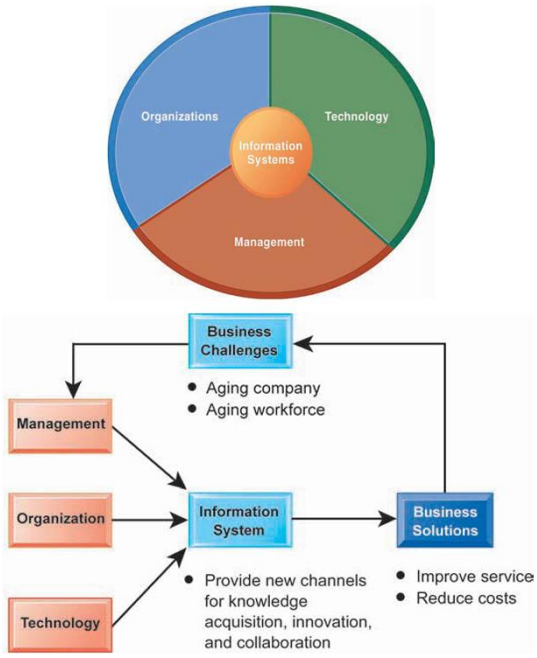


Figure 3. Information system, management, organization, technologies [11].

- **The Role Player** is responsible for making the exercise as realistic as possible. In accordance with the scenario, based on incidents and events that are created in advance and that they receive from the MEL/MIL team, using the JCATS system and simulating, in the best possible way, everything that happens in reality (civil organizations, police, "situation forces" international governmental and non-governmental organizations, media, protection and rescue, fire-fighters, etc.) they initiate events towards the training group and monitor actions of the training group.
- **Team of Instructors and Controllers** (IC Team) supervises the regularity of the training group as a whole and all components of the training group. This is done in coordination with the MEL/MIL team, the role player team, as well as the simulation systems monitoring team. In accordance with the elaborated scenario, incidents and events, before the start of the exercise, it prepares forms for evaluation or monitoring of the achieved results of the training group.
- **The Training Audience** is, as a rule, the participants of the training, in this case, the Operations Planning Course. They make up the headquarters of the battalion or brigade (25 to 30 trainees), and in the case of international and NATO exercises, the headquarters are often at higher levels. Often, the training group does not have the opportunity to see the simulation model at all but receives information from the simulation model (in this case, JCATS) from the operators of the JCATS system (Members of the Workstation Team), through information and communication systems.

TABLE I. COMPONENTS OF SCENARIOS AND MEL/MIL

Modules	Sub - Modules
Geo – Strategic situation	Map
	Historical background
	Political, military, economic, humanitarian, legal condition
Theatre of Operation	Map dataset
	Theatre data
	Country books
	Friendly and enemy order of battle
Strategic Initiation Module	Road to crisis
	Strategic Military assessments
	Initiating directive
	Strategic Planning Guidance
Crisis response planning information	Intelligence summary
	Friendly forces
	Civil/Military, environmental assessments
	Reconnaissance reports
Force activation and deployment information	Allied force list
	Force balancing results
	SOFA/MOU/TA
	Deployment plan
	Intelligence summary
	Joint Target list
	Rules of Engagement Authorization
Execution information	Road to war
	Intelligence summary
	Assessment reports
	Order of battle
	Situation reports
	Common operating pictures for C2 system
	MEL/MIL database

- **Response Cells -** this is the only team that directly interacts with the simulation model (in this case JCATS). So, the members of that team are at their workstations - computers where they observe what is happening on the simulation model and transfer to the simulation model the decisions made by themselves and/or the superior level, i.e., the headquarters of the brigade/battalion, or the Commander of the brigade/battalion.
- Based on the information provided, **The Simulation Team** creates databases before the start of the exercise (completes the scenario of the exercise and creates individual incidents or events in the exercise, etc.) and enters doctrinal models of behaviour (of all participating parties - it is not necessary to include only two so-called "red" parties and "blue", as a rule, there are several sides in the simulation model). In addition to the simulation model used in the preparation and implementation of the exercise, there is also GIS (Geographic Information Systems) support [5]. In addition, artificial intelligence (AI) can be used in the development of possible "versions of action" as well as in some other phases

of the preparation and implementation of the exercise [12].

- **PR** (Public Relations and Information), the team responsible for simulating the media: written media - press, television, radio, internet (social networks...), i.e., participates in the exercise as "Role Players".
- **EXEVAL** - The evaluation of the exercise is always carried out by persons from a higher hierarchy than the one that is the "training group" in the composition of the EXEVAL team and there are always different experts, depending on the topic/scenario on which the exercise is carried out. As a rule, EXEVAL teams also include experts who deal with technical support and simulation models (Figure 4: Transfer of information to CAX).
- Other teams are not specified because the name of the team itself suggests the task they have.

After implementation of the exercise, there is a (first) analysis after the exercise (First Impression Report - FIR). During the exercise, all communications (telephone, mobile phone, VoIP, internet, e-mail...) between the exercise participants are recorded/memorized and at the end of the exercise, what was planned, what happened during the exercise and how and why the differences arose are analysed.

Technology:

Here are listed only the most important components of the system that support the preparation, implementation, and analysis of the exercise (NATO concept).

- Joint Exercise Scenario Tool (JEST)
- Map Unit Builder (MUB)
- Database Development System (DDS)
- Terrain Modification Utility
- Joint Theater Level Simulation (JTLS)
- Joint Contact and Tactical Simulation (JCATS)
- Joint Multiresolution Model (JMRM)
- Web Hosted Interface Program (WHIP)
- Tool for operational planning, force activation and simulation (TOPFAS)
- Allied deployment and movement system (ADAMS)

- The simulation system on which the exercise is carried out is, as a rule, a constructive simulation model JCATS
- In addition to the above systems, the "e-learning" system ILIAS (or MERLIN in the Republic of Croatia) is also used in preparation for implementation and analysis.

Information system components

The exercise command or exercise director is the person who, to the greatest extent, from the beginning of the preparations for the exercise to the final analysis of the conducted exercise, creates the information management system during the preparation and implementation of the exercise and coordinates with all information systems that support the exercise.

- Four components of the information systems in CAX:

1. Information system for exercise planning and management. These tools can be used to automate the process, manage and share information for the preparation of exercise specification (EXPEC) and exercise plan (EXPLAN) documents. They can help prepare scenarios as well as major events and incidents (MEL/MIL - list). Through these interfaces, the data collected during the planning process of the exercise can be directly entered into the simulation, as well as into the command and control software (Command & Control - C2).

2. Constructive Simulation Systems (JCATS) and Support Tools: These are the simulation systems and software required to run the simulation, e.g. database preparation tools, user interfaces, etc.

3. The information system that connects C2, the operational planning system and the simulation system (JCATS) on which the exercise is conducted. The training group (training participants) has only the systems it would have in a "real situation", i.e. in the actual implementation of the operation.

4. Tools for experimentation and analysis: These are programs used to design and manage experiments using CAX data and to compile and present data collected by the simulation system as well as derive information from this data.

- The training group (training participants) can only use the second and third information systems. The training is focused on the headquarters, as a rule, of the battalion/brigade and the components of that unit.
- EXCOM and other teams in the exercise (which are not the "Training Group") can use all information systems, but they mainly use the first, second and fourth information systems and, as a rule, only monitor the third

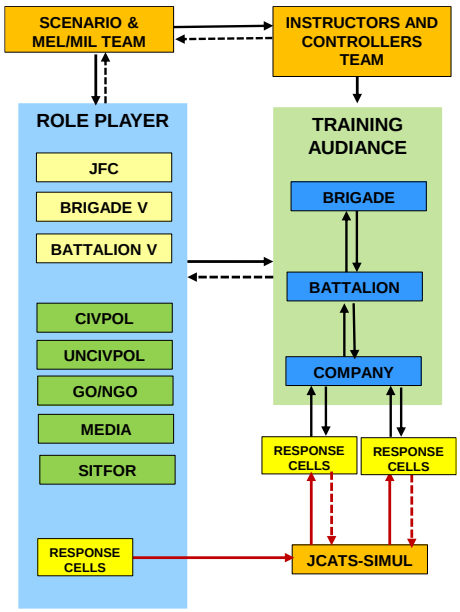


Figure 4. Transfer of information in CAX

system used by the "Training Group". Only the members of the Management Staff of the exercise have access to all information, and they, together with members of the other teams, manage and moderate the exercise.

Before the start of the training/CAX, the organizer of the exercise creates all essential elements for the exercise, so that the databases can be formed on time. All this needs to be done for the "opposite side" - that is, more "sides" if the scenario of the exercise foresees that. To carry out an exercise like this, you need a minimum of: 40 workstations for JCATS, 30 VoIP phones and 15 "Admin" computers.

- Management of the information systems in CAX

For successful management during the preparation, implementation and analysis of the exercise, it is necessary to design a joint information system for managing the CAX exercise (Joint CAX management system - JCMS), the basic components of the JCSM database are (Figure 5):

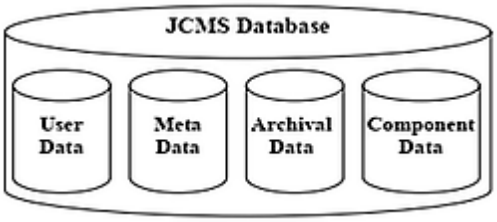


Figure 5. JCMS database

- User Data - defines users, user data and their user rights
- Meta Data - Data that define characteristics of the data available in the database in JCSM. For example, the name, type and size of the scenario that is stored in the database and which may correspond to creation of a new scenario.
- Archival Data - these are all possible data from the preparation, implementation and analysis of past exercises.
- Component Data - these are the data needed to implement the planning process of the exercise. These data are in generic form. They may include: field data, information about individual countries (political system, military, economy, culture, humanitarian and legal issues, significant persons from political life, etc.), events and incidents (MEL-MIL list) that were created in in accordance with the scenario, data on different types of units, weapon systems, logistics and supply systems, forms on all other data for the planning process of preparing the exercise.

IV. ASSESSMENT OF KNOWLEDGE

"If you want to change how students learn, change the methods of knowledge assessment, and keep in mind that the learning objectives, chosen teaching strategy, teaching methods and context determine which method and which tool is most suitable for knowledge assessment." [13]

Bearing in mind: the learning outcomes at the course level that the participants must achieve (p.2), the duration of the course, the structure of the participants (reserve officers, aged between 30 and 45), in the first six weeks to achieve the learning outcomes of the most important subjects in the course (tactics, operations planning, joint operations, crisis response operations,..) the emphasis would be on formative assessment, through: forums, "on line" discussions, short tests and essays and seminars. Teachers act as moderators of discussions, conduct critical reviews of test results, and choose the most adequate knowledge assessment methodology to achieve learning outcomes. In the last two weeks of the course, the summative assessment would be applied, which would rank the participants and provide a formal assessment of the quality of the learning outcomes. In the last two weeks during the implementation of the exercise - CAX teachers are mostly in the role of supervisors and those who manage the exercise.

V. CONCLUSION

A new concept of education for members of the armed forces is proposed in a brief description, which is based on computer-supported education with an emphasis on the preparation and implementation of a computer-supported simulation exercise - CAX, through which, at the end of the training, participants synthesize all the knowledge they have from planning operations and should achieve the given learning outcomes. In this form of training, computers are used in almost all phases of training implementation. This work proposes a modern approach, a new model, and methods of training reserve officers in the armed forces, which will help solve one of the most pressing problems in the field of defence and security in the member states of NATO and the EU in a better and faster way, which is education of the reserve components of the armed forces.

REFERENCES:

[1] NATO Science & Technology Organization, Office of the Chief Scientist, Science & Technology Trends 2020-2040, Exploring the S&T Edge, Brussels, Belgium, 2020. http:\www.sto.nato.int (accessed: August 22, 2022)

[2] National Defence Office of the Commande Canadian Defence Academy, Generic Officer, Professional Military Education, Reference Curiculum, Station Forces Kingston, ON K7K7B4, Canada, 2011.

[3] Hrvatski sabor, Strategija nacionalne sigurnosti Republike Hrvatske, Narodne novine, br.73/2017. https://narodnenovine.nn.hr/clanci/sluzbeni/2017_7_73_1772.html (accessed: November 16, 2022)

[4] J. E. Hannay, O. M. Mevassvik, A.Skjeltorp, K.Brathen, LVC Simulation for Land Operations Training, Norwegian Defence Research Establishment (FFI), P.O. Box 25, NO-2027 Kjeller, Norway, 2017.

[5] Sumari A.D.W, Ahmad A. S; The Application of Multiagent Collaborative Computation to Military Operation Planning and Execution, The Jurnal of Information Technology and Electrical Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia, 2009.

[6] B.Pollock, E.Winer, S.Gilbert, LVC interaction within a mixed reality training system, Iowa State University, Ames, IA, Julio de la Cruz, Army RDECOM – STTC, Orlando, FL, USA, 2015.

[7] CARNet, Priručnik, E-učitelj – suvremena nastava uz pomoć tehnologije, Zagreb, 2016. https://www.e-skole.hr/wp content/uploads/2016/12/Prirucnik_e-Ucitelj.pdf (accessed: November 10, 2022)

[8] J. Lave, E. Wenger, Situated Learning Legitimate Peripheral Participation, Cambridge University Press, NY, USA, 2008.

[9] M.Laal, Collaborative learning: what is it ?, Tehran University of Medical Sciences, Iran: file:///C:/Users/COC3/Downloads/science5.pdf (accessed: November 10, 2022)

[10] T. Anderson, The Theory and Practice of Online Learning, Athabasca University, Canada, 2011.

[11] K.C. Laudon, J.P. Laudon, Management Information Systems - Managing the Digital Firm (Thirteenth Edition), New York University, USA, 2014.

[12] M. Bistron, Z. Piotrowsk, Artificial Intelligence Applications in Military Systems and Their Influence on Sense of Security of Citizens, Faculty of Electronics, Military University of Technology, Warsaw, Poland, 2021.

[13] H. Douglas Brown, Teaching by Principles, Second Edition, San Francisco State University Pearson Education, USA, 2007.

Design, Development and Control of a Ball-on-Beam Control System Ising Industrial Equipment

N. Muškinja
University of Maribor, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia
nenad.muskinja@um.si

***Abstract* - This paper presents the design and implementation of a ball-on-beam system (BBS) using Mitsubishi industrial equipment. The system was implemented both mechanically and in simulations. An industrial servo motor with servo drive ensures the angle of the beam. An industrial PLC is used to control the system and an HMI display is used for the user interface. A dynamic mathematical model of the ball-on-beam system has been developed. The model was used in simulations to design the corresponding control algorithms and simulations of the designed algorithms were performed in the Matlab/Simulink environment. These algorithms were then implemented in the PLC, the hardware configured, and the on-screen HMI user interface created. We have measured the performance of the control on a real system and compared it with simulations. The device we have built allows us to control the position of the ball along the beam. Different control algorithms are available, and the controller parameters can be changed via the HMI interface. The developed system can be used as an alternative to commercial systems to learn control theory and at the same time to learn about industrial hardware and software.**

Keywords - ball on beam system, programmable logic controller, servo motor control, HMI

I. INTRODUCTION

A ball on beam system (BBS) is a type of control system that involves balancing a ball on a beam using control inputs. This system is often used as a teaching tool to demonstrate the principles of control theory and the use of feedback to stabilize a system. In BBS, the beam is typically mounted on a base that allows it to pivot or rotate. The ball is placed on top of the beam and is free to roll back and forth along its length. The goal of the control system is to maintain the ball in a stable position on the beam, despite external disturbances or perturbations that may cause the ball to move off balance. To achieve this, the control system typically includes sensors that measure the position of the ball on the beam, as well as actuators that can apply forces to the beam to move it and keep the ball balanced. These sensors and actuators are connected to a control algorithm, which processes the sensor data and generates control signals that are used to drive the actuators. The control algorithm in a BBS can be implemented in a variety of ways, such as using a programmable logic controller (PLC), a microcontroller, or a personal computer. The specific design of the control system will depend on the requirements of the application and the level of performance that is needed. Fig. 1 illustrates the ball on beam system.

Recent research and applications of ball on beam systems present a comparison between different types of controllers, validated by simulation and experimental results [2-6].

The aim this work is to develop and manufacture such a system using industrial equipment and to implement algorithms for stabilizing and controlling the position of the ball along the length of the beam. Most of the equipment used was manufactured by Mitsubishi. A servomotor is used to incline the beam. The system is controlled using an industrial Human-Machine Interface (HMI) and the entire control program is implemented on a Mitsubishi Programmable Logic Controller (PLC). A three-dimensional (3D) model created in Autodesk Fusion 360 environment is used for the mechanical design. More complex mechanical components are 3D printed. From the mechanical part, the differential equation of the system and the geometrical connections are derived. A mathematical model of the system is built using the equations in Matlab and Simulink and simulations are performed. Different control algorithms are developed based on the model. Using Mitsubishi MELSOFT GXWorks3 software, a complete program is implemented for measuring the value of the sensors and controlling the servomotor. The designed control algorithms are implemented within this program. A touch screen HMI user interface is designed to display the operating status and to allow user control.

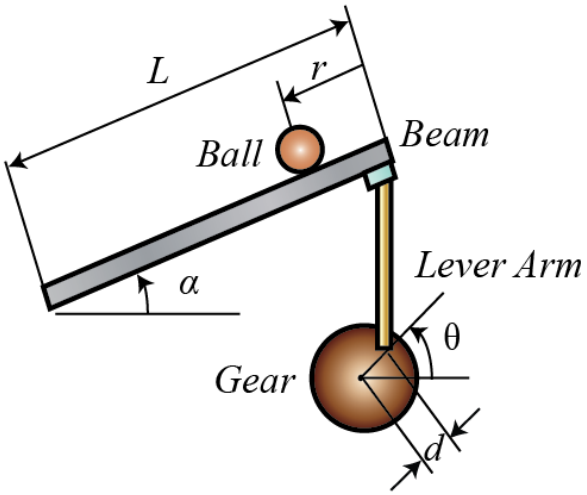


Figure 1. Ball on beam illustration [1].

INEA RBT d.o.o. and Mitsubishi Automation.

II. Components of the Ball on Beam System

For our system, we used components used in industrial applications as well as a linear diaphragm potentiometer and self-build components.

A. Servo Motor, Servo Controller, PLC and HMI

The beam was inclined using a servomotor. The servomotor was chosen to be an HG-KN- 13J motor [7]. Fig. 2 and Table I shows the servomotor and its technical characteristics. The selected servomotor can be controlled using a servo amplifier. We have chosen Mitsubishi's MELSERVO MR-JE-10C [8]. The controller allows different motor control modes: speed, position and also torque mode. It supports both analogue and pulse-width modulated (PWM) input for the desired motor speed via the CN3 connector. It also supports the possibility of control via CC_Link communication, which is done via an Ethernet connector. The amplifier parameters can be set using the MR Configurator2 software package. It can be connected to a PC via USB or Ethernet. The amplifier is shown in Fig. 3. We chose a Mitsubishi MELSEC iQ-F series controller called FX5UC-32MT/DSS-TS [9] to control the whole system. The controller has 16 inputs and 16 outputs. It allows communication via Ethernet or RS-485 bus. It shall have a connection point for a standard SD card to which data can be logged during operation. An overview of the controller is shown in Fig. 4.



Figure 2. Servomotor HG-KNJ [7].

TABLE I. Servomotor Technical Data

Nominal power	100 W
Nominal torque	0,32 Nm
Maximum torque	0,95 Nm
Nominal speed	3000 min^{-1}
Maximum speed	5000 min^{-1}
Nominal current	0,8 A
Maximum current	2,4 A
Incremental encoder	17-bit, 131072 pulses per revolution



Figure 3. Digital AC-Servo Amplifier MR-JE [8].



Figure 4. PLC FX5UC-32MT/DSS-TS [9].

TABLE II. PLC Analogue Expansion Cards

FX5-4DA-ADP	FX5; Analog Output Module; 4 channels
FX5-4AD-ADP	FX5; Analog Input Module; 4 channels



Figure 5. HMI interface [10].

Figure 6. SoftPot Membrane Potentiometer [11].

Table II shows the two analogue expansion cards that were also attached to the PLC.

We used the GT2107-WTSD HMI [10] to interact with the system (Fig. 5). The interface has a seven-inch color touchscreen. It allows communication via Ethernet, RS-232 and RS-422/485 connectors and can also be connected to a PC via a USB port. Graphical elements can be created in the GT Designer3 software environment to allow monitoring of the system operation and modification of global variables in the PLK program.

B. Sensors

To measure the position of the ball on the beam, we chose the SoftPot SEN-08681 linear diaphragm potentiometer [11]. The sensor has a measuring range of 500 mm and a measured resistance of 12 kΩ. The sensor consists of a lower resistive layer, an air gap and an upper conductive membrane which contacts the lower layer when pressure is applied at the ball location. It works in much the same way as an ordinary potentiometer, except that the ball acts as a slide. The voltage across the upper membrane thus represents the position of the ball. The sensor is shown in Fig. 6.

C. Self-buid Conponents

A 3D model of the mechanical system was drawn in Autodesk Fusion 360. The basic element of the beam is an aluminum U-profile. We have designed plastic bearing supports at each end of the profile. Bearings were used in the pivots to minimize friction in the system.



Figure 7. 3D model of the ball and beam mechanical system.

We have designed a plastic arm for incline the beam. We have adjusted the length of this so that when the system is in the equilibrium position, the pivot of the lever with the beam is directly above the servomotor axis. This made it easier to derive the mathematical relationship between the angle of the beam at and the angle of the servomotor. We also designed a railing to ensure that the ball stays on the beam and moves only within the measuring range of the position sensor. Fig. 7 shows a drawing of a beam mechanical system.

III. Mathematical Modelling

The mathematical model was designed from a 3D model of the system, where we could easily measure all the necessary dimensions. A sketch of the system is shown in Fig. 8. The motor angle ϑ varies between -90° and 90°, the horizontal position of the motor arm is 0°. The movement of the beam angle α is physically limited by the length of the motor arm b and the length of the beam L. The ball is located at a distance from the pivot of the beam r. When the motor increases the angle ϑ, the angle α increases at the same time, which causes the ball to roll downhill due to the gravitational acceleration and its potential energy is converted into kinetic energy. In our example, we have neglected frictional forces.

We used the Euler-Lagrange method of mathematical model derivation [12], [13]. First, we determined the kinetic energy of the ball. This consists of the translation of the ball and the rotation of the ball. J represents the moment of inertia of the ball and ω represents the angular velocity of the ball rotation about its own axis. The Lagrangian equation is as follows:

$$\mathcal{L} = K - U, \tag{1}$$

where K is the kinetic energy and U is the potential energy of the BBS. The total kinetic energy consists of the kinetic energy of the ball K_1 and the kinetic energy of the beam K_2.

The kinetic energy of the ball K_1 consists of translational and rotational kinetic energy:

$$K_1 = \frac{1}{2}mv^2 + \frac{1}{2}J\omega^2, \tag{2}$$

where m is mass of the ball and v is the velocity of the center of the ball.

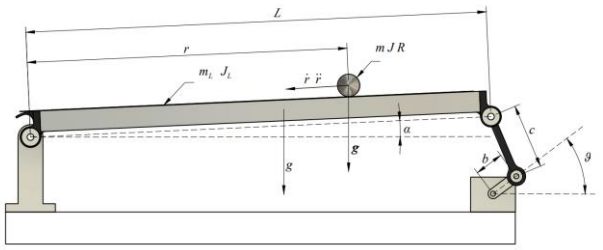


Figure 8. A sketch of the ball on beam system.

Assume that the ball is just rolling on the surface and not sliding. Thus, the angular velocity of the rotation of the ball can be written as:

$$\dot{r} = R\omega, \tag{3}$$

where R is radius of the ball.

Inserting (3) into (2), we obtain:

$$K_1 = \frac{1}{2}mv^2 + \frac{1}{2}J\left(\frac{\dot{r}}{R}\right)^2. \tag{4}$$

To calculate the translational kinetic energy, we need to calculate the velocity v of the ball. This is expressed in Cartesian coordinates:

$$v^2 = \dot{x}^2 + \dot{y}^2. \tag{5}$$

Converting them from radial to Cartesian coordinates gives:

$$\begin{aligned} x &= r \cdot \cos(\alpha) \\ y &= r \cdot \sin(\alpha) \end{aligned}. \tag{6}$$

After inserting (5) into (4) and further derivation, we get:

$$K_1 = \frac{1}{2}\left(\frac{J}{R^2} + m\right)\dot{r}^2 + \frac{1}{2}mr^2\dot{\alpha}^2. \tag{7}$$

The kinetic energy of the beam is:

$$K_2 = \frac{1}{2}J_L\dot{\alpha}^2, \tag{8}$$

where J_L represents the moment of inertia of the beam.

The potential energy of the ball is:

$$U_1 = mg\sin(\alpha)\, r, \tag{9}$$

and a potential energy of the beam is:

$$U_2 = \frac{1}{2}Lm_Lg\sin(\alpha), \tag{10}$$

where m_L is mass of the beam and g is a gravitational acceleration constant.

After summarizing all energies:

$$K = K_1 + K_2$$
$$U = U_1 + U_2, \quad (11)$$

and after solving Lagrange equation, we get final BBS mathematical model:

$$\left(m + \frac{J}{R^2}\right)\ddot{r} + mg\sin(\alpha) - mr\dot{\alpha}^2 = \tau, \quad (12)$$

where τ is the torque produced by the motor applied on the end of the beam.

Once we derived the differential equation of the system, we used this to design the mathematical model. The model was built in MATLAB's Simulink software package, which allows graphical system design. First, the relationships between the angles were determined using the physical constants of the system (Table III).

We have determined how the angle of the motor arm ϑ and the angle of the beam α depend on each other. Assume that the end of the beam moves only vertically, neglecting angular motion due to small angles. The value y_{off} represents the vertical distance between the beam pivot and the motor axis. In our system it is 0,078 m.

A sketch of the system (Fig. 9) shows the angle dependencies.

TABLE III. FIZICAL BBS CONSTANTS

Acceleration of gravity	g	9,81 $\frac{\text{m}}{\text{s}^2}$
Length of the beam	L	0,555m
Mass of the beam	m_L	0,2 kg
Moment of inertia of the beam	J_L	0,037 kg m^2
Motor arm length	b	0,036 m
Length of baem handle	c	0,078 m
Radius of the ball	R	0,0135 m
Mass of the ball	m	0,082 kg
Moment of inertia of the ball	J	5,96 kg m^2

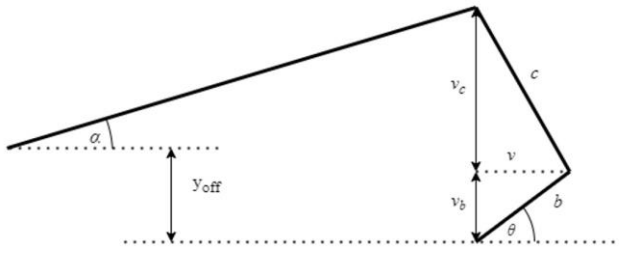


Figure 9. The BBS angle dependencies.

After some derivations and simplifications, we get:

$$\alpha = a\sin\left(\frac{h}{L}\right) = a\sin\left(\frac{b\sin(\vartheta) - y_{off} + \sqrt{c^2 - b^2\cos(\vartheta)^2}}{L}\right). \quad (13)$$

IV. DESIGN OF CONTROL ALGORITHMS

We used a cascaded proportional-integral-differential (PID) algorithm for control as proposed in [13]. PID is the most used control algorithm in industry because it is simple to implement, allows experimental parameter determination, and has good control characteristics. We used the PID-PD cascade structure as shown in Fig. 10. In the following, the name cascaded PID is used for this controller. For the determination of the controller parameters, we used the IMC-PID (Internal Model Control) procedure [14]. First, we need to measure the open-loop step response of the inner ball velocity loop to determine the process model, and then we can calculate the parameters of the inner PD ball velocity controller, as shown in Table IV. Fig. 11 shows the closed-loop response of the ball velocity control.

After the successful implementation of the internal PD velocity controller, an external PID position controller was added to the system. The input is the difference between the desired and actual ball position, and the output is the desired ball velocity. First, we need to measure the open-loop step response of the outer loop with the inner velocity loop closed. Then we can calculate the parameters of the external PID ball position controller as shown in Table V.

TABLE IV. VELOCITY PD CONTROLLER PARAMETERS

Process transfer function	K_p	T_D[sec]
$\frac{-0{,}48}{s} \cdot e^{-0{,}1 \cdot s}$	4,9	0,04

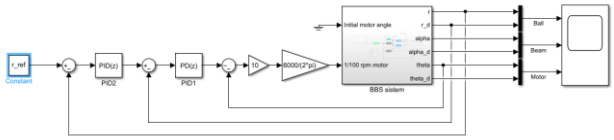


Figure 10. Cascaded PID BBS.

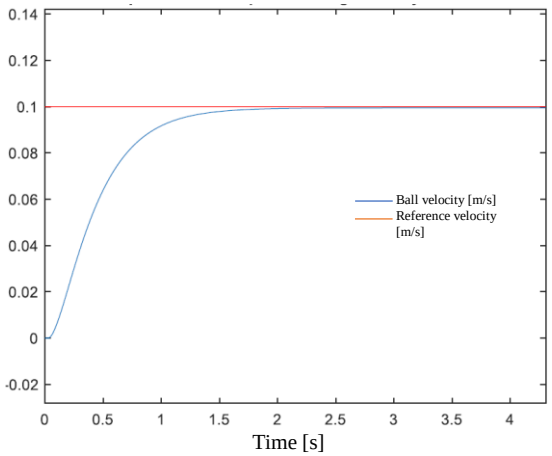


Figure 11. Closed-loop response of controlling the ball velocity.

TABLE V. POSITION PID CONTROLLER PARAMETERS

Process open loop transfer function	K_p	T_I[sec]	T_D[sec]
$\frac{1}{s} \cdot e^{-0,5 \cdot s}$	1,15	3,5	0,23

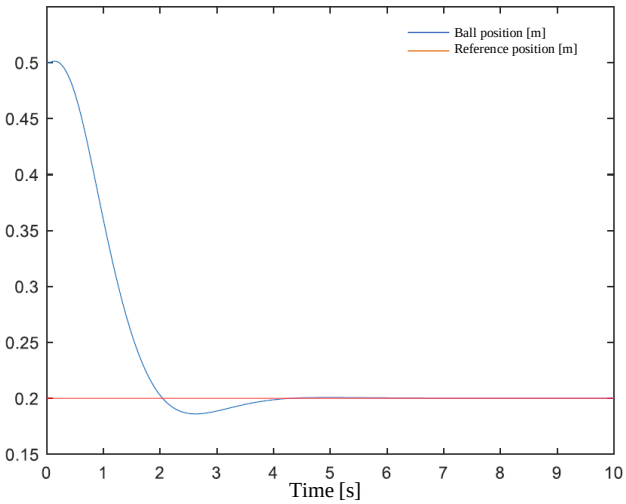


Figure 12. Closed-loop response of controlling the ball position.

Fig. 12 shows the closed-loop response of the ball position control.

In this section, we present the mathematical modelling of our BBS system and the design of the cascaded PID controller, as well as the simulation results of the ball position control. In the next section we will present the implementation of the algorithms in the PLK software and the real-time BBS experiments.

V. IMPLEMENTATION AND REAL-TIME EXPERIMENTS

A trolley made of extruded aluminum profiles was used to assemble all the elements. At the back of the trolley there are higher profiles where we mounted the electrical components. Most of the electrical components were mounted on DIN-rails, but the larger components were mounted directly on the mounting plate, for which we used a wooden furniture panel. Before starting to assemble the components, we created an approximate layout of the components in the Fusion 360 environment (Fig. 13).

When implementing the cascaded PID controller structure in the PLC, the same values of position and velocity PD controller parameters were used. Fig. 14 shows the response of the position control of the BBS ball.

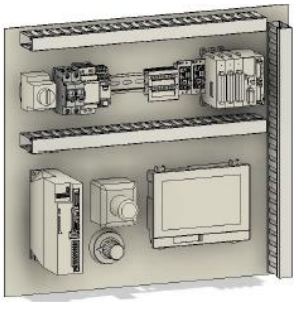


Figure 13. 3D model of the panel with electrical components.

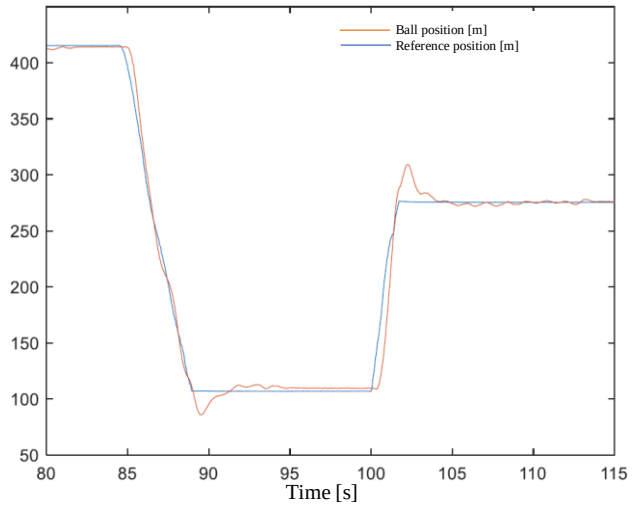


Figure 14. Closed-loop response of controlling the ball position on real BBC.



Figure 15. A real ball position control system.

VI. CONCLUSION

A ball-on-beam system was designed using industrial equipment (Fig. 15). The system was implemented both mechanically and programmatically. We designed the user interface for easy operation and at the same time added the possibility to change the control parameters. The developed system can be used as an alternative to commercial systems for learning control theory and at the same time for learning about industrial hardware and software.

The system successfully stabilizes the position of the ball using implemented algorithm. The position can be determined by using a potentiometer, by using a virtual slider and by using a changing reference that varies in the form of a PWM signal. The system is mounted on an aluminum trolley, which allows easy transport. It operates on its own local network, so there is no need for an external network router.

The Advanced Control System curriculum includes the study of complex control algorithms that are often abstract and difficult to understand, and many aspects of control of electronics. Therefore, in order to help students better understand them, it is necessary to implement laboratory exercises with real control systems and hardware implementation.

The practical implementation of a ball on a beam system involves a wide variety of integrated field such as:

- Control system components
  - Servo Motor and Servo drives
  - Potentiometer Sensor
  - Controller
- Feedback control systems
  - Reading Ball Position from Potentiometer Sensor
  - Derivative Filtering
  - System modeling
- System Identification Methods
  - Modeling of Actuator
  - Obtaining Transfer Function of Plant
- Real time measurements
  - Time Domain Characteristics
  - Steady State Response and Steady State Error
- Control system design
  - Time Response Characteristics
  - PID Controller and Fuzzy Logic Controller
  - Step Response and Steady-State Error
- Control system verification
  - Frequency Response Analysis
  - Experimental Bode Diagrams
  - Cut-Off Frequency Determination

Our ball on a beam system allows students to practice the theory of automatic control in a closed-loop experiment. Controlling the position of a ball on a beam is one of the classical problems of control theory. The ball and a beam system can be applied to a wide range of control system design implementations, from basic linear controllers to advance nonlinear methods. Students can understand approaches to system design using the Ball on a Beam method of learning by doing. They can use the curriculum to distinguish the effects of linearization's, assumptions, and modelling errors due to differences between simulations and real-world experiments.

REFERENCES

[1] "Control Tutorials for MATLAB and Simulink - Ball & Beam: System Modeling." https://ctms.engin.umich.edu/CTMS/index.php?example=BallBeam&section=SystemModeling (accessed Dec. 21, 2022).

[2] I. M. Mehedi, U. M. Al-Saggaf, R. Mansouri, and M. Bettayeb, "Two degrees of freedom fractional controller design: Application to the ball and beam system," *Measurement*, vol. 135, pp. 13–22, Mar. 2019, doi: 10.1016/j.measurement.2018.11.021.

[3] S. Zaare and M. R. Soltanpour, "The position control of the ball and beam system using state-disturbance observe-based adaptive fuzzy sliding mode control in presence of matched and mismatched uncertainties," *Mech. Syst. Signal Process.*, vol. 150, p. 107243, Mar. 2021, doi: 10.1016/j.ymssp.2020.107243.

[4] Z.-H. Pang, G. Zheng, and C.-X. Luo, "Augmented state estimation and LQR control for a ball and beam system," in *2011 6th IEEE Conference on Industrial Electronics and Applications*, Beijing, China, Jun. 2011, pp. 1328–1332. doi: 10.1109/ICIEA.2011.5975793.

[5] W. Yu, "Nonlinear PD Regulation for Ball and Beam System," *Int. J. Electr. Eng. Educ.*, vol. 46, no. 1, pp. 59–73, Jan. 2009, doi: 10.7227/IJEEE.46.1.5.

[6] S. Krafes, Z. Chalh, and A. Saka, "A Review on the Control of Second Order Underactuated Mechanical Systems," *Complexity*, vol. 2018, pp. 1–17, Dec. 2018, doi: 10.1155/2018/9573514.

[7] "HG-KN13J - Mitsubishi Electric Factory Automation." https://si.mitsubishielectric.com/fa/products/drv/servo/mr_je/rotary-servo-motor/hg-kn13j.html (accessed Dec. 21, 2022).

[8] "MR-JE-10C - Mitsubishi Electric Factory Automation." https://si.mitsubishielectric.com/fa/products/drv/servo/mr_je/servo-amplifier/mr-je-10c.html (accessed Dec. 21, 2022).

[9] "FX5UC-32MT/DSS-TS - Mitsubishi Electric Factory Automation." https://si.mitsubishielectric.com/fa/products/cnt/plc/plcf/cpu-module/fx5uc-32mt-dss-ts.html (accessed Dec. 21, 2022).

[10] "GT2107-WTSD | Human Machine Interface | HMI | Catalogue | Mitsubishi Electric Europe e-shop." https://mitsubishi-electric-eshop.mee.com/mee/FA_IA/en/EUR/Catalogue/HMI/Human-Machine-Interface/GT2107-WTSD/p/000000000000311489 (accessed Dec. 21, 2022).

[11] "SoftPot Membrane Potentiometer - 500mm - SEN-08681 - SparkFun Electronics." https://www.sparkfun.com/products/8681 (accessed Dec. 21, 2022).

[12] M. Keshmiri, A. F. Jahromi, A. Mohebbi, M. H. Amoozgar, and W.-F. Xie, "Modeling and Control of Ball and Beam System using Model Based and Non-Model Based Control Approaches," *Int. J. Smart Sens. Intell. Syst.*, vol. 5, no. 1, pp. 14–35, Jan. 2012, doi: 10.21307/ijssis-2017-468.

[13] P. V. M. Maalini, G. Prabhakar, and S. Selvaperumal, "Modelling and control of ball and beam system using PID controller," in *2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, Ramanathapuram, India, May 2016, pp. 322–326. doi: 10.1109/ICACCCT.2016.7831655.

[14] B. W. Bequette, *Process control: modeling, design, and simulation*. Upper Saddle River, N.J: Prentice Hall PTR, 2003.

Using the Matlab Programming Environment in the Course of Student Training in the Electrical Machines Discipline

V. S. Ruseva *, A. Hr. Krasteva **
* Angel Kanchev University of Ruse, Ruse, Bulgaria
** Angel Kanchev University of Ruse, Ruse, Bulgaria
e-mail address: vruseva@uni-ruse.bg; akrasteva@uni-ruse.bg

***Abstract* – This article presents the results of the use of simulation models in the MATLAB software environment in the course of student training in the Electrical Machines module, which is included in the curriculum for acquiring the Bachelor's educational – qualification degree in Electrical Power Engineering and Electrical Equipment. The processes for implementing innovative educational technologies gained momentum during the Covid-19 pandemic and are currently increasingly being adopted. The students, pursuing this course, use various computer-based tools such as MS Excel and MATLAB in their training to plot the performance and mechanical characteristics of a DC motor with parallel excitation. The article presents an approach for plotting these characteristics using a virtual laboratory bench and analyses the students' opinions on its use in the Electrical Machines module.**

***Keywords* – DC motor; Matlab; simulation.**

I. INTRODUCTION

The COVID - 19 pandemic presented us with new challenges along with a purpose to search for solutions regarding the performance of virtual laboratory exercises in the Electrical Machines module as part of the training of students enrolled in the degree programme of Electrical Power Engineering and Electrical Equipment. In the course of teaching the curriculum material, two methods of learning - synchronous and asynchronous - were employed [1]. Asynchronous communication is appropriate in cases where students prepare individually on a given topic, with the e-Learning Shell system (https://e-learning.uni-ruse.bg) being used for this purpose at Ruse University. The materials published in the system help students enhance their knowledge of the topics of the lectures, to practice and solve tests. Synchronous online communication at our university is conducted mainly through the use of the BigBlueButton system. Synchronous learning in the Electrical Machines module allows students to use the MATLAB programming environment to develop virtual models that simulate the operation of real laboratory benches with sufficient accuracy. The virtual laboratory benches provide a solution to the distance learning problem when students do not have the opportunity to work in a real laboratory setting. The virtual laboratory exercises are carried out using the same methodology as the actual experiments, in order to compare the results obtained.

There are numerous projects in which the processes in real machines are compared with the results obtained with the Matlab-synthesized model. By linking well-selected blocks and finished DC motor models in **MATLAB**, virtual test benches can be obtained for studying and analysing the performance of DC machines [2, 3, 4]. It is possible to create such benches for a specific motor whose parameters must be calculated in advance according to the methodology described in [4]. In [5, 6], the operating modes of a DC motor with independent excitation using a starting rheostat are modelled, while in [7], other methods for modelling such systems are analysed. In [8], a new topology of a power electronic circuit providing the operation of a series excitation DC motor in a regenerative braking mode is proposed. A refined SimPowerSystem model of a real motor operating in a Matlab/Simulink environment has been created.

Through virtual laboratory studies, data can be obtained to plot various characteristics when students do not have the opportunity to take the readings in a real laboratory setting, which is the case in online learning.

In this paper, the structure of the Electrical Machines module, included as a compulsory module on the curriculum of the students from the Electrical Power Engineering and Electrical Equipment programme leading towards a Bachelor of Science degree, is presented. The aim of the paper is, first, to present methodological guidelines for creating a virtual bench in the MATLAB programming environment that can simulate with sufficient accuracy the operation of a real DC motor with parallel excitation used in conducting face-to-face laboratory exercises of the module; and second, to identify an approach for plotting the operating characteristics through a virtual laboratory bench and to analyse the students' opinions on its use for online learning in the Electrical Machines module.

II. EXPLANATION

A. Aim of the Electrical Machines module, structure and expected learning outcomes

The aim of the training in the module is to provide students with theoretical knowledge of DC machines and transformers, and practical skills for testing them. The

module content covers the structure, operating principles, characteristics and applications of DC machines and transformers, as well as their test methods.

The learning outcomes achieved by the students in the Electrical Machines module include:

Knowledge: of the structure and principle of operation of electrical machines and transformers, the characteristics of generators, the operating and mechanical characteristics of DC motors, the parallel operation of single-phase and three-phase transformers, their modes of operation, and the methods of voltage regulation of transformers. Students must independently interpret the acquired knowledge, using it in the implementation of electrical circuits and the analysis of the results obtained.

Skills: proficiency in methods and tools available for solving complex problems; demonstrating innovative thinking and creativity in solving non-standard problems; ability to think logically in analysing the processes occurring in DC machines and transformers.

Competences: a personal approach in analysing the processes taking place in DC machines and transformers, ability to apply the acquired knowledge and skills in laboratory or production settings, constant improvement and learning, communication in performing creative work. In this way, students acquire professional and specific competences.

Training in the Electrical Machines module is conducted via lectures and laboratory exercises. Most of the laboratory exercises are organised in a cycle on the topics shown in Tab. I. The exercises are concluded with the preparation of an individual paper.

TABLE I. Topics of the laboratory exercises in the Electrical Machines module

№	***LABORATORY EXERCISES***	***Credit hours***
1.	Familiarisation with the rules of work in the Electrical Machines laboratory. Instruction in occupational safety. Familiarisation with the main parts of DC machines and transformers and the methods of testing electrical machines.	3
2.	Changing the excitation type of a DC machine, the direction of rotation of a DC motor and the polarity of a DC generator.	3
3.	Detecting faults in the armature winding of a DC machine.	3
4.	Plotting the magnetic field curve in the air gap of a DC machine.	3
5.	Drawing up the armature winding diagram of a DC machine.	3
6.	Studying the characteristics of a DC generator with separate excitation.	3
7.	Studying the characteristics of a DC generator with mixed excitation.	3
8.	Studying the characteristics of a DC motor with parallel excitation.	3
9.	Studying the characteristics of a DC motor with series excitation.	3
10.	Studying the characteristics of a DC tachogenerator.	3
11.	Studying the characteristics of a single-phase transformer.	3
12.	Studying the characteristics of a three-phase transformer.	3
13.	Determining the winding terminals and connection group of a three-phase transformer.	3
14.	Studying the characteristics of an autotransformer.	3
15.	Familiarisation with manufacturers' catalogues of DC machines and transformers.	3
	Total	**45**

B. Work procedure of studying a particular DC motor using a virtual bench in the MATLAB programming environment

Creating a virtual bench in the Matlab programming environment to study the operating characteristics of a particular DC motor with parallel excitation.

The operating characteristics of DC motors with parallel excitation are the functional dependencies of the rotational frequency n, the effective torque M_2, the current I consumption from the supply network, the power P_1 consumption and the efficiency η on the effective power P_2 delivered by the motor shaft, with constant values of the supply voltage and the excitation current, i.e. n, M_2, I, P_1, $\eta = f(P_2)$, U = const. and I_f = const. For the students to set up a virtual bench to study a DC motor with parallel excitation within a synchronous training session, it is necessary for them to:

- become familiar with the blocks and the parameters thereof, needed to build the virtual bench, from the **Simulink Library Browser/ SimPowerSystems** library;
- create a virtual stand to simulate the operation of a DC motor with parallel excitation in the **Matlab/Simulink** programming environment under the guidance of the lecturer;
- set up the model blocks;
- simulate the operation of the motor and plot the curves defined as the objective of the specific exercise they are performing.

The virtual bench is created used the Simulink library, which contains a large number of functional blocks. The required ones are selected and moved into the window of a newly created file. The necessary links are placed. By modifying the parameters of these blocks during the process of modelling, the behaviour of the system being tested can be observed.

From the laboratory exercise topics shown in Tab. 1, the students work on the topic: Studying the characteristics of a DC motor with parallel excitation.

Fig. 1 shows the schematic of a virtual bench created in the Matlab programming environment to study the operating characteristics of a particular DC motor with parallel excitation. The following appropriate basic blocks of the Matlab/Simulink programming environment which were selected, are presented in Tab.II.

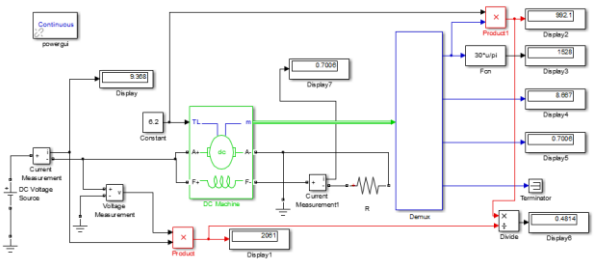


Figure 1. A shematic of a virtual bench to study the operation of a DC motor with parallel excitation

TABLE II. Blocks used in Figure 1

№	Names of the blocks	Purpose
1.	DC_Machine	A model of a DC machine
2.	Constant	For setting the shaft load (the moment of resistance)
3.	DC Voltage Source	A source of constant voltage
4.	Voltage Measurement	To measure the supply voltage value
5.	Current Measurement	To measure the value of the current I, consumed by the motor
6.	Current Measurement 1	To measure the value of the current through the field winding
7.	Product	A block for calculation of the power P_1 consumption
8.	Product 1	A block for calculation of the effective power P_2 delivered by the motor shaft
9.	Divide	A block for calculation of the efficiency η
10.	Fcn	A block for calculation of the rotational frequency n
11.	Display	A block for taking the readings of current I consumption from the supply network
12.	Display 1	A block for taking the readings of consumption effective power
13.	Display 2	A block for taking the readings of the effective power of the machine
14.	Display 3	A block for taking the readings of the rotational frequency
15.	Display 4	A block for taking the readings of the current Ia through the armature circuit
16.	Display 5	A block for taking the readings of the current through the excitation circuit
17.	Display 6	A block for taking the readings of the efficiency factor η
18.	Display 7	A block for taking the readings of the value of the current through the field winding

The block Product performs a calculation procedure according to the formula:

$$P_1 = U.I \quad (1)$$

where U is the supply voltage of the armature and field windings, and I is the motor current consumption.

The block Product 1 performs a calculation procedure according to the formula:

$$P_2 = M \,.\, \omega \quad (2)$$

where M is the moment of resistance applied to the motor, and ω is the motor angular velocity.

The block Divide performs a calculation procedure according to the formula:

$$\eta = \frac{P_2}{P_1} \quad (3)$$

where P_2 is the effective power of the machine;

P_1 – consumption effective power.

In the F_{cn} – block, a calculation procedure was performed using the following formula, which was entered in the Expression field:

$$n = \frac{30.\omega}{\pi} \quad (4)$$

The input value in the F_{cn} block was the angular velocity ω of the motor shaft, and the output value was the rotational frequency n.

All Display blocks are connected to block Demux. The resistance R (fig. 1) introduced in the excitation coil serves to regulate the excitation current.

Adjustment of the basic blocks of the model in Fig. 1.

The object of study was a DC motor with the following technical parameters: rated power P_n = 3 kW, rated voltage Un = 220 V and rotational frequency n = 1500 min^{-1}. In the dialogue box shown in Fig. 2, in the fields for setting the motor parameters, the following values had to be pre-entered: R_a - resistance of the armature winding; L_a- inductance of the armature winding; R_f and L_f - resistance and inductance of the field winding; L_{af} - mutual inductance factor; J - moment of inertia applied to the armature; B_m - friction coefficient; T_f - friction torque and initial rotational speed of the shaft [4].

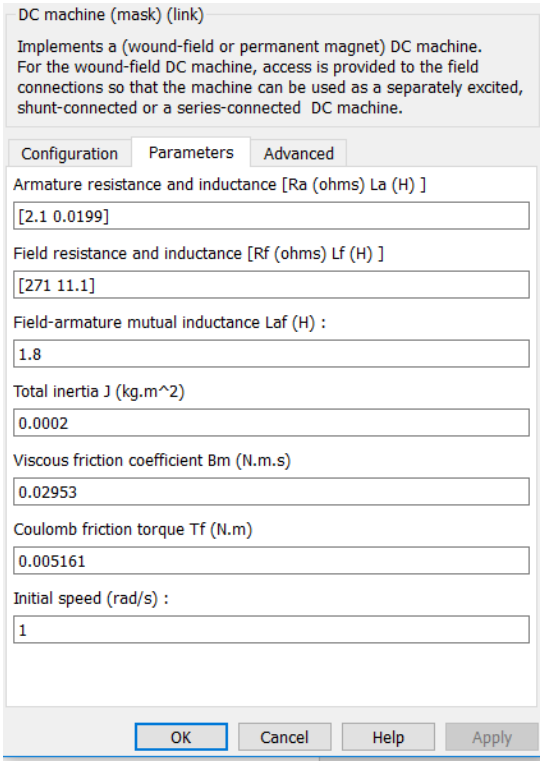


Figure 2. Dialogue box for setting the parameters of a DC motor

The resistance and inductance of the armature and field windings of the motor were pre-measured using a *Sverker* 760 - Fig. 3. This is an instrument designed to measure voltage, current, time, impedance, resistance, reactance, apparent power, active power, reactive power, phase angle and power factor. Its appearance can be seen in Figure 3.

The values of the resistance, the reactance and the other parameters were calculated according to the methodology given in [4].

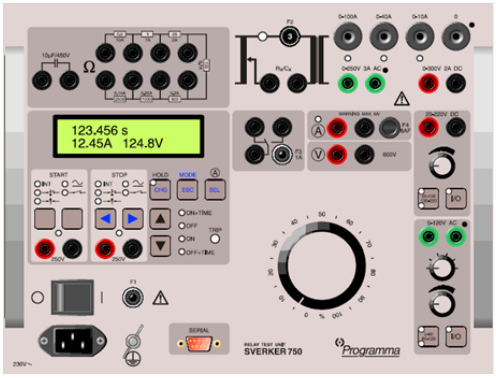


Figure 4. Front panel of *Sverker* 760

In addition to the DC Motor block, the DC Voltage Source, Demux, and Fsn blocks needed to be adjusted before simulating the operation in order to obtain data to plot the operating characteristics of the DC motor with parallel excitation.

In the Amplitude field of the DC Voltage Source block, the value of the rated supply voltage was set for the armature and field windings, respectively, which for the studied motor was Un = 220 V. In the dialogue box of the Demux block, the number of observed values was specified.

Simulating the operation of the DC motor with parallel excitation under study and collecting data to plot its operating characteristics.

Before the simulation, the process simulation time must be preset in the Simulation/ Parameters control panel in the Matlab/Simulink programming environment [2]. The simulations for the motor studied were conducted at a rated supply voltage of U=220V and rated excitation current value $I_f = 0,7$ A. The shaft load (moment of resistance) of the motor studied was entered in the **Constant value** field of the **Constant** block. Data for plotting various characteristics using the virtual bench in Fig. 1 could be obtained by setting values of the moment of resistance M in the range (0...1,1) M_n. For each value of the moment of resistance, a new simulation was run, and the data read from the **Constant** (for the moment of resistance) and Display3, Display, Display1, Displey6, Display2, blocks (for the rotational frequency n, the current consumption I from the supply network, the power consumption P_1 and the efficiency η on the effective power P_2 delivered by the motor shaft, respectively) were used to plot the operating characteristics.

The simulations were performed in such a way that the study corresponded to real experiments. Thus, it was possible to compare the results obtained from the simulation with those obtained in a laboratory setting. The obtained operating characteristics are presented in Fig.5.

As with real experiments, it is recommendable to monitor the anchor current in simulations in virtual experiments as well.

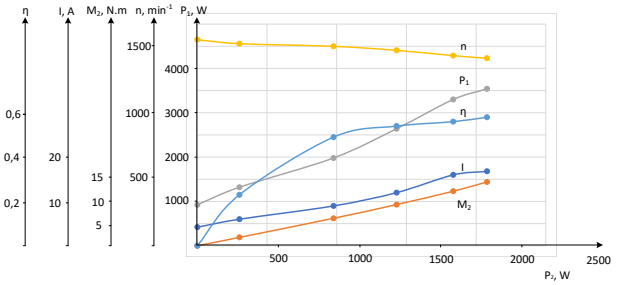


Figure 5. Operating characteristics of DC motor with parallel excitation, obtained via the virtual bench

C. Calculating the errors of the studied parameters obtained by the virtual bench presented in this paper and those obtained in real laboratory settings.

The errors were calculated at different unit load values (UL), shown in Tab. III. The unit load was calculated using the expression:

$$UL = \frac{P_2}{P_{2n}} . 100, \% \quad (5)$$

where the rated effective power of the machine is P_{2n} = 3 kW.

The relative error of the studied parameters was calculated by the formula:

$$\delta = \left|\frac{(X_2 - X_1)}{X_1}\right| . 100, \% \quad (6)$$

where X_1 is the magnitude obtained by analytical calculations and X_2 is the one obtained by simulating the operation of a DC motor with parallel excitation using the virtual bench.

TABLE III. Data about the relative errors obtained for the studied parameters at different values of the unit load

UL	δ_I	δ_n	δ_{P1}	δ_η	δ_{P2}
%	%	%	%	%	%
10	1,5	3,8	1,4	0,0	0,0
33	4,4	1,8	3,9	2,0	2,1
48	2,3	1,3	2,3	3,7	3,6
62	6,7	1,7	7,1	7,0	6,7
70	4,5	1,5	4,7	5,8	5,4

The greatest relative error in the conducted study – 7,1%, was obtained for the power consumption P_1 at a unit load of 62 %.

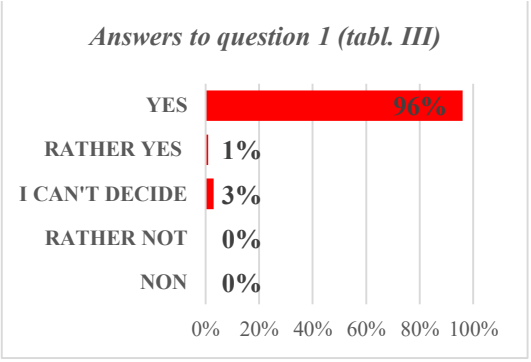
D. Students' opinion

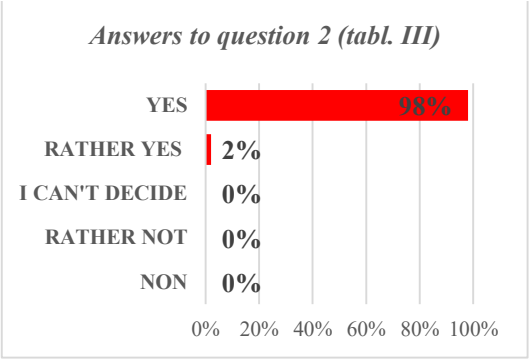
At the end of the semester in which the Electrical Machines module was taught, a survey was conducted among the students regarding synchronous learning and the use of the Matlab/Simulink programming environment in carrying out the laboratory exercises. The questions are presented in Table IV, and the answers to them are displayed in the same table. All of the students who were questioned were 100 in count.

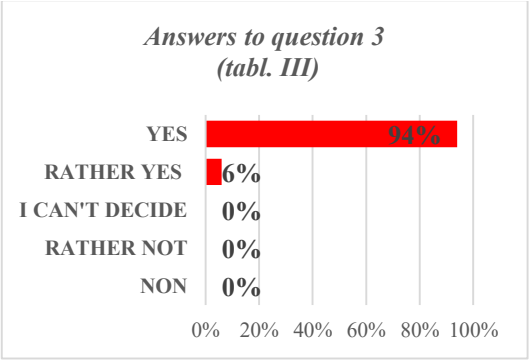
Table IV. Questions and answers for the survey conducted during the 2022 academic year

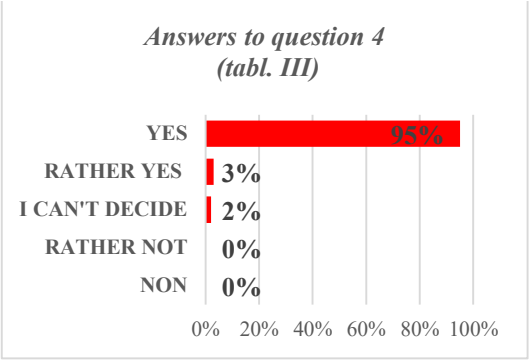
№	QUESTION	ANSWERS				
		NON	RATHER NOT	I CAN'T DECIDE	RATHER YES	YES
1	Is the content of the module up to date?	0	0	3	1	96
2	Has your interest in the subject of this module increased while you studied it?	0	0	0	2	98
3	Did you manage to create a virtual bench in the Matlab programming environment in the synchronous class?	0	0	0	6	94
4	Were you able to set up, under the guidance of the lecturer, the basic blocks of the virtual bench in an online learning format of the Electrical Machines module?	0	0	2	3	95
5	Did you master the method of calculating the resulting errors between the values of the studied parameters from the virtual bench and those obtained in real laboratory settings?	0	0	5	4	91
6	Do you evaluate in a positive way the use of the Matlab/Simulink programming environment in online learning for creating virtual benches, on which to simulate the operation of electrical machines, in cases where real laboratory settings cannot be used?	0	0	0	0	100

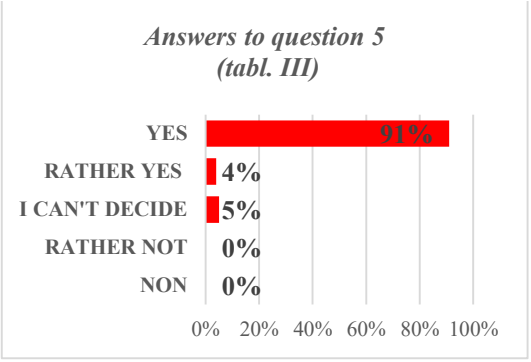
The analysis of the results obtained shows that more than 96% of the students identified the content of the module as relevant. Electrical machines are among the most ubiquitous in industry and the knowledge obtained about their structure, principle of operation and application is very important for future electrical engineers, such as our students.

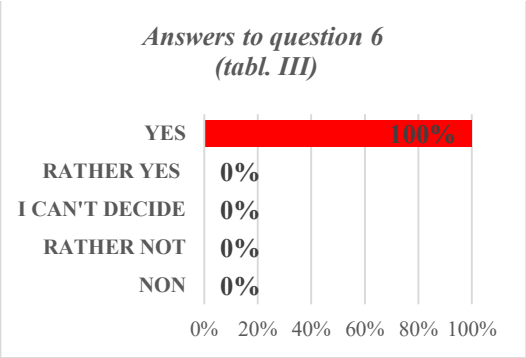












One hundred percent of the respondents answered that their interest in the subject had increased during their study, with 98% answering "yes" and 2% answering "rather yes". This, in our opinion, is associated with the way the curriculum material on Electrical Machines is taught using the Matlab/Simulink programming environment, a conclusion that is supported also by the answers to the sixth question in Table III, which is visualized above.

Again, 100% of the students approve of the use of Matlab/Simulink programming environment for online learning in cases where it is not possible to conduct face-to-face laboratory exercises.

The positive responses to question four: 95% "yes", 3% "rather yes", and 2% "do not know" indicate that the students received the necessary assistance from the lecturer during the online classes.

A survey was conducted amongst the teachers from Electrical Power Engineering and Electrical Equipment department, Angel Kanchev University of Ruse, regarding the value of the developed simulation models in the MATLAB software environment. All questioned teachers were unanimous in such models being useful for educational purposes, especially in times of global pandemic and distance learning.

III. Conclusion

This paper presents an approach used in synchronous training for conducting laboratory exercises in the Electrical Machines module for students from the Electrical Power Engineering and Electrical Equipment degree programme at the Angel Kanchev University of Ruse. The structure of the study module is analysed, as well as the knowledge, skills and competences that the students acquire.

For the purpose of the study, a virtual bench was set up in the Matlab programming environment to explore the operating characteristics of a particular DC motor with parallel excitation, which simulated with the required accuracy the operation of a real DC motor with parallel excitation, used for conducting in-person laboratory exercises in the Electrical Machines module. Methodological guidelines for carrying out the virtual exercise were summarized. The operation of the studied DC motor with parallel excitation was simulated and data were collected for plotting its operating characteristics. The relative errors for the studied parameters obtained by the virtual bench presented in the paper and those obtained in real laboratory settings were calculated. The performance viability of the model was proven.

The results of the survey conducted among the students who studied synchronously the discipline of Electrical Machines are presented. These results confirm the appropriateness of the applied training approach using the Matlab programming environment.

References

[1] Daniel, S.J. (2020). Education and the COVID-19 pandemic, *Prospects*, 49, 91–96. doi: 10.1007/s11125-020-09464-3.

[2] Hristova M., D. Dimov, A. Krasteva. Electrical drives - a manual for laboratory exercises, Printing facilities at Angel Kanchev University of Ruse, Ruse, 2013.

[3] Lurie S., O. Lurie., Simulation modelling in laboratory practice, course and diploma design. Textbook – Part Krasnoyarsk, 2005.

[4] Chernykh I. V., Simulation of electrical devices in MATLAB, SimPowerSystems and Simulink, DMK Press, Moscow 2008.

[5] Semenov A.S., V.M.Khubieva, M.N. Petrova Mathematical modelling of DC motor operation modes in the Matlab environment. Journal of Fundamental Research. - 2015. - No. 10 (Part 3) - pp. 523-528.

[6] Semenov A.S., Modeling of rheostat start of DC motor with independent excitation. Journal of Applied and Fundamental Research No. 9, 2014, pp. 29-34.

[7] Bebikhov Yu.V., A.S.Semenov, M.N.Semenova, I.A.Yakushev. Analysis of methods of simulation of technical systems in the Matlab environment. Modelling, optimisation and information technologies. Scientific Journal, Vol. 7, No. 3 http://moit.vivt.ru/, 2019.3

[8] Tsankov D., E. Yonchev, T. Yonkov. Recuperative impulse converter for DC motor with serial excitation. Annual of the Technical University, Sofia, vol. 66, book 2, 2016, pp. 61-71

Exploring Pre-scoring Clustering for Short Answer Grading

L. Petricioli, K. Skračić, J. Petrović and P. Pale
University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{Lucija.Petricioli, Kristian.Skracic, Juraj.Petrovic, Predrag.Pale}@fer.hr

***Abstract* – Automatic short answer grading is a topic that has gained significant popularity recently, especially due to developments in natural language processing. While automated grading in computer supported assessment tasks traditionally imposed significant restrictions on the answer format (e.g., multiple choice questions), automated short answer grading could enable assessment scalability with very few answer format limitations and thereby increase the assessment tasks' validity. Here, 'short answer' refers to a text of up to, approximately, 10 sentences. However, automatic solutions require a lot of pre-graded material. In this paper, several pre-trained machine learning models were utilized to explore pre-scoring clustering for short answer grading of text in Croatian. The aim of this approach is to shorten the process of manual short answer grading by clustering similar answers, facilitating the development of automatic grading solutions. The described approach was evaluated on a dataset containing graduate students' answers in Croatian to six questions related to cyber security topics. The obtained results are promising and show how increases in cluster purity, normalized mutual information, Rand index, and adjusted Rand index measures can be achieved by finetuning a pre-trained model.**

Keywords – automatic short answer grading; ASAG; semi-automated short answer scoring; short answer grading; short text; short answer; automatic grading; natural language processing

I. INTRODUCTION

The ability to assess the knowledge and skills of a large number of students in a remote, quick, and scalable way has been a topic of interest for quite some time [1]–[3], and it has gained additional attention since the outbreak of the COVID-19 pandemic, when a lot of formal education had to be abruptly shifted into online environments. Online assessments are already used in massive open online courses, various paid platforms, distance learning, hybrid courses that are a part of formal education, and many others. The importance of assessment in learning can hardly be overestimated since it can be used to assist learning through not only getting students to perform knowledge retrieval [4] or through enabling timely feedback on assessment outcomes [5] (formative assessment) but also through evaluating the students' achieved learning outcomes for course completion or certification (summative assessment).

Automatic assessment grading has been a topic of interest for decades already. Whether an assessment item can be automatically graded, however, depends on the assessment item type. Items with more constrained answer types [6], like multiple choice questions, are easy to grade automatically. However, their ecological validity and ability to assess higher learning outcomes are, at times, debated – some authors point out that multiple choice questions do not support the same kind of knowledge retention and synthesis as, for example, short answer questions do [7].

Short answer questions are questions that require a short response written in one's own words [8]. Answering a short answer question can, therefore, require understanding, whereas answering multiple-choice questions can be based on solely recognizing the correct answer [1], [2]. Unsurprisingly though, this question type requires much more effort to grade, not only manually but also automatically. Automatic Short Answer Grading (ASAG) is based on Natural Language Processing (NLP) and is certainly a difficult problem to solve. Even automation does not absolve the human graders of their task – namely, all automatic grading solutions require pre-graded content to extract features from. Machine Learning (ML) solutions are especially notorious for requiring large amounts of data [8]–[10].

Since manually scoring answers to short answer questions is time-consuming and exhausting, and fully automating short answer scoring a difficult NLP task that cannot really be performed in absence of any manual scoring, a possible reasonable trade-off could be found in semi-automation [1], [8]–[15]. Semi-automated short answer scoring techniques aim to reduce the time human graders need to grade short answers, either for directly grading exams or for annotation needed for further automatization [11].

This paper explores several pre-trained machine learning models and how finetuning one of them improves the results of pre-scoring clustering for grading short answers in Croatian. The goal of this is to reduce the time needed to score short text answers manually and the approach taken provides promising results. This paper is organized as follows: in section II, related semi-automated approaches to short answer grading are presented; section III describes the explored approach to pre-scoring short answers, the dataset and evaluation procedures used; section IV showcases the results, and section V presents the authors' conclusions.

II. Related work

Many of the approaches dealing with short answer scoring semi-automation are based on the same premise: lexically similar answers should likely be scored similarly [8], [11], [13]. The data sets used in each of the following methods are pre-graded to enable testing the various approaches.

Basu et al. [1] propose an approach in which they cluster answers based on a similarity metric, where each cluster is subdivided into subclusters, resulting in a cluster hierarchy. The authors' idea is that a grader can grade a cluster as a whole, thereby significantly reducing the time needed for manual grading. The subclusters are there to help with a finer grading granulation – if a subcluster does not really fulfil the criteria for the score of the cluster, the subcluster can be assigned a different grade easily. Furthermore, clustering answers could help teachers determine whether their students have a common misconception regarding a question, which can help them rectify it quickly [1], [8].

Horbach et al. [11] experiment with clustering answers, grading one answer per cluster and then propagating the grade throughout the cluster. According to their analyses, grading the answer closest to the centroid of a cluster and propagating the grade to the whole cluster yields good results. They can achieve an accuracy of 85% by manually grading just 40% of the answer data set.

Wolska et al. [15] take a slightly different approach to answer clustering – their clustering is aimed at speeding up answer grading by showing the graders similar answers one after the other. Their theory is that reducing "context switching" (jumping from grading correct answers to incorrect ones and back) reduces the time needed for grading all the student answers. Ultimately, they get mixed results. Pado et al. [13] present a very similar idea to the one in [15] and, consequently, get similar results.

Zesch et al. [14] are mostly interested in quicker annotation of training examples for automatic short answer scoring solutions based on ML. They test several clustering solutions and conclude clustering is best suited for very short answers (phrases of up to three words) that are subsequently easy to separate from other phrases. Longer answers can have a lot of lexical overlap but have completely different meaning, making them difficult to cluster in a meaningful manner.

Horbach et al. [9] switch to Active Learning (AL) as their method of choice. They adopt an iterative approach: the teacher is presented with a number of answers to grade. The resulting graded answers are used to train a sequential minimal optimization classifier which is then run on the test data set (pre-graded), and the results of this experiment are then evaluated. This sequence of steps can be repeated until the whole data set is graded or up to a certain threshold. The authors find this AL approach works well for *some* questions; they especially note that questions that have (at least some) clearly separable answer classes are good candidates.

Horbach et al. [8] return to clustering once again. This time they propose manual grading before and during clustering to achieve better results. Namely, the teacher needs to grade some answers before anything is clustered. The graded answers are then used for feature extraction – the authors argue this can reduce the noise in the features, thereby improving clustering results. They also "reuse" the graded answers during clustering in two ways: firstly, they use them as seeds for the clustering algorithm, and, secondly, they construct relational constraints between them (e.g., whether two answers should or should not be part of the same cluster). Finally, they continue with the centroid grade propagation approach from [11]. They show that their multifaceted approach yields promising results.

Mieskes et al. [12] propose an ensemble of automatic graders as the base for their solution. They automate the grading process with three graders; one based on random forests, one on decision trees, and one on support vector machines. Manual grading is employed in those cases where the automatic graders disagree. With this approach, the authors achieve a reduction of effort needed for manual grading of up to 75% in the case of answers with a binary class (correct/incorrect), and up to 40% for those answers that have a more complex grading scheme.

On the other hand, Tashu et al. [10] opt for locality sensitive hashing as a way of both speeding up and improving answer clustering attempts. They start with a small subset of graded answers that can be considered seeds for clustering. They then hash each new answer and calculate its distance to currently graded answers – the new answer is then assigned the grade of the answer it is closest to. If there are multiple graded answers at the same minimal distance to the new answer, their grades are averaged, and the resulting score is assigned to the new answer.

Generally, many of the authors test their approaches on data sets commonly used for Automatic Short Answer Grading (ASAG). Two of the most used data sets are the PowerGrading (PG) [1], [2], [8], [14] and the Automated Student Assessment Prize (ASAP) [2], [8], [10], [12], [14], [16] data set. However, some of the researchers note that the PG data set is lacking – it has very short answers that often consist of a single phrase and resemble fill-the-gap exercises [8], [14], making them perfectly suitable for clustering approaches, but providing little value to any purported innovative approaches tested solely on it. On the other hand, the ASAP data set has proved to be a clustering challenge with its high number of tokens (median around 48 and maximum of up to 66 per answer [2]) and considerable lexical variance.

III. Pre-scoring clustering approach

This section describes the data set used for the research presented in this paper, as well as the experimental setup and the explored approach.

A. Data set

The data set used in this research consists of six questions in the field of Cyber Security (CS) and their corresponding answers. The answers were provided by graduate students studying either electrical engineering or computing. There are 72 answers per question with an average number of tokens of around 18. All the answers were graded by a single grader and scored with 0-2 points or 0-3 points, depending on the question. The language of

both the questions and the answers is Croatian. The questions the students had to answer were the following (translated to English; the point range for each question is listed in brackets next to it):

Q1: What does a hacker do? (0-3)

Q2: What is cryptography? (0-2)

Q3: What is phishing? (0-3)

Q4: What is social engineering? (0-3)

Q5: What is the difference between phishing and spam (in the context of e-mail)? (0-3)

Q6: How would you check the credibility of information found on the Internet? (0-2)

It is important to note that the choice of language does limit the number of applicable approaches.

B. Experimental setup

In order to enable the calculation of distances between answers and, therefore, their subsequent clustering, their text needs to be vectorized (as naïve approaches like Levenshtein distance would not yield usable results on text of up to or approximately 10 sentences in length [17]). Three standard text vectorization approaches were used for benchmarking:

- **FastText**[1], an open-source library often used for learning text representations and classifiers. It offers pre-trained word vectors for 157 languages, one of which is Croatian. These vectors were used as a baseline in this experiment, since many existing approaches use static word vectors and n-grams [2], [8], [9], [11], [12], [14], [15].
- Two models from TensorFlowHub – a 12-layer model [2] (**XLM-R-12**) and a 24-layer model [3] (**XLM-R-24**). Both models were pre-trained on the Croatian language. These two models are based on Bidirectional Encoder Representations from Transformers (BERT) [18], specifically the RoBERTa architecture [19] used for cross-lingual representation learning [20].

In addition to those three models, the described XLM-R-12 was modified to attempt to achieve better results. This model was chosen for two reasons: the dataset available is relatively small, so using a larger model would most likely not yield better results, and, more importantly, a larger model would require significantly more resources for finetuning than the researchers had available.

Changes to the standard XLM-R-12 model included adding two more layers: a fully connected layer of 512 neurons with ReLU activation and 10% dropout to prevent overfitting, and a softmax output that predicts the grade of the answer using a one-hot encoding vector. Both were added to the pooled [CLS] token output. Five random answers per question were selected from the pre-graded data set for finetuning. Once the process was complete, the softmax output layer was stripped from the model, leaving the fully connected 512-neuron layer for vectorizing the answers.

Once all the answers were vectorized with the different methods, an elbow analysis was performed to work out how many clusters would yield the best distribution of the data. The analysis showed that six clusters would be appropriate.

Finally, the vectors were all clustered using scikit-learn's [21] spectral clustering implementation[4], with the number of clusters parameter set to six and the affinity parameter set to nearest neighbours clustering. K-means and mean shift were also explored as alternatives to spectral clustering, but with weaker results. In spectral clustering, data points are regarded as nodes of a connected graph and clusters are found by partitioning this graph based on its spectral decomposition. Word embeddings are compared using the cosine distance measure because the cosine distance can measure the semantic similarity of the words. By using the pooled embedding of the answer, the answers are effectively grouped based on their semantic closeness.

Once the various methods' vectors were clustered, the results were evaluated using several cluster quality measures. The comparisons between two clusterings are done between the graded data set and the clustered data sets. The grade that is given to an answer in the data set is seen as the "cluster" the answer is assigned to (albeit manually). The cluster quality measures used were:

- **Cluster purity.** Cluster purity measures how often the same class is part of the same cluster. In the case of this paper, it measures how often answers that have been manually given the same grade are part of the same cluster. The higher the percentage of answers of the same grade in a cluster, the purer the cluster.
- **Normalized Mutual Information.** Normalized Mutual Information [5] (NMI) is a normalized measure of how dependent two variables, or, in this case, clusters are. Logically, if two clusters coincide perfectly, they are completely dependent, and their NMI score is 1. If they are independent, their score is 0.
- **The Rand Index.** The Rand Index[6] (RI) measures the agreement between two clusterings. For each pair of elements in both clusterings, it calculates whether the two clusterings agree on whether the elements are part of the same cluster or a different cluster. If all the pairs are distributed across clusters in the same way in both clusterings, the

[1] https://fasttext.cc/
[2] https://tfhub.dev/jeongukjae/xlm_roberta_multi_cased_L-12_H-768_A-12/1
[3] https://tfhub.dev/jeongukjae/xlm_roberta_multi_cased_L-24_H-1024_A-16/1
[4] https://scikit-learn.org/stable/modules/generated/sklearn.cluster.SpectralClustering.html
[5] https://scikit-learn.org/stable/modules/generated/sklearn.metrics.normalized_mutual_info_score.html
[6] https://scikit-learn.org/stable/modules/generated/sklearn.metrics.rand_score.html#sklearn.metrics.rand_score

index equals 1. If the clusterings always disagree, the index is 0.

- **The Adjusted Rand Index.** The Adjusted Rand Index[7] (ARI) is essentially the RI that has been adjusted for chance. The RI does not account for those pairs of elements that have the same distribution in both clusterings by chance, whereas the ARI does. The implementation used in this paper has a range of –0.5 to 1, where negative numbers indicate disagreement, scores around 0 pairings purely by chance, and 1 complete agreement between the two clusterings on which pair of elements belongs to the same (or different) cluster.

IV. Results and Discussion

The obtained results are displayed in Table 1. The results have been grouped by question and sorted by cluster purity in a decreasing order. For each of the four selected measures a higher number means a better score. The best results are indicated with bold letters. In all cases, the approach based on the modified XLM-R-12 model yields the best results. The high purity of clusters obtained using the vectors constructed by the finetuned model means that, if a grader were shown answers grouped according to the results of the clustering, they would be shown very similar answers one after the other. It has been shown that such grouping can speed the process of grading up [13], [15]. Furthermore, very few answers need to be used for finetuning, which means the grader would not need to do much grading beforehand.

Interestingly, fastText representations mostly yield better results than the non-finetuned cross-lingual models, even though the models have a greater capability of encoding context. This could be a result of Croatian being a low resource language [21], [22], which makes it less likely the model has seen many examples that fall within the CS domain during its pre-training phase. Furthermore, the fairly consistent last place the XLM-R-24 takes may be due to the size of the model and the aforementioned scarcity of examples.

Table 1: Cluster quality scores for all methods used in this paper.

Answer	Model	Cluster purity	NMI	RI	ARI
Q1	XLM-R-12-modified	**0,647887324**	**0,550325943**	**0,741247485**	**0,432060567**
	FastText	0,563380282	0,489899417	0,687323944	0,319812449
	XLM-R-12	0,535211268	0,488588322	0,669617706	0,31468908
	XLM-R-24	0,507042254	0,470907847	0,655130785	0,30095057
Q2	XLM-R-12-modified	**0,732394366**	**0,624871623**	**0,798390342**	**0,559145192**
	FastText	0,577464789	0,48728793	0,675251509	0,327944656
	XLM-R-12	0,535211268	0,488588322	0,669617706	0,31468908
	XLM-R-24	0,507042254	0,470907847	0,655130785	0,30095057
Q3	XLM-R-12-modified	**0,774647887**	**0,607358355**	**0,773843058**	**0,561388136**
	FastText	0,563380282	0,489899417	0,687323944	0,319812449
	XLM-R-12	0,535211268	0,488588322	0,669617706	0,31468908
	XLM-R-24	0,507042254	0,470907847	0,655130785	0,30095057
Q4	XLM-R-12-modified	**0,746478873**	**0,652056433**	**0,826156942**	**0,60922951**
	FastText	0,577464789	0,48728793	0,675251509	0,327944656
	XLM-R-12	0,535211268	0,488588322	0,669617706	0,31468908
	XLM-R-24	0,507042254	0,470907847	0,655130785	0,30095057
Q5	XLM-R-12-modified	**0,774647887**	**0,731058037**	**0,889336016**	**0,743018779**
	FastText	0,591549296	0,499113707	0,687323944	0,347823003
	XLM-R-12	0,535211268	0,488588322	0,669617706	0,31468908
	XLM-R-24	0,507042254	0,470907847	0,655130785	0,30095057
Q6	XLM-R-12-modified	**0,690140845**	**0,591833147**	**0,756539235**	**0,481332382**
	XLM-R-24	0,61971831	0,494363828	0,686519115	0,382209113
	FastText	0,577464789	0,482160491	0,679275654	0,32754582
	XLM-R-12	0,535211268	0,425642942	0,641448692	0,279889356

[7] https://scikit-learn.org/stable/modules/generated/sklearn.metrics.adjusted_rand_score.html#sklearn.metrics.adjusted_rand_score

The results indicate that researchers have an important decision to make when constructing (semi-)automatic answer grading solutions. If it is important for the solution to be general and easily transferrable to a data set with different content, fastText and similar static encoding solutions are a better option. However, it is important to note it is unlikely such a solution will provide state-of-the-art results, since it cannot encode the subtleties of the specific content. On the other hand, if the goal is to develop a solution that would give better results, a finetuned cross-lingual model is the way to go. Naturally, this sort of solution is highly specific to the problem and cannot be used on different content (at least not meaningfully). This trade-off currently seems inescapable.

V. Conclusion

In this paper, pre-scoring clustering approaches based on several pre-trained cross-lingual representation models and an additionally modified one have been explored for short answer grading. All models were evaluated on a dataset containing textual answers in Croatian to six questions related to cyber security topics. The obtained results are promising and show improvements regarding the cluster purity, normalized mutual information, Rand index, and adjusted Rand index measures. The improvements are dependent on the properties of each question. The results also indicate there is a trade-off between the power of generalization and high-quality clusters (clusters that are pure), since better results are achieved by finetuning the model for a specific context.

References

[1] S. Basu, C. Jacobs, and L. Vanderwende, 'Powergrading: a Clustering Approach to Amplify Human Effort for Short Answer Grading', Trans. Assoc. Comput. Linguist., vol. 1, pp. 391–402, 2013.

[2] A. Horbach and T. Zesch, 'The Influence of Variance in Learner Answers on Automatic Content Scoring', Front. Educ., vol. 4, no. April, 2019.

[3] J. Lun, J. Zhu, Y. Tang, and M. Yang, 'Multiple data augmentation strategies for improving performance on automatic short answer scoring', AAAI 2020 - 34th AAAI Conf. Artif. Intell., pp. 13446–13453, 2020.

[4] J. D. Karpicke, 'Retrieval-Based Learning: Active Retrieval Promotes Meaningful Learning', Curr. Dir. Psychol. Sci., vol. 21, no. 3, pp. 157–163, 2012.

[5] B. Wisniewski, K. Zierer, and J. Hattie, 'The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research', Front. Psychol., vol. 10, no. January, pp. 1–14, 2020.

[6] K. Scalise and B. Gifford, 'Computer-based assessment in E-learning: A framework for constructing "intermediate constraint" questions and tasks for technology platforms', J. Technol. Learn. Assess., vol. 4, no. 6, pp. 3–44, 2006.

[7] R. C. Anderson and W. B. Biddle, 'On asking people questions about what they are reading', Psychol. Learn. Motiv. - Adv. Res. Theory, vol. 9, no. C, pp. 89–132, 1975.

[8] A. Horbach and M. Pinkal, 'Semi-supervised clustering for short answer scoring', Lr. 2018 - 11th Int. Conf. Lang. Resour. Eval., pp. 4065–4071, 2019.

[9] A. Horbach and A. Palmer, 'Investigating active learning for short-answer scoring', Proc. 11th Work. Innov. Use NLP Build. Educ. Appl. BEA 2016 2016 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL-HLT 2016, pp. 301–311, 2016.

[10] T. M. Tashu, D. Szabó, and T. Horváth, 'Reducing annotation effort in automatic essay evaluation using locality sensitive hashing', Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11528 LNCS, pp. 186–192, 2019.

[11] A. Horbach, A. Palmer, and M. Wolska, 'Finding a tradeoff between accuracy and Rater's workload in grading clustered short answers', Proc. 9th Int. Conf. Lang. Resour. Eval. Lr. 2014, pp. 588–595, 2014.

[12] M. Mieskes and U. Pado, 'Work Smart - Reducing Effort in Short-Answer Grading', Proc. 7th Work. {NLP} Comput. Assist. Lang. Learn., vol. 2018, no. Nlp4call, pp. 57–68, 2018.

[13] U. Pado and C. Kiefer, 'Short Answer Grading: When Sorting Helps and When it Doesn't', Proc. 4th Work. NLP Comput. Assist. Lang. Learn. NODALIDA, pp. 42–50, 2015.

[14] T. Zesch, M. Heilman, and A. Cahill, 'Reducing annotation efforts in supervised short answer scoring', 10th Work. Innov. Use NLP Build. Educ. Appl. BEA 2015 2015 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL-HLT 2015, pp. 124–132, 2015.

[15] M. Wolska, A. Horbach, and A. Palmer, 'Computer-assisted scoring of short responses: The efficiency of a clustering-based approach in a real-life task', Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8686, pp. 298–310, 2014.

[16] M. Heilman and N. Madnani, 'The impact of training data on automated short answer scoring performance', 10th Work. Innov. Use NLP Build. Educ. Appl. BEA 2015 2015 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL-HLT 2015, pp. 81–85, 2015.

[17] S. J. Greenhill, 'Levenshtein distances fail to identify language relationships accurately', Comput. Linguist., vol. 37, no. 4, pp. 689–698, 2011.

[18] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, 'BERT: Pre-training of deep bidirectional transformers for language understanding', NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf., vol. 1, no. Mlm, pp. 4171–4186, 2019.

[19] Y. Liu et al., 'RoBERTa: A Robustly Optimized BERT Pretraining Approach', no. 1, 2019.

[20] A. Conneau et al., 'Unsupervised cross-lingual representation learning at scale', Proc. Annu. Meet. Assoc. Comput. Linguist., pp. 8440–8451, 2020.

[21] F. Pedregosa et al., 'Scikit-learn: Machine learning in Python', J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

Identifying Higher Software Engineering Education's Design-Reality Gaps in Rural India

Paul Spiesberger*, Lukas Bürstmayr*, Raoul Vallon*, Thomas Grechenig*
*Research Group for Industrial Software, Technische Universität Wien, Vienna, Austria
paul.spiesberger, lukas.buerstmayr, raoul.vallon, thomas.grechenig
@inso.tuwien.ac.at

***Abstract*—Over the last decade one of the major challenges of the Indian higher education system has been a discrepancy between the skills of computer science graduates and the needs of industry. The majority of India's private higher education sector is composed of over 38,000 affiliated colleges, most of which are based in rural India. We selected a college in rural India as a representative case to analyze the local computer science curriculum's implementation using a Design-Reality Gap Analysis. Our examination considers both students' and professors' points of view by conducting interviews and observations. Key findings suggest that inadequate professors skills, improperly implemented teaching methods and a missing curriculum evaluation, among others, play a major role in the Indian computer science graduates' skill deficiency.**

***Index Terms*—Higher Education, Software Engineering Education, Computer Science Curriculum, Curricula, Design-Reality Gap Analysis, Rural Higher Education, ICT4D, India**

I. Introduction

Due to the increasing demand for technically skilled workers in the Information & Communication Technology (ICT) sector in India since the 1980s, Higher Education Institutions (HEIs) responded by increasing the number of graduates in engineering disciplines. In order to cope with this growth while decreasing the investments from the government into the Higher Education (HE) sector, several state governments and the government of India urged the private sector to establish semi-self-financed engineering HEIs [8, 12].

Although the number of engineering graduates exponentially grew from 1998 to 2008, the quality of education students received was insufficient, leading to a mismatch between the needs of industry and graduates' skills [8, 12]. In 2012, a survey of the World Bank confirmed this mismatch and pointed out that "*64% of employers are only somewhat satisfied or worse with the quality of engineering graduates' skills*" [3]. In 2018, the National Employability Report by Aspiring Minds examined a sample of more than 170,000 engineering students from more than 750 engineering colleges across multiple Indian states which graduated [20]. According to this report, 52.5% are not able to write functionally correct code and 37.7% are not able to write compilable code. Candidates of colleges located in lower populated cities are overall less employable than from colleges of higher populated cities, regardless of the examined job role [20]. In 2019, more than 90% of the HEIs in India were private, so-called affiliated colleges, which implement curricula designed by nationally recognized state universities, also referred to as affiliating universities [21, 23]. The affiliating universities define curricula, which are taught at affiliated colleges and conduct the standardized examinations at the end of each semester. An affiliated college is not allowed to adapt the given syllabi of the courses or parts of the curricula [22]. The implementation of the curricula and teaching methods are in the hands of the affiliated colleges. According to the Annual Report 2018-2019 by the University Grant Commission (UGC), 40,489 such colleges exist in India and 60.53% of these colleges are located in rural areas [23].

This paper aspires to contribute to the Sustainable Development Goal (SDG) 4.4 by providing insights into underlying problems of higher Software Engineering (SE) education in rural India [24]. The specific goal of this paper is to examine the previously described key characteristic of the private higher education sector in the context of rural India. In order to achieve the specific goal, this paper takes a closer look at Computer Science (CS) curricula that are implemented at Indian HEIs of the private higher education sector in order to determine whether and to what extent gaps exists between the implemented curricula and their intended design.

The following Research Question (RQ) is addressed:

RQ. What gaps exist between a designed CS curriculum and their implementations at Higher Education Institutions (HEIs) in rural India?

II. Context

The partnership between a South rural Indian college and a European university was formed after a mutual desire to cooperate in 2016.

A. South Indian Partner

The research has been carried out in collaboration with an affiliated college in South rural India located in Andhra Pradesh. The partner institution was chosen as it is located in one of the states in the bottom 25 percentile of employability percentage of graduates [20]. It is 1 out of over 80 affiliated colleges that implement the same curriculum designed by a nationally accredited state university. The students are examined by standardized tests developed by the affiliating university. The college offers five bachelor programs, including CS and was founded in 2015.

B. European Partner

The European partner is a research group and part of a major European university of technology. The university has a history in engineering disciplines and finds international recognition in research as well as in teaching. The research group involved generally focuses on various large-scale software engineering applications, but also seeks to provide a practically useful and competitive educational pattern for software engineering students at their local university.

III. Related Work

In the past Bass and Heeks conducted research to find gaps between the reality and the implementation of curricula by using the Design-Reality Gap Analysis (DRGA) [2, 11]. Masiero extended the DRGA by a diagnostic model to identify the root causes of design–reality gaps, which carries relevance for future investigations [17].

In India the graduates' skills do not match the industry's requirements, according to various studies over the last 15 years [3, 8, 15, 16, 18]. The following is a summarized overview of (software) engineering graduates skills that are categorized as missing or insufficient:

- Communication skills (oral & written) [1, 3, 10, 13, 16]
- Ability to analyze and solve problems [3, 10, 13, 16]
- Collaboration, teamwork and social skills [1, 3, 13, 16]
- Creativity [1, 3]
- Reliability [3, 16]
- Self-motivated [3, 16]
- Willingness to learn, take new directions [1, 3, 10]
- Technical skills [1, 3]
- Use of modern (software) tools [3, 13]
- Management skills [1]

Furthermore, there is evidence in literature that India is not the only country facing these problem as Duell illustrates in her work focusing on Indian, Indonesian and Thai universities [9].

IV. Methodology

The following research methods were used to collect data, which were subsequently used to answer the RQ defined in Section I. Research was conducted with professors as well as CS students of the South Indian partner institution in order to gather data about the actual implementation of the CS curriculum and the everyday life of professors and students.

A. Semi-Structured Interviews

The authors chose semi-structured interviews as an appropriate interview type because of the significant importance of establishing a personal contact to subjects as well as the exploratory character [19, p. 324]. As is outlined in Table I, five professors of the partner institution holding one or more courses as part of the implemented CS curriculum were interviewed. Four students of the 4th semester were interviewed to gather data about their everyday life and the challenges students can face. The students were chosen as they had recently completed a practice orientated programming courses from the curricula partly taught by the interviewed professors. The interviews took place in a face-to-face setting and were conducted with the help of an online conferencing tool. With the consent of the respective interview partner, the semi-structured interviews were recorded, transcribed and analyzed using a thematic analysis (see Section IV-C).

Table I
Interviewed persons, their role and level of experience

Pseudonym	Role	Sex	Experience
Frankie	Prof.	Female	$<$1 year
Glenn	Prof.	Male	1-3 years
Jordan	Prof.	Female	1-3 years
Morgan	Prof.	Male	$>$3 years
Riley	Prof.	Male	$>$3 years
Nuru	Student	Female	4th Semester
Rory	Student	Female	4th Semester
Sam	Student	Male	4th Semester
Taylor	Student	Male	4th Semester

B. Observations

The observations conducted are uncontrolled and participant. The uncontrolled setting is meant to ensure that professors and students are behaving as naturally as possible and therefore, existing problems are not covered by different behavior. The authors were participant observers by observing professors conducting a course or project review sessions as well as working with groups of students on their final project while revealing the role of the authors as researchers [19, p. 294].

Guidelines by Bortz and Döring as well as Saunders, Lewis, and Thornhill were taken into account when planning and conducting the observations [4, 19]. The observations took place in three different settings between March and June 2021:

1) Professors were observed teaching a class by recording a video call in which the professor was visible and audible in front of the whiteboard throughout the lesson. A regular course and a laboratory course were observed.
2) Two groups of four students each were observed working on their graduation project. The authors acted as assistant professors. To this purpose, regular "office" hours were held and a chat software workspace was set up where the group members could ask questions at any time.
3) Professors and all final year students were observed in project review sessions, in which the groups of students reported on the progress of their project and received feedback from the faculty.

An overview of the conducted observations is presented in Table II. The researchers conducted two observations as observers and two as direct participants. Both, the work with the groups and the project review sessions were not recorded but the authors took notes.

C. Thematic Analysis

Braun and Clarke outline a guide for conducting a thematic analysis and the method's 6 phases were consecutively imple-

Table II
OVERVIEW OF THE OBSERVATIONS, THEIR SETTING AND SCOPE

Observations	Conducted as	Scope
A	Observer	Curricular Course Sessions
B	Observer	Project Review Sessions
C	Participant	Assisting Final Project Group 1
D	Participant	Assisting Final Project Group 2

mented [5]. Statements and notes from interviews and observations were analyzed in terms of how the real implementation of the CS curriculum actually takes place as well as regarding respective problems and challenges a person or a group of persons expressed. The conclusions drawn from the thematic analysis regarding a student's or professor's reality were then used within the DRGA.

D. Design-Reality Gap Analysis

The DRGA framework was originally designed by Heeks and is used as a tool to analyze the deviation as well as the extent of deviation of the resulting reality from its planned design [11]. It can be applied to the evaluation of an implemented IT/IS system, for example, to e-government projects, but also to a curriculum or its modified version [2, 6, 7, 14]. Heeks adapted the DRGA and used the OPTIMISM model (Figure 1) to consider the dimensions of a curriculum or the modification of it.

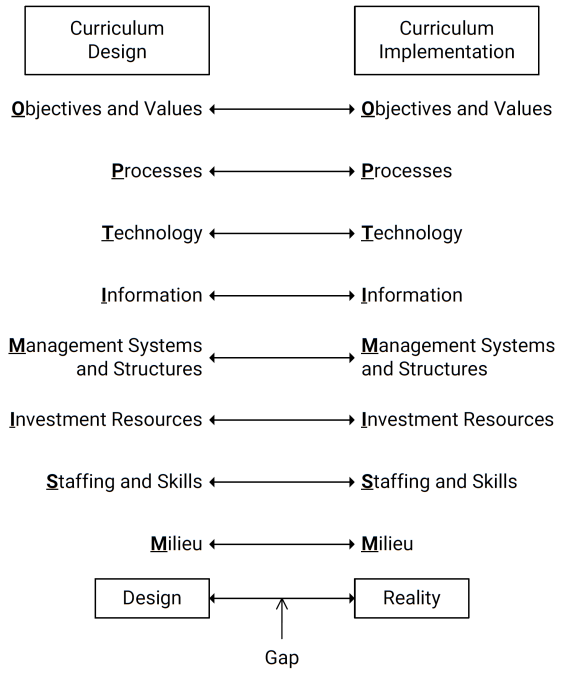


Figure 1. Adapted OPTIMISM model of Bass and Heeks [2, p. 6]

The following DRGA is structured according to these OPTIMISM dimensions. For each of these dimensions, the Design part is based on the CS curriculum itself and information that the affiliating university communicates through its websites. The majority of the curriculum consists of the course descriptions. The course description includes the "Course Objectives", "Course Outcomes", and the contents of each course unit. The Design part is complemented with the National Board of Accreditation (NBA) and National Assessment and Accreditation Council (NAAC) guidelines and criteria. The Reality part consists of the thematically analyzed interviews and observations of professors and students, as well as information that the partner institution communicates through its website.

The Design-Reality Gaps can be found in Section V. To illustrate the methodology, the following part IV-E exemplifies the DRGA in the dimension "Objectives and Values". The same approach was carried out for each dimension.

E. DRGA Example: Objectives and Values

This dimension considers the objectives of the curriculum and the values it wants to transport through its implementation.

1) Design: The NBA states that an institution must have a vision and a mission and a curriculum possesses multiple Program Outcomes (PO) and Program Specific Outcomes (PSO). The NAAC lists "Institutional Vision and Leadership" and "Institutional Values and Social Responsibilities" in its set of Key Indicators (KIs). The affiliating university and the partner institution are both communicating their individual vision and mission as well as the curriculum's POs and PSOs with the support of their websites. Regarding the POs and PSOs, all curricula offered on the affiliating university's website do not contain general or specific objectives, but they do explain within their methodology page which teaching methods they use to achieve certain objectives, such as preparing students to *"meet project requirements of the industry"* [source: website].

The CS curriculum defines objectives and learning outcomes for each course of the curriculum. In addition, each course is divided into individual units whose learning outcomes are listed.

2) Reality: The vision and mission of the partner institution are outlined on their website. They published a vision statement and mission statements for their curriculum implementation: *"Excellence in creating globally competent professionals and leaders in the field of computer science & engineering"*. They list PSOs for their curriculum implementation such as *"The ability to understand, analyze and develop computer programs in the areas related to system software, multimedia, web design, big data analytics, and networking for efficient design of computer-based systems of varying complexity"* [source: website].

It was found that courses and their objectives change almost every two years due to a revision by the affiliating university. According to professors, there is a tendency for courses that combine the basics with new topics to replace the existing courses. Professors' main objective is that as many students as possible pass the final exams. Professors fully trust the end exams by the affiliating university to be able to verify the actual learning success and those students who have passed all the exams are then also expected to achieve the goals, listed skills and learning outcomes of the curriculum. However,

some students feel that in this curriculum implementation they mainly study for the exams, but not to acquire practical knowledge. Students are very focused on becoming software developers or engineers.

Students and professors confirmed that the curriculum implementation is only useful in combination with internships. They confirmed that placements (future employment) are seen as a crucial part of the curriculum's implementation as they help students to enter the respective labour market in the first place.

3) Gap Analysis: As required by both accreditation organisations, the affiliating university formulates objectives and learning outcomes for the individual courses. The POs and PSOs are not listed in the curriculum, but the program-specific visions, missions and objectives are described analogously on the partner institution's website. The students' goal to become a software engineer aligns with the program-specific vision statement of the partner institution in terms of becoming a competent professional. The frequent changes to the curriculum and its courses are in line with the recommendations of the accreditation bodies regarding the continuous improvement of a curriculum.

V. Results: Design-Reality Gaps

As exemplified in Section IV-D, it was examined whether and to what extent the design of a CS curriculum deviates from the implementation in reality. The following section summarizes all discovered gaps and challenges. The authors noticed that there is thematically analyzed data which, strictly speaking, would not be taken into account by the DRGA, since no opposing design part existed to certain realities. These data, which therefore do not represent gaps between design and reality in the strict sense, are thus denoted "challenges" instead of gaps.

A. Gap 1 - Incomplete Student Skills

This gap results from the analysis of the differences between design and reality in the dimensions *Objectives and Values* and *Staffing and Skills*. The subject areas listed in the PSOs, in which students are supposed to be proficient at the end of their studies, correspond to the contents and objectives of the courses in the curriculum. Furthermore, it is of foremost importance to professors that as many students as possible pass the standardized final examinations and not whether the students actually achieve the course objectives and acquire the respective skills, which was also confirmed by students. There are indications that there is a big gap between the skills that students should have by design at the end of their studies and those that they actually have in reality. It was found that students did not possess certain skills, such as sufficient programming skills, that they should have acquired through several courses of the curriculum.

B. Gap 2 - Unmet Employability Goal

This gap is based on findings of the *Objectives and Values* dimension. It was found that the partner institution attempts to achieve the employability goal of the curriculum specifically through extra curricular courses. Therefore, the employability goal is not achieved through the curriculum. It is also noted that the resulting students' employability is limited to those companies that carry out the corresponding extra curricular course.

C. Gap 3 - Improperly Implemented Teaching Methods

This gap is based on findings of the *Processes* dimension. Although both the curriculum design and its implementation consider teaching methods such as Project-Based Learning (PBL), they are not properly implemented. This gap between PBL in theory and its actual implementation at the partner institution was found, for example in the graduation project course. It has been observed that both students and professors tend to copy a ready-made solution from somewhere else instead of developing their own solution to a problem. This implemented interpretation of PBL is not sufficient for students to actually achieve the intended learning objectives and acquire the respective skills that they should have according to the curriculum. Therefore, the actual implementation of the teaching methodologies constitutes a large gap between design and reality. From the point of view of the authors, the lack of experience of professors described in Gap 5 (Section V-E) is related to this Gap 3.

D. Gap 4 - Missing Curriculum Evaluation

The *Processes* as well as the *Management Systems and Structures* dimensions of the curriculum were examined with regard to the curriculum's evaluation. It was found that no processes exist to evaluate or revise the implementation of the curriculum. However, feedback is solicited from students in which they evaluate their professors and their abilities at the end of each semester. No management systems or structures were found that deal specifically with quality assurance and thus the evaluation and revision of curricula.

E. Gap 5 - Inadequate Professors Skills

Through the *Staffing and Skills* dimension it was found that professors lack both teaching knowledge and professional experience. Most professors have no professional experience before they start teaching, and some of them have no teaching experience either. In addition, professors are not informed by the curriculum or the partner institution about potential teaching methods and the teaching methods applied.

The lack of professional experience was particularly evident in those situations where the professors had to review students' projects. It was found that none of the professors were able to recognize even obvious errors or inconsistencies in presumably elaborated small software applications.

Furthermore, professors were not able to properly implement other teaching methods that were not teacher-centered. For example, it was found that professors in different courses provide students with solutions in advance instead of letting students work out the solutions themselves (see Section V-C).

F. Challenge 1 - Difficulties in Self-Learning

This challenge results from the *Technology* dimension: Students find it very difficult to work independently on curriculum-related topics, to study for them or to educate themselves. One reason for this is that only very few students have their own computer or laptop available. Another reason is the lack of space available for students. In addition, students are only allowed to stay outside the campus until a certain time in the evening, which means that potential places outside the campus are also only available to a limited extent.

G. Challenge 2 - Excessive Student Workload

According the findings of the *Investment Resources* dimension, another very significant challenge is the workload that students have. Due to the combination of curricular and extra-curricular courses, students are usually busy from early in the morning until 07:00 PM in the evening, with the exception of some smaller breaks and a longer lunch break, continuously for 6 days per week. This also leaves students relatively little time for self-study of content, even if a computer and a suitable study space were available to them.

H. Challenge 3 - Ever-Changing Course Assignments

This challenge concerns the constant reassignment of professors to courses that was found in the *Staffing and Skills* dimension. Professors get courses assigned according to the staffing needs of the faculty. However, it frequently happens that professors with no previous experience in the respective subject areas of a course have to teach this course by themselves. From the point of view of both the professors and the students, this situation constitutes a new challenge every term.

VI. Conclusion & Future Work

This paper investigates through a DRGA if and which differences exist between a curriculum designed by an affiliating university and its implementations at an affiliated college in rural India. Semi-structured interviews and observations were conducted with students and professors. The gathered data is thematically analyzed along the OPTIMISM dimension (see Figure 1) and rated with *small – medium – large* in Table III. The results then served as the reality part for the DRGA. The design is compared with the reality along each OPTIMISM dimension of the curriculum. The identified gaps were elaborated based on the dimension findings and summarized as follows:

- **Gap 1:** Incomplete student skills **(V-A)**
- **Gap 2:** Unmet employability goal **(V-B)**
- **Gap 3:** Improperly implemented teaching methods **(V-C)**
- **Gap 4:** Missing curriculum evaluation **(V-D)**
- **Gap 5:** Inadequate professors skills **(V-E)**

It was found that students and professors face challenges when implementing the curriculum that, strictly speaking, are not to be considered as gaps, because the curriculum does not contain design counterparts:

- **Challenge 1:** Difficulties in self-learning **(V-F)**
- **Challenge 2:** Excessive student workload **(V-G)**
- **Challenge 3:** Ever-changing course assignments **(V-H)**

Table III
Summary of Design-Reality Gaps with Rating

OPTIMISM Dimension	Rating	Comment
Objectives & Values	large	Gap 1, 2
Processes	large	Gap 3, 4
Technology	small/medium	Challenge 1
Information	–	–
M. Systems & Structures	medium/large	Gap 4
Investment Resources	medium	Challenge 2
Staffing & Skills	large	Gap 1, 5; Challenge 3
Milieu	–	–

Limitations:

1) This paper evaluates the implementations of a curriculum in relation to their underlying design. The curriculum's design and whether it is suitable for increasing the number of CS graduates with sufficient skills is not examined.
2) The results of this paper are based on one implementation of one curriculum design. This paper provides first insights, examining different implementations of the same curriculum or adapted versions of it are scheduled for future work.

A. Future work

The curriculum itself should be evaluated in further work to determine whether the structure and content are fundamentally suitable for equipping students with sufficient skills at the end of their studies. The Institute of Electrical and Electronic Engineers (IEEE) & Association for Computing Machinery (ACM) guidelines for CS curricula could be used in the evaluation of the curriculum's design and the skills to be acquired through the curriculum.

It was beyond the scope of this work to investigate the actual root causes of the particular gaps and challenges. For instance, it is not known whether the teaching methods are used incorrectly due to insufficient or against better knowledge. Future work will investigate the root causes of the respective design-reality gaps based on the work of Masiero to develop counter measurements to narrow down or even close the discovered gaps [17].

References

[1] Nurul Ezza Asyikin Mohamed Almi et al. "Software engineering education: The gap between industry's requirements and graduates' readiness". In: *ISCI 2011 - 2011 IEEE Symposium on Computers and Informatics*. 2011. DOI: 10.1109/ISCI.2011.5958974.

[2] Julian M. Bass and Richard Heeks. “Changing Computing Curricula in African Universities: Evaluating Progress and Challenges via Design-Reality Gap Analysis”. In: *The Electronic Journal of Information Systems in Developing Countries* (2011). DOI: 10.1002/j.1681-4835.2011.tb00341.x.

[3] Andreas Blom and Hiroshi Saeki. “Employability and Skill Sets of Newly Graduated Engineers in India: Policy Research Paper, The World Bank”. In: *IUP Journal of Soft Skills* (2012).

[4] Jürgen Bortz and Nicola Döring. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Vol. 3. Springer, 2016, pp. 157–165. DOI: 10.1007/978-3-642-41089-5.

[5] Virginia Braun and Victoria Clarke. “Using thematic analysis in psychology”. In: *Qualitative Research in Psychology* (2006). DOI: 10.1191/1478088706qp063oa.

[6] Hyeri Choi et al. “Rethinking the assessment of e-government implementation in developing countries from the perspective of the design–reality gap: Applications in the Indonesian e-procurement system”. In: *Telecommunications Policy* (2016). DOI: 10.1016/j.telpol.2016.03.002.

[7] S. I. Dasuki et al. “Evaluating the implementation of international computing curricular in African universities: A design-reality gap approach”. In: *International Journal of Education and Development using Information and Communication Technology (IJEDICT)* (2015).

[8] Amlendu Dubey et al. “Reforms in technical education sector: evidence from World Bank-assisted Technical Education Quality Improvement Programme in India”. In: *Higher Education* (2019). DOI: 10.1007/s10734-018-0343-1.

[9] Nicola Duell. “Skills shortages and mismatch in the ICT labour market: Challenges for skill development in India, Indonesia and Thailand”. In: *Digital Skills Insights 2019*. Ed. by Suella Hansen. CH-1211, Geneva, Switzerland, 2020, pp. 51–59. ISBN: 978-92-61-29331-4.

[10] Kirti Garg and Vasudeva Varma. “People issues relating to software engineering education and training in India”. In: *Proceedings of the 2008 1st India Software Engineering Conference, ISEC’08*. 2008. DOI: 10.1145/1342211.1342235.

[11] Richard Heeks. “Information systems and developing countries: Failure, success, and local improvisations”. In: *Information Society* (2002). DOI: 10.1080/01972240290075039.

[12] Ganesh Anant Hegde. “Privatization in higher education in India: A reflection of issues”. In: *A Global Perspective on Private Higher Education*. 2016. DOI: 10.1016/B978-0-08-100872-0.00009-4.

[13] Jeevamol Joy and V. G. Renumol. “Activity oriented teaching strategy for software engineering course: An experience report”. In: *Journal of Information Technology Education: Innovations in Practice* (2018). DOI: 10.28945/4116.

[14] Lemma Lessa, Solomon Negash, and Mesfin Belachew. “Steering e-government projects from failure to success: Using design-reality gap analysis as a mid-implementation assessment tool”. In: *International Business: Concepts, Methodologies, Tools, and Applications*. 2016. DOI: 10.4018/978-1-4666-9814-7.ch086.

[15] Rupa Mahanti and P. K. Mahanti. “Software engineering education from indian perspective”. In: *Proceedings - 18th Conference on Software Engineering Education and Training, CSEE and T 2005*. 2005. DOI: 10.1109/CSEET.2005.30.

[16] K. S. Mann, M. Narendra Kumar, and Harvinder Singh Saini. “The major challenges for Engineering institutions to fulfill industry demands in India”. In: *Proceedings of the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education, MITE 2015*. 2016. DOI: 10.1109/MITE.2015.7375282.

[17] Silvia Masiero. “The Origins of Failure: Seeking the Causes of Design–Reality Gaps”. In: *Information Technology for Development* (2016). DOI: 10.1080/02681102.2016.1143346.

[18] S. Pavai Madheswari and S. D. Uma Mageswari. “Changing paradigms of engineering education - An Indian perspective”. In: *Procedia Computer Science*. 2020. DOI: 10.1016/j.procs.2020.05.034.

[19] Mark Saunders, Philip Lewis, and Adrian Thornhill. *Research Methods for Business Students 5th Ed.* 2008. DOI: 10.1007/s13398-014-0173-7.2.

Web References

[20] Varun Aggarwal, Siddharth Nithyanand, and Malvika Sharma. “National Employability Report - Engineers”. In: (2019). URL: https://www.aspiringminds.com/research-reports/national-employability-report-for-engineers-2019/ (visited on 05/30/2020).

[21] Bureau of Planning, Monitoring and Statistics, Department of Higher Education, Ministry of Education, Government of India. *All Indian Survey on Higher Education Instruction Manual*. 2020. URL: http://aishe.nic.in/aishe/viewDocument.action?documentId=271 (visited on 12/06/2020).

[22] Centre for Civil Society. *Policy Brief: Regulatory Structure of Higher Education in India*. URL: https://ccs.in/sites/default/files/research/policy-brief-regulatory-structure-higher-education.pdf (visited on 11/08/2020).

[23] University Grants Commission. *Annual Report 2018-2019*. New Delhi, India. URL: https://www.ugc.ac.in/page/Annual-Report.aspx (visited on 11/08/2020).

[24] United Nations. *Sustainable Development Goal 4. Target 4*. URL: https://sdgs.un.org/goals/goal4 (visited on 01/10/2023).

Understanding the Teamwork Challenges of Software Engineering Students

L. Dorić*, N. Luburić*, J.Slivka* and A. Kovačević*
*Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
{luka.doric, nikola.luburic, slivkaje, kocha78}@uns.ac.rs

***Abstract* - Developing collaborative skills in students is nontrivial. The fact that students work in teams does not mean they become skilled in teamwork. Students face varied challenges when working in teams that harm their skill development and attitude towards teamwork. To prepare students for the collaboration-intensive workplace, we researched and designed a catalog of challenges present in the teamwork of undergraduate software engineering students on 3-month projects.**

We created an initial catalog of 10 challenges by examining the literature, surveying 15 teaching assistants, and coding their opinions regarding the problems faced by student teams. Using the catalog, we crafted a survey for students nearing the end of their team project to assess which challenges were present in their teamwork. We surveyed students from multiple contexts, including teams of 3, teams of 4, and teams of 16 students.

We analyzed 155 answers to determine the prevalence and intensity of the 10 challenges in student teams. We discuss our findings and best practices for resolving the most prevalent challenges. The catalog and recommendations are directly valuable for software engineering educators and can inform the broader community of collaborative learning researchers and instructional designers.

***Keywords* - teamwork challenges, collaboration, computing education, software engineering, higher education**

I. INTRODUCTION

The need for collaboration is a natural consequence of the growing complexity of our society, where modern projects surpass the individual's capabilities [1][2]. Collaboration skills, such as communication, team task management, and solution brainstorming, are crucial aspects of most contemporary professions [3]. Software engineering is one such profession [4][5]. Software engineering involves code reviews, software design brainstorming, and meetings with engineers, managers, and clients. Consequently, the *ACM Computing Curricula* [6] and more general learning engineering literature [7-9] emphasize the need to develop collaborative skills in young software engineering professionals.

Many software engineering programs require students to develop software in teams [10] with the goal of developing their collaborative skills. Developing collaborative skills in students can be challenging [11][12], and the fact that students work in teams does not mean that they become skilled in teamwork. These challenges prevent teamwork skill development and degrade students' attitudes towards them [13-16].

The first step in solving this problem is understanding the challenges students face when working in teams. In this paper, we research and define a catalog of challenges present in the teamwork of software engineering students on projects that last 3 months. We frame our methodology using the empirical cycle of design science [17]. First, we surveyed 15 teaching assistants using open-ended questions to collect their opinions on the problems in student teamwork. We coded their responses to create quantitative data and combined the findings with related work on teamwork challenges [13][14] to create an initial challenge catalog. Using the catalog, we created a survey with 25 questions and distributed it to 155 students. The students came from three different contexts, including (1) teams of 3 students, (2) teams of 4 students, and (3) teams of 16 students. Finally, we analyzed the results to determine the least and most prevalent challenges.

Section 2 lists the related work. Section 3 explains our methodology in detail. Section 4 presents the results. Section 5 discusses the findings and limitations of our study. Section 6 summarizes our conclusions.

II. RELATED WORK

Much research has been conducted examining what challenges students encounter in teamwork [13][15][16].

Kazemitabar et al. [13] examined how participants collaborate in teams for two days in the context of a hackathon. The goal of this hackathon was for each team to develop a new computer program to demonstrate a physics phenomenon of their choice artistically. The participants were undergraduate or graduate students with expertise in physics, computer science, software, electrical, mechanical, or civil engineering. Teams consisted of 2 to 4 participants. Their results revealed 16 general challenges that hamper teamwork in a hackathon.

Järvenoja et al. [15] explored what type of challenges trigger team-level emotion regulation. First-year teacher education students attended a six-week course on mathematics and worked on six collaborative mathematical tasks in three-to-four-member teams. The

This research was supported by the Science Fund of the Republic of Serbia, Grant No 6521051, AI-Clean CaDET.

results showed that in collaborative learning situations, a wide range of challenges emerge, including cognitive, motivational, and emotional issues and different socially- and contextually-oriented challenges.

Piia et al. [16] explored cognitive, motivational, and socio-emotional challenges experienced in collaborative learning. The participants were 22 adult students attending an education science program at a Finnish university. The participants worked in teams of four to five students. Groups were required to complete a project by the end of the semester (3 months). Overall, all teams experienced challenges in their collaborative tasks. Socio-emotional challenges were exhibited the most frequently, followed by motivational and cognitive challenges.

The aforementioned research had some limitations. They did not examine teams of more than five members in terms of long-term collaboration spanning over two months. Their research subjects were students from various fields, whereas our focus was on software engineering students.

III. METHODOLOGY

Our knowledge goal is to define a catalog of challenges present in the teamwork of software engineering students. We decompose this goal into several research questions we seek to answer to provide a more comprehensive catalog. The questions include:

1. Which challenges manifest themselves in students' teamwork, and to what degree? Answering this question determines the challenge's likelihood of occurrence, helping educators prioritize their attention.
2. How do students perceive the challenge's impact on teamwork quality? Answering this question determines the challenge's significance when it occurs, helping educators prioritize their attention.
3. How do familiarity and cohesion between students in one team affect these challenges? While indirectly related to the catalog, we explore the correlation between familiarity and challenge occurrence. As some student teams are made from a group of friends and others from strangers, we hypothesize that this difference can have an impact on challenge manifestation.

To find answers to these questions, we follow the empirical cycle of design science [17]. Our methodology entails two phases. The first phase consists of consulting the literature and the opinions of the teaching assistants to compile the initial catalog of challenges. Based on this catalog, we created a survey for the second phase, which involved surveying students and analyzing their responses.

A. Creating the initial challenge catalog

At our university, the teaching assistants working with software engineering undergraduates are in direct contact with the student teams and follow their progress on projects regularly. We created a survey that consisted of a brief introduction and a single open-ended question "What are the problems you observed in student teamwork?" 15 teaching assistants gave comprehensive answers.

We coded [18] their answers based on a schema of teamwork challenge categories defined in Adaptive Instrument for Regulation of Emotions (AIRE) [14]. AIRE is an instrument aimed at accessing students' experiences of individual and socially shared regulation of emotions in a socially challenging learning situation. We explored which categories of challenges were perceived by teaching assistants to determine if the AIRE schema was sufficient or if we needed to remove or add some challenge category. The results of this phase are presented in detail in section 4.D.

The output of this phase was an initial set of challenge categories relevant to our contexts, which include:

a. Team members have different project goals,
b. Communication is inefficient because members prioritize not offending other members,
c. Members have different work habits/styles,
d. Communication is inefficient because a member dominates the conversation with their opinions,
e. Members do not maintain friendly relations,
f. Members do not complete their weekly tasks,
g. Members have different work quality standards,
h. Members do not know how to resolve conflicts due to differences in opinions,
i. Members misunderstand tasks, and
j. Members have trouble organizing due to personal or non-faculty-related commitments.

B. Creating the student survey

We used the initial challenge catalog to create a survey for the second phase. The survey consists of 24 questions divided into three sections, each related to one of our research questions.

The first section contains one Likert-scale question for each challenge (10 questions in total), where students mark if a challenge was: not present in their team (0), if it presented a minor issue (1), if it was a notable issue (2), if it was a significant issue (3), or if it presented an overwhelming issue (4).

The second section has one Likert-scale question for each challenge (10 questions in total), where students denote their opinion on the impact that a challenge has on teamwork results on a scale of 1 (has little impact) to 5 (has a major impact).

The third section has one Likert-scale question that determines the students' overall satisfaction with their team, one Likert-scale question to denote how frequently they had business meetings, and one Likert-scale question to define how frequently they organized teambuilding activities. The final question is an open-ended question where students can highlight any additional challenges in collaboration they may have encountered or observed.

C. Administering the survey

The surveyed student population was composed of software engineering students from two study programs (we denote them as Program A and Program B), working in teams to develop a software solution over three months.

Program A students are third-year undergraduate students attending a study program that focuses on software engineering. Working in teams of three members, Program A students were tasked with developing a large software solution through five courses during a single semester.

Program B students are fourth-year undergraduate students attending a study program that covers computer engineering, control engineering, and software engineering. Working in teams of four, Program B students developed a software component through one course during a single semester. Additionally, every four teams formed one larger team (i.e., a team of sixteen members) that integrated the four components into a large software solution.

Despite their differences, students from both programs possessed similar software engineering knowledge and skills (e.g., programming, software design, and working with databases). Also, students from both study programs have experience in working in teams. The resulting solutions were web applications with roughly the same size and complexity (16 people x 1 course ~= 3 people x 5 course). Students in both programs received roughly the same amount of feedback on their work, which translates to roughly 6 minutes per student per week[1].

We examine the challenges in student collaboration in the three different contexts:

- The experience of Program A students working in teams of three (Context A),
- The experience of Program B students working in teams of four (Context B1),
- The experience of Program B students working in teams of sixteen (Context B2).

We collected responses from 155 students. 66 students from Program A supplied answers relevant to Context A. 89 students from Program B supplied answers relevant to contexts B1 and B2[2].

D. Processing the answers

To answer our first research question (*Challenge occurrence likelihood*), we use the Likert-scale answers to calculate the average value, the median, and the mode [19]. We derive the same values for the second research question (*Challenge impact perception*).

To answer our third research question (*Team cohesion and challenge occurrence correlation*), we calculate Spearman's rho correlation coefficient between each challenge occurrence value and the three values given for the questions in the third section (i.e., team satisfaction, work meeting frequency, teambuilding frequency).

Finally, we encoded [18] the students' answers to the open-ended question to determine if they encountered or observed any additional challenges, apart from the initial 10 categories.

[1] In both programs, 16 students attended sessions dedicated to monitoring their progress once per week. The session lasted for one and a half hour, which roughly translates to 6 minutes per student.

[2] The answers from students to the survey can be found at the following link - https://zenodo.org/record/7602921#.Y90SX3bMLGJ.

IV. Results

Here we present the results of our survey and coding efforts, grouped around our three research questions.

A. To what degree is each challenge present in a team?

Table I presents the mean, median, and mode values of students' responses from the survey based on the initial catalog of challenges. We observe that challenges *c* and *g* are minor issues in context A (median = 1), while the others are not present (median = 0). In context B1, challenges *a, c, g, and j* have a minor presence, while the others are not present. In context B2, challenges *b, e, f,* and *j* are minor issues, challenges *g* and *c* are moderate

TABLE I. Mean, median, mode values of students' responses (the presence of challenges)

Context	*Challenge*	*Mean*	*Median*	*Mode*
A	a. Different goals	0.45	0	0
	b. Avoiding offending	0.5	0	0
	c. Different work styles	0.86	**1**	0
	d. Dominating member	0.45	0	0
	e. No friendship	0.32	0	0
	f. Tasks not completed	0.61	0	0
	g. Different quality standard	0.77	**1**	0
	h. Poor conflict resolution	0.41	0	0
	i. Task misunderstanding	0.35	0	0
	j. Work-life organization	0.7	0	0
B1	a. Different goals	1	**1**	0
	b. Avoiding offending	0.84	0	0
	c. Different work styles	1.52	**1**	1
	d. Dominating member	0.60	0	0
	e. No friendship	0.42	0	0
	f. Tasks not completed	0.75	0	0
	g. Different quality standard	1	**1**	0
	h. Poor conflict resolution	0.51	0	0
	i. Task misunderstanding	0.42	0	0
	j. Work-life organization	1.11	**1**	0
B2	a. Different goals	1.55	**2**	2
	b. Avoiding offending	1.06	**1**	0
	c. Different work styles	1.93	**2**	2
	d. Dominating member	0.82	0	0
	e. No friendship	0.96	**1**	0
	f. Tasks not completed	1.27	**1**	1
	g. Different quality standard	1.62	**2**	2
	h. Poor conflict resolution	0.9	0	0
	i. Task misunderstanding	0.54	0	0
	j. Work-life organization	1.07	**1**	0

issues (median = 2), and the others are not present.

B. How impactful is each challenge to students?

Table II presents the mean, median, and mode values of students' responses to the survey. Since the students from Program B gave the same answers to this part of the survey for Contexts B1 and B2, we group the data around Program A and B. We observe that students from Program B rated the impact of all challenges on teamwork quality with a median of 4 or 5. The students from Program A gave similar answers, rating challenges *c e* and *j* with a median value of 3.

C. How does team cohesion affect challenge occurrence?

To measure associations among Likert item answers, we used the nonparametric Spearman correlation. According to [19], this is an appropriate way to analyze Likert item questions.

Table III presents Spearman's rank-order correlation matrix. Statistically significant correlations (significance level $\alpha = 0.05$) are bolded. Correlations higher than 0.29 indicate a medium-to-strong relationship between items [20].

The results show a moderate-to-strong negative correlation between team satisfaction and challenges *a, b, c, e, f, g,* and *h* in all three contexts. Additionally, in all three contexts there is a moderate-to-strong negative correlation between work meeting frequency and the occurrence of challenges *a*, *e, f* and *g*. Finally, in all three contexts, there is a moderate-to-strong negative correlation between team building frequency and challenge *e*.

TABLE II. MEAN, MEDIAN, MODE VALUES OF STUDENTS' RESPONSES (IMPACT OF EACH CHALLENGE)

Context	*Challenge*	*Mean*	*Median*	*Mode*
A	a. Different goals	4.3	5	5
	b. Avoiding offending	4.09	4	5
	c. Different work styles	3.55	3	3
	d. Dominating member	3.65	4	4
	e. No friendship	3.27	3	3
	f. Tasks not completed	4.3	5	5
	g. Different quality standard	3.71	4	4
	h. Poor conflict resolution	3.73	4	4
	i. Task misunderstanding	3.73	4	5
	j. Work-life organization	3.38	3	4
B1, B2	a. Different goals	4.35	5	5
	b. Avoiding offending	4.12	4	5
	c. Different work styles	3.93	4	5
	d. Dominating member	3.84	4	5
	e. No friendship	3.63	4	5
	f. Tasks not completed	4.4	5	5
	g. Different quality standard	4.13	4	5
	h. Poor conflict resolution	3.76	4	4
	i. Task misunderstanding	3.83	4	5
	j. Work-life organization	3.79	4	4

D. Coding results

Based on the challenge categories from AIRE, we coded the responses from teaching assistants on a single open-ended question "What are the problems you observed in student teamwork?".

TABLE III. SPEARMAN CORRELATIONS BETWEEN TEAM SATISFACTION, MEETING FREQUENCY, AND CHALLENGE PRESENCE

Context	*Challenge*	*Team satisfac.*	*Work meeting freq.*	*Team building freq.*
A	a. Different goals	**-0.461**	**-0.331**	-0.166
	b. Avoiding offending	**-0.466**	-0.118	**-0.363**
	c. Different work styles	**-0.378**	-0.127	-0.225
	d. Dominating member	**-0.474**	**-0.454**	**-0.267**
	e. No friendship	**-0.402**	**-0.383**	**-0.510**
	f. Tasks not completed	**-0.527**	**-0.371**	**-0.393**
	g. Different quality standard	**-0.546**	**-0.313**	**-0.408**
	h. Poor conflict resolution	**-0.323**	**-0.261**	-0.051
	i. Task misunderstanding	-0.165	0.042	-0.059
	j. Work-life organization	**-0.319**	-0.140	-0.181
B1	a. Different goals	**-0.609**	**-0.482**	**-0.225**
	b. Avoiding offending	**-0.578**	**-0.469**	**-0.309**
	c. Different work styles	**-0.572**	**-0.414**	**-0.577**
	d. Dominating member	**-0.379**	**-0.234**	**-0.379**
	e. No friendship	**-0.427**	**-0.401**	**-0.377**
	f. Tasks not completed	**-0.493**	**-0.529**	**-0.227**
	g. Different quality standard	**-0.524**	**-0.367**	**-0.265**
	h. Poor conflict resolution	**-0.416**	-0.184	**-0.214**
	i. Task misunderstanding	-0.147	-0.122	-0.114
	j. Work-life organization	**-0.404**	**-0.495**	**-0.287**
B2	a. Different goals	**-0.502**	**-0.269**	-0.79
	b. Avoiding offending	**-0.262**	**-0.276**	-0.178
	c. Different work styles	**-0.437**	**-0.293**	**-0.211**
	d. Dominating member	-0.110	-0.033	-0.158
	e. No friendship	**-0.504**	**-0.332**	**-0.309**
	f. Tasks not completed	**-0.346**	**-0.235**	-0.183
	g. Different quality standard	**-0.392**	**-0.240**	-0.027
	h. Poor conflict resolution	**-0.402**	-0.190	-0.139
	i. Task misunderstanding	-0.007	-0.005	-0.154
	j. Work-life organization	-0.181	-0.110	0.24

We identified 10 responses that belong to category *a* (Our goals for the project were different), 11 responses that belong to category *c* (We seemed to have incompatible styles of working), 1 response that belongs to category *d* (We seemed to have different styles of interacting), 5 responses that belong to category *e* (People in our team did not connect very well with one another), 2 responses that belong to category *f* (One/some people were not fully committed to the team project) and 5 responses that belong to category *h* (Team members were not equal). Other categories had 0 answers that belonged to them.

Based on the challenge categories from AIRE, we have composed our survey for students. The first step was to eliminate challenge categories from AIRE that are not relevant to our context or to expand existing ones.

1. We removed the category "different priorities" because it relates to a context where students collaborate over a shorter period. An example of such a challenge would be "some people were more interested in socializing than getting on with the task.",
2. We combined the categories "some members were rather shy and others very outspoken" and "some members' opinions were not taken into account" into *b. Communication is inefficient because members prioritize not offending other members* and added another category, *d. Communication is inefficient because a member dominates the conversation with their opinions* because we have noticed this kind of behavior in student collaboration in previous projects,
3. We removed the category "some people were easily distracted." because it relates to a context where students collaborate over a shorter period. An example of such a challenge would be "students made and received phone calls on their mobiles during meetings.".

The final question in the students' survey was an open-ended question where they could highlight any additional challenges in collaboration they may have encountered or observed. After coding the students' responses, we didn't identify any new categories of challenges.

V. Discussion

Here we discuss our findings, present implications for practice, and describe the limitations of our study.

A. Results

Based on the results listed in Table I (Presence of challenges), we observe that, in general, challenges occur more frequently in larger teams. Table I shows that mean, median, and mode values rise with the number of team members (e.g., Students from Context A with teams of 3 members have significantly fewer challenges than students from Context B2 with teams of 16 members).

We observe that challenge *c* (*Different work styles*) changes across all contexts (Context A – Median 1, Mode 0; Context B1 – Median 1, Mode 1; Context B2 – Median 2, Mode 2). We conclude that as the number of team members grows, so does the number of different work styles. For example, in software engineering, different work styles can be related to code size per commit or frequency and schedule for PR reviews. Interestingly, students do not perceive that this challenge significantly impacts team results relative to other challenges (Context A – Median 3, Mode 3; Context B1 and B2 - Median 4, Mode 5). However, Table III shows a moderate-to-strong negative correlation between the level of teamwork satisfaction and different work styles which means that teams with varying work styles had members that were less satisfied with their teamwork. This finding, coupled with the perception that most challenges significantly impact results (Table II), indicate that our students were not great evaluators of team performance.

We also observe that challenges *g* (*Different quality standards*) changes across all contexts (Context A – Median 1, Mode 0; Context B1 – Median 1, Mode 0; Context B2 – Median 2, Mode 2). One explanation for this result is the lack of an accurate metric to determine quality. When paired with different coding styles, work styles, and poor communication, team members might perceive the work of others as of low quality simply because it differs from their way of developing software.

Finally, challenge *a* (*Different goals*) changes across all contexts (Context A - Median 0, Mode 0; Context B1 - Median 1, Mode 0; Context B2 - Median 2, Mode 2). This result can be a consequence of the fact that students from both B1 and B2 contexts have four courses apart from this project, and students from context A have just one additional course. Also, students from contexts B1 and B2 are in the final year of their studies and, in large part, have started internships thus their focus is divided.

According to students in all contexts, the two challenges perceived as most impactful on team results were *a* (*Different goals* - Median 5, Mode 5) and *f* (*Tasks not completed* - Median 5, Mode 5). These results align with the correlation of these challenges with overall team satisfaction (Table III) and the broader literature on shared goals and accountability as a necessary component of strong teams [21].

Challenges *i*, *h*, and *d* have median 0 and mode 0 in all three contexts. Given that students have received clearly written tasks throughout the project, it is not unusual for the challenge *i* to not have manifested. However, the results for challenges *h* and *d* are unusual, considering that the students have not been formally trained in conflict resolution and similar soft skills.

B. Implications for practice

The results of the study we conducted on the students of the third and fourth years can aid higher education teachers who incorporate team projects as part of their courses. Since the study primarily refers to software engineering students that are nearing the end of their studies, it has the potential to serve software vendors that

hire and train junior engineers, as well as to manage their current software engineering teams.

We recommend educators and team leads become aware of the general challenges teams face and share our findings with their teams. An awareness of hurdles goes a long way in resolving them. Furthermore, we recommend educators examine the more prevalent challenges and those that correlate with overall team satisfaction when developing teamwork training materials. Finally, educators can employ our survey or its derivative to examine which challenges their trainees face and make a step towards resolving these issues.

C. Limitations

In terms of construct validity, we do not guarantee that we have covered all the challenges students may encounter. We utilized existing tools (i.e., AIRE) and adapted them to our context. Additionally, we employed open-ended questions on teachers and students to discover any additional categories and mitigate this issue. However, a more systematic approach is needed to create a comprehensive catalog of teamwork challenges.

In terms of internal validity, it is possible that some of the students did not understand all the survey questions. To reduce their chance of misunderstanding, we provided an example with each question. It is also possible that there was an error in the open-ended question coding. To mitigate this threat, two authors examined the answers to verify the coding validity.

In terms of external validity, we believe that the results of this research are applicable to any domain where people work in teams and have prior experience with collaboration.

VI. Conclusion

In this paper, we examined challenges present in the context of third and fourth-year software engineering students working on 3-month projects. We covered three contexts: teams of 3 members, teams of 4 members, and teams of 16. The research results are the quantified presence of challenges in each context (median, mean, frequency), students' opinions on the impact of each challenge (median, mean, frequency), and the correlations between team satisfaction and challenges. This can be significant for teachers and IT companies. In future work, we plan to utilize the enumerated challenges and their presence to help students overcome them using an automated intelligent tutoring system.

Acknowledgment

We are supported by the Science Fund of the Republic of Serbia, Grant No 6521051, AI-Clean CaDET.

References

[1] Griffin, P. and Care, E. eds., 2014. Assessment and teaching of 21st century skills: Methods and approach. Springer.

[2] von Davier, A.A. and Halpin, P.F., 2013. Collaborative problem solving and the assessment of cognitive skills: Psychometric considerations. *ETS Research Report Series*, *2013*(2), pp.i-36.

[3] Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W. and Hesse, F.W., 2018. Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest*, *19*(2), pp.59-92.

[4] Lingard, R. and Barkataki, S., 2011, October. Teaching teamwork in engineering and computer science. In *2011 Frontiers in education conference (FIE)* (pp. F1C-1). IEEE.

[5] Mason, S., 2020, August. Collaborative learning in computing education: Faculty perspectives and practices. In *Proceedings of the 2020 ACM Conference on International Computing Education Research* (pp. 136-146).

[6] Lunt, B.M., 2021, October. CC2020: An IT-Centric Overview. In Proceedings of the 22st Annual Conference on Information Technology Education (pp. 5-8).

[7] Iniesto, F., Sargent, J., Rienties, B., Llorens, A., Adam, A., Herodotou, C., Ferguson, R. and Muccini, H., 2021, October. When industry meets Education 4.0: What do Computer Science companies need from Higher Education?. In *Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21)* (pp. 367-372).

[8] Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M. and Rumble, M., 2010. Draft white paper 1: Defining 21st century skills. *Available online also at: http://cms. education. gov. il/NR/rdonlyres/19B97225-84B1-4259-B423-4698E1E8171A/115804/defining21stcenturyskills. pdf [accessed in Kota Kinabalu, Malaysia: April 15, 2015]*.

[9] Häkkinen, P., Järvelä, S., Mäkitalo-Siegl, K., Ahonen, A., Näykki, P. and Valtonen, T., 2017. Preparing teacher-students for twenty-first-century learning practices (PREP 21): a framework for enhancing collaborative problem-solving and strategic learning skills. *Teachers and Teaching*, *23*(1), pp.25-41.

[10] Raibulet, C. and Fontana, F.A., 2018. Collaborative and teamwork software development in an undergraduate software engineering course. *Journal of Systems and Software*, *144*, pp.409-422.

[11] Le, H., Janssen, J. and Wubbels, T., 2018. Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration. *Cambridge Journal of Education*, *48*(1).

[12] Stenberg, M., Mangrio, E., Bengtsson, M. and Carlson, E., 2021. Formative peer assessment in higher healthcare education programmes: a scoping review. *BMJ open*, *11*(2), p.e045345.

[13] Kazemitabar, M.A., Lajoie, S.P. and Li, T., 2022. A classification of Challenges encountered in Complex Teamwork Settings. *International Journal of Computer-Supported Collaborative Learning*, pp.1-23.

[14] Järvenoja, H., Volet, S. and Järvelä, S., 2013. Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), pp.31-58.

[15] Järvenoja, H., Näykki, P. and Törmänen, T., 2019. Emotional regulation in collaborative learning: When do higher education students activate group level regulation in the face of challenges?. *Studies in Higher Education*, *44*(10), pp.1747-1757.

[16] Näykki, P., Järvelä, S., Kirschner, P.A. and Järvenoja, H., 2014. Socio-emotional conflict in collaborative learning—A process-oriented case study in a higher education context. *International Journal of Educational Research*, *68*, pp.1-14.

[17] Wieringa, R.J., 2014. Design science methodology for information systems and software engineering. Springer.

[18] Seaman, C. B., 1999. Qualitative methods in empirical studies of software engineering. IEEE Transactions on software engineering, 25(4), 557-572.

[19] Subedi, B.P., 2016. Using Likert type data in social science research: Confusion, issues and challenges. *International journal of contemporary applied sciences*, *3*(2), pp.36-49.AA

[20] Cohen, J., 2013. *Statistical power analysis for the behavioral sciences*. Routledge.

[21] Coyle, D., 2018. The culture code: The secrets of highly successful groups. Bantam.

Stimulating (Open) Data Literacy at the Basis of Society: Approaches for Active Learning and Teaching to Young Children

I. Bosnić*, A. Kuveždić Divjak** and B. van Loenen***
* University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
** University of Zagreb, Faculty of Geodesy, Zagreb, Croatia
*** Delft University of Technology, Department of Architecture and The Built Environment, Delft, the Netherlands
ivana.bosnic@fer.hr, akuvezdic@geof.hr, b.vanloenen@tudelft.nl

***Abstract* - Children in the primary educational level learn about fundamental reading, writing and mathematics skills as a basis for future learning. In some European Union countries, they are also taught basic informatics and computer usage. However, only a few efforts are made to start teaching them about (open) data literacy. When teaching about data, multiple stages of the data value chain can be tackled, adapted to children's capabilities; from data production or collection to data usage or impact. Such teaching can be more theoretical, with already prepared examples, or a complete hands-on experience, where the data is collected from the environment, further analyzed, visualized and the conclusions about data are made. In this paper, we will present the currently available research on the approaches to teaching children about data. We will also discuss the potential benefits and the main issues impeding such an effort. Finally, we will propose the way forward for informal active learning about data for lower-level primary school children, based on the hands-on workshops using environmental data.**

Keywords – education; (open) data; data literacy; children; ISCED level 1

I. INTRODUCTION

Digital literacy is commonly recognized as essential for today's and our future societies. The digital literacy strategy of the European Union [1] announced and introduced "Strengthening Skills Intelligence" and "Increasing STEM graduates, fostering entrepreneurial and transversal skills" as two of the twelve actions in the skills revolution that is needed "to ensure people can thrive in the green and digital transitions, and to help in the recovery from the coronavirus pandemic" [2].

As a next step, the Digital Education Action Plan [3] was developed to support Europe's "Digital Skills for All" ambition. The Action Plan aims to "support the development of robust digital competences and organizational capabilities in education and training systems while fully harnessing the potential of emerging technologies, data, content, tools and platforms to make education and training fit for the digital age". Data literacy is a key element of digital literacy. It focuses on both the technical and social aspects of data. Data literacy encompasses activities related to data management, including data collection, data storage, data processing, data visualization, data quality, data dissemination and data use [4].

However, despite the increasing role of data in our lives, and the ambitions of the European Union to advance basic data skills at all levels, from early childhood to elderly, data education in primary and secondary school levels is underdeveloped, if developed at all. Known barriers to introduce data education are teachers' lack of technical expertise, teachers' lack of understanding of (open) data, and lack of understanding of how to integrate (open) data sets in school subjects (see [5], [6]).

According to the standard framework for categorizing cross-nationally comparable education statistics, primary education is classified in the International Standard Classification of Education (ISCED) Level 1 [7]. This type of education (also called elementary or basic education) typically starts at the age of 5-7 years and lasts for six years. ISCED level 1 education programs are typically designed to provide pupils with basic reading, writing, and mathematics skills that for the foundation for further learning. Other subjects such as history, geography, natural science, social sciences, art and music are acknowledged as fundamental and should be taught at this level [7].

In this paper, we aim to present a first step towards promoting data literacy of pupils taking part in the primary education (ISCED level 1). Most particularly the emphasis is on the lower-level primary school education, the first three grades (age 5-8 year). We will present the currently available research on the approaches to teaching children about data. We will also discuss the potential benefits and the main issues impeding such an effort. Finally, we will propose the way forward for informal active learning about data for lower-level primary school children, based on the hands-on workshops using environmental data.

II. LITERATURE REVIEW

A. Teaching general data literacy to lower-level primary school children

Data literacy is closely related to and dependent on other informatics skills and competences, especially those of digital literacy [8]. The European Commission (see [8],

p. 22) identified ten core learning areas of informatics in primary and secondary education: 1. Data and information, 2. Algorithms, 3. Programming, 4. Computing systems, 5. Networks, 6. People–system interface, 7. Design and development, 8. Modelling and simulation, 9. Awareness and empowerment, and 10. Safety and security.

Here, we focus on the data and information category. The European Commission report does not specify any categories of data and information skill's categories. Rather it provides example learning outcomes on data and information on different levels of education (primary and secondary school school levels). Applying the six cognitive learning levels in the Bloom's taxonomy (1. Remember, 2. Understand, 3. Apply, 4. Analyze, 5. Evaluate, 6. Create) to the European Commission report we see that the focus in primary school data and information education in some European countries is at the lower Bloom levels. For example, in the Czech Republic, pupils in primary school need to understand "the flow of information from its generation, storage on a medium, transfer, processing, retrieval by search and practical use", and in Slovenia they "should understand the binary system for representing data and know that the data can be compressed lossless and with loss of information". Only one example was provided for the lower primary school level (in Poland): "the learner lists the risks associated with widespread access to technology and to information and describes ways of avoiding them. The learner recognizes and respects the right to privacy of data and information and the right to intellectual property."

Research on data/information education in the lower levels of primary schools is scant. Some lessons may be learned from the almost as scarce research on data/information education in the higher level of primary school (grade 5 and up) [9]. For these groups, Saddiqua et al. [5] highlighted the importance of using up-to-date data or preferably hands-on experiences with data collection to better understand the concept of data, and how to use (open) data. Hands-on experiences would entail having the pupils collect and interpret the data themselves. In this respect, Saddiqa et al. [10] added that the data activities should be related to the world of the pupils. One may think of exercises about the pupils' height, the color of their eyes and generating some basic statistics about the collected data (average, mean, etc.). They also found that data education should focus on local, easy to recognize, situations such as measuring the height of a tower next to the school or counting how many cars pass their school. They also found that "simple open data visualizations representing students' municipalities captured their attention and encouraged discussions and reflections. These include topics such as water quality, forests, crimes, pollution sources, traffic, plastic pollution, lakes (locations, length, etc.)". Finally, in another study of the same authors they refer to the majority of teachers in the study that argue that pupils can develop basic data handling skills if they work with small datasets [11].

In addition to the type of learning activity research also identified prerequisites/issues for successful data/ information education. Some of those issues are at the education providers/teachers' side (limited data/technology skills). They may, for example, not be familiar with data cleaning and preparing concepts. Practical issues such as limited access to data, dataset language (not in native language e.g., Croatian), and the dataset size being too large to process by the teacher's laptop are mentioned as well as the need for an interface tailored to educational purposes.

Finally, Olsen [12] researched adapting a data literacy game to teach data literacy, initially intended for children aged 10-12, to children aged 8-10. To better suit younger children, one needs to be very careful in copy pasting educational methods of higher education to lower levels. This conclusion may also apply to the above: best practices on higher primary school levels may not necessarily apply to lower primary school levels.

However, one domain developed and implemented data/information learning outcomes for lower levels of primary school: the domain of (geo)spatial data.

B. Teaching (geo)spatial data to lower-level primary school children

Spatial data is any type of data that describes different instances through space and time. Spatial can refer to the study of any type of features be it outer space, the world, the spatial layout of the human body, the spatial layout of a room. *Geospatial data* is a term used to describe data about objects, events, or phenomena that directly or indirectly references a specific geographical area or location on the surface of the Earth, but the terms spatial and geospatial are often used interchangeably.

Spatial literacy is a fundamental form of literacy and refers to the ability to use the properties of space to communicate, reason, and solve problems. When examining geographic and environmental scale, spatial literacy is referred to as *geospatial literacy* to identify the spatial subset and scale under consideration.

Following the definition and use of the term literacy in an educational context, geospatial literacy refers to the mix of knowledge, skills, and dispositions required to communicate and understand geospatial information through representation [13]. Geospatial literacy is the key to transforming geospatial data into information and then into personal knowledge that can be used to make decisions, empower, ask broader questions, communicate, and report to society.

There is growing recognition that spatial literacy is important for mastering STEM content. The U.S. National Research Council [14] has found that spatial reasoning, a key component of spatial literacy, "…is at the heart of many great discoveries in science. It underpins many of the activities of the modern workforce, and it pervades the everyday activities of modern life". Indeed, spatial skills in childhood have been associated with success in several STEM fields [15], including engineering, chemistry, medicine, physics, geology, and computer programming.

Several recent studies (e.g., [16], [17]) have found that spatial skills are malleable and can be improved through early training, which is especially beneficial for children in the primary educational level. Early spatial skills training that includes interactive components such as hands-on exploration and feedback from the environment (e.g., visual cues) is thought to have positive effects [18].

The National Research Council [14], referred to spatial reasoning as a “missing link” in the (primary) school curriculum. The report stated that a “concerted effort to better integrate spatial thinking and bring it into schools could enable students to engage with their learning at a deeper level and also help to raise a new generation of spatially literate students”. Increasingly, demands for spatial literacy are being incorporated into education efforts in the UK, the US, and Australia [19].

In terms of classroom-based hands-on spatial training, a literature search revealed many suggestions for how educators can "spatialize" their teaching in the primary educational level. Spatializing mathematics curriculum, for example, includes adapting geometry lessons to include dynamic visualizations that incorporate 2D and 3D shapes [20]. Pupils can practice spatial reasoning from 2D to 3D spatial thinking by constructing 3D geometric figures out of paper. Simply drawing a flat sketch of a 3D object, such as a box, pencil, or chair, trains spatial thinking.

As part of geography curricula, map reading and navigation are great ways to train spatial skills. For example, by creating models and maps of their classroom, school, and/or community, pupils use these skills to deepen their knowledge of ecological, physical, and human systems and to develop a sense of place [21], [22]. Pupils of all ages can engage in treasure hunt activities, where they use a map of their school backyard (provided by the teacher or created by the pupils) to find hidden treasure. As with map skills, aerial and satellite imagery can be effectively used in classroom-based spatial training to help pupils in visualizing the physical or human phenomena that is present [23].

A geographic information system (GIS) is also a useful resource for developing pupils' spatial literacy, helping them navigate locally and globally, and combining data sets to identify relationships between them [24]. Although there is a general perception that GIS should be taught in the upper grades due to the more technical aspects of spatial modelling, there are systems that are accessible to pupils at the primary educational level. For example, an interactive website ‘Linsey the GIS Professional’ [25] teaches young children how to collect geospatial data and create a map of a park and builds children's spatial analysis and critical thinking skills.

C. Data literacy educational resources aimed at lower-level primary school children

Over the years, there have been attempts to involve children in data literacy with the help of various activities, both offline and online. A small handful of websites exist offering different approaches of teaching data to kids. Here we will present several examples and focus on the type of activity offered.

Kids in data [26] is a learning platform that teaches young children basic data concepts. The recommended age range is not specified, but the target audience is referred to as "the young children". The data used is the children's own data, generated by playing the game "Space Invaders": number of games played, score, ranking, etc. The platform serves as the basis for workshops in which children come together, play the game, and then analyze their own game data using a special online data visualization tool that was initially an open-source solution, but is now a paid service. In addition, the website includes short learning modules that teach kids the basics of data theory, such as the difference between quantitative and qualitative data, or the concepts of tables and charts (bar, line, and pie charts). While the website still works, it is not updated since 2020, and most theoretical content is missing with the “Coming soon” message. The website offers a very interesting idea, data gathering through joint game-playing. However, it would be better if the data visualizations and manipulations were provided directly in the web application, without the need to download data and use specialized (commercial) applications.

DIY Data Art: Creative data literacy for public libraries [27] is a set of “short self-guided activities that use personal data as an entry point to data science”, aimed to youth. The simple, yet highly engaging online and offline activities tackle particular data examples, such as network usage of their mobile phone applications, using emojis for visualization, exploring networks of connected people, exploring the community buildings in the neighborhood, etc. The activities are described in the Toolkit & Activity guide, which includes more examples, modifications and ideas.

A fresh squeeze on data [28] is a set of resources, including a free children's book and lesson plans, to teach young children how data is the basis for many problem-solving activities and life decisions, including the professional ones. Resources include introductions to data gathering, data analysis, data bias, etc. The available activities in the lesson plan are targeted to 3rd-5th grade of primary education.

Tableau Data Kids [29] is an educational program from the commercial data analytics and visualization web platform, Tableau. Their program for kids includes a variety of activities, including a full “Data Kids Educator Kit” with lesson plans intended for children ages 6-13. The demonstrations are made in the Tableau Public [30], a free platform to explore, create, and publicly share data visualizations online. Depending on the activity, data is obtained either from external sources (Spotify songs, Pokemon rules, etc.), from children surveying others, or from children’s own experience (favorite songs, the speed of room cleaning depending on the “messy” level, the number of toy bricks grouped by color, simple statistics of sprouting seeded vegetables, etc.). Data manipulation and visualization can be done both on paper or at the computer, in spreadsheet software and specialized applications. Often the emoji symbols are extensively used in different contexts and lesson plans.

Web platform **Youcubed** [31], provided by Stanford Graduate School of Education offers a set of resources for learning and teaching data science for different ages, from elementary school to high school students. The lessons include a community data collection project for engaging children with “citizen science” [32] or an elementary school outdoor activity “Data Tells Us about Ourselves” [33]. Extensive lesson plans for teachers are provided.

DataBasic.io [34] is a set of web tools that introduce data concepts to pupils of age 12+. The four available tools

focus on finding the most used words and phrases, analyzing the data in a CSV file, showing the similarity of text files and analyzing the networks of connected data. Each tool contains the activity guide for teachers, several example data sources, and the possibility to import the custom data.

III. LESSONS LEARNED

Several lessons can be learned from our primary and secondary source literature review. The first lesson is that it is very well possible to teach 'data' to lower-level primary school pupils. However, there are clearly a number of aspects that must be considered in order to make it a success. For example, our study shows that data education should be connected to the children's world. Data activities that relate to the canal next door, the street around the corner, or the pupil's home are much more engaging to pupils than activities at a general level such as a city or country. The literature review also reveals that to be effective pupils should collect and interpret the data themselves: they should have hands-on experience with the tools and the data. They can create their own map of their surroundings, collect data about the number of cars in their street, or generate the average length of pupils in their class.

More advanced activities may include using a GIS to integrate collected data (number of cars per street) with existing base layers (aerial imagery or topographic maps) and perform basic spatial analysis of the data, creating new information (e.g., insight in car density per street) and visualizing this. If it is not possible for pupils to collect the data themselves, at least up-to-date data should be used. Another take away is (serious) gamification. This can be an effective strategy for our target group. For example, go for a geo treasure hunt (using GPS device to find the treasure), or have them analyze their own behavior for example on the time spent on playing a computer game, or map and analyze their route from home to school. For a sustainable approach to this type of teaching, we recommend the use of open data and open educational resources in the classroom, otherwise it is possible that the data or resources will soon become legally or technically unusable or obsolete.

In addition to the lessons learned for what and how to teach data to pupils at lower-level primary schools, also attention should be paid to the teachers: they need to be trained in data skills and equipped with sufficient user-friendly support systems so they can easily access up-to-date data and tools/apps, but also to share the data they collect and new insights from the data with the rest of the world (e.g., by connecting to citizen science platforms).

IV. OUTLOOK TOWARDS A NEW APPROACH FOR DATA EDUCATION FOR YOUNG CHILDREN

Based on the literature review on formal and informal ways to teach children (open) data skills, we propose a new approach for lower primary school children. In it, we aim to address all stages of the data value chain, from data production/collection to data use, adapted to children's abilities. We propose to prepare a set of workshops with a hands-on approach where children collect data from their local environment, analyze it, visualize the data, and draw conclusions based on the data, ask further questions, or think about the future possible extensions. In doing so, they would need to think critically not only about the data they collect, but also about the methods and conclusions. Although this may seem overwhelming at first for children at such an early age, we argue that this approach is feasible and that young children would gain the basic introduction to the key stages of the data value chain and improve their data literacy. The approach would include both offline and "online" activities, starting with various pen-and-paper approaches to arrive at initial findings about data without technical issues, but quickly moving to data literacy in the digital world.

The workshop topics relate to the children's world, but also involve their relation to the environment. Examples of such topics would be:

- counting the things in their environment (how many cars or bicycles pass in the street)
- mapping their environment (where are the benches in the park)
- measuring the noise levels in different areas near the school
- reporting on the plants found in the neighborhood
- collecting data from children's own lives – time they spend doing certain tasks, data about their daily habits, etc.

Depending on the workshop type, it is also possible to introduce more data sources or environment properties to find possible relations, such as whether the number of cars is higher and the number of bicycles is lower on a rainy day than on a sunny day. Also, a part of the workshops would have the (geo)spatial component, where the visualizations would be made on a map, from the geocoded data collected.

Each workshop should – at least partially, depending on the topic – enable children to go through the following phases (Figure 1):

Data collection – the main data source should be obtained by the children, from the environment, from their own daily lives, or from making the basic surveys with the other people. Depending on the context and the type of data source, children can use pen and paper, mobile phones, or some other devices. Various applications can help them in data collecting, starting from the basic ones such as data timers, thing counters, text editors to more advanced ones, such as sound recorders, camera apps, noise measuring tools, geotagging apps.

Data processing – the data obtained needs to be imported to an application, checked for possible errors and cleansed if needed. In most cases, if the data is obtained manually (e.g., counting, surveying) children need to understand the point of data processing. This can be done in workshops where the data is gathered using other methods and apps. For this step, we should ensure that the data export from the device to the application is done automatically and as simple as possible, so that neither the teachers nor children have additional technical issues to solve, besides the ones that directly impact our learning outcomes.

Data analysis – after the data is imported and presented in some "children-readable" form, a simple data analysis

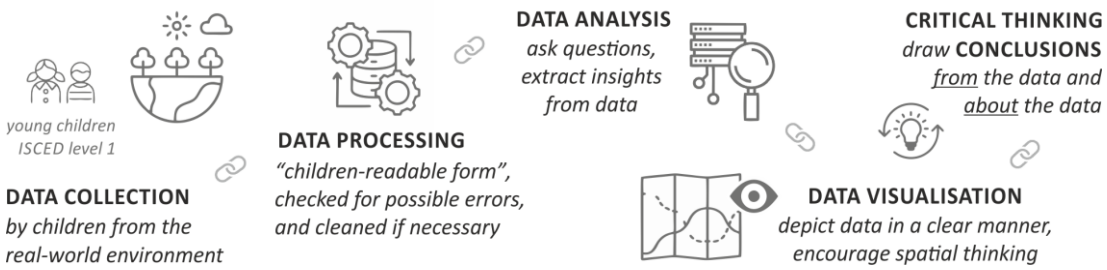


Figure 1. Data value chain approach to teaching (open) data literacy.
Adapted to children's abilities, key stages can be approached – from data collection to data use or impact

can be performed: finding averages, minimum, maximum, asking questions about what could be found out from the data, etc.

Data visualization – along with the data analysis, data visualization helps children to understand the raw data and draw conclusions in the next step. The visualizations should be clear and accessible to as many children possible. The way data is visualized is especially important for geospatial workshops, where data is visualized on maps, to help them improve their spatial thinking.

Conclusions based on data – after the basic data analysis and visualizations are performed, we'd like to engage in a more in-depth step: drawing conclusions not only from the data, but also about the data. Our goal is to try to engage pupils in critical thinking about the data – how was it obtained, were there some errors, why is the data different, how else they could have approached the task of data gathering and analysis. This step is one of the highest-ranked in Bloom's taxonomy of learning outcomes, but we feel it should be tried out with as much effort as possible, depending on the workshop type and pupils' characteristics, as it would also gain the most benefits, if successful.

From a practical view, one may think of a first series where the basic data skills are trained for each step in the data chain (Bloom level: remember and understand), followed by series that cover higher levels of Bloom, up to the design/create level. An alternative may be to educate the basic skills for the first step of the chain (understanding data collection), then do the data processing step at the Bloom level of apply and analyze, and the evaluate and design level of Bloom for the data visualization step.

The educational resources developed for running the workshops should be:

(i) extensive – teachers should be able to obtain the full and detailed versions of lesson plans / activity guides, which could be followed without the need for improvisation. These should in particular include the instructions for data gathering, the technical instructions for data manipulation, the methods for data visualization, and place an emphasis on the tools/applications needed to perform these tasks in the digital world.

(ii) ready to use – in order to facilitate the teachers in running the workshops, any content which might be of use during the workshop should be provided in a reusable manner before the workshop. For example, any sort of examples, demos, template sheets, etc. should be ready in advance, to leave the teachers worry-free as much as possible.

(iii) pedagogically sound – the workshop resources should – besides practical and technical instructions – provide the full set of pedagogical goals and methods, such as the learning outcomes based on Bloom's taxonomy, with higher levels of the taxonomy included.

(iv) up-to-date and published under an open license – in order to ensure the sustainability of the educational content provided, but also to provide flexibility for teachers to adapt the workshops as needed, any instructions or lesson plans should be published under the open license, such as Creative Commons, with as open version as possible. The best option would be to publish them under CC BY license, which retains the authorship, but allows free adaptations and commercial use. In case some example dataset is used (although our goal is to create/obtain the dataset at the beginning of each workshop), it should always be up-to-date, having timely and accurate data, to provide for decent conclusions.

(v) accessible and created with the universal design in mind – as the workshops should be accessible and inclusive for most of the children, they should be designed universally, respecting the principles of "universal design" and accessibility guidelines as much as possible. This includes the color schemes of graphs, large text, high contrast, simple instructions, different content formats, etc.

V. CONCLUSION

In this paper, we presented a first step towards promoting data literacy of pupils taking part in the International Standard Classification of Education (ISCED) level 1: primary education. Although there is limited research available on data education for 5-7 years old pupils, the currently available research on the approaches to teaching children about data reveals that it is very well possible to teach 'data' to lower-level primary school pupils. There is clearly a number of aspects that need to be taken into account in order to make it a success: (1) data education should be connected to the world of the children, (2) pupils should collect and interpret the data themselves: hands on experience with the tools and the data, and (3) use up-to-date data. To be effective, teachers must also be trained in data skills and equipped with sufficient user-friendly support systems to provide the data education that is urgently needed in our lower-level primary school systems in Europe.

For data education for lower-level primary school we propose a data education approach that builds on the concept of the data value chain. We propose to prepare a series of workshops with a hands-on approach, where the children would collect the data from their local environment, analyze it, visualize the data, and reflect on the results. In this way our approach contributes to the development of a data education approach for ISCED level 1 pupils that is needed to fulfil the ambition of the European Union to advance basic data skills at all educational levels, including those at lower-level primary schools.

ACKNOWLEDGMENT

This work was supported by the EU Horizon2020 project Twinning Open Data Operational 2019-2022 (Grant Agreement no. 857592).

REFERENCES

[1] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *European Skills Agenda for sustainable competitiveness, social fairness and resilience*, COM(2020) 274 final, 2020. [Online]. Available: https://op.europa.eu/s/xOLT. (Accessed: Feb. 15, 2023).

[2] European Commission, Directorate-General for Employment, Social Affairs and Inclusion, *European skills agenda : skills for jobs. Publications Office*; 2020. [Online]. Available: https://doi.org/10.2767/978543. (Accessed: Feb. 15, 2023).

[3] European Commission, Publications Office of the European Union, *Digital Education Action Plan (2021-2027)*, 2022. [Online]. Available: https://op.europa.eu/s/xOLZ. (Accessed: Feb. 15, 2023).

[4] Rivera Perez, A., Emilsson, C. and Ubaldi, B. "OECD Open, Useful and Re-usable data (OURdata) Index: 2019", OECD Policy Papers on Public Governance No. 1, OECD Publishing, Paris, Mar. 2020. [Online]. Available: http://www.oecd.org/gov/digital-government/policy-paper-ourdata-index-2019.htm. (Accessed: Feb. 15, 2023).

[5] M. Saddiqa, M. Kirikova, R. Magnussen, B. Larsen, and J. M. Pedersen, "Enterprise Architecture Oriented Requirements Engineering for the Design of a School Friendly Open Data Web Interface," *Complex Systems Informatics and Modeling Quarterly*, no. 21, pp. 1–20, Dec. 2019, doi: 10.7250/csimq.2019-21.01.

[6] T. Coughlan, "The use of open data as a material for learning," *Educational Technology Research and Development*, vol. 68, no. 1, pp. 383–411, Feb. 2020, doi: 10.1007/s11423-019-09706-y.

[7] OECD/Eurostat/UNESCO Institute for Statistics, *ISCED 2011 Operational Manual: Guidelines for Classifying National Education Programmes and Related Qualifications*, OECD Publishing, Paris, 2015. [Online]. Available: https://doi.org/10.1787/9789264228368. (Accessed: Feb. 15, 2023).

[8] European Commission, European Education and Culture Executive Agency, "Informatics education at school in Europe", *Publications Office of the European Union*, 2022. Available: https://data.europa.eu/doi/10.2797/268406. (Accessed: Feb. 15, 2023).

[9] J. Altenas and L. Havemann, "Education", in *The State of Open Data: Histories and Horizons*, T. Davies, S. Walker, M. Rubinstein, and F. Perini, Eds. Cape Town and Ottawa: African Minds and International Development Research Centre, 2019, pp-91-102.

[10] M. Saddiqa, B. Larsen, and R. Magnussen, "Open Data Visualization in Danish Schools: A case Study," in Computer Science Research Notes, 2019. doi: 10.24132/CSRN.2019.2902.2.3.

[11] M. Saddiqa, R. Magnussen, B. Larsen, and J. M. Pedersen, "Open Data Interface (ODI) for secondary school education," *Comput Educ*, vol. 174, p. 104294, Dec. 2021, doi: 10.1016/j.compedu.2021.104294.

[12] E. Olson, "Data literacy and young children Design suggestions for a game intended to teach data literacy to children 8-10 years of age", BSc. thesis, Linköping University, Sweden, 2021.

[13] L. Moorman, "The Evolution and Definition of Geospatial Literacy", In *GIScience Teaching and Learning Perspectives*, S. Balram, J. Boxall, Eds. Advances in Geographic Information Science. Springer, Cham., 2019, pp. 9–36. doi: 10.1007/978-3-030-06058-9_2.

[14] *Learning to Think Spatially*. Washington, D.C.: National Academies Press, 2006. doi: 10.17226/11019.

[15] J. Wai, D. Lubinski, and C. P. Benbow, "Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance.," *J Educ Psychol*, vol. 101, no. 4, pp. 817–835, Nov. 2009, doi: 10.1037/a0016127

[16] D. H. Uttal et al., "The malleability of spatial skills: A meta-analysis of training studies.," *Psychol Bull*, vol. 139, no. 2, pp. 352–402, Mar. 2013, doi: 10.1037/a0028446.

[17] W. Yang, H. Liu, N. Chen, P. Xu, and X. Lin, "Is Early Spatial Skills Training Effective? A Meta-Analysis," *Front Psychol*, vol. 11, Aug. 2020, doi: 10.3389/fpsyg.2020.01938.

[18] W. Yang, H. Liu, N. Chen, P. Xu, and X. Lin, "Is Early Spatial Skills Training Effective? A Meta-Analysis," *Front Psychol*, vol. 11, Aug. 2020, doi: 10.3389/fpsyg.2020.01938.

[19] S. W. Bednarz and K. Kemp, "Understanding and nurturing spatial literacy," *Procedia Soc Behav Sci*, vol. 21, pp. 18–23, 2011, doi: 10.1016/j.sbspro.2011.07.004..

[20] Z. Hawes, J. Moss, B. Caswell, S. Naqvi, and S. MacKinnon, "Enhancing Children's Spatial and Numerical Skills through a Dynamic Spatial Approach to Early Geometry Instruction: Effects of a 32-Week Intervention," *Cogn Instr*, vol. 35, no. 3, pp. 236–264, Jul. 2017, doi: 10.1080/07370008.2017.1323902.

[21] J. S. Smith, "Integrating Maps in the K-12 Curriculum," *The Geography Teacher*, vol. 1, no. 1, pp. 26–32, Oct. 2004, doi: 10.1080/19338341.2004.11089832.

[22] M. R. Petty and A. C. Rule, "Effective Materials for Increasing Young Children's Spatial and Mapping Skills," *Journal of Geoscience Education*, vol. 56, no. 1, pp. 5–14, Jan. 2008, doi: 10.5408/1089-9995-56.1.5.

[23] M. R. Petty and A. C. Rule, "Effective Materials for Increasing Young Children's Spatial and Mapping Skills," *Journal of Geoscience Education*, vol. 56, no. 1, pp. 5–14, Jan. 2008, doi: 10.5408/1089-9995-56.1.5.

[24] J. J. Kerski, "The role of GIS in Digital Earth education," *Int J Digit Earth*, vol. 1, no. 4, pp. 326–346, Dec. 2008, doi: 10.1080/17538940802420879.

[25] T. Danielsen. "Lindsey the GIS Professional". https://learn.arcgis.com/en/esripress/lindsey/ (accessed Feb. 15, 2023).

[26] "Kids in Data – an interactive data literacy workshop platform," https://kidsindata.com/ (accessed Feb. 15, 2023).

[27] "DIY Data Art," https://itsliteracy.wordpress.com (accessed Feb. 15, 2023).

[28] "A fresh squeeze on data," https://freshsqueezekids.com (accessed Feb. 15, 2023)

[29] "Tableau Data Kids," https://www.tableau.com/academic/data-for-kids (accessed Feb. 15, 2023).

[30] "Tableau Public," https://public.tableau.com (accessed Feb. 15, 2023).

[31] "Youcubed," https://www.youcubed.org/resource/data-literacy (accessed Feb. 15, 2023).

[32] "Youcubed Unit 5 – A Community Data Collection Project," https://www.youcubed.org/wp-content/uploads/2020/05/Unit-5.pdf (accessed Feb. 15, 2023).

[33] "Data Tells Us about Ourselves," https://www.youcubed.org/wp-content/uploads/2020/07/Data-Tells-Us-about-Ourselves.pdf (accessed Feb. 15, 2023).

[34] "Databasic.io," https://databasic.io (accessed Feb. 15, 2023).

The Relationship between Extracurricular Activities and Student Achievement in the Affective Domain: The Case of a Vocational Electrical Engineering School

D. Delač*, D. Purković**
* Electro Industrial and Crafts School Rijeka, Rijeka, Croatia
** University of Rijeka/Study of Polytechnics, Rijeka, Croatia
dino.delac@eios.hr

***Abstract* - Student achievements in the affective domain is important to the learning process, as adopted values and attitudes can promote the acquisition of knowledge and the development of skills. This paper examines the effects of extracurricular activities on student achievements in the affective domain and provides an analysis and comparison with the achievements of a similar group of students in traditional classes. The research was conducted through the extracurricular activity "Young Innovators", in which 18 third grade students of a vocational high school participated. The aim of activity is to develop new and innovate existing solutions in the field of electrical engineering and computer science. The achievements in the affective domain was observed and analysed in the students who participated in the extracurricular activity and in a similar group of students who did not participate. The results of the study indicate that students who participate in extracurricular activities communicate and cooperate better with other participants, and show a higher level of responsibility for completing classroom tasks than students who participate only in the traditional class. These findings suggest that there is a need to develop extracurricular activities that give every student the opportunity to participate in something that interests them.**

Keywords* - *affective domain, engineering education, extracurricular activities, technology education, vocational education

I. INTRODUCTIONS

The education process is one of the fundamental processes that affects the development of the individual, but ultimately contributes to the entire society or community. In addition to the traditional teaching in the classroom by presenting and lecturing theoretical knowledge, the practical experience of students that complements theoretical knowledge is also very important for the learning process. Learning through extracurricular activities, outside of formal education in the classroom, additionally develops cognitive and intellectual abilities [1]. Following the scientifically verified Bloom's taxonomy, as a system of classification of educational goals, it is determined that it is correct to observe the achievements of students in the cognitive and psychomotor domains, but it is also necessary to study the achievements of the affective domain. As stated by [2], the affective domain describes the level of adoption of socially acceptable beliefs and attitudes and related emotions on which the motivation for a certain action depends. According to the revised Bloom's taxonomy, the behaviour patterns of individuals are divided according to the degree of complexity, from basic listening and responding through evaluation and organization in accordance with priorities to the most complex pattern of characterization within which the individual cooperates in solving problems as a team and reconsiders his opinion while simultaneously appreciating the people in his environment [3]. The goal of education is for students, through primary and secondary education, to constantly improve their professional knowledge and skills, accompanied by the development of personal attitude and character. In order to achieve this, it is necessary to enable students to participate in extracurricular activities, because in this way each individual has the opportunity to apply his professional knowledge and comment on it with the rest of the participants he is surrounded by. Students in the current elementary school education system of the Republic of Croatia have very few options and opportunities to choose added teaching content or extracurricular activities in support of technical culture [4]. By integrating into programs of extracurricular activities, each individual chooses the activities that best enable him to achieve his goals, with work dynamics adapted to his own abilities [5]. Previous research has proven that this type of education indicates favourable opportunities in the segments of academic achievements and integration in communities [6], and also the quality of teachers as one of the factors of quality education. Analysis of the teaching process, in which extracurricular activities are integrated, showed that students aspire to knowledge in the fields of technology, engineering and mathematics [7]. The problems of today's education indicate that constant monitoring of the educational process is necessary since the environment in which we live is constantly changing. This is precisely why the educational field of science has the task of improving the educational process, i.e. achieving the synergy of the traditional form of teaching and learning based on work. One of the first examples was the People's Republic of

China, which focused its policy on a sustainable economy based on innovation, as opposed to the previous manufacturing economy, since its goal is to achieve the status of a world innovator by 2050 [8]. This is why today, in addition to traditional ways of learning and teaching, modern strategies and approaches in the field of engineering and technology, such as project learning, are increasingly being applied. Such an approach to learning and teaching enables participants to build their competencies and professional knowledge through learning based on the work process. Individuals in the education system, who are given the opportunity to learn through project-based learning, stand out with a higher degree of engagement and generic competencies [9], which is reflected in better communication and more productive cooperation through teamwork. Project learning enables every individual, who wants it, to prepare more easily for the future working environment, because knowledge is acquired through traditional teaching in which project learning is integrated, i.e. learning based on work, i.e. practice.

The situation in the educational system is further complicated by the fact that clear visions of education have not been defined and harmonized by those who shape educational policies [10]. The current situation indicates a necessary need for changes in education curricula. But despite the very expressed need for changes, they are still lacking [11], since extracurricular forms of work require professional individuals who are capable of performing extracurricular work activities, and the problem is also the unwillingness of teachers to maintain extracurricular forms of work.

In today's curricula and programs, which must be harmonized with the modern age, situations from the real environment must be integrated into the teaching contents, which will prepare students for life challenges in the current reality and the coming future. Through education by attending extracurricular activities, students are given the opportunity to strengthen their current abilities and to build their future profession according to their wishes [12].

II. THE PROBLEM AND AIM OF THE RESEARCH

Achievements of students in the affective domain, involved in extracurricular activities, are very poorly researched and do not provide concrete results for activities in the STEM field of science. Although there are studies that analyse the affective achievements of students for the teaching of natural sciences, there is still a lack of concrete evidence of the influence of extracurricular activities in the field of electrical engineering and computing on the affective domain of student achievements in the STEM field of education, which represents the synergy of cognitive and affective achievements of students. This research deals with the problem of determining the affective achievements of students that have been improved through the participation of students in the extracurricular activity "Young Innovators". The aim of this research is to determine as clearly and precisely as possible the difference in the achievement of affective achievements of students who are involved only in a traditional concept of education and students who combine a traditional concept of education that is imbued with extracurricular forms of work. The research question posed by this research is: *What is the link between the extracurricular activity "Young Innovators" and student achievement in the affective domain? Does it really improve student's affective achievements?* In this research, it is an extracurricular activity that represents the synergy of all components of the STEM field of education, from electrical engineering and computing (programming of works/manufactured devices, wiring of components) to mathematics and natural sciences that enable the application of electrotechnical laws and sustainable development).

III. METHODS, PROCEDURE AND SAMPLE

A. Methods

The research was conducted in the form of a quasi-experimental design, which was carried out after the process of monitoring students during their education from the second grade of high school to the fourth grade of high school. Two groups of students were included in the research, one group of students who were involved in an extracurricular activity and another group of students who were not involved in an extracurricular activity but were involved in the educational process only through the traditional form of teaching. The participants of the groups were monitored during three years of secondary education, and the students of both groups were equalized according to the criteria of general success and assessment of proficiency in previously completed secondary education classes and participation in competitions of technical disciplines in secondary education. The initial assumption that was made at the beginning of the research is that the affective achievements of the selected group of students who participate in the program of extracurricular activities do not differ statistically significantly compared to the achievements of students who do not participate in these activities, but attend classes according to the traditional teaching model.

B. Research procedure

The students of both groups are participants of the same high school qualification "Mechatronics technician" lasting four years. All are educated according to an equivalent curriculum that includes general education and vocational subjects. One group of students, in addition to the traditional form of teaching, was involved in the extracurricular activity "Young Innovators", through which students create technical and technological creations with which they participate in national, state and international competitions and work reviews. The curriculum and program of the extracurricular activity "Young Innovators" is imbued with vocational content in the fields of electrical engineering and computing, but it allows students to choose the vocational field in which they believe they can best prove themselves (automation, energy, renewable energy sources, etc.). Students independently develop their concept of the work/creation they will create, then make a specification of tools and materials and a conceptual sketch of the work. Then the teacher, who is the leader of the extracurricular activity, examines the students' conceptual versions of the works, and corrects and/or supplements them if necessary. The students then acquire the necessary material and tools (if there is a specific need, with the

support of the teacher) and then make the work/creation by applying different techniques of joining/connecting the elements and as a final step, perform tests and tests the correctness of the final product. In all stages of the creation of the finished work, the teacher - the leader of the extracurricular activity - is available to the students as support. The finished works are registered by the head of extracurricular activities for presentation at national and/or state and/or international exhibitions and reviews of works, and some of them are: "Inova - mladi", "Inova – Sajam inovacija i tehničkog stvaralaštva", "mladi@inovacije ", " Sabor mladih inovatora ", "SOELA". At these exhibitions, students present their works to visitors and other participants of the exhibition to an expert committee that evaluates the exhibits, and depending on the placement, students receive a medal and/or recognition for their work.

C. Sample

The research was conducted on a sample (N=36) of fourth-grade students of a secondary vocational school in Rijeka who were studying for the professional qualification Mechatronics Technician. The students of one group (N_A=18) were educated according to the officially prescribed curriculum, while the students of the second group (N_B=18) were educated according to the same curriculum with the difference that they were involved in the extracurricular activity "Young ". innovators", in addition to the traditional learning and teaching process. All students acquired prior knowledge in the field of general education subjects and vocational content in the sector of electrical engineering and computing. The members of both observed groups are equal in terms of general success, which is an exact indicator, but also in terms of interest and motivation in adopting new teaching content, communication with the environment and cooperation with peers, the degree of responsibility towards obligations and independence when completing tasks and the level of systematicity when documenting in teaching process.

D. Data collecting

Since the research paper focused on differences in the affective achievements of students involved in extracurricular activities and students who do not participate in extracurricular activities, data on achievements was collected during the fourth year of schooling, i.e. after two full years of monitoring students during secondary education. For this purpose, instruments were used to assess students' affective achievements. The elements of evaluation were: interest in adopting new teaching content, communication with other participants in the teaching process, cooperation with other participants in classes, timely and complete performance of obligations, keeping notes in a notebook related to the studied teaching content, and independence in solving problem tasks. For each evaluation element, points were assigned to each student, in accordance with the notes in the e-Diary student success administration system and the monitoring of work and progress during the final year of high school, on a scale from 1 to 5, where 1 is not satisfactory, 2 is mostly not satisfies, 3 – partly satisfies, 4 – satisfies to a greater extent and 5 – completely satisfies. The points collected in this way were used for further statistical analysis and determination of differences in achievements between groups.

TABLE 1. RESULTS OF DESCRIPTIVE STATISTICS: MEAN VALUE (M), DOMINANT VALUE (MOD), STANDARD DEVIATION (SD), COEFFICIENT OF VARIATION (CV)

Group	*Achievements*	*N*	*M*	*Mod*	*SD*	*CV*
A + B	Interest	36	4,17	4	0,81	19,46
	Communication	36	4,23	4	0,80	18,87
	Cooperation	36	4,28	5	0,74	17,32
	Responsibility	36	4,00	4	0,93	23,14
	Documentation	36	4,23	5	0,87	20,50
	Independence	36	4,13	4	0,83	20,14

IV. RESULTS

The descriptive analysis of achievements (Table 1) for both groups (N = 36) indicates the results from which it is evident that the achievement of cooperation *(M = 4.28)* has the highest mean score, while responsibility is slightly less pronounced. Looking at the total sample, it is evident that the achievements of independence *(M = 4.17)* and interest *(M = 4.13)* in the vocational content of the STEM field of education should be improved.

The results of the descriptive analysis of the achievements of each group individually (Table 2) indicate that affective achievements are more developed among students of the group that participates in extracurricular activities. Among such students, the higher mean value of the assessment of cooperation stands out (M = 4.67) and communication (M = 4.67) is highlighted, which is very pronounced in their actions within extracurricular activities through cooperation and exchange of knowledge with other individuals of the group, although other achievements also show a high-level development. Students of the group not involved in extracurricular activities have the least developed achievement of responsibility (M = 3.50), since such students are not conditioned by deadlines and obligations that they must fulfil within a certain time frame, but only participate in the performance of regular obligations of the traditional way of schooling.

TABLE 2. RESULTS OF DESCRIPTIVE STATISTICS: MEAN VALUE (M), DOMINANT VALUE (MOD), STANDARD DEVIATION (SD), COEFFICIENT OF VARIATION (CV)

Group	*Achievements*	*N*	*M*	*Mod*	*SD*	*CV*
A	Interest	18	4,56	5	0,62	13,52
	Communication	18	4,67	5	0,49	10,39
	Cooperation	18	4,67	5	0,49	10,39
	Responsibility	18	4,50	5	0,62	13,74
	Documentation	18	4,50	5	0,71	15,72
	Independence	18	4,61	5	0,50	10,88
B	Interest	18	3,78	4	0,81	21,40
	Communication	18	3,78	4	0,81	21,40
	Cooperation	18	3,89	4	0,75	19,50
	Responsibility	18	3,50	4	0,92	26,39
	Documentation	18	3,94	5	0,94	23,78
	Independence	18	3,67	4	0,84	22,92

A statistical comparison of the results of both groups A and B was performed using the t-test for independent samples.

A comparison of the results of the assessment of achievements between the group of students who attend extracurricular activities and students who do not attend extracurricular activities (Table 3) shows that the differences in all achievements, except for the achievement of documentation, are statistically significant at the level of statistical significance $p < 0.05$. The above result clearly indicates that the achievements of interest, communication, cooperation, responsibility and independence are more successful for students involved in extracurricular activities.

TABLE 3. RESULTS OF T-TEST COMPARISON OF ACHIEVEMENTS BETWEEN GROUPS A AND B

Achievements	*t*	*df*	*p*
Interest	3,247	34	0,002622
Communication	4,000	34	0,000324
Cooperation	3,665	34	0,000835
Responsibility	3,817	34	0,000545
Documentation	2,007	34	0,052743
Independence	4,094	34	0,000247

V. DISCUSSION

The results of the research show that significantly more successful achievement of the observed affective achievements is achieved by students who, in addition to attending classes according to the traditional model, also participate in extracurricular activities with an emphasis on activities in the STEM field of education. Through participation in extracurricular activities, students strengthen their actions through cooperation with other activity participants, develop additional interest in professional knowledge, but also improve their independence and responsibility because they take responsibility for the execution of tasks and obligations. This fact is also confirmed by research related to project learning [10], which showed that students should be exposed to work situations and tasks in which their personal reflection on the problem of the task and innovation come to the fore. As proven by previous research, learning imbued with practical work results in more advanced knowledge than learning through traditional teaching principles [13, 14]. The development of the affective domain of achievement increases the quality of the education system, but also the mutual cooperation of all stakeholders in society [15]. One of the possible solutions for modernizing the educational system and enabling the selection of extracurricular activities is the transition from a linear model of learning to learning using puzzles [16], with the help of which the participants very successfully complete the assigned problem-solving tasks, but also spread enthusiasm

VI. RESEARCH LIMITATIONS

The limitations of this research are manifested in the fact that the groups are not completely uniform through appropriate testing, it is a quasi-experimental study because the participants of the groups are uniform according to the general success in the previously completed years of education, according to the grade of governance in the previously completed school years and the criterion of participation in associations that conduct workshops in the STEM field of education. In addition, different teachers approach the activities differently, so the limitation may be related to the teacher's personality and expertise.

The research produced results that indicate the fact that students involved in extracurricular activities achieve better results in the domain of affective achievements, but there is a possibility that the results of the research were influenced by factors that were not taken into account during the implementation of the research, such as the environment in which the members of the sample are surrounded, the influence parents, level of education of parents, interest groups in which members of the sample reside, as well as innate characteristics of the individual within the affective domain of achievement. From the above, it is concluded that the research should be expanded with included factors that can influence the research results, with the aim of obtaining more credible research findings.

VII. CONCLUSION

The conducted research clearly indicates a difference in the affective achievements of students involved in extracurricular activities from students who are not involved in extracurricular activities, with an emphasis on more developed affective achievements in students whose education combines a traditional approach to teaching and extracurricular activities. Significant differences are present in all analysed achievements: interest, communication, cooperation, independence and responsibility. This research indicates that the degree of development of affective achievements is much more emphasized in students who have the opportunity to participate in team work, who are more dedicated to work-based learning and have the opportunity to research depending on their interest, and all activities are accompanied by a time frame for fulfilling obligations and working tasks. The initial assumption that was made at the beginning of the research that the selected affective achievements of the group of students who participate in the extracurricular activities program do not differ statistically significantly compared to the achievements of students who do not participate in these activities, but attend classes according to the traditional teaching model, is also rejected. As previously stated, students who participate in extracurricular activities have a statistically significant development of affective achievements, which greatly helps them in further education, but also in participation in the labour market.

REFERENCES

[1] Simmons, Denise; Creamer, Elisabeth; Yu, Rongrong, (2017), "Involvement in out-of-class activities: A mixed research synthesis examining outcomes with a focus on engineering students", Journal of STEM Education, 18(2), pp.10−16.

[2] Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). Taxonomy of educational objectives: The classification of educational goals, Hand book II: Affective domain. New York: David Mckay Company In corporated.

[3] Anderson, L. W., Krathwohl, D. R. (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives, New York: Longman.

[4] Purković, D., Delač, D. i Kovačević, S. (2022). Interests of Croatian primary school pupils about elective technology teaching and school activities. Metodički ogledi, 29 (1), 167-189. https://doi.org/10.21464/mo.29.1.6

[5] Rennie, Leonie; Venville, Grady; Wallace, John (2018), "Making STEM Curriculum Useful, Relevant, and Motivating for Students", in: Jorgensen, R.; Larkin, K. (eds.), STEM Education in the Junior Secondary, Springer, Singapore. https://doi.org/10.1007/978-981-10-5448-8_6

[6] Schuepbach, Marianne (2015), "Effects of extracurricular activities and their quality on primary school-age students' achievement in mathematics in Switzerland", School Effectiveness and School Improvement, 26(2), 279-−95, https://doi.org/10.1080/09243453.2014.929153

[7] Katherine P. Dabney, Robert H. Tai, John T. Almarode, Jaimie L. Miller-Friedmann, Gerhard Sonnert, Philip M. Sadler & Zahra Hazari (2012) Out-of-School Time Science Activities and Their Association with Career Interest in STEM, International Journal of Science Education, Part B, 2:1, 63-79, DOI: 10.1080/21548455.2011.629455

[8] Han, Xueying; Appelbaum, Ritchard (2018), "China's science, technology, engineering,and mathematics (STEM) research environment: A snapshot", PLoS ONE, 13(4): e0195347, https://doi.org/10.1371/journal.pone.0195347

[9] Differences in the Students' Achievements between Traditional and Project-based Learning of Basic Engineering Competencies: A Quasi-experimental Study D. Purković and M. Prihoda, "Differences in the Students' Achievements between Traditional and Project-based Learning of Basic Engineering Competencies: A Quasi-experimental Study," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 1514-1519, doi: 10.23919/MIPRO48935.2020.9245303.

[10] Changtong, Nidawan; Maneejak, Nantakarn; Yasri, Pratchayapong (2020), "Approaches for implementing STEM (Science, Technology, Engineering & Mathematics) activities among middle school students in Thailand", International Journal of Educational Methodology, 6(1), pp. 185−198. https://doi.org/10.12973/ijem.6.1.185

[11] Tumasheva, Olga; Shashkina, Mariya; Shkerina, Lyudmila; Valkova, Yu (2020), "Elective courses for training the mathematics teachers to realise STEM approach", Journal of Physics: Conference Series, 1691 (2020) 012225, https://doi.org/10.1088/1742-6596/1691/1/012225

[12] The Relationship Between Extracurricular Activities, Student Engagement, and Academic Outcomes, Pollard, Kylah. University of Georgia ProQuest Dissertations Publishing, 2021. 27831180.

[13] Hammond, Linda; Flook, Lisa; Cook-Harvey, Channa; Barron, Brigid; Osher, David (2020), „Implications for educational practice of the science of learning and development", Applied Developmental Science, 24(2), pp. 97-140, https://doi.org/10.1080/10888691.2018.1537791

[14] Extracurricular Management in Improving Student Creativity. Source: Review of International Geographical Education Online . 2021, Vol. 11 Issue 8, p1779-1790. 12p. Author(s): Rukajat, Ajat; Abas, Totoh Tauhidin; Gusniar, Iwan Nugraha

[15] Ashley Casey & Javier Fernandez-Rio (2019) Cooperative Learning and the Affective Domain, Journal of Physical Education, Recreation & Dance, 90:3, 12-17, DOI: 10.1080/07303084.2019.1559671

[16] Changtong, Nidawan; Maneejak, Nantakarn; Yasri, Pratchayapong (2020), "Approaches for implementing STEM (Science, Technology, Engineering & Mathematics) activities among middle school students in Thailand", International Journal of Educational Methodology, 6(1), pp. 185−198. https://doi.org/10.12973/ijem.6.1.185

Academic Motivation of Sophomore and Junior Electrical Engineering Students

A. Gero and B. Catz
Technion – Israel Institute of Technology, Haifa, Israel
gero@technion.ac.il, betoc@campus.technion.ac.il

***Abstract* - Given the importance of motivation in engineering education, the research described here analyzed the factors driving sophomore and junior students to study electrical engineering. The study involved 399 students (192 sophomore students and 207 junior students) at a leading technical university in Israel. At the beginning of the relevant year, the participants completed an anonymous Likert-like scale, based on self-determination theory. According to the results, both sophomore and junior students were mainly driven by autonomous factors. It was also revealed that the intrinsic motivation of junior students was significantly higher than that of sophomore students. Based on the curriculum, the paper provides a possible explanation for these findings.**

Keywords - motivation; electrical engineering; sophomore students; junior students.

I. INTRODUCTION

It is common to analyze learning from different perspectives which are often intertwined. The cognitive point of view deals with thinking processes [1]. The social aspect focuses on interactions between the learner and his/her environment [2], and the affective domain deals, among other things, with academic motivation, i.e., the individual's desire to learn [3].

Motivation plays a central role in higher education, including engineering education. Its importance is validated in view of the ongoing lack, in many parts of the world, of engineers [4]. Therefore, mapping the factors driving engineering students to peruse their studies, despite the difficulties involved, has both theoretical and practical values. Such an analysis may even constitute a preliminary step to reduce dropout.

In light of the above, studies have characterized motivational factors among different groups, such as high-school students majoring in engineering [5], students at two-year technological colleges [6] and mechanical engineering students [7]. The current study focused on analyzing the motivational factors of electrical engineering students, at a leading technical university, at the beginning of their sophomore and junior years. The theory of motivation that served as the framework for the analysis was self-determination theory [8], which is considered one of the most important theories today.

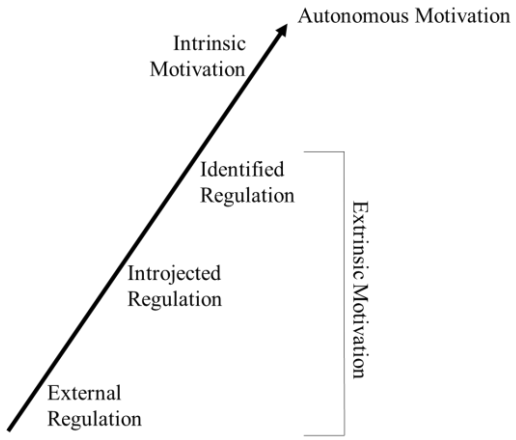


Figure 1. Major motivational factors (self-determination theory)

II. SELF-DETERMINATION THEORY

Self-determination theory suggests that the person strives to satisfy three basic needs [8]:

- Autonomy refers to the person's need to feel that his/her behavior stems from free will.
- Competence refers to the person's need to feel that he/she is capable of meeting objectives.
- Relatedness refers to the individual's need to be in meaningful contact with others.

When some (or all) of the needs are not met, the individual is at a lower level of autonomous motivation, which does not allow for self-actualization. On the other hand, fulfilling these needs leads the person to a higher level of autonomous motivation, which enables self-actualization.

In addition to identifying the basic needs, self-determination theory classifies the factors that drive the individual according to the level of autonomy that characterizes them (Fig. 1). The motivational factor characterized by a maximum level of autonomy is intrinsic motivation. This factor originates from the interest or enjoyment inherent in the activity and is considered long-term [9]. The other motivational factors are extrinsic factors, the most important of them are listed below:

- Identified regulation – a factor arising from the recognition of value associated with the activity. This factor is closest to intrinsic motivation.
- Introjected regulation – a factor stemming from the individual's wish to avoid feelings of guilt or to receive positive feedback from his/her environment.
- External regulation – a factor originating from the person's desire to receive material rewards from the environment or fear of punishment. This factor is the furthest from intrinsic motivation.

To assess the individual's degree of autonomous motivation, it is customary to use the Relative Autonomy Index (RAI) [10]. This measure takes into account the four factors described earlier with appropriate coefficients (suggested by literature) reflecting the degree of autonomy that characterizes each [11]. Therefore, intrinsic motivation and identified regulation have a positive coefficient, while external regulation and introjected regulation have a negative one. Additionally, since the distance between intrinsic motivation and external regulation is the maximum, the coefficient, in absolute value, of each of them is higher than the coefficient, in absolute value, of identified regulation and introjected regulation. The measure is:

$$RAI = 3S_{intrinsic} + S_{identified} - S_{introjected} - 3S_{external} \quad (1)$$

In (1), S_i stands for the score of motivational factor i, measured by a validated tool.

III. Research Goal

The aim of the research was to analyze the motivational factors of electrical engineering students, at a leading technical university, at the beginning of their sophomore and junior years.

IV. Research Setting

The Technion – Israel Institute of Technology is the leading technical university in Israel, and is ranked among the top 100 universities in the world [12]. The Faculty of Electrical Engineering offers a four-year program (BSc) that provides in-depth training in diverse subjects, such as computers, communication, control, nanoelectronics, photonics, electromagnetism and power systems.

TABLE I. Electrical Engineering Program

Year	Main Topics
1	Basic courses: • Mathematics • Physics
2	Introductory courses: • Electronic devices & circuits • Signals & systems • Electromagnetism
3	
4	Elective courses: • Computers • Communication • Control • Nanoelectronics • Photonics • Electromagnetism • Power systems

The freshman student focuses on mathematics and physics (Table I). In the second and third years of study, the student takes introductory courses on electronic devices and circuits, signals and systems, and electromagnetism. In the senior year, the student specializes in elective topics that reflect the broad areas of practice of electrical engineering detailed above. Along with in-depth theoretical training, the students receive practical training in teaching laboratories.

V. Methodology

Three hundred ninety-nine electrical engineering students at the Technion – Israel Institute of Technology (192 sophomore students and 207 junior students) were involved in the study.

At the beginning of the relevant academic year, the participants completed an anonymous questionnaire, designed to assess their motivational factors for studying electrical engineering. This five-level Likert-like scale, ranging from "strongly disagree" to "strongly agree", was based on the concept of Self-Regulation Questionnaire – Academic (SRQ-A) [13] adapted to engineering students.

The instrument consisted of twenty statements that reflected the four motivational factors mentioned above. For instance, the statement "I am studying electrical engineering because I think the studies are interesting" represents intrinsic motivation. As for the three extrinsic factors, the statement "I am studying electrical engineering because I think working in electrical engineering would be a good job for me" reflects identified regulation, the statement "I am studying electrical engineering because my parents want me to do so" expresses introjected regulation, and, finally, the statement "I am studying electrical engineering because I do not have a choice" reflects external regulation. Two engineering education experts validated the statements. As for reliability, Cronbach's alphas were $0.78 \leq \alpha \leq 0.86$, thus indicating acceptable internal consistency.

The data were statistically analyzed and the RAI was calculated. Based on (1) and the fact that the questionnaire was a five-level scale, it turns out that $-16 \leq RAI \leq 16$. It should be noted that since the correlation between some of the motivational factors is high [7], MANOVA was not performed. Instead, *t*-tests with Bonferroni correction were conducted.

VI. Findings

Table II displays the RAI (mean m and standard deviation s) for students at the beginning of their sophomore and junior years. It can be concluded that in both groups of students the average value of the measure is close to the third quartile, i.e., students' autonomous motivation is relatively high. According to a *t*-test (unequal variances), no significant difference between the two groups is observed ($t = 1.08$, $p > 0.05$).

TABLE II. RELATIVE AUTONOMY INDEX

Year	*m*	*s*
2	7.81	3.37
3	8.21	4.04

A more subtle analysis appears in Fig. 2. It shows the distribution of the motivational factors for students at the beginning of their sophomore and junior years. It can be seen that in both cases the highest score is assigned to intrinsic motivation, and that identified regulation is ranked slightly lower, far from introjected regulation (third place) and external regulation (last place).

A *t*-test (equal variances) reveals a significant difference between the two groups of students regarding intrinsic motivation ($t = 3.07$, $p < 0.01$). As for identified regulation ($t = 0.17$, $p > 0.05$), introjected regulation ($t = 0.88$, $p > 0.05$) or external regulation ($t = 1.16$, $p > 0.05$), no significant difference is observed. The effect size associated with the significant gap is small-medium ($d = 0.31$).

As an example, a histogram of responses to the statement "I am studying electrical engineering because I think the studies are interesting" is given in the Appendix.

VII. DISCUSSION

The study shows that the autonomous motivation of electrical engineering students at the beginning of their second and third years of study is relatively high and that they are mainly driven by intrinsic motivation and identified regulation. Other studies also reveal that sophomore and junior engineering students are primarily motivated by these factors [7, 11]. An explanation for this may originate from the fact that students who were mainly driven by introjected regulation and/or external regulation, which are short-term factors, dropped out of their studies during the first year.

In addition, a significant gap (small-medium effect size) was found in intrinsic motivation in favor of the junior students. An explanation for this finding may stem from the characteristics of the curriculum. As described in Section IV, the freshman year focuses on mathematics and physics, and only in the sophomore year the student begins to study the basics of electrical engineering. Therefore, the interest (reflected in intrinsic motivation) of the average student may increase during the second year and manifest itself at the beginning of the third year. An alternative explanation, according to which students who were primarily motivated by extrinsic factors dropped out of their studies during the sophomore year, is not in line with this year's lower dropout rates compared to the first year.

The study had one main limitation due to the fact that only one program was examined. However, the research has both theoretical and practical contributions. The former is the characterization of the motivational factors of electrical engineering students at the beginning of their sophomore and junior years. The latter may be reflected in the application of the findings for the purpose of fostering students' autonomous motivation and reducing dropout. These contributions are becoming more important in view of the ongoing shortage of engineers in many parts of the world [4].

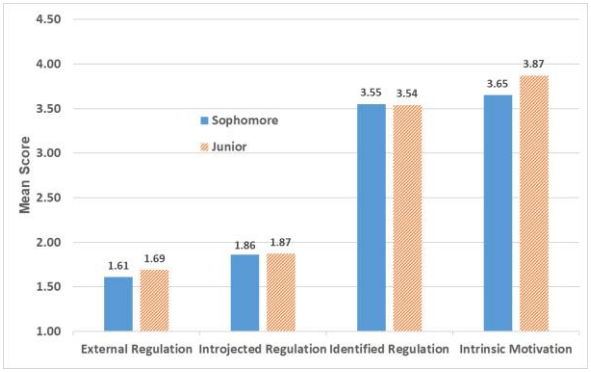


Figure 2. Motivational factors

VIII. CONCLUSION

Electrical engineering students at the beginning of their second and third years of study are mainly driven by autonomous factors. Moreover, the intrinsic motivation of junior students is significantly higher than that of sophomore students.

REFERENCES

[1] J. G. Greeno, A. M. Collins, and L. B. Resnick, "Cognition and learning," Handbook of Educ. Psychol., vol. 77, pp. 15–46, 1996.

[2] R. J. Amineh and H. D. Asl, "Review of constructivism and social constructivism," J. Soci. Sci. Lit. & Lang., vol. 1, pp. 9–16, 2015.

[3] D. H. Schunk, "Self-efficacy and academic motivation," Educ. Psychol., vol. 26, pp. 207–231, 1991.

[4] L. Paku, M. Jensen, and R. Evans, "Attracting future engineers through practical and collaborative initiatives," Proceedings of the 29th Australasian Association for Engineering Education Conference, p. 521, 2018.

[5] A. Gero and O. Danino, "High-school course on engineering design: enhancement of students' motivation and development of systems thinking skills", Int. J. Eng. Educ., vol. 32, no. 1(A), pp. 100–110, 2016.

[6] A. Gero and S. Mano-Israeli, "Analysis of the factors motivating students at a two-year technological college to study electronics", Int. J. Eng. Educ., vol. 33, no. 2(A), pp. 588–595, 2017.

[7] C. Koh, H. S. Tan, K. C. Tan, L. Fang, F. M. Fong, D. Kan, S. L. Lye, and M. L. Wee, "Investigating the effect of 3D simulation-based learning on the motivation and performance of engineering students," J. Eng. Educ., vol. 99, pp. 237–251, 2010.

[8] E. L. Deci and R. M. Ryan, "Self-determination theory," in Handbook of Theories of Social Psychology, P. A. M. Van Lange, A. W. Kruglanski, and E. T. Higgins, Eds., Sage Publications, 2012, pp. 416–436.

[9] L. Yuan, K. Hau, and X. Zheng, "Do both intrinsic and identified motivations have long-term effects?," J. Psychol., vol. 153, pp. 288–306, 2019.

[10] G. Roth, A. Assor, Y. Kanat-Maymon, and H. Kaplan, "Autonomous motivation for teaching: how self-determined teaching may lead to self-determined learning," J. Educ. Psychol., vol. 99, pp. 761–774, 2007.

[11] B. Catz, N. Sabag, and A. Gero, "Problem based learning and students' motivation: the case of an electronics laboratory course", Int. J. Eng. Educ., vol. 34, no. 6, pp. 1838–1847, 2018.

[12] Academic Ranking of World Universities https://www.shanghairanking.com/institution/technion-israel-institute-of-technology

[13] R. M. Ryan and J. P. Connell, "Perceived locus of causality and internalization: examining reasons for acting in two domains," J. Pers. Soci. Psychol., vol. 57, pp. 749–761, 1989.

APPENDIX

A histogram of responses (sophomore students) to the statement "I am studying electrical engineering because I think the studies are interesting" is given in Fig. 3. The five-level Likert-like scale ranges from "strongly disagree" (1) to "strongly agree" (5).

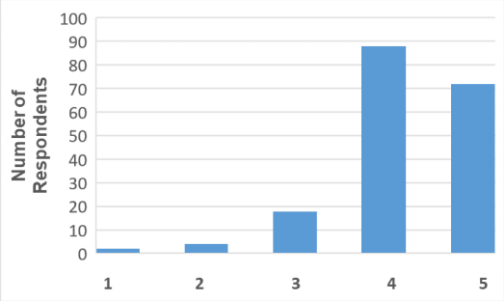


Figure 3. Histogram (sophomore students)

Teaching and Assessing Ethics in a Specialized Professional Skills Course

B. Pejcinovic*
* Portland State University, Electrical and Computer Engineering dept., Portland, OR, USA
pejcinb@pdx.edu

***Abstract* – Ethical considerations are featured prominently in the professional codes of ethics of various engineering professional societies. Arguably, having a code of ethics is a requirement for a field to be recognized as professional. The importance of ethics is further exemplified by accreditation requirements from, e.g., ABET. One of seven ABET Student Outcomes states: "4. an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts". The teaching of engineering ethics would ideally be done across the curriculum and in a proper engineering context. However, due to many practical considerations, this is done in a few, often specialized, courses. In this paper, we report on the implementation of ethics education in one such specialized course at the senior (fourth year) level. We discuss the course structure and how we use rubrics to assess students' attainment of course outcomes related to ethical reasoning. We also provide guidance on scaffolding teaching materials and constructing assignments.**

Keywords* – *accreditation, assessment, engineering education, engineering ethics, social reading.

I. INTRODUCTION

Engineering is not immune to scandals involving various levels of unethical behavior: from cheating on emissions testing [1] to deadly design flaws in aircraft design [2]. Many other examples can be found in books, e.g., [3], but many of those are centered on failures in the area of civil engineering. While these cases present a nice introduction to the problems involved, they do not translate easily into areas of electronics and similar fields where failures and ethical misbehavior may not have as grave a consequence. Nonetheless, the need for recognizing ethical problems and acting accordingly is common across engineering disciplines.

While the need is recognized, there seems to be a debate about whether it is possible to teach ethics to undergraduate students [4]. Several objections to teaching ethics can be put forward, such as: a) ethics is an inborn trait or it is possibly a function of early childhood training, and b) ethical decision-making is not based on reasoning but intuition. While these may be serious philosophical questions there is evidence that teaching ethics does work which justifies our attempts to teach it [4].

Ethical decision-making in engineering has long been recognized as an important part of the profession and individual responsibility. In the following sections, we will briefly describe what constitutes engineering ethics, and provide the curriculum and course context in which we teach engineering ethics. This is followed by the description of assignments and assessment techniques we use with some concluding remarks at the end.

II. ENGINEERING ETHICS

Ethics is variously defined, from a simple dictionary statement such as "a set of moral principles" or "the principles of conduct governing an individual or a group" [5], to more descriptive statements such as "Ethics or moral philosophy is a branch of philosophy that involves systematizing, defending, and recommending concepts of right and wrong behavior" [6]. Ethics is often used interchangeably with morality, but most descriptions will agree that ethics "provides a framework for understanding and interpreting right and wrong in society" [7]. Therefore, ethics deals with the interpretation and application of rules to determine right and wrong behavior or actions.

Most professional societies will have codes of conduct that attempt to describe what these rules are and how to interpret them. Examples include the National Society of Professional Engineers [8], the American Society of Civil Engineers [9], and IEEE [10].

Traditionally, ethics has been taught by philosophy faculty, but business and engineering departments have become involved more recently. For the engineering programs, this was driven in part by the ABET accreditation requirements. In the most recent version of the Student Outcomes, ABET requires that engineering programs demonstrate student attainment of this outcome:

"4. an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts" [11].

This requirement presents difficulties in curriculum design and implementation. Ideally, discussion of ethics would be present across the curriculum, but engineering curricula are already full and ever-expanding. On the positive side, however, there are many resources available to instructors, such as books, e.g., [12], institutions dedicated to the promotion of engineering ethics, e.g., [13], and a variety of websites and professional organizations that have already been mentioned.

III. Curriculum and Course Structure

The Electrical Engineering and Computer Engineering Bachelor of Science programs at Portland State University are quarter-based, 4-year long, and ABET accredited [11]. The quarters are 11 weeks long with one week dedicated to final exams. Like all ABET accredited programs, ours have a culminating design experience in the so-called Capstone course. Questions related to the interaction between engineering design and ethics come up most naturally during the Capstone course. In our curriculum, however, the Capstone course is done in collaboration with local industry and each project has different goals and constraints [14]. This has led us to separate out the discussion of ethics and other professional issues into a required 2-credit course entitled ECE 424 Engineering Professional Practice. The course is taught twice a year in the senior year and the total enrollment varies between roughly 70 and 90 students.

Ethics is specifically targeted in two additional courses in our curriculum:

- ECE 101 Exploring Electrical Engineering which is a freshman course that introduces students to a variety of subfields and some basic engineering skills. Students also learn how to recognize some basic ethical issues through watching videos and in-class activities [15].
- Introduction to Design Processes. As part of the discussion of the design process for the product they are developing, students are asked to consider the potential ethical issues related to their product in global, economic, environmental, and societal contexts (as prescribed by ABET) [16].

Further discussion of ethics may happen in the general studies segment of the curriculum, but this is dependent on the student's selection of specific courses.

A. Engineering Professional Practice Course

The goal of the course is to provide students with skills they likely did not learn in technical courses. This includes effective communication skills, teamwork, intellectual property, career planning, and similar. Specific topics covered in this class include strategies for getting a job, the importance of communication skills (written and oral), teamwork, engineering ethics, global and societal considerations, legal issues such as intellectual property, patents and compliance, and career planning. A full list of course outcomes includes:

1. Explain and follow professional standards of written communication, including email (ABET-3)
2. Describe the purpose and limits of patents and intellectual property rights (ABET-7)
3. Perform a preliminary patent search at uspto.gov (ABET-7)
4. Describe the basic expectations of engineering ethics and the IEEE code of ethics (ABET-4)
5. Demonstrate basic knowledge of and ability with ethical reasoning through the discussion of case studies (ABET-4)
6. Describe the role of ethics in design decisions (ABET-4)
7. Describe current social, environmental, political and economic factors influencing the development and use of technology (ABET-4, 7)
8. Describe how considerations of sustainability affect engineering decisions (ABET-4)

Underlined items directly address engineering ethics and are covered over roughly three weeks.

One problem facing instructors of this, and similar courses, is how to frame the discussion and lead students through the process of making informed ethical decisions. To make the process transparent, we like using an algorithmic approach to explain the decision-making process [17]:

1. State the problem
2. Gather and assess relevant facts in the case
3. Identify the stakeholders
4. Develop a list of at least five options
5. Test options
6. Make a tentative choice
7. Review and make the final choice

Assessment of the course outcomes is discussed in the following section.

IV. Assignments And Assessment

Assignments used in the assessment of student performance include readings, exam questions, team presentations, and surveys. We will discuss only the first three assignments.

A. Readings

In this course we rely on assigned readings to provide additional information and depth of analysis that is not possible to do during the lecture time alone. In the past, there was no reliable way to determine whether students did the reading and how engaged they were with it. As discussed in [18], new tools like Perusall [19] have made this activity visible and subject to more qualitative and quantitative assessment. We have used Perusall since 2017 in courses ranging from freshman introduction to engineering courses to graduate-level research seminars and we believe it has had a positive impact on student learning.

There are three readings:

1. Start with a practical and more conversational approach to describing the types of ethical issues that are likely to arise in the workplace. This could be accomplished through case studies, but they tend to be longer and more technical. Given the time constraints, we use a selection of short chapter(s) from books like [20]. To illustrate the difference between legal liability and social responsibility at a more detailed level, students are asked to read an essay on that topic such as [21]. To help frame the reading, students also read the IEEE Code of Ethics.
2. The second reading deals with moral frameworks and we usually take excerpts from books on the topic or websites with appropriate content. This

provides students with a more complete framework that can be used to analyze cases or practical dilemmas.

3. Finally, we use a contemporary topic to illustrate the nature of the ethical problems likely to come up during their careers. One recent example would be a discussion of ethics in the context of Artificial Intelligence [22].

It is hoped that this progression helps students develop their own understanding of ethical issues but in the context of social reading, where they can discuss readings with their peers who very likely have similar questions and doubts.

Figure 1 illustrates the instructor's interface to Perusall. The central column has a list of assignments, while the details of the assignments are given in the right column. This includes overall class results for the quality of the submitted comments. In this example, 86.8% of students "Completed with maximum score", 0% had "Completed but not maximum score", further 2.6% had "Some work submitted", and 10.5% had "No work submitted." This gives the instructor immediate and easy-to-understand feedback on how well students are engaging with and performing on this assignment.

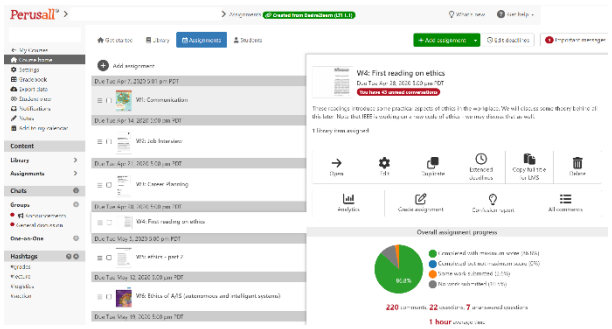


Figure 1. Example of instructor's interface to Perusall. Assignments are listed in the middle in chronological order and details of one of the assignments are on the right.

Perusall reported the following percentages for the reading assignments in one representative section of the ECE 424 course:

TABLE 1. PERCENTAGE OF STUDENTS WHO COMPLETED THE READINGS WITH THE MAXIMUM SCORE.

Reading	1	2	3
% Completed with max score	86.8	78.9	76.3

The overall performance is satisfactory, demonstrating that students are engaging with the material. The drop-off in the later assignments can be attributed to the increased complexity and length of the assignments but it is also likely that students will generally have more other coursework later in the term.

B. Exam questions

Several questions dealing with ethical issues are included in the final exam for the course. Given that the exam is roughly two hours long, the questions cannot be very detailed. Some questions address a basic understanding of the course material, such as basic facts of moral frameworks or the IEEE Code of Ethics. On the other end of the spectrum, we provide a brief case description that can fit into roughly one paragraph of text. Students are then asked to pick one of the multiple-choice answers and justify their selection. This gives the instructor more insight into student thinking and how they arrived at their answers. The dilemmas usually address safety or conflict of interest because these are deemed to be most easily relatable to student experiences.

The following assessment was done by analyzing four questions on the final exam. Here are some overall observations:

- Students have difficulty precisely identifying the ethics issue which then leads to mistaken identification of the applicable segment of the IEEE Code of Ethics, but
- students can identify problems related to safety relatively easily.
- Students have difficulties analyzing problems related to the responsibility to the public or dealing with unethical requests from superiors. This is expected given that these are more complex issues relative to, for example, safety concerns.

This can be summarized by stating that students have a harder time analyzing and providing justification for problems with more ambiguity. This indicates that further training is required. However, time constraints are the limiting factor, and we will have to consider expanding the course hours or eliminating some of the other content.

C. Team Presentations

Students are organized into teams of four and tasked with preparing and delivering a 15-minute presentation about a news item that is relevant to our discussion of ethics. Teams take up a variety of topics, from technical ones such as brain-computer interface, to privacy issues related to Amazon Ring. Presentations are judged by using rubrics that are shared with the class.

Rubrics can be designed in different formats and goals. For example, rubrics may be focused on scoring a particular assignment where the distribution of scores across different criteria is important because that determines the final score, or equivalently the grade for the assignment. Other rubrics may be designed for providing qualitative feedback to students and they are more formative than summative.

The rubric style given in Table 2 is more useful in qualitatively assessing student performance and can be used for providing feedback as well as documenting student attainment for accreditation purposes. Note that in Table 2 we have skipped level 2 for brevity. The number of Performance Levels typically varies between 3 and 5.

TABLE 2. OUTLINE OF A GENERIC RUBRIC IN A TABLE FORMAT.

Criteria	Performance level 3 (e.g. Proficient)	Performance level 1 (e.g. Beginning)
Criterion A. (e.g., Identifies ….)	Performance **indicator 1** at level 3	Performance **indicator 1** at level 1
Criterion B. (e.g. Develops …)	Performance **indicator 2** at level 3	Performance **indicator 2** at level 1

The rubric format given in Table 3 is more appropriate for scoring specific assignments. Giving the "Target characteristic" informs students about what the instructor is looking for and what they should strive to accomplish. It also gives the instructor more flexibility in assigning scores for each criterion while keeping the format simple and easy to navigate. If more than one person is grading the assignments, it may be necessary to break the criteria down into more detailed performance metrics but at the expense of generality.

TABLE 3. OUTLINE OF A GENERIC RUBRIC IN A TABLE FORMAT BUT LISTING ONLY THE TARGET PERFORMANCE INDICATOR.

Criteria	Target characteristics (Excellent or Proficient)	Score
Criterion A. (e.g., Identifies ….)	Performance **indicator 1** at the highest level	(e.g., 60)
Criterion B. (e.g. Develops …)	Performance **indicator 2** at the highest level	

In our previous paper [23], we have reported on the design of the ethics rubric in addition to several others. Table 4 is a slightly edited version of the ethics rubric published there but in the rubric style emphasizing the target characteristics.

TABLE 4. CRITERIA AND PERFORMANCE INDICATORS FOR ABET OUTCOME 4 ON ETHICS.

Criteria	Target characteristic
A. Recognition of ethical dilemma or issue	Clearly identifying the type of ethical dilemma or issue
	Focused, clear, and detailed framing of a dilemma or issue, no irrelevant digressions
	Clearly identifies stakeholders (those affected by the dilemma or issue)
B. Providing relevant information for the case at hand	Produces well-crafted arguments based on new information; justifies assumptions and brings information from their own experiences
	Identifies critical issues & components of the new knowledge
C. Analysis and fairness	Provides clear arguments
	Considers multiple points of view in their analysis; global view with perspectives from e.g. employers, professions, and society
D. Ethics in professions	Familiar with the value and importance of professional codes
	Understands personal, professional, and wider social consequences of violations of codes

Student team presentations were assessed in Fall 2019 using criteria in Table 4 with the results presented below in Figure 2. Five out of 13 teams did not approach a proficient level of performance. This is due to a couple of reasons:

- Ethical issue or dilemma was not properly identified, i.e., it is not focused enough, or it is not explicitly stated.
- Some topics dealt strictly with technical issues which were very informative and relevant but had no obvious ethical component.
- If the core ethics issue is poorly identified then criteria B. Providing relevant information for the case at hand, and C. Analysis and fairness, become difficult to decouple and assess independently.
- The logistics of the class interfere with student performance, so some teams are asked to deliver presentations before we have an in-depth, in-class discussion of ethics-related issues.

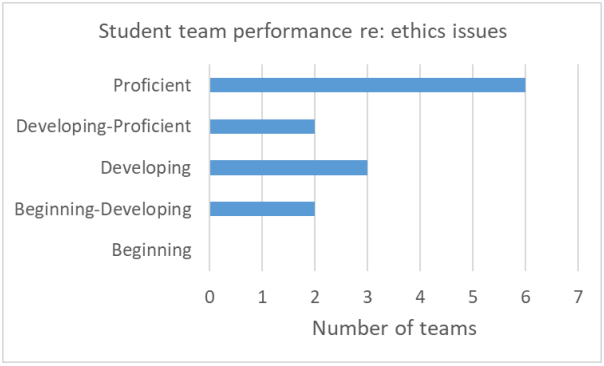


Figure 2. Assessment of team performance on the news presentation assignment. The assessment was done using the ethics rubric.

In the future, the news assignment can be improved by

- Directing teams to be even more explicit about identifying stakeholders and ethical dilemma or issue.
- Emphasizing during in-class activities that all perspectives need to be considered.
- Have teams propose their topic first and present an outline to the instructor before digging in.

V. CONCLUSIONS AND FUTURE WORK

We have presented a rationale for introducing a course dealing with engineering ethics during the senior year of undergraduate electrical and computer engineering programs. We have also provided details of how such a course can be implemented. The main constraint is the relatively low number of contact hours (lectures) which can be compensated somewhat by activities outside of the classroom. One such activity is reading additional literature on topics of interest. Students have shown good engagement with this type of social reading using the Perusall website.

As expected, students find it difficult to handle ambiguity stemming from more complex ethical dilemmas. Students perform satisfactorily on the exam questions but there is room for improvement. Performance on the team presentation assignment can also be improved but there is an additional complexity due to potential problems related to teamwork and scheduling. The latter will be addressed by adjusting the course schedule and the former by giving more direct instruction on how to approach the assignment with examples of good and bad presentations and team organization.

Ultimately, however, we have to look for ways to introduce ethical issues in other courses in the third and fourth years of study. This would not only decrease the time pressure to cover various topics but would also place the issues in their proper engineering context making it easier for students to realize their importance.

REFERENCES

[1] Volkswagen's Emisions Evasion, web page and video, Ethics Unwrapped, Accessed March 23, 2023. [Online]. Available: https://ethicsunwrapped.utexas.edu/video/volkswagens-emissions-evasion

[2] J. Kerkert, J. Borenstein, and K. Miller, "The Boeing 737 MAX: Lessons for Engineering Ethics," Sci Eng Ethics, vol. 26, no. 6, pp. 2957-2974, 2020. doi: 10.1007/s11948-020-00252-y.

[3] H. Petroski, To engineer is human: The role of failure in successful design, St Martins Press, 2018.

[4] G. R. Johnson, "Can Engineering Ethics Be Taught," in The Bridge, Engineering Ethics, vol. 47. No. 1, National Academy of Engineering, 2017. Accessed: Mar. 12, 2023. [Online]. Available: https://www.nae.edu/168649/Can-Engineering-Ethics-Be-Taught

[5] Merriam-Webser dictionary, definition of "ethic", web page, Accessed on March 13, 2023. [Online]. Available: https://www.merriam-webster.com/dictionary/ethic .

[6] Ethics, Wikipedia, Wikimedia Foundation, Accessed March 13, 2023. [Online]. Available: https://en.wikipedia.org/wiki/Ethics

[7] Ethics – Ethics Unwrapped, web page, Accessed March 13, 2023. [Online]. Available: https://ethicsunwrapped.utexas.edu/glossary/ethics

[8] Code of Ethics, National Society of Professional Engineers, Accessed March 13, 2023. [Online]. Available: https://www.nspe.org/resources/ethics/code-ethics

[9] Code of Ethics, America Society of Civil Engineers, Accessed March 13, 2023. [Online]. Available: https://www.asce.org/career-growth/ethics/code-of-ethics

[10] IEEE Code of Ethics, Accessed March 13, 2023. [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html

[11] ABET Criteria for Accrediting Engineering Programs 2020-2021, Accessed March 16, 2022. [Online]. Available: https://www.abet.org/wp-content/uploads/2020/09/EAC-Criteria-2020-2021.pdf

[12] C.E. Harris Jr., et al., Engineering Ethics: Concepts and Cases, 6th edition, Cengage Learning, 2018.

[13] National Institute for Engineering Ethics (NIEE), web page, Accessed March 13, 2023. [Online]. Available: https://www.niee.org/

[14] M. Faust, A. Greenberg, and B. Pejcinovic, "Redesign of senior capstone program in electrical and computer engineering and its assessment," Frontiers in Education Conference Proceedings, Oct. 2012, pp. 1–6. doi: 10.1109/FIE.2012.6462390.

[15] P. Wong, M. Holtzman, B. Pejcinovic, and M. Chrzanowska-Jeske, "Redesign of Freshman Electrical Engineering Courses for Improved Motivation and Early Introduction of Design," ASEE Annual Conference and Exhibition, Vancouver, Canada, 2011, p. 22.1224.1-22.1224.13. [Online]. Available: https://peer.asee.org/18541

[16] B. Pejcinovic, "Introducing Cornerstone Courses into Electrical and Computer Engineering Curriculum," 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2019, pp. 1427–1432. doi: 10.23919/MIPRO.2019.8756847.

[17] M. Davis, Ethics and the University, Routledge, London, 1999, pp. 166-167.

[18] B. Pejcinovic, "Use of social reading in engineering education," in 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), May 2022, pp. 1336–1340. doi: 10.23919/MIPRO55190.2022.9803672.

[19] Perusall collaborative annotation tool, Accessed on March 1, 2022. [Online]. Available: https://perusall.com/

[20] Cal Selinger, Stuff You Don't Learn in Engineering School, IEEE, 2004.

[21] Engineering Design: Literature on Social Responsibility Versus Legal Liability (Essay #4), [Online]. Available: https://ethics.tamu.edu/wp-content/uploads/sites/7/2017/04/ethics.tamu_.edu_ethics_essays_design.pdf

[22] Ethically Aligned Design, Version 1, IEEE, Accessed March 13, 2023. [Online]. Available: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v1.pdf

[23] B. Pejcinovic "Design of Rubrics for Student Outcomes in 2019-2020 ABET Criteria," 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020, DOI: 10.23919/MIPRO48935.2020.9245228.

Levels and Differences in Professional Identity in Engineering Undergraduates

V. Putarek and J. Petrović*
*University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
vanja.putarek@gmail.com

***Abstract* - Professional identity is a form of social identity which is developed with the adherence to perform competently and legitimately in profession. It is the basis of professional functioning and socio-cultural capital for employability. Since engineers play a crucial role in today's rapidly accelerating technological development, it is important for them to have a stable and strong professional identity. The aim of this research was to examine the level and differences in professional identity depending on sociodemographic characteristics of Croatian electrical engineering and computing students. Research participants were 431 freshmen students (73.1% male; age: *M* = 18.67, *SD* = 0.57). Participants completed an online survey containing questionnaires measuring three professional identity dimensions (knowledge about future profession, direct experience with future profession, and self-efficacy related to future profession) and socio-demographic characteristics. The results showed that the self-efficacy dimension of professional identity was higher than the other two dimensions. Male students, students informed about the study program, students who have engineer(s) among close family members and who were raised in the place of study, as well as students who work student job have higher levels of some dimensions of professional identity.**

Keywords - engineering education; professional identity; engineering students; engineering undergraduates; self-efficacy

I. INTRODUCTION

Social identity represents the ways in which a person's self-concepts are based on their group membership [1]. According to the Social identity theory [2], [3], people have multiple social identities (e.g., identity related to gender, friends, family, occupations, sports, religion) and a group is defined and it is meaningful only in terms of how it is related or compared to other groups. Social identification shifts between different groups, depending on which social identity is currently most salient [4]. Social identities have an impact on the individual's perceptions, emotions, and behavior [5]. For instance, they can affect choices which people make, the perception of people in other groups (e.g., stereotypes), the norms and values which will be followed [6].

As a form of social identity, professional identity is one of many social identities which people have and usually is viewed as a key component of one's overall sense of identity [7]. Tan et al. [8] define professional identity as "the self that has been developed with the commitment to perform competently and legitimately in the context of the profession, and its development can continue over the course of the individuals' careers". It serves as a cognitive structure that assimilates and integrates information about self and profession [9].

Professional identity begins to develop mostly in adolescence [10]. During this period, most adolescents achieve a mature level of professional identity, but professional identity development for most individuals continues through adulthood. Professional identity helps individuals to make rational career decisions [11] and predicts a successful transition from school to work [12].

A. Dimensions of professional identity

Professional identity is not a unitary construct and there are various approaches to conceptualize its multidimensionality. On the one hand, in some research, professional identity is operationalized via variables, such as self-efficacy and motivation [13]. On the other hand, research on students defined dimensions of professional identity. Tan et al. [14] revealed five dimensions of professional identity. First dimension, "knowledge about professional practices", consists of "knowing that" and "knowing how". Namely, this dimension is related to domain knowledge, as well as the appropriate application of competences in practice. Second dimension, "having the professional as a role model", is described as the contact with people who work in the profession of individual's interest, in order to observe how to behave and reason appropriately in some working role. Powerful role models are teachers who have worked or are still working in the respective industries, family members or relatives and friends. "Experience with the profession" is the third dimension, which depicts students' authentic contact with the future profession. Through that contact students can realize what they really have learnt in the classroom and what the responsibilities of practitioners are in their future profession. Fourth dimension, "preference for a particular profession", describes the personal identification with learning to become the professional, which leads to higher levels of grit and motivation during learning. Finally, "professional self-efficacy" is the fifth dimension that is related to the individual's perception that (s)he has the competences to adequately perform actions required in some situation.

In Croatian validation of Tan's et al. [14] dimensions of students' professional identity, three dimensions were obtained: "knowledge about professional practices", "experience with the profession", and "perception of the role model and professional future" [15]. First two dimensions correspond to Tan's et al. [14] dimensions of the same name, while the third dimension consists of Tan's et al. [14] dimensions labelled as "having the professional role" and "professional self-efficiency". The difference in

the number of obtained dimensions can be related to the differences in samples, that is in [15] study there were more nursing students, while in Tan's et al. [14] study participants were from various fields. Considering that professional fields differ in practices and requirements, it can be assumed that dimensionality of professional identity should be conceptualized for the specific profession.

B. Engineering identity

Rapid globalization and changes in contemporary society are largely stimulated and depend on technological development. In the 21st century science, technology, and innovation are the crucial factors for state and global economy, as well as social well-being [16]. Consequently, education policy makers try to encourage pupils to choose STEM (Science, Technology, Engineering, Mathematics) fields for their career path [17]. Moreover, engineers are highly sought-after employees in many countries [18], but the retention of employees in various professions nowadays is quite challenging [19]. Taking into account that higher levels of professional identity are related to the lower intention to job turnover [9], it would be important to examine the professional identity of engineers.

Engineers primarily base their professional identity on their technical competences [20]. Other important aspects of engineers' professional identity are creativity and innovation, while non-technical aspects, such as communication, ethics, social impact, and lifelong learning, are not seen by engineers as crucial for their professional identity [21]. These aspects represent technical/social dualism. That is, one part of engineers' professional identity relies on the technical aspects of engineering profession, while the other part is related to more heterogeneous, managerial, social and human aspect of abovementioned profession [22].

Considering the relevance of engineers and their professional identity, in our research we were interested in the better understanding of engineers' professional identity, with the emphasis on the professional identity in future engineers. More specifically, we examined the professional identity of engineering undergraduate students in Croatia. We were focused on the level of professional identity dimensions, defined by the Croatian version [15] of Tan's et al. [14] dimensions of professional identity. That is, we examined students' knowledge about engineers' practices, experience with engineering, and perception of the role model and professional future. To our knowledge, there is no research in Croatia regarding the professional identity of engineers, with the focus on engineering undergraduate students. Better understanding of the professional identity of future engineers may help educators and employers to build stable and strong attachment to the profession, which may result in positive professional outcomes and well-being [9].

C. Differences in professional identity

There are numerous personal and contextual factors which shape professional identity and depending on the exposure to those factors, some individual differences in professional identity can emerge [9]. The personal factor, which is extensively explored in scientific research, as a potential influence on professional identity, is gender. Recent studies demonstrated little or no gender differences in the professional identity development or the outcomes of professional identity [9]. These results are explained by the relevance of the working role in overall identity in women nowadays. Moreover, in order to balance the personal and professional life, as well as to solve the conflict between various roles more than men, it is assumed that women need to develop a clear and straightforward career path, interests, and identity.

However, one field, in which gender differences in professional identity are repeatedly demonstrated, is engineering. Over centuries, engineering is described and perceived as a masculine discipline [23]. Women are often seen in the workplace more as women and invisible as engineers [24]. According to [20], there are four ways in which interactions marginalize women's professional identity as an engineers and overly emphasize their gender identity: 1) gender is overly amplified, 2) gendered expectations are imposed, 3) women are tuning out, 4) doubts about technical abilities.

Differences between male and female engineers were observed in the perceived importance of some aspects of work more than in the perceived skill development [25]. That is, women prefer face-to-face communication compared to written communication, which is favored by men. Women place a higher value on social skills and relations than men counterparts. In other words, in terms of technical/social dualism [22], men are more technically oriented, while women are more socially oriented.

Research on engineering graduates demonstrated that male engineering graduates form their professional identities more easily compared to the female ones [26]. Moreover, female engineering graduates perceive the transitions from the study to the workplace a little bit more difficult than their male counterparts, because women struggled more to construct the engineering 'selves'.

Considering previous findings about gender differences in engineers' professional identity, in our research we examined whether gender differences in professional identity could be found as early as in the beginning of the engineering study.

Contextual factors also have an important influence on professional identity development and can yield to individual differences in professional identity. Impact of family characteristics is extensively explored and results are inconclusive [9]. On the one hand, it seems that dysfunctional family characteristics endanger professional identity development. On the other hand, some research (e.g., [27], [28]) demonstrated that family characteristics, such as the quality of family relationships or the degree of control, have very small or no effects on professional identity development. However, one type of influence from family on professional identity, which was not explored, can be rooted in the profession of family members. As Tan et al. [14] suggested, one dimension of professional identity is "having the professional as a role model" and family members can serve as an important work role model. Therefore, we decided to examine whether having engineers among close family members is related to higher levels of engineers' professional identity.

Participation in leisure and organized extracurricular activities or community services can facilitate identity development. However, high school students who are employed usually do not form a more mature professional identity and are not more decisive regarding professional choices compared to their peers who do not have jobs [29]. These results are attributed to the type of work of high school students. That is, high school students usually have unskilled manual jobs unrelated to their professional interests and future profession. However, employment in adolescence can have a postponed effect, helping individuals in psychological transition to work role (e.g., encouragement of responsibility, self-confidence) [30]. In higher education students have more opportunity to have jobs important for their professional development and aspired profession. Taking into account the key role of authentic work experiences for professional identity development in higher education [31], it would be important to explore whether students who work some student job would have higher levels of professional identity.

Finally, one important determinant of professional identity among higher education students can be the place of growing up. That is, students who have to move to the place of study have more mental health problems than students who did not have to move to the place of study [32], [33]. They also have to adapt their identity to a new living context and build a lot of new skills. There is no research on the influence of moving to the place of study on professional identity, but it would be important to explore whether individuals who move to the place of study are at risk to have lower professional identity, which can lead to dropout [34].

D. The current research

The aim of this research was to examine the level and differences in professional identity of Croatian undergraduate engineering students. We conceptualize professional identity as a three-dimensional construct, using the Croatian version [15] of Tan's et al. [14] dimensions of professional identity. That is, we defined professional identity through 1) students' knowledge about engineers' practices, 2) experience with engineering and 3) perception of the role model and professional future. In our research we were interested in the level of professional identity dimensions, as well as whether there are any differences in the level between these dimensions. To our knowledge there are no studies which explored those questions, therefore we are not able to define any specific hypothesis. In Gusar's et al. research [15] the trend of the lowest scores on the dimension "experience with the profession" can be seen, but differences between dimensions were not tested and study was not conducted on the engineering students. Therefore, we can assume a similar trend in our research, but the data on which we make such an assumption are not strong enough to form a specific hypothesis, so we assumed a null hypothesis that there were no differences between professional identity dimensions.

We also examined gender differences in professional identity dimensions. Considering previous research (e.g., [20]), we can hypothesize that female engineering students would have lower levels of professional identity than male engineering students. This hypothesis is supported by the findings that male engineering students have a stronger sense of identification with their future profession even before starting university compared to the female ones [26].

We also explored whether differences in professional identity depend on how informed students were about the study program before enrolling in the study, having engineers among close family members, and working a student job. Engineering professional interests and career plans are relatively stable over the high school period [35]. That is, by the end of high school, students usually have a clear picture that they want to become engineers and they have basic information about their future profession and what they need to master to perform it. Therefore, we can assume that students with higher levels of professional identity would be those students who informed themselves more about the study.

Moreover, considering previously mentioned research [9], [14], we assumed that students who have engineers among close family members and who work at student job would have a higher level of professional identity.

Finally, we examined whether students who had to move to the place of study (i.e., did not grow up in the same place as the place of study) have different levels of professional identity compared to students who grew up in the same place where they started their studies. We can assume that students who had to move to the place of study would have lower professional identity compared to students who did not have to move to the place of study, because they might have identity struggles and consequently their professional identity may not be so in focus. However, there is no research on this question so we cannot define any specific hypothesis, so we assumed a null hypothesis that there were no differences in professional identity between students who moved and those who did not move in the place of study.

In short, our research questions were:

- RQ1: Examine the level of professional identity and whether the dimensions of professional identity have different levels among undergraduate engineering students.
- RQ2: Examine whether the professional identity among undergraduate electrical engineering students differs depending on gender, how informed students were about the study program before they got enrolled in it, having engineers among close family members, working a student job, and place of growing up (is the place of growing up the same as the place of study or not).

II. METHODS

A. Participants

A convenience sample of 431 (73.1% male; age: M = 18.67, SD = 0.57) students of University of Zagreb Faculty of Electrical Engineering and Computing participated in this research. Students were mostly enrolled in the first year (98.6%), and some were enrolled in the second (1.2%)

or the third (0.2%) year of undergraduate study during the academic year of 2022/2023. Most students (75.6%) had achieved the highest possible grade on the final exam in mathematics at the end of high school. Only a few students (5.3%) reported working part-time jobs.

B. Procedure

The Ethics Committee of the University of Zagreb Faculty of Electrical Engineering and Computing had given approval of the study. The survey was conducted online. At the beginning of the survey, it was emphasized that students' participation was voluntary, and that information related to students' identity will only be used to assign them a symbolic amount of assignment credits in the course within which they were invited to participate in the survey. Students were also informed that they were allowed to terminate their participation at any time. Students completed the survey in 15-20 minutes.

C. Measures

In this section we will describe measures that were used. The factor structure obtained in confirmatory factor analysis (CFA), as well as the reliability of the measures in our research, will be shown in section *Data analyses*.

Professional Identity Five-Factor Scale (PIFFS [14]) was used to measure students' professional identity. It consists of 25 items divided in five factors: *knowledge of professional practice* (six items; "I know the nature of the work I will do in my future profession."), *professional experience* (six items; "I know personally some people who work in my future profession."), *role model in profession* (five items; "I admire professionals who are already working in my future work environment."), *self-efficiency* (six items; "I'm confident that I can do an excellent job in the future."), and *affinity to certain profession* (two items; "I am already pretty sure what kind of profession I will enter after completing the polytechnic or university education."). Response scale was from 1 ("definitely not correct") to 5 ("definitely correct"). Tan et al. [14] reported the reliability of the original five subscales $\alpha = 0.65–0.85$. PIFFS was validated in Croatia by [15]. In their research a three-factor solution was obtained, with factors Knowledge ($\alpha = 0.84$), Experience ($\alpha = 0.79$) and Perception of the role model and professional future ($\alpha = 0.84$).

D. Data analysis

In order to examine potential differences in professional identity t-test for independent samples was used, while for testing the differences in the level of the dimensions of professional identity repeated-measures ANOVA was used.

Prior to the aforementioned analysis, instrument exploratory (EFA) and confirmatory factor analysis (CFA) were performed for PIFFS in order to examine whether indicators were loaded to some extent on a latent variable (i.e., factor structures of our measures were examined). The results of EFA and CFA will be shown in this section, while the results that will answer our research questions will be shown in the *Results* section.

For EFA, SPSS statistical package, version 20 was used. Factors were extracted using principal axis factoring, with direct oblimin rotation (delta = 0). Eigenvalues greater than one, scree plot test and parallel analysis were used as the criteria for factor interpretation.

The R statistical package, version 4.1.3 was used to conduct CFA. To interpret models fit, several goodness-of-fit measures were taken into account: (a) χ^2 test (statistical insignificance confirms a good fit) [36]; (b) χ^2/df (good fit: values of less than 5) [37]; (c) comparative fit index (CFI) and Tucker-Lewis index (TLI) (good fit: values greater than .95) [38]; (d) a root mean square error of approximation (RMSEA) (good fit: values of less than .06) [38]; and (e) a standardized root mean square residual (SRMR) (good fit: values of less than .08) [37].

Professional Identity Five-Factor Scale in [14] obtained a five-factor structure, but with one scale with only two items, which represents fewer items than suggested (e.g., [39]). Their study participants were students of polytechnic in Singapore. Croatian validation study [15] obtained three-factor structure on students in different study areas, not only technical studies. Considering differences in previous research in factor structures and samples [14], [15], we conducted EFA using the random half of the sample. Prerequisites for factor analysis were met: KMO = 0.825; Bartlett's test of sphericity $p < .001$ with $\chi^2(120) = 1162.85$). EFA suggested a three-factor structure. The first factor explained 26.86% of variance, the second factor explained 9.58% of variance, and the third factor explained 7.57% of variance. A total of five items with cross-loadings and loadings lower than .3 – .4 [36] were not included in final solution. First factor, labelled Knowledge has seven items (e.g., "I know the nature of the work I will do in my future profession.") and examines students' knowledge of their future profession. Second factor, labelled Experience, has five items (e.g., "I am part of an interest group related to my computing/engineering.") and examines how many direct experiences with their future profession students already have had. Third factor, labelled Self-efficacy, has four items (e.g., I'm confident that I can do an excellent job in the future.") and examines students' level of self-efficacy related to their future work. Using the other random half of the sample, we conducted CFA for two-factor solution and obtained good model fit: $\chi^2 = 177.27$, $p < .001$, $df = 100$, $\chi^2/df = 1.77$, CFI = .927, TLI = .913, RMSEA = .063, RMSEA 90% CI [.047 .077], SRMR = .061. Modification indices indicated that one correlation between errors of items should be included because of items' wording. In our research the result on the subscales was calculated as the mean of all participant ratings. Internal consistency was .85, .68 and .74 for Knowledge, Experience, and Self-efficacy subscales respectively.

III. RESULTS

Regarding the RQ1, the obtained results displayed in the Table 1 suggest students' levels of professional identity vary with the highest level of professional identity being related to the dimension of Self-efficacy, followed by Knowledge, and finally Experience. The data demonstrated that the level of Knowledge dimension was around the average point of the response scale. The level of Self-efficacy dimension was above the average point,

while the level of Experience dimension was below the average point of the response scale.

Regarding the RQ2, the results presented in Table 2 demonstrated that male students score significantly higher than female in dimensions of Knowledge and Experience. Students informed about the study had significantly higher levels of all three dimensions of professional identity. Furthermore, students who have engineers among close family members have significantly higher levels of Experience dimension. Students who are working student job score significantly higher in dimensions of Knowledge and Experience. Finally, students whose place of growing up is the same as the place of study score significantly higher in dimensions of Experience and Efficacy.

TABLE 1. DIFFERENCES IN THE LEVELS OF THREE DIMENSIONS OF PROFESSIONAL IDENTITY (N = 431)

	$M(SD)$	F	$df1,df2$	η^2	$\Delta_{M1,M2}(SE)$	$\Delta_{M1,M3}(SE)$	$\Delta_{M2,M3}(SE)$
1. Knowledge	3.32 (0.80)	822.53	1.89, 812.02	0.66	1.05 (0.04)***	-0.77 (0.4)***	-1.82 (0.05)***
2. Experience	2.27 (0.87)						
3. Self-efficacy	4.09 (0.74)						

Note: ***$p < .001$.

TABLE 2. DIFFERENCES IN PROFESSIONAL IDENTITY

			Knowledge	Experience	Efficacy
Gender	Women (N=115; 26.93%)	$M(SD)$	3.08 (0.84)	1.86 (0.77)	4.06 (0.83)
	Men (N=312; 73.07%)	$M(SD)$	3.40 (0.77)	2.42 (0.86)	4.10 (0.72)
		$t(df)$	3.80 (425)***	6.51 (224.66)***	0.44 (425)
		Cohen's d	0.40	0.69	
Informed about study	Yes (N=202; 46.87%)	$M(SD)$	3.58 (0.78)	2.57 (0.89)	4.25 (0.66)
	No (N=229; 53.13%)	$M(SD)$	3.10 (0.75)	2.01 (0.77)	3.94 (0.78)
		$t(df)$	6.48 (429)***	6.92 (399.70)***	4.42 (428.22)***
		Cohen's d	0.63	0.67	0.43
Engineer(s) in family	Yes (N=170; 39.44%)	$M(SD)$	3.35 (0.78)	2.42 (0.88)	4.06 (0.75)
	No (N=261; 60.56%)	$M(SD)$	3.30 (0.82)	2.17 (0.85)	4.10 (0.74)
		$t(df)$	0.609 (429)	3.04 (429)**	0.541 (429)
		Cohen's d		0.29	
Lived in the place of study	Yes (N=198; 45.94%)	$M(SD)$	3.28 (0.82)	2.16 (0.85)	3.97 (0.78)
	No (N=233; 54.06%)	$M(SD)$	3.36 (0.79)	2.36 (0.87)	4.19 (0.70)
		$t(df)$	0.96 (429)	2.35 (429)*	2.96 (429)**
		Cohen's d		0.23	0.30
Student job	Yes (N=23; 5.34%)	$M(SD)$	3.70 (0.99)	2.83 (1.24)	4.23 (0.87)
	No (N=408; 94.66%)	$M(SD)$	3.30 (0.79)	2.24 (0.84)	4.08 (0.74)
		$t(df)$	2.31 (429)*	2.24 (23.14)*	0.92 (429)
		Cohen's d		0.56	

Note: *$p < .05$, **$p < .01$, ***$p < .001$.

IV. DISCUSSION

The aim of this research was to examine the levels and differences in professional identity of Croatian undergraduate engineering students. The level of professional identity ranges from below average to above average, depending on the professional identity dimension. Students did not have a lot of direct experience with the future profession (e.g., being part of an interest group related to the profession; interaction with professionals in the industry). Their knowledge about future profession, rules and regulation in the industry or roles and responsibilities of future job was on average level, while their self-efficacy about their future professional competences was above average. Differences in levels between three dimensions supported this trend. Just like in [15], Experience dimension is the lowest-scored dimension of professional identity, which is not surprising given that for most students this was their first semester and they didn't have much time to gain experience related to their future profession. Students' level of self-efficacy was the highest and above average. Considering that self-efficacy is related to achievement and occupational aspirations choice [40], it is promising for future employers that first-year students believe in their professional competences. Moreover, self-efficacy is related to students' grit [41], so we can assume that first-year students would have a lower inclination to drop out from higher education. However, students' institution as well as higher education institutions in general should invest in further professional identity development, because recent studies have highlighted that identity status is a dynamic and flexible process of exploration and commitment, rather than a static commitment to some identity [9], [42]. Professional identity in higher education can particularly be fostered by active engagement in activities related to the future profession and authentic experiences [31].

Regarding the differences in professional identity, our data are mostly in line with the hypothesis and previous research results (e.g., [20]) suggesting that male students would have higher levels of professional identity than female students. This difference, however, was not

obtained for the self-efficacy dimension. On the one hand, previous research demonstrated lower self-efficacy, in terms of using knowledge of engineering to solve relevant problems, in female students, including first-year students [43]. Previous research had also showed that, while academic self-efficacy in the context of engineering was higher in male students, female students had higher levels of career self-efficacy [44]. This finding is in line with the interviews with female engineering students who reported high career aspirations at the beginning of their studies [26], which may suggest that they were confident in their abilities. Unfortunately, during studies, their career aspirations declined, which may be attributed to gendered interactions and increase in interest in parenthood. Despite that, it seems that career self-efficacy amplifies during study [45]. Altogether, for future career paths, it is encouraging that female engineering students in our research had a high level of self-efficacy and that this level is as high as the self-efficacy level of their male counterparts. However, as suggested in [31], it would be important to construct activities fostering direct experience with the profession, and female students may profit more from those activities, given their lower knowledge and experience regarding their future profession.

As we assumed, students who were more informed about the study have higher levels of all three dimensions of professional identity. Engineering identity and interest towards engineering are mostly formed by the end of high school [35]. Higher interest is related to in depth engagement in activities associated with the subject of interest [46]. That is, students with higher levels of interest in engineering may search for more information related to their future profession, including study programs. Therefore, more information about the study program may reflect students' already formed professional identity, which has led them to collect more information about it. However, the relationship can be reversed: more information about the study program may lead to higher professional identity because students know what to expect in their future profession and that can facilitate professional identity development. To solve this question, longitudinal studies are recommended and needed.

Students who have engineers among close family members had a higher level only of Experience dimension of professional identity. Similar to previous research [27], [28], it seems that family has a small influence on professional identity. From adolescence onwards, individuals are inclined to question family members' work preferences rather than introject their opinions and work roles [47]. Therefore, it is possible that students in our research who have engineers as close family members wanted to build a career path, collecting knowledge, and fostering self-efficacy, on their own. However, they profit from exposure to engineering in the family, which was apparent from our data. That is, students who have engineers among family members could have easy access to engineering activities, interest groups, and direct experiences related to engineering.

Moving to the place of study is an enormous stressor [32], [33]. It is not surprising that students who had to move to Zagreb to study engineering had lower levels of self-efficacy and experience related to their future profession. These students have to adapt to new surroundings, separate from family and some peers, and promptly acquire some adult skills. Therefore, their self-efficacy related to future profession can be blocked by numerous stressors and skills to which they need to devote themselves. Some students may come from small places in which they perhaps do not have enough opportunity for direct experience with engineering. However, it seems that their professional aspirations and interests lead them to acquire as much declarative knowledge about their future profession as their counterparts from Zagreb.

Finally, our hypothesis regarding differences in professional identity depending on whether student work a student job was partly supported. That is, students who had a student job had a higher level of Experience dimension of professional identity. The partial support of our hypothesis can be explained by the characteristics of our sample. Namely, only 23 students (5.3 %) have a student job, and from those who had a student job, only 10 of them (43.5 % of students who had a student job) work in the field of engineering. A larger sample would be needed to provide support for our hypothesis.

V. LIMITATIONS AND FUTURE DIRECTIONS

Our research had a correlational research design, which makes it impossible to draw causal conclusions. Furthermore, all measures were self-reports, so students could respond socially desirable. The generalization of the results is limited because the participants were students of only one engineering faculty in Zagreb.

For future research it would be important to include longitudinal research design in order to determine possible changes in professional identity during study. It would also be important to have more students who have a student job in engineering field to make stronger conclusions about the relationship between student jobs and professional identity.

VI. CONCLUSIONS

The aim of this research was to examine the professional identity levels and differences in Croatian undergraduate engineering students. Research results were obtained from a sample of 431 electrical engineering and computing undergraduate students. Professional identity differences and levels were assessed using an online survey measuring three professional identity dimensions: knowledge about future profession, direct experience with future profession, and self-efficacy related to future profession, as well as socio-demographic characteristics.

Research results showed that self-efficacy dimension of engineering students' professional identity was higher than knowledge about future profession and direct experience with future profession dimensions. Male students, students informed about the study program, students who have engineer(s) as family members, students who were raised in the place of study, as well as students who have a student job, have all reported a higher level of direct experience with future profession and, in some cases, also one of the other two dimensions of professional identity. In line with [48], our results suggest that pre-professional socialization including having engineer(s) as family members and

having acquired information about the study program positively affects formation of professional identity, especially regarding the Experience dimension.

The obtained results have implications for higher education institutions' policies and practices. More emphasis should be placed on supporting professional identity formation, especially for female students. It might be of best interest for higher education institutions to familiarize their students with their programs, even before they enroll, both for fostering their professional identity, but also to better prepare them for challenges that await them. Finally, opportunities which would enable students to gain practical experience with their future profession would make a valuable addition to higher education institutions to foster professional identity formation, both on institution level, connecting students with their future employers, or on course level, providing students with insights into their future profession.

REFERENCES

[1] H. Tajfel, Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations. European Association of Experimental Social Psychology by Academic Press, 1978.

[2] H. Tajfel, The social psychology of intergroup relations. European Association of Experimental Social Psychology by Academic Press, 1978.

[3] H. Tajfel and J. C. Turner, "An integrative theory of intergroup conflict," in The social psychology of intergroup relations, W. G. Austin and S. Worchel, Eds., Monterey, CA: Brooks/Cole, 1979, pp. 33–37.

[4] T. Mussweiler, S. Gabriel, and G. V. Bodenhausen, "Shifting social identities as a strategy for deflecting threatening social comparisons," J. Pers. Soc. Psychol., vol. 79, no. 3, pp. 398–409, Sep. 2000, doi: 10.1037//0022-3514.79.3.398.

[5] N. Ellemers, R. Spears, and B. Doosje, "Self and social identity," Annu. Rev. Psychol., vol. 53, pp. 161–186, 2002, doi: 10.1146/annurev.psych.53.100901.135228.

[6] S. D. Reicher, R. Spears, and T. Postmes, "A Social Identity Model of Deindividuation Phenomena," Eur. Rev. Soc. Psychol., vol. 6, no. 1, pp. 161–198, Jan. 1995, doi: 10.1080/14792779443000049.

[7] J. Kroger, Identity development: Adolescence through adulthood, 2nd ed. in Identity development: Adolescence through adulthood, 2nd ed. Thousand Oaks, CA, US: Sage Publications, Inc, 2007, pp. xiv, 305.

[8] C. P. Tan, H. T. Van der Molen, and H. G. Schmidt, "A measure of professional identity development for professional education," Stud. High. Educ., vol. 42, no. 8, pp. 1504–1519, Aug. 2017, doi: 10.1080/03075079.2015.1111322.

[9] V. B. Skorikov and F. W. Vondracek, "Occupational identity," in Handbook of identity theory and research, Vols. 1 and 2, New York, NY, US: Springer Science + Business Media, 2011, pp. 693–714. doi: 10.1007/978-1-4419-7988-9_29.

[10] O. Malanchuk, E. E. Messersmith, and J. S. Eccles, "The ontogeny of career identities in adolescence," New Dir. Child Adolesc. Dev., vol. 2010, no. 130, pp. 97–110, 2010, doi: 10.1002/cd.284.

[11] D. Saunders, P. Gary, J. Sampson, and R. Reardon, "Relation of Depression and Dysfunctional Career Thinking to Career Indecision," J. Vocat. Behav., vol. 56, pp. 288–298, Apr. 2000, doi: 10.1006/jvbe.1999.1715.

[12] R. Lapan, "Career development across the K-16 years: Bridging the present to satisfying and successful futures.," Jan. 2004.

[13] E. T. Canrinus, M. Helms-Lorenz, D. Beijaard, J. Buitink, and A. Hofman, "Self-efficacy, job satisfaction, motivation and commitment: exploring the relationships between indicators of teachers' professional identity," Eur. J. Psychol. Educ., vol. 27, no. 1, pp. 115–132, Mar. 2012, doi: 10.1007/s10212-011-0069-2.

[14] C. P. Tan, H. T. Van der Molen, and H. G. Schmidt, "A measure of professional identity development for professional education," Stud. High. Educ., vol. 42, no. 8, pp. 1504–1519, Aug. 2017, doi: 10.1080/03075079.2015.1111322.

[15] I. Gusar, A. Tokić, and R. Lovrić, "Psychometric Analysis of the Professional Identity Questionnaires in Croatian Nursing Student Context," SAGE Open Nurs., vol. 7, p. 23779608211044600, 2021, doi: 10.1177/23779608211044600.

[16] "Science, Technology and Innovation in the New Economy." OECD, Sep. 2020. [Online]. Available: https://www.oecd.org/science/inno/1918259.pdf

[17] M. Caprile, G. Dente, R. Palmén, and P. Sanz, Encouraging STEM studies: labour market situation and comparison of practices targeted at young people in different Member States. LU: Publications Office of the European Union, 2015. Accessed: Apr. 24, 2023. [Online]. Available: https://data.europa.eu/doi/10.2861/519030

[18] T. J. Kennedy and M. R. L. Odell, "Engaging Students in STEM Education," Sci. Educ. Int., vol. 25, no. 3, pp. 246–258, 2014.

[19] "Ageing and Employment Policies: Retaining Talent at All Ages." OECD, 2023.

[20] D. M. Hatmaker, "Engineering identity: Gender and professional identity negotiation among women engineers," Gend. Work Organ., vol. 20, pp. 382–396, 2013, doi: 10.1111/j.1468-0432.2012.00589.x.

[21] J. R. Morelock, "A systematic literature review of engineering identity: definitions, factors, and interventions affecting development, and means of measurement," Eur. J. Eng. Educ., vol. 42, no. 6, pp. 1240–1262, Nov. 2017, doi: 10.1080/03043797.2017.1287664.

[22] K. L. Tonso, "Engineering Identity," in Cambridge Handbook of Engineering Education Research, A. Johri and B. M. Olds, Eds., Cambridge: Cambridge University Press, 2014, pp. 267–282. doi: 10.1017/CBO9781139013451.019.

[23] J. Jorgenson, "Engineering Selves: Negotiating Gender and Identity in Technical Work," Manag. Commun. Q., vol. 15, no. 3, pp. 350–380, Feb. 2002, doi: 10.1177/0893318902153002.

[24] W. Faulkner, "Doing gender in engineering workplace cultures. II. Gender in/authenticity and the in/visibility paradox," Eng. Stud., vol. 1, pp. 169–189, Nov. 2009, doi: 10.1080/19378620903225059.

[25] J. Naukkarinen and S. Bairoh, "Gender differences in professional identities and development of engineering skills among early career engineers in Finland," Eur. J. Eng. Educ., vol. 47, no. 1, pp. 85–101, Jan. 2022, doi: 10.1080/03043797.2021.1929851.

[26] V. Papafilippou and L. Bentley, "Gendered transitions, career identities and possible selves: the case of engineering graduates," J. Educ. Work, vol. 30, no. 8, pp. 827–839, 2017, doi: 10.1080/13639080.2017.1375088.

[27] B. K. Hargrove, A. G. Inman, and R. L. Crane, "Family Interaction Patterns, Career Planning Attitudes, and Vocational Identity of High School Adolescents," J. Career Dev., vol. 31, no. 4, pp. 263–278, 2005, doi: 10.1177/089484530503100404.

[28] B. Hargrove, M. Creagh, and B. Burgess, "Family Interaction Patterns as Predictors of Vocational Identity and Career Decision-Making Self-Efficacy," J. Vocat. Behav., vol. 61, pp. 185–201, 2002, doi: 10.1006/jvbe.2001.1848.

[29] M. J. Zimmer-Gembeck and J. T. Mortimer, "Adolescent Work, Vocational Development, and Education," Rev. Educ. Res., vol. 76, pp. 537–566, 2006, doi: 10.3102/00346543076004537.

[30] J. Mortimer, M. Zimmer-Gembeck, M. Holmes, and M. Shanahan, "The Process of Occupational Decision Making: Patterns during the Transition to Adulthood," J. Vocat. Behav., vol. 61, pp. 439–465, Dec. 2002, doi: 10.1006/jvbe.2002.1885.

[31] F. Trede, R. Macklin, and D. Bridges, "Professional identity development: a review of the higher education literature," Stud. High. Educ., vol. 37, no. 3, pp. 365–384, 2012, doi: 10.1080/03075079.2010.521237.

[32] R. Hechanova-Alampay, T. A. Beehr, N. D. Christiansen, and R. K. Van Horn, "Adjustment and strain among domestic and international student sojourners: A longitudinal study," Sch. Psychol. Int., vol. 23, pp. 458–474, 2002, doi: 10.1177/0143034302234007.

[33] I. Živčić-Bećirević, S. Smojver-Ažić, M. Kukić, and S. Jasprica, "Akademska, socijalna i emocionalna prilagodba na studij s obzirom na spol, godinu studija i promjenu mjesta boravka," Psihol. Teme, vol. 16, no. 1, pp. 121–140, 2007.

[34] D. Verdín, A. Godwin, A. Kirn, L. Benson, and G. Potvin, "Understanding How Engineering Identity and Belongingness Predict Grit for First-Generation College Students," Sch. Eng. Educ. Grad. Stud. Ser., Apr. 2018, [Online]. Available: https://docs.lib.purdue.edu/enegs/75

[35] P. M. Sadler, G. Sonnert, Z. Hazari, and R. Tai, "Stability and volatility of STEM career interest in high school: A gender study," Sci. Educ., vol. 96, no. 3, pp. 411–427, 2012, doi: 10.1002/sce.21007.

[36] T. A. Brown, Confirmatory factor analysis for applied research. in Confirmatory factor analysis for applied research. New York, NY, US: The Guilford Press, 2006, pp. xiii, 475.

[37] S. G. West, A. B. Taylor, and W. Wu, "Model fit and model selection in structural equation modeling," in Handbook of structural equation modeling, New York, NY, US: The Guilford Press, 2012, pp. 209–231.

[38] L. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," Struct. Equ. Model., vol. 6, pp. 1–55, 1999, doi: 10.1080/10705519909540118.

[39] R. B. Kline, Principles and practice of structural equation modeling, 3rd ed. in Principles and practice of structural equation modeling, 3rd ed. New York, NY, US: Guilford Press, 2011, pp. xvi, 427.

[40] A. Bandura, C. Barbaranelli, G. V. Caprara, and C. Pastorelli, "Self-efficacy beliefs as shapers of children's aspirations and career trajectories," Child Dev., vol. 72, no. 1, pp. 187–206, 2001, doi: 10.1111/1467-8624.00273.

[41] E. L. Usher, C. R. Li, A. R. Butz, and J. P. Rojas, "Perseverant grit and self-efficacy: Are both essential for children's academic success?," J. Educ. Psychol., vol. 111, pp. 877–902, 2019, doi: 10.1037/edu0000324.

[42] A. Brown, S. Kirpal, and F. Rauner, Eds., Identities at Work. Dordrecht: Springer Netherlands, 2007. doi: 10.1007/978-1-4020-4989-7.

[43] K. M. Whitcomb, Z. Y. Kalender, T. J. Nokes-Malach, C. D. Schunn, and C. Singh, "Comparison of self-efficacy and performance of engineering undergraduate women and men," Int. J. Eng. Educ., vol. 36, no. 6, pp. 1996–2014, 2020.

[44] C. Burger, J. Raelin, R. Reisberg, M. Bailey, and D. Whitman, "Self-efficacy in female and male undergraduate engineering students: Comparisons among four institutions," Jan. 2010.

[45] R. Marra, K. Rodgers, D. Shen, and B. Bogue, "Women Engineering Students and Self-Efficacy: A Multi-Year, Multi-Institution Study of Women Engineering Student Self-Efficacy," J. Eng. Educ., vol. 98, Jan. 2009, doi: 10.1002/j.2168-9830.2009.tb01003.x.

[46] M. Ainley, "Students' interest and engagement in classroom activities," in Handbook of research on student engagement, New York, NY, US: Springer Science + Business Media, 2012, pp. 283–302. doi: 10.1007/978-1-4614-2018-7_13.

[47] P. Fadjukoff, L. Pulkkinen, and K. Kokko, "Identity Processes in Adulthood: Diverging Domains," Identity, vol. 5, no. 1, pp. 1–20, Jan. 2005, doi: 10.1207/s1532706xid0501_1.

[48] M. Tomlinson and D. Jackson, "Professional identity formation in contemporary higher education students," Stud. High. Educ., vol. 46, no. 4, pp. 885–900, 2019, doi: 10.1080/03075079.2019.1659763.

Virtual and Augmented Reality in Mechanical Engineering Education

G. Ivanova*, A. Ivanov** and L. Zdravkov*
* University of Ruse/Department of Computing, Ruse, Bulgaria
** University of Ruse/Department of Machine tools and manufacturing, Ruse, Bulgaria
giivanova@uni-ruse.bg , akivanov@uni-ruse.bg

Abstract **- Nowadays innovative technologies are entering the educational process with the aim of learners easily understanding and learning the material provided. The technologies used and/or based on virtual and augmented reality in the educational process affects today's digital generation and motivation of learners, giving them the opportunity to use their smartphones in the learning process. The aim of the paper is to show the main positive aspects of 3D Virtual Reality environments and 3D Augmented Reality Learning environments in mechanical engineering education. An interactive Mobile Apps for 3D Augmented Reality Learning environment and 3D Virtual Reality Learning environment for setting up and operating a variety of Grinding Devices for Cutting tool sharpening are presented. The Mobile Apps are using Mobile Virtual Glasses suitable for smartphones. The use of 3D Virtual learning and 3D Augmented learning reality in the training of Cutting Tools has been tested with students from the Faculty of Mechanical and Manufacturing Engineering of the University of Ruse. With the Mechanical engineering students, comparative research was conducted to determine the effectiveness of the learning and students' experience level in a 3D Augmented and Virtual Reality Learning environment versus that of a real learning laboratory.**

Keywords – virtual reality, augmented reality, 3D educational process, 3D laboratory, 3D virtual reality, 3D augmented reality

I. INTRODUCTION

Nowadays, technologies and methods of obtaining information that are used require a different level, and new approaches for training in universities to prepare students for the intellectual challenges of the new era. There is a need to develop new educational tools that improve the understanding of the taught material and at the same time speed up the learning process and teach students in engineering specialties with innovative educational technologies. The ideas that are most frequently used to improve the effectiveness of the educational process include replicating the content in an interactive form, if possible, most frequently by making three-dimensional models of the things that are included in the learning material. The popularity of these methods is due to the fact that 80% of learning new material is visual, while only 20% is verbal, i.e. these technologies use the way the human cognitive apparatus is designed for maximum information absorption [1]. Modern educational practices incorporate cutting-edge technologies, such as virtual and augmented reality [2, 3], to make it easier for students to comprehend and absorb contemporary learning content in a fun and engaging way. Today's digital generation students are motivated and more engaged in using technologies based on virtual and augmented reality in the educational process.

One example of a 3D educational program using virtual and augmented reality is AR Liver Viewer, which is utilized in medicine [4]. It displays in real-time accurate anatomical models of the liver. The app was created by a team of anatomy experts, certified medical illustrators, animators, and programmers exclusively for the iPad and New iPad using real CT data and the most precise 3D modeling technologies. The program has two modes of operation. In virtual mode, the user can view the model from all angles, zoom in and out, and focus on a different highlighted area thanks to the 3D display of the model on the application screen. In augmented reality mode, the model is visualized using a marker and the iPad's camera, placing it in actual space in front of the tablet. The ability to rotate the model along all axes, which enables analyzing the appearance and various regions of this organ, expresses the contact with the object.

The usage of markers for viewing three-dimensional objects is noticeable when looking at the current augmented reality solutions applied in the educational process [5]. Due to how easy it is to create applications and show models using this method, marker-based visualization has become the standard method for visualizing 3D models in augmented reality. The fact that some information about the studied objects is visualized in both applications under consideration, whether directly in the program or beforehand in a textbook, leaves an impression. The two augmented reality applications both run on smartphones, but they are designed for different operating systems: one is for iOS and the other is for Android.

InCell VR is an educational game focused on virtual reality technology [6]. The game is of a competitive type and in it the learner is placed in a computer-generated world representing a human cell. The interaction in the game of the learner with the world inside is achieved through the so-called control through gaze (gaze input) and the phone's gyroscope. Initially, the student is provided with an introductory-level option (tutorial). The educational game presents the organelles and other components necessary for human cells in an intriguing

way. For instance, proteins have a positive impact on the player's characteristics by giving him more life points and accelerating his movement during the race. In contrast, macromolecules harmful to the cell, in this case viruses, have a negative impact by lowering the player's hit points and slowing down his movement.

For the Cardboard and Daydream platforms, Google built Google Expeditions, which uses augmented and virtual reality for educational purposes. The software offers learners the chance to explore 3D models and 360-degree scenes from various locations around the world. Both iOS and Android devices can use the software. Use the phone's screen or, if one is available, a joystick to select or modify a scene. The user can explore fully produced settings in virtual reality applications, and in the case of Google Expeditions, they are obligated to fully investigate the virtual world. Because both apps are designed for smartphones, it is simple to flip the learner 360 degrees. Since a smartphone is used, these applications are made to work well with them and can be used by many people. Information regarding the objects under consideration is provided by both apps [7].

II. VIRTUAL AND AUGMENTED REALITY FOR MECHANICAL ENGINEERING EDUCATION

A. Comparison of different types of virtual reality helmets

Virtual reality has grown in popularity over the past few years as a result of better picture quality, more user-friendly programs, and more cheap virtual reality helmets. The broad adoption of virtual reality helmets among consumers was also made possible by the hardware advancement of computer components, particularly the video card, a crucial component in virtual reality games and applications.

The helmet, which covers the user's eyes and visually hide the outside environment, is a crucial element when utilizing virtual reality computer programs. In this way, the user is focused on the computer-generated reality, and depending on how realistically the objects are rendered, and how well the helmet and displays are made, the user of the application will have a particular experience. This is called the level of immersion, and in general, in virtual reality, due to the presence of helmets that "hide" the real world from the user, the level of immersion is high.

There are two main types of virtual reality helmets: mobile based virtual reality helmets using mobile phones and computer-based virtual reality helmets, which are more expensive. The technology and software of smartphones have reached a point of development that enables them to run virtual reality apps. The virtual reality helmets used with mobile phones have a distinct structural design from the ones used with desktop or laptop PCs. The helmets, which are used with personal computers, have built-in displays in them. There are two displays total in the "goggles," one for each eye. Without using a common display, each eye receives a separate simulation of the perspective of the created world in this manner.

When smartphone helmets are used, things are different. In most cases, the user sees the smartphone's single display, which is split into two sections for the left and right eyes, respectively. Each eye can focus exactly on the area of the smartphone screen where the image of the computer-generated reality is being displayed thanks to the magnifiers in these "glasses". Of course, for now, the level of "immersion" in computer-generated reality is far higher with virtual reality "glasses" used with personal computers than it is with smartphones since the resolution of LCD displays and their refresh ability is much higher than that of the smartphone display. However, the refresh rate of modern smartphone screens, which is more than 60Hz, is sufficient for them to work effectively with computer-generated content, and the resolution increases with each new generation of these devices. Table 1 presents a comparison between different VR helmets.

TABLE I. A COMPARISON BETWEEN VR HELMETS – TYPE 1

Name	**View angle, °**	**Resolution, pixel**	**Refresh rate, Hz**
Oculus rift	110	2160 x 1200 (1080 x 1200 per each eye)	90
HTC Vive	110	2160 x 1200 (1080 x 1200 per each eye)	90
Daydream	100	Depending on the smartphone	Depending on the smartphone
Gear VR	96	2560 x 1440	60
Zeiss VR ONE	100	Depending on the smartphone	Depending on the smartphone
FreeflyVR Beyond	120	Depending on the smartphone	Depending on the smartphone

As can be seen from the table, the refresh rate of smartphones is lower than that of displays in helmets that connect to a computer, but not by much - on the order of 30Hz.

In order to use virtual reality helmets with a personal computer, it is required to use USB and HDMI cables between the glasses and the computer. As a result, the user's ability to move and rotate 360 degrees in the virtual environment is limited by the presence of cables that, when rotated, have a certain length, and have the potential to twist. This poses a risk of damage to the glasses and/or the PC as well as the user's safety while wearing the glasses.

Since cellphones are used with virtual reality headsets, this issue is avoided because there is no need for wires because the device is housed inside the "glasses." This permits 360-degree panning and rotation, allowing for the viewing and exploration of every point in the computer-generated world.

The following Table 2 compares several virtual reality helmets based on the device type — wired or wireless — and other factors that are taken into consideration while picking eyewear.

The table shows that smartphone-based helmets allow the operator more mobility because they don't require a cable connection.

The sensors required to function in the virtual world, such as a gyroscope, etc., are another component of the table that need our attention. Such sensors are not included in smartphone helmets because they are part of the phone itself. Therefore, the same sensors can also be utilized in mobile device virtual reality applications.

TABLE II. A COMPARISON BETWEEN VR HELMETS - TYPE 2

Name	Type	Brand	Price less than, €	Software bundle	Sensors
Oculus rift	Cable	Oculus (Facebook)	700	Oculus SDK, SteamVR SDK	Accelerometer, gyroscope, magnetometer
HTC Vive	Cable	HTC	730	SteamVR SDK	Accelerometer, gyroscope
Daydream	Wireless	Google	120	Daydream/ Cardboard SDK	From the smartphone
Gear VR	Wireless	Samsung, Oculus	110	Oculus SDK	From the smartphone
Zeiss VR ONE	Wireless	ZEISS	150	Daydream/ Cardboard SDK	From the smartphone
FreeflyVR Beyond	Wireless	FreeFlyVR	65	Daydream/ Cardboard SDK	From the smartphone

Different virtual reality helmet manufacturers offer different software packages for application creation, with smartphones primarily adopting the Daydream/Cardboard technologies' Daydream software package.

The price of the smartphone is not considered as can be seen from the table and the price difference between the two types of helmets because these devices are typically used for various purposes and their price value is also influenced by other factors and usage scenarios. Because of this, computer helmets do not factor in the cost of personal computers. It is important to note that PCs and smartphones with varying price ranges will perform differently in virtual reality programs, accordingly.

B. Design Process of Virtual and Augmented Reality Applications for Mechanical Engineering Education

The main purpose of the paper is to present 3D virtual reality and 3D augmented reality apps for educational purposes and more specific for mechanical engineering education. Both applications were created at the University of Ruse and are used to educate students in the course of Cutting Tools. Students learn about various cutting tools with extremely complex shape and cutting geometry [8], their parameters and characteristics, how they work [9], as well as how to repair them if they become worn out.

The Cutting Tools textbook's figures are recognized as markers by the augmented reality software, which then launches corresponding 3D files of the studied machine tools, orienting devices, or cutting tools. In this way, students can visually get to know the studied objects in much bigger details. The second application is for a virtual reality. The students even in pandemic situation can enter a virtual laboratory room, where the sharpening processes of various cutting tools, the cutting tools themselves and the orienting devices can be examined in details. For both applications students can simply use their smartphones.

There are many software solutions for building three-dimensional models, which at the same time offer the possibility of optimizing the models. The same programs also offer the option of texturing the objects, but there are also specially created applications for this purpose. Regarding augmented reality technology – there is also a choice of technological solutions serving to implement this possibility in mobile applications. With smartphone virtual reality, we have a narrower choice of technologies provided to implement such applications.

Programs for creating 3D models can be divided into two groups - those that create accurate technical models (CAD systems), but at the same time do not create visually attractive models. The second type of programs are those used in making games, movies, advertisements, etc. These programs create the so-called CGI (Computer-generated Imagery) and they aim for a better visual representation of the object, as the technical parameters remain in the background. However, the programs through which CGI is created also allow an accurate description of the dimensions of the modeled object, which, together with the greater possibility of visual modeling and presentation of the objects, is the preferred choice for creators working in the field of augmented reality and not only.

Regarding the creation of augmented and virtual reality applications, the Unity 3D computer game creation software is mainly used. To implement the two applications, implementing the two different technologies, a solution should be chosen from the available software packages, which will best support and facilitate the development of the two applications and the possibilities it offers to enable the realization of the set goals.

For the needs of augmented and virtual reality applications, models of machines used in practical training in the discipline "Cutting Tools" were selected. Some of the models that are related to the figures in the textbook were selected specifically in the context of the teaching material of the printed textbook. Machine models are created in a CAD modeling computer program. The peculiarity here is that with these programs accurate engineering models are created, in the context of three-dimensional CAD programs. These programs are used by engineers to create accurate models that are then, most often, created in the real world. The models are mathematically accurate and consider all the characteristics of the object - size, the material used to manufacture it, distances between parts, etc.

Because of this detail required in engineering designs, the skeleton of 3D CAD models consists of multiple faces to make the model as detailed as possible. For this reason, when visualizing such models in a professional setting, computers with powerful and expensive CPUs and video chips are used, which allow the three-dimensional models to be visualized and interacted with visually, such as being viewed from different sides, in a fast and smooth way.

Visualizing such models using commercially available PCs and smartphones is a labor-intensive task due to CPUs with lower frequency and cores and GPUs with significantly fewer cores, clock speeds, and memory. In this case, the smartphone will be used to work with the

applications - a device whose hardware capabilities are even lower than those of personal computers. Therefore, these models need to be optimized in order to run the applications correctly.

The main method that was used is called retopology. This is a method in which the skeleton of the three-dimensional model (mesh) is recreated in an optimal way without losing visual information. Mainly, retopology consists of reducing the number of faces and converting them from triangles and polygons to squares or rectangles. The optimized model, whose surface now consists of fewer quadrilateral faces, is easier to work with in texturing programs and requires less device hardware resources to render.

Machine tools 3D models were optimized following the following steps:

- Removing duplicate vertices from the model skeleton. Three or more of these vertices form the face of the model skeleton. When you delete one of these vertices, the face collapses and is removed from the model. This is the first method that is used;
- The next step is to convert some of the triangular faces from the skeleton into quadrilaterals by merging the adjacent triangles into a quadrilateral. This again reduces the number of faces, and already quadrilateral faces, as mentioned earlier, are more convenient for texturing and visualization;
- This is followed by removal of geometry that is not needed for the presentation of the model, but remained after the work on it. Such geometry is, for example, angles without length, quadrilaterals with zero area value, etc. Although this geometry does not exist visually, it is respected by the program and added to the model, therefore it is taken into account when rendering it by the application and occupies graphics memory;
- The number of vertices and faces is reduced by a method called "decimate". Through various parameters, such as face-vertices ratio, groups of vertices per area, etc. the number of faces in the model is reduced;
- Finally, a method is resorted to, by which persons are united to the so-called boundary angles, i.e. angles separating different planar regions of the model geometry. Thus, in different areas of the model, faces are minimized, generally reducing the number of faces in the model geometry without leaving gaps in the model skeleton.

Following these steps visually optimizes model geometry and reduces the number of faces, ensuring proper texture placement and rendering with fewer hardware resources in augmented and virtual reality applications. The optimization was performed using the Blender programming environment. The ultimate goal is for models to have a geometry face count of up to 30,000. This number takes into account the hardware capabilities of modern and older devices so that models are rendered smoothly and interact smoothly without traffic interruption.

C. Design of Virtual Reality Application for Mechanical Engineering Education

The virtual environment for the virtual reality application is a laboratory room with a table in front of the user's "eyes" on which the models will be visualized. The lab was created entirely in Blender 3D modeling software. The final appearance of the virtual room is presented on Figure 1.

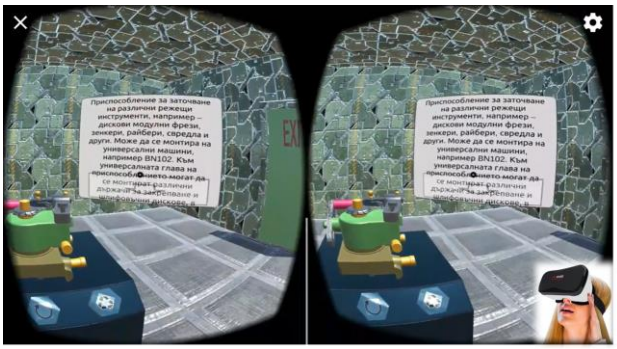


Figure 1. Virtual Reality Laboratory room

Model of an orienting device for sharpening a disc modular cutter

The model is an orienting fixture device holding firmly various cutting tools still during sharpening or resharpening the cutting edges, for example – disc modular cutters, countersinks, reamers, drills, etc. The orienting device can be mounted on universal sharpening machines, for example BN102. Various clamping holders and grinding discs can be used / shown, depending on the tool to be sharpened. The sharpening machine also allows making various settings, such as limiting the longitudinal movement of the table. The model was initially built from 45,941 faces with varying numbers of face angles. Although the number of faces is acceptable, optimization is needed to avoid problems in creating the texture of the model. After optimization, most of the faces of the model are in the form of quadrilaterals, and the number of faces is reduced to 17,062. Thus, after optimization, the geometry of the model is significantly cleaner, i.e. the shape and area of the polygons is optimized, which is a prerequisite for smooth texturing of the model, and the smaller number of faces means relatively less hardware load on the learner's device. The model before and after transformation is presented on Figure 2.

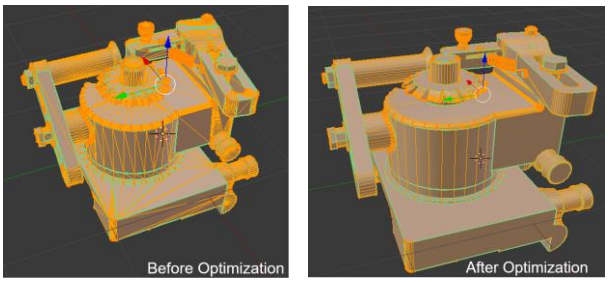


Figure 2. 3D model of orienting device for sharpening a disc modular cutter before and after optimization

Model of an orienting device for sharpening a worm modular cutter (hub)

Initially, the geometry of the sharpening machine model consisted of 83,386 faces. The model consists of

multiple polygons with more than 4 corners, which when texturing would create a problem in the texturer. The number of faces of the model before optimization would also create a problem when rendering the model and applying transformations, related to more work on the CPU and graphics chip and the resulting heat dissipation and battery drain. After applying the optimization steps explained above, the model retained its appearance without loss of visual information. Now most of the polygons are quadrilaterals and the faces of the model geometry are now 22,359. So the model will be easier to texture and the smartphone hardware will be significantly less loaded when rendering and transforming the model.

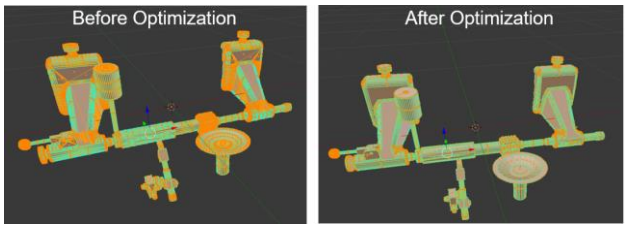


Figure 3. 3D model of the orienting device for sharpening a worm modular cutter (hub) before and after optimization

Unity Game Engine along with Google Cardboard SDK is used to create the virtual reality application. The application works with a single scene, which is essentially the virtual environment. The app also has built-in text visualization objects so that the information needed for each model can be visualized in the virtual lab (Figure 4).



Figure 4. Text visualization functionality

Worm modular cutter (hub)

The hub cutter model consists of 20,712 faces. This number of faces fits within the initially established upper limit, but we have polygons from the model skeleton with a large number of corners and this would cause difficulties when applying the texture.

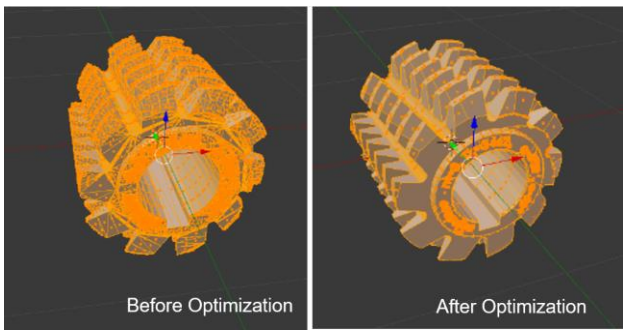


Figure 5. 3D model of the hub cutter before and after optimization

After optimization, most of the faces of the geometry were converted to quadrilaterals and the number of faces was reduced to 3,314, without loss of visual information. The model before and after optimization is presented on Figure 5.

Additional challenge which we faced with was the way users work in the Virtual reality environment. The presented Virtual reality application is providing "Rotation" functionality around two of the axes which are horizontally positioned. Using Rotation around only two of the coordinate axes gives simplicity in the way the users operate and work with the Virtual models, the functionality is shown on Figure 6.

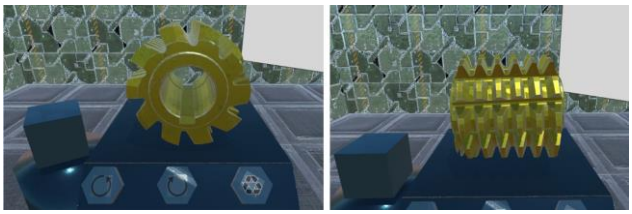


Figure 6. Two axis rotation of the hub cutter 3D model

D. Design of Augmented Reality Application for Mechanical Engineering Education

Since the augmented reality application is marker-based, markers need to be identified and embedded to be scanned by the application and the three-dimensional model to be visualized accordingly. Due to the fact that this application will be based on an already existing textbook that is used in face to face education, it was decided that the markers will actually be figures from the textbook, Figure 7. This way, there is no need to create an entirely new marker system to embed in the textbook and reprint it in large quantities. Using ready-made figures from the textbook to create the markers saves a huge amount of financial and time resources that would otherwise have to be booked to creating a textbook to work with the application.

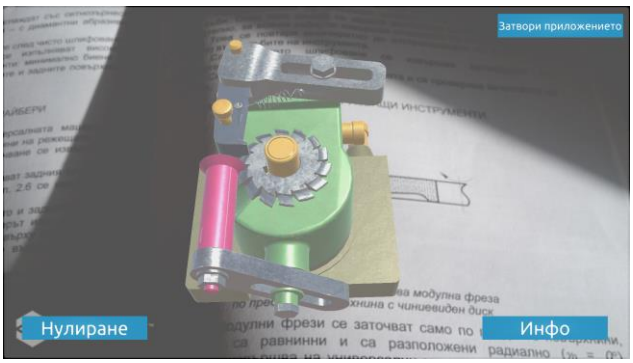


Figure 7. Augmented Reality App visualize 3D model on textbook

The following steps were used to create the tags for augmented reality application:

- First, the textbook figures for the required models were photographed using a camera. The photographs were then copied to a personal computer;
- Second, by using Vuforia's built-in tagging system, the photos were uploaded to the web

application Vuforia Developer Portal, where their system converted them into augmented reality tags;

- Finally, after all the shapes have been converted to markers, the package created by the Vuforia system is downloaded, ready to be embedded in Unity.

Following the completion of all these stages, the resulting tags are embedded into the development environment using the unity package that was obtained from the Vuforia system, allowing the user to decide which model is visualized by each tag (shape/marker).

The augmented reality application is implemented in the Unity Game Engine using the Vuforia SDK. The app works with one scene where all the models and targets are embedded.

E. Students Feedback and Survey Results

It is crucial for the apps' design process to receive feedback from students. The apps were iteratively tested with students in University of Ruse. The tests showed that the applications work according to the intended logic and some recommendations were made to improve better performance. The "Reset" button was added to the augmented reality app, and the font and size of the content in the same app were altered based on user feedback. The format of the visualized information was altered in the virtual reality application, specifically by reducing the area where the text is displayed and by including additional buttons for a kind of information scrolling. The details for each model include test questions and answers.

An online survey was made with students to assess the level of student satisfaction using the applications for augmented and virtual reality. Students must respond to multiple questions regarding various aspects of apps functionality. Survey results shows that 91 % of students think that the interaction with the 3D models in the augmented reality application and the virtual environment in the virtual reality application is intuitive and easy to learn and use. 86 % answers that interaction with the apps are fun and more engaged. 89 % of the survey attendants who used the apps agreed that the apps would be important asset of the Mechanical engineering education and prefer to use them in their studies.

III. CONCLUSION

To meet the current state of technical growth and the intellectual capacity of modern learners, education, like nearly all aspects of social life, is susceptible to modernization. The introduction of new technologies such as virtual and augmented reality helps to show the items in the studied disciplines in a unique, understandable, and fascinating way, changing the mechanical engineering educational content and the ways of its assimilation.

The 3D visualization of models of the cutting tools and orienting devices used in training of students through augmented and virtual reality technologies is an innovative and interesting approach that supports the effectiveness of learning by providing the opportunity for learners to enter to the virtual learning laboratory or simply 3D visualize a given machine tools or cutting tools from the standard textbook.

As a conclusion some features, positive and negative sides of virtual and augmented reality educational tools can be summarized:

• The use of marker-based visualization of models is a preferred approach when the dimensions of real objects represented by 3D models are relatively small;

• The use of smartphones in the classroom demonstrates a new paradigm for students where these portable devices can be used as effective learning tools.

• The usage of smartphones or other cordless portable devices, like a tablet, allows greater flexibility in the use of the applications because the augmented reality application doesn't need any extra hardware. The required equipment is also simple for the virtual reality application-just a helmet for smartphones which is not expensive;

• As portable devices offer greater freedom of use of the learning tool it is crucial to choose a method of developing the application that can be easily implemented in the two most popular mobile operating systems, Android and iOS;

• In the application implementing virtual reality technology, it is important to create a complete virtual environment that can be viewed by rotating 360 degrees.

REFERENCES

[1] E. Bobek, B. Tversky, “Creating visual explanations improves learning,” Cognitive research: principles and implications 1, 2016, pp.1-14.

[2] R. Radoeva, E. Petkov, T. Kalushkov, D. Valcheva and G. Shipkovenski, “Overview On Hardware Characteristics Of Virtual Reality Systems," 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 01-05, 2022.

[3] K. Videnov, Krum, V. Stoykova, and Z. Kazlacheva, “Application of augmented reality in higher education,” ARTTE Applied Researches in Technics, Technologies and Education 6, no. 1, pp. 1-9, 2018.

[4] I. Sural, “Mobile augmented reality applications in education,” In Mobile technologies and augmented reality in open education, pp. 200-214. IGI Global, 2017.

[5] G. Molnar, Z. Szuts, and K. Biro, “Use of augmented reality in learning,” Acta Polytechnica Hungarica 15, no. 5, pp 209-222, 2018.

[6] T. De Souza Silva, E.C.R. Marinho, G.R.E. Cabral, and K.S. Da Gama, “Motivational impact of virtual reality on game-based learning: Comparative study of immersive and non-immersive approaches,” 2017 19th Symposium on Virtual and Augmented Reality (SVR), pp. 155-158. IEEE, 2017.

[7] S. Ebadi, and M.Ebadijalal, “The effect of Google Expeditions virtual reality on EFL learners’ willingness to communicate and oral proficiency,” Computer Assisted Language Learning, 35(8), pp.1975-2000, 2022.

[8] T. Avramova, E. Lefterov, “Regarding the strength calculations of complex combined tools,” International Conference - Modern Technologies in Industrial Engineering (ModTech) 2018, IOP Conference Series: Materials Science and Engineering, ModTech 2018, Vol. 400, pp. 1-7, 2018.

[9] D. Vasileva, E. Lefterov, “Change from one Method to Another for Processing by Cutting a Complex Rotary Surfaces,” IOP Conference Series: Materials Science and Engineering, 2021.

Video Creation as a Catalyst of Value Change in Engineering Education

Riadh Habash[1] and Nabil M. Abdulkadir[2]
1 School of Electrical Engineering and Computer Science, University of Ottawa, Canada (rhabash@uottawa.ca)
2 Universiti Brunei Darussalam, Brunei (nabil.abdulkadir@ubd.edu.bn)

***Abstract*— Today educational videos have become a significant element in enhancing teaching and learning. They represent an important learning value to promote knowledge creation and translation for in-class, blended, and online learning. Effective creation and sharing of videos are enriched when instructors consider several principles: how to promote digital literacy in the class environment; how to maximize student engagement in making and sharing videos; how a video can make a tangible difference in the education landscape, and how to utilize this process to promote value creation in education? This article reviews literature relevant to each of these principles and presents a case of making educational videos in the context of value creation in teaching ethical professional practice in engineering and computing. The case advocates several feasible approaches instructors may embrace to assimilate knowledge creation, digital engagement, critical thinking, and growth mindset that allow students to take ownership of their learning.**

***Index Terms*—Video creation, reflection practice, knowledge creation, active learning, flipped classroom, digital literacy, value change.**

I. INTRODUCTION

Teachers know that their students have different learning styles depending on their life and needs. These may include learning by seeing and hearing, reflecting critically and culturally, reasoning logically and intuitively, memorizing and visualizing, drawing analogies, and building mathematical models [1-3]. All the above styles promote interest and open minds in classrooms, however learning through reflection is a practice of engagement that empowers learners to explore, examine, and realize their feeling and reasoning [4] as well as the value of their learning.

Sfard [5] labeled the "acquisition" metaphor for learning as something which results in the personal acquisition of knowledge and a well-rounded skill set, much like possessing an object. This metaphor is a conventional way of thinking about learning in many educational institutions, in particular, where knowledge is transferred from teacher to student. In this regard, attention may be paid to other concepts of learning, including "learning as participation" or "learning as knowledge creation". For example, if learning is realized as a matter of attaining content only, then tools should be available for students to acquire that content. However, if learning is perceived as a matter of participating in a practice, tools are there to be grasped, as mechanisms of that practice [6]. On the other hand, if the process of learning is understood as an idea of collaboration in knowledge creation, new methodologies may be designed mostly by teachers who should evolve into resourceful educational leaders within an active learning environment.

The infusion of information and communication technologies (ICT) in education is a primary driver behind active learning. Engineering as an inherently collaborative field responded to the explosion in knowledge by incessantly developing new learning tools. As more of these tools evolve, the challenges of the new learning environment intensify. What makes this particularly important today is the range of opportunities that digital transformation provides for creating new kinds of learning activities and experiences. The question is in identifying the right tools and using them intelligently. To create a competitive advantage, learners should be enabled to establish an innovative environment of knowledge collection and creation through digital engagement. A successful environment provides the learners both a chance to reach their potential as learners and the opportunity to engage in beyond pure technical activities that they may benefit from after graduation.

Engaging in learning entails effective content, knowledge-sharing platforms, services, tools, tasks, and assessment strategies. In education, knowledge creation and translation pedagogies have not been completely applied to various learning settings [1]. As education is becoming more competitive, digital transformation is turning into a necessary instrument of endurance as the new digital world requires educators to adapt and adopt digital mindsets and pedagogies. Capabilities such as simulation, animation, and visualization offer new alternative concepts in teaching by embedding ubiquitous learning into flipped or inverted classroom approach [7] to deliver experiences to users at anytime, anywhere, and anyway, mostly with individualized guidance. This may mean delivering the content outside the classroom and keeping learning activities in the classroom. The phenomenon of a "flipped classroom" combined with digital transformation may be considered a promoter for change within the typical learning (in-class and/or hybrid) environment.

The process of using educational videos as a learning experience has a long tradition in education, but its functionality has radically changed in the last few years. What was like a lot of work in the past, is nowadays feasible for every educator with suitable software [8]. Web 2:0 technologies, including minimally edited videos, have multiplied the number of participants in the production of knowledge, and they greatly extend the reach and circulation of digital objects, video, sound clips, text, graphics, and so forth [9].

Today, very little is known about the production of educational videos and their role in knowledge creation. This is identified as a major gap in the research that requires more

investigation. That being said, how do the roles of instructors and learners evolve and overlap as the learning environment moves beyond a focus on passively watching videos to actively making and sharing them?

This article explores the above research question by introducing a methodology in knowledge creation to make and share educational videos as a changing pedagogy on reflective practice and redesign of learning responsibilities within an active learning environment. It presents a case of teaching a course on ethical professional practice in engineering and computing. This methodology recommends creating project educational videos among other activities like research cases and technical articles, which instructors may adopt to develop a growth mindset that integrates knowledge creation, digital engagement, and critical thinking to enhance learners' well-rounded skill sets and advance their careers.

II. Knowledge Creation and Active Learning

Knowledge creation denotes the process of generating new ideas or insights through exploration. On the other hand, active learning is a student-centered pedagogy that involves engagement and collaboration. Both concepts are intertwined and vital for critical thinking skills as well as creativity and innovation.

Today, the amount of knowledge that students need to acquire and understand is increasing faster than the ability of curricula to include it or the students to learn it just by attending classes, memorizing content, and writing exams [1]. To understand, students need to think about what they learn, reflect on their past experiences, and apply it to what they think is valuable and probably part of themselves [10]. This reality demands reconsidering a learner-centered pedagogy by moving from the possession of knowledge to the creation and translation of knowledge, which comes about through reflection and engagement in an active learning environment.

Active learning is an instructional strategy that involves students in the learning process beyond listening and passive note-taking. It develops skills and higher-order thinking through activities that might include reading, writing, and discussion [12], but goes beyond the classroom to "learning by doing". It guides better performance in tasks beyond the passive mode of classroom lectures. It involves maximizing opportunities for students to experiment and collaborate. However, using active learning strategies does not require abandoning lecturing. Instead, including activities may make lecturing more relevant to student learning especially when teaching about an invisible phenomenon, where a design and simulation become essential [13, 14].

In knowledge creation, there are two dimensions to consider: epistemological (tacit and explicit knowledge) and ontological (knowledge creation at the individual, group, and organizational levels) [1, 15]. Tacit knowledge implies ideas and hands-on skills that are hard to articulate because they are learned largely through experience (know-how), competence, and dedication. However, explicit knowledge refers to information that is easy to articulate, write down, and share such as clear-cut facts (know-what) including books, software, and videos. Learning, in general, has both a tacit and an explicit dimension [1, 16]. This four-dimensional structure is understood as a catalyst of disruptive innovation or proactive educational pedagogy to change the learning process.

Another important notion for the active learning environment is knowledge translation. It is a relatively new term, first proposed by the Canadian Institute of Health Research (CIHR). It is an active multistep process that is focused on associating the "know-how-what-do gap" between knowledge creation and its implementation [17]. To develop a successful knowledge-translation approach, it is needed to understand how different types of knowledge are communicated, developed, and disseminated.

After all, teaching that encourages active learning invites learners to make or produce something, which may serve to facilitate and evaluate their understanding [18]. Therefore, students learn more effectively when they are vigorously involved in creating and translating their knowledge [19], the process that fosters innovation.

III. Video-based Learning (VBL)

VBL has become a conduit for a more interactive learning experience than just the provision of information. This involves the action of doing; it is about thinking while doing. Schon [20] terms this activity "reflection-in-action" and encourages students to learn "a kind of reflection-in-action that goes beyond statable rules". He suggests two modes of action: telling and listening; demonstrating and imitating. The first mode is typical in engineering education, while the second mode requires the instructor to reveal how to design something followed by simulation by the student. This entails the collaboration of the instructor and students in solving open-ended problems that are new to both of them.

Regardless of modality, blending videos with active learning strategies encourages students to create and view videos and make connections between new concepts and prior knowledge. Video design and creation may also help instructors maximize their effectiveness in the classroom by considering three components: cognitive load, student engagement, and active learning [21].

One of the primary considerations when developing educational videos is cognitive load. Cognitive load theory, initially articulated by John Sweller [21], implies that memory has several elements including sensory, working, and long-term. Sensory memory is temporary, gathering information from the environment to be chosen for temporary storage and processing in working memory. This process is a requirement for training in the long-term memory.

Although VBL is still top-down and teacher-centered [22], there is still an increasing inclination toward linking both teacher-centered and learner-centered pedagogies [23]. This occurs when students make videos rather than just view videos in addition to sharing the videos they produce with others. Figure 1 shows the transition from the old pedagogy where the teacher is the only source of knowledge to the learner-centered pedagogy where the teacher plays the role of facilitator [11]. This involves shifting the focus from the teacher who is the authority figure and the source of knowledge to the students who are the active participants in their learning environment. Accordingly, VBL as a process of making and sharing videos represents one of the vehicle of migration from the old pedagogy to the new one. However, both pedagogies have their place in education.

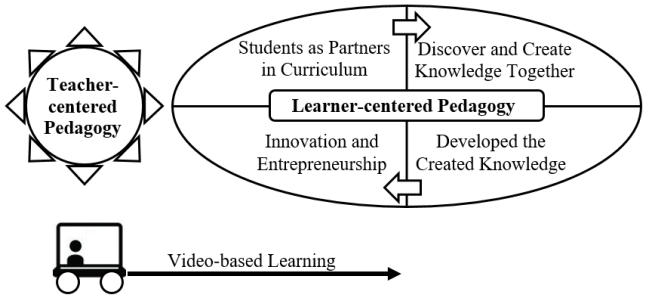


Figure 1. The transition from old pedagogy to new pedagogy.

Today, videos are valuable in many blended and online courses, but ensuring that videos can help student learning involves taking a greater look at video design and production. Creating video content may be an effective way to assist students in their learning. The key is, to begin with, reliable instructional objectives and following a good plan on the storyboard to insure the core message is presented and have the students effectively engaged. A script for the educational video will help shorten and organize its content. To keep the attention during watching, the use of signaling [24] is advised. This may be realized with arrows that point to a specific location on the screen or with colors, which highpoint the part of the figure or the picture that is being clarified.

To understand how people learn with videos, it is essential to think of the principle of multimedia learning as shown in Figure 2. This affirms that humans learn more deeply from a pattern of words and pictures (illustrations, photos, animation, and video) than from words only. Multimedia training is therefore outlined as "presenting words and pictures that are intended to advance learning" [25]. However, this requires prior knowledge that is gained independently of an experience as a result of abstract or logical reasoning alone.

The most important result of learning with digital means (ubiquitous) is that learning with a combination of text and pictures is more effective than learning with just a text. This supports the flipped classroom approach which consists of two distinct parts: direct online instruction at the student's own time and pace, and interactive group learning activities in scheduled classes [26]. This combination is critical for developing a growth mindset that inspires cognitive learning, seeks out new challenges, increases access to practical explanations, and importantly flips learning spaces: the prospect of flexible learning anywhere at any time.

IV. Knowledge-sharing Platforms

Knowledge-sharing platforms are online services and tools that offer a variety of educational resources and allow students to share information and knowledge with individuals, groups, and communities. These platforms may exist in different forms including blogs, social media, learning resources, and podcasts.

An engaging platform like Brightspace is an excellent learning management system for communities like colleges and universities. It provides a range of features and tools for the creation and delivery of digital content and activities. It is effectively employed for teaching the course under consideration in this article. Brightspace analytics tools may be used to monitor student engagement with the video content.

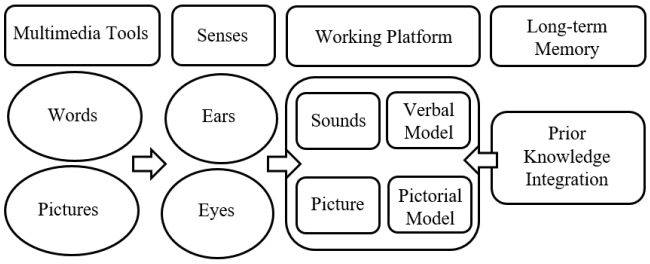


Figure 2. Cognitive theory of multimedia learning.

For students, the free video streaming platform YouTube and Vimeo are the most popular on the web with many channels dedicated to various topics. They offer a range of high-quality features and customization options for video creators within many different genres that exist across diverse art forms.

Another reinforcing platform to facilitate access to information and knowledge creation is the establishment of open educational resources (OER). The OER definition is coined by UNESCO [27] as an "open provision of educational resources, enabled by pervasive ICT, for consultation, use, and adaptation by a community of users for noncommercial purposes." To accomplish the above goal, an OER, namely "g9toengineering.com" was developed by the first author in 2007 to support teaching several engineering courses and to provide open access to knowledge creation in several topics [28]. This OER contains a set of activities and learner-generated content around the creation and sharing of knowledge. The main objective of the OER is to provide a wide range of learners with open access to educational resources and content.

V. Case of Teaching Ethical Professional Practice

As a continuing process of developing activities for educational reflective practice, a systematic process for knowledge creation and translation was formed several years ago for a high-enrollment undergraduate course, "Professional Practice in Engineering and Computing". To achieve the goal, case-based learning (CBL), project-based learning (PBL), and VBL are usually implemented in teaching the course while the instructor acts as a facilitator in the learning process. The aim is to teach the course material on a "need to know" basis where the class is flipped into a community of learning and practice, researching and gathering knowledge, then deciding from among the best solutions. CBL, PBL, and VBL employ the 4Cs principle: critical thinking, communication, collaboration, and creativity, with outcomes tied to learning goals.

The course learning components are divided into four categories namely multiple-choice question-based midterm and the final exam, four research case studies, and a project of three tasks. The 10-week case studies and project are carried out in groups of two students, each based on authentic open-ended real-world problems with clear instructions. The topics as a whole make the experience worthy of reflection. The case studies present students with ethical situations that

involve the application of their knowledge to decide what they know that is relevant to the cases, what other information they may require to obtain, and what impacts their decision may have. Topics include professionalism, ethics and equity in engineering, intrapreneurial mindset, and ChatGPT ethicality. The case narratives are accompanied by supporting information and tools that facilitate knowledge creation including the course reference book [29] and various research articles. For each case, the students develop an "e-Portfolio" that includes the investigation details and a summary of the results. The e-Portfolio reflects the student's advancement in a self-directed knowledge translation approach with a cumulative "e-Poster" to propagate outcomes.

A list of project topics with clear research questions was given to the students for investigation. Among the topics are the ethical practice of robotics, artificial intelligence, self-driven vehicles, Internet-of-Things (IoT), 5G technologies, technology in classrooms, virtual reality, and sustainability. Students are also given the option to pick up their topic ideas. These ideas are collaborative tasks that are largely based on the triangular concept of "values", "ethics", and "morals" as shown in Figure 3. Values represent the principles, models, and standards, which may help in making the judgment of what is important in life. Ethics refers to the guidelines for conduct that address questions about morals. Morals are the system of beliefs that emerge out of core values and ethics. The course project involves three tasks including a proposal, a two-minute educational video, and an article. The above three project activities reflect the entire tasks of the project.

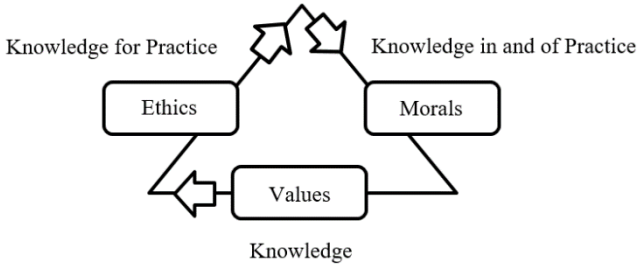


Figure 3. The triangularity of values, ethics, and morals.

The educational videos combine reflection on what students have learned; deliberately analyze decision-making, depict knowledge content, and connect it to what they do and learn in the classroom. Undoubtedly the key idea on how to use videos in education is the learning by design methodology. In this process, students are active designers of the learning content. They write storyboards in a team, look for pertinent materials, and produce their videos artistically.

VI. Assessment and Student Experience

The course under consideration is taught with the goal of knowledge creation in mind. The use of videos in education is usually associated with challenges where all students watch the produced videos together in class and/or online while individual learning pace or prior knowledge should be integrated to enhance reflection on the learning content. At the end of the semester, several video presentation sessions were organized. The class 192 groups were invited to watch their videos together in several sessions. This is the step in which the outcomes of learning are exchanged where students and teachers interact together by watching videos and realizing the newly created knowledge.

VBL may be an effective tool for assessing learning outcomes, however, the great challenge for the instructor is how to evaluate the videos within the content and cognitive context in mind. Before viewing the videos, the instructor and teaching assistant made a series of subjective ratings: judgment of learning, presentation style, engagement, and interest rating. Videos were marked on a scale of 0 to 10. The top five videos were selected for an award.

For the online anonymous survey, the instructor thought of four specific questions about the effectiveness of the course learning tasks including the four case studies, project, and multiple-choice questions-based midterm and final exam. Students' satisfaction concerning the course learning components and results was measured by analyzing the data collected from Questions 1 to 4 (Table 1). The data from Question 1 of the survey shows that most of the students support learner-centered (research cases and project) over teacher-based (midterm and final exam) pedagogy. In particular, data from Question 4 indicates that about 83% of students "strongly agreed and/or agreed" with the video creation exercise. Despite the satisfaction, this process might involve student uneasiness by being responsible for their learning tasks.

An important aspect of the assessment outcome is student fulfillment, which is correlated to but different from success as verified by student feedback and opinion about the video-making experience:

"*I believe the creation of educational videos has a large impact on the learning objectives of students. While working on this project video, I learned much more about my topic and was able to condense it in a way my peers could understand. It informally improved my speaking skills and allowed me to connect with the audience, like a YouTube video."*

"*When showing my friends, the video, they easily understood my research questions and became interested in the subject. Overall, I think letting students make videos in an ethics class is great as it aids the creator in learning and allows students to become aware of issues around them in an engaging manner. Personally, creating the project video was the best aspect of the project as it allowed me to showcase my creative skills and combine them in a way to educate others on a topic I am very interested in*."

According to the student feedback, it is obvious that teachers have an opening to increasingly motivate their students to influence their perception of learning.

A significant driver to pioneer the application of educational videos for student engagement is by making them. The idea is straightforward: if students do not make videos, they cannot learn simply by passively watching them. Engaging students to digitally reflect on a project task yields deeper levels of knowledge. Because students were working on different project tasks, they cannot all be expected to discover the same knowledge; however, each student possesses something unique from this engaging experience. One major takeout is to keep videos brief and targeted at learning goals as well as avoiding an in-person presentation style. The length of the video is a key factor to keep students

watching until the end. To deepen engagement and maximize learning outcomes, the top five videos were awarded a bonus of five marks in addition to the complete graded marks.

TABLE 1 Course anonymous survey outcomes

Survey Questions	**Student Response**			
I learnt for future practice from the course content overall, however, I enjoyed working on	Cases 63.77%	Project 57.97%	Midterm 21.74%	Final Exam 20.29%
Among the four given assignments, I enjoyed working on	Case 1 53.62%	Case 2 43.48%	Case 3 49.28%	Case 4 33.33%
I learnt for future practice from the course project, however, I enjoyed in particular	Proposal 27.54%	Debate 49.28%	Article 26.09%	Video 47.83
Videos improve students' understanding of difficult concepts when compared with a course where the videos are not used in conjunction with the project	Strongly agree 42.03%	Agree 40.58%	Neutral 13.04%	Disagree 4.35%

VII. Reflection Practice and Value Creation

In education, reflective practice plays a crucial role in value creation when learners acquire knowledge and skill set that apply to their lives [1]. The whole process is about learners' participation in a continual cycle of self-observation and self-evaluation to realize their actions and the reactions they stimulate in themselves. To effectively recognize this value, several changes in the education system should take place where classes will be more open to different learning experiences and teachers will continue as facilitators and organizers of services and tools [30]. However, this break from the traditional system is not without dispute. Students often may be uninterested to leave their comfort zone if they do not assure ownership or see the value that deserves more effort.

Initially, the teaching experience under consideration has specifically centered on exploring the flipped class practice, and instructional video production to adapt content to forthcoming learners. It exposes students to the pace of change by bringing research, professional practice, and teaching cultures close together while continuing to place the students at the heart of the learning process. In addition, it increases the usage of digital technology to reshape and reinvent the active learning model to shift to a more learner-centered experience that is suitable for large cohorts of students.

In the case under consideration, educational videos are used as an integral part of the learning activities provided in teaching ethical professional practice particularly when they are part of an engaging task like the course project. As an outcome of the survey and observation, students valued the videos and post-video discussions as though they were resourceful in promoting student understanding. It was found that students value the process of making and sharing the videos as a promotor for improving the deeper realization of complex concepts when compared to a course without video making. Student feedback and data indicate that integrating engaging activities with digital media where students receive traditional teaching and VBL combined may considerably improve student analytical thinking and reflective judgment.

The pervasive ICT, simulation tools and cloud capacity can promote student understanding and help articulate several key lessons taught by the experience under consideration. It also assists educators to facilitate reflection on learning practices by creating and translating knowledge more effectively.

Re-thinking and re-designing course content in response to the existing digital transformation for students who have different learning experiences and expectations may in itself have a substantial impact on improving teaching. Changing learning away from knowledge possession toward knowledge creation involves thorough consideration of the students as learners. Because the effectiveness of an approach lies in transformative learning outcomes. The focus should shift toward keynote lectures, research tasks, interaction in the classroom, new assessment strategies, and rewards. Further, instructors that wish to provide effective transformative learning environments may create a structure of incentives and motivation that encourage students to put in the efforts needed. As the learning journey continues, the above experience may promote a powerful catalyst of value creation in education to bear in mind.

References

[1] R. Habash, "Unleashing knowledge creation and sharing in a reflective open education," *Proc. 2018 Canadian Engineering Education Association (CEEA-ACEG 18) Conf*, pp. 1-5.

[2] R. M. Felder, and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering Education*, vol. 78, no. 7, pp. 674-681, 1988.

[3] G. K. Tillman, "Facilitating the creation of knowledge within an e-Learning framework to develop critical thinking skills," *International Journal of Information and Education Technology*, vol. 6, no. 2, pp. 166-169, 2016

[4] R. Habash, M. Hasan, J. Chiasson, and M. Tannous, "Phenomenon- and project-based learning through the lens of sustainability," *International Journal of Education*, vol. 38, no. 1, pp. 1-7, 2022.

[5] A. Sfard, "On two metaphors for learning and the dangers of choosing just one," *Educational Researcher*, vol. 27, no. 2, pp. 4-13, 1988.

[6] NAE, *The Engineer of 2020: Visions of Engineering in the New Century*, National Academy of Engineering. https://inside.mines.edu/UserFiles/File/Assessment/Engr2020.pdf.

[7] M. J. Lage, G. J. Platt, and M. Treglia, "Inverting the classroom: a gateway to creating an inclusive learning environment," *The Journal of Economic Education*, vol. 31, no. 1, pp. 30–43, 2000.

[8] J. Buchner, "How to create educational videos: From watching passively to learning actively," *Open Online Journal for Research and Education*, Special Issue #12, September 2018, ISSN: 2313-1640.

[9] T. Warger, and G. Dobbin, Learning environments: where space, technology, and culture converge," *EDUCAUSE Learning Initiative*, 2009, 1-14.

[10] S. Daud, R. E. Abdul Rahim, and R. Alimun. "Knowledge creation and innovation in classroom," *World Academy of Science, Engineering and Technology International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, vol. 2, no. 4, pp. 440-444, 2008.

[11] R. Habash. *Green Engineering: Innovation, Entrepreneurship, and Design*, Boca Raton, FL: CRC Tylor and Francis, 2017.

[12] Active learning, Vanderbilt University Center for Teaching. https://cft.vanderbilt.edu/guides-sub-pages/active-learning/

[13] R. Habash, Phenomenon-based Learning for Age 5.0 Mindsets: Industry, Society, and Education, IEEEEduCon

(Global Engineering Education Conference), Tunis, Tunisia, March 28-31, 2022.
[14] C. J. Brame, Effective Educational Videos, https://cft.vanderbilt.edu/guides-sub-pages/effective-educational-videos, 2015.
[15] F. Neumann, *An Integrated Descriptive Model of Knowledge Creation Applied to Mechatronic Product Development*, PhD Thesis, Otto-von-Guericke-University, Lehrstuhl für Maschinenbauinformatik, 2014.
[16] I. Nonaka, and H. Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*? New York, NY: Oxford University Press, 1995.
[17] P. Govender, and K. Mosterd, "Making sense of knowing: Knowledge creation and translation in student occupational therapy practitioners," *African Journal of Health Professions Education*, vol. 11, no. 2, pp. 38-40, 2018.
[18] J. Handelsman, S. Miller and C. Pfund, *Scientific Teaching*, New York, NY: W.H. Freeman, 2007.
[19] R. E. Mayer, "Applying the science of learning: Evidence-based principles for the design of multimedia instruction," *American Psychologist*, vol. 63, no. 8, pp. 760-769, 2008.
[20] D. A. Schon, *Educating the Reflective Practitioner: Toward a New Way of Teaching and Learning in the Profession*, San Francisco: Jossey-Bass Publishers, 1987.
[21] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learning and Instruction*, vol. 4, pp. 295-312, 1994.
[22] A. Kirkwood, and L. Price, "Examining some assumptions and limitations of research on the effects of emerging technologies for teaching and learning in higher education", *British Journal of Educational Technology*, vol. 44, no. 4, pp. 536–54, 2013.
[23] R. Kay, and I. Kletskin, "Evaluating the use of problem-based video podcasts to teach mathematics in higher education", *Computers and Education*, vol. 59, no. 2, pp. 619–627, 2012.
[24] C. J. Brame, "Effective educational videos: principles and guidelines for maximizing student learning from video content," *Life Science Education*, vol. 14, no. 6, pp. 1-6, 2016.
[25] K. R. Butcher, "The multimedia principle," in *The Cambridge Handbook of Multimedia Learning*, Second edition, pp. 174–205), Cambridge, UK: Cambridge University Press.
[26] D. P. M. Comber, and M. Brady-Van den Bos, "Too much, too soon? A critical investigation into factors that make flipped classrooms effective," *Higher Education Research and Development*, vol. 37, no. 4, pp. 683-697, 2018.
[27] UNESCO, Forum on the *Impact of Open Courseware for Higher Education in Developing Countries*: Final report, UNESCO. Design, July, pp. 1-3, 2002.
[28] R. Habash, "g9toengineering: a virtual community of practice in knowledge creation," *2019 IEEE Global Engineering Education Conference (EDUCON)*, 978-1-5386-9506-7/19/$31.00.
[29] R. Habash, *Professional Practice in Engineering and Computing: Preparing for Future Career*, Boca Raton, FL: CRC Taylor and Francis, 2019.
[30] R. Habash, "An experiential engineering learning model for knowledge and state of flow creation", *Athens Journal of Technology and Engineering*, vol. 10, pp. 1-17, 2023.

Ensuring Global Perspectives within Reading Lists to Increase Students' Engagement

G.J. Collins
University College London, Department of Computer Science, London, United Kingdom
g.collins@ucl.ac.uk

***Abstract* - University College London attracts students worldwide and provides students with the latest research opportunities to meet global challenges. The Department of Computer Science within the university has won recognition for its diversity and inclusion initiatives. During 2022 a series of AI ethics lectures were developed for the software engineering professional practice module to help students prepare for their research projects. Teaching was through the viewpoints of power imbalances relating to AI and green technologies, including resource extraction. The AI ethics series was designed as online flipped lectures. Students were encouraged to read research papers beforehand, which were subsequently discussed during the online lecture. Qualitative data analysis of the reading list provided insight to optimize content and reduce cognitive load. During 2022/23, students completed redesigned questionnaires to help the lecturer understand their knowledge of the topics and interests. Students outlined their interests in leading-edge AI concepts and how AI can reduce energy consumption and protect the environment and biodiversity. Analysis of questionnaires and discussions indicates students' increased engagement in sustainability related topics. When perspectives from different regions of the world, reflecting the class's global community, were introduced to the reading list, students were more engaged with the module.**

Keywords - AI ethics, cognitive load, curriculum development, engineering education, reading lists, student engagement

I. INTRODUCTION

This paper advances the approach of diversifying reading lists [1] to engage with students who register for a software engineering course from countries worldwide. It covers the further adaptation of flipped online lectures and the development of questionnaires to establish students' interests and prior knowledge.

University College London (UCL) welcomes students from around the world. UCL recognizes the importance of attracting students worldwide to ensure the right mix of leading research talent. Each department emphasizes the university's equality, inclusion, and diversity (EDI) initiatives. The Department of Computer Science recognizes the importance of supporting females in research and won several awards in recognition. The department won the Athena Scientific Women's Academic Network (SWAN) award in 2015 and 2019 for advancing female career opportunities for students and staff. Within the research-intensive universities in the UK, the department has one of the highest female ratios of students and staff, with over 20% female staff [2]. The postgraduate taught module is no exception. Female students comprise over 30% of the intake.

There is an increasing emphasis on interdisciplinary research to solve some of the most challenging global problems within all departments. Student research is focused on and supported via the UCL Grand Challenges initiative [3], which seeks to bring UCL's academic expertise to address these global problems by integrating knowledge across disciplines. This initiative is a collaborative venture across the university and with industry, research institutes, and associated clinical partners such as University College Hospital. The themes within this initiative address: Global Health, Sustainable Cities, Cultural Understanding, Human Wellbeing, Equality and Justice, and Transformative Technology. These cover the spectrum of the 17 United Nations (UN) Sustainability Development Goals [4]. Students' research projects, therefore, align with both these frameworks.

International conferences such as the 27th Conference of the Parties to the UN Framework Convention on Climate Change 2022 (COP27) and the 2022 UN Biodiversity Conference (COP15) seek to address some of the world's most significant problems. These conferences call for more progress toward addressing climate change and biodiversity loss. The World Economic Forum's (WEF) Global Risks Report 2023 [5] outlines that the most urgent problems faced by humanity relate to climate change and the destruction of the environment. Their report cites biodiversity loss and ecosystem collapse as the fastest-growing risk to the planet over the next ten years. Students are aware of these calls for action and increasingly want to be involved with projects that create a more sustainable future.

Since 2003 the lecturer (author) has, in collaboration with colleagues, developed the postgraduate professional practice module within the Department of Computer Science. From its inception, this module has included industry guest speakers providing an alternative perspective to the academic content taught by research staff at UCL. The initial development of the AI ethics

lectures during 2021/22 [1] provided an opportunity to adapt these lectures to students' interests in bias and creating a fairer society. The lecturer refined this series during 2022/23 based on a more detailed questionnaire to incorporate students' interests and views on leading-edge AI developments and how AI can reduce environmental impact and biodiversity loss and mitigate the climate crisis.

Software engineering students can take machine learning modules which focus on the technical aspects. The technical examples introduced within the AI ethics series during 2023 provide a background for students to discuss the ethical implications of the increasingly complex models: machine learning, neural networks, and foundation models. The lecturer has included these examples to help students understand approaches to interpreting data and how AI can reduce energy consumption. Developing and updating the AI ethics lectures ensured they aligned with students' interests and the university's EDI and sustainability initiatives.

This paper builds on the development of the online flipped learning AI lectures for the spring term of 2022 [1] and outlines further refinement based on student questionnaires and student feedback. The reading list mainly consisted of research papers. During 2023 the lecturer diversified the pre-reading to include sections of textbook chapters, industry blogs, and online articles from NGOs and Indigenous peoples. The module emphasizes different perspectives for each topic and allows adequate time for discussions before, during, and after lectures so that students can reflect on various perspectives.

II. Curriculum Development

A. *Diversifying the Reading Lists*

During the lecturer's previous classes, students have often commented on projects within their own experience or projects that have affected their communities. Where there has been coverage of associated topics in previous years, students have shown greater engagement, as indicated by the number of discussions and questions raised in class and online within Moodle, their online module resource. As the postgraduate students within the computer science department register from over 50 countries worldwide, students were asked through questionnaires whether teaching should cover a range of projects from around the world. The majority, 86%, of the students[1] during the academic year 2022/23 participating in these optional questionnaires either "strongly agreed" or "agreed" that projects covered within these lectures should include examples from different world regions. The reading list, to support the AI ethics online lectures, was diversified to represent a range of authors with different perspectives from around the world.

Reading lists can contribute to making universities more equitable. Reading lists can also provide a voice to marginalized authors. Students are increasingly demanding that there are more diverse sources of knowledge [6]. One way of achieving this is via the reading list: ensuring there are perspectives from around the world. Encouraging students to contribute to these resources is another way to collaborate and engage in the process.

All students benefit from a plurality of views. By diversifying reading lists, students from minority groups see that their opinions and communities are represented [7]. The lecturer followed the suggestion of the SOAS, the School of Oriental and African Studies (University of London) toolkit [8], to widen sources within the reading list to be more representative of the Global South. In addition to recommended sources from around the globe, the lecturer considered the perspectives of each source. The lecturer recorded each author's university or institution and location. Gender was also recorded by reference to personal pronouns within biographical notes at the end of the paper or other articles they had written.

The initial design of the lectures only included research papers from the United Kingdom, the United States of America, and Australia. These resources were not representative of the students' registrations. Postgraduate students in the department 2021/22 registered from many countries, including Austria, Germany, South Africa, China, Hong Kong, the Islamic Republic of Iran, Japan, Saudi Arabia, Singapore, South Africa, and the United Kingdom. Following students' opinions recorded within the 2022/23 questionnaire, the lecturer has widened the reading list further by ensuring different perspectives for each topic and including research papers from other world regions, Fig. 1. The

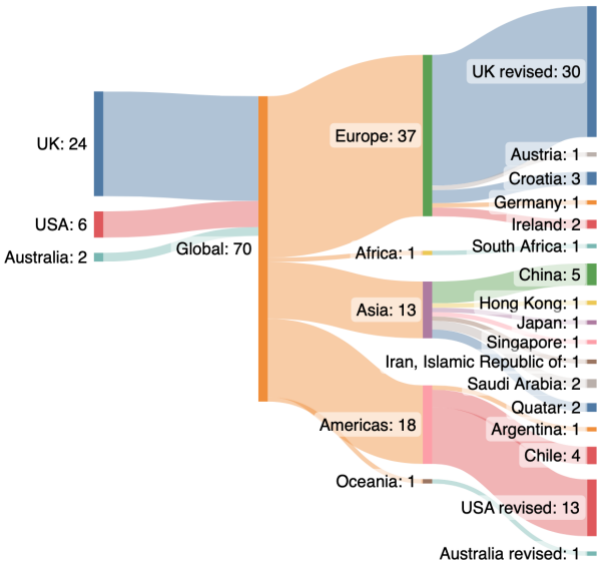


Figure 1. Geographic distribution of authors' locations for the AI ethics reading list. The left-hand side represents the authors' locations for the documents initially selected in 2021. The right-hand side indicates the authors' locations for the academic year 2022/23. Adapted and updated from G.J. Collins [1].

[1] Questionnaires were anonymous. Class size (*N*) 2022/23 =*32*.

reading list better reflects the diversity of UCL software engineering student registrations.

B. Questions to Understand Students' Prior Knowledge

The pre-lecture questionnaire is in two parts; the first section is designed to understand students' knowledge of machine learning and AI ethics. The second part is designed to understand students' interests (Section C).

Questioning can provide an enhanced understanding of a subject and can be a valuable teaching tool [9]. The lecturer can ascertain the student's knowledge by asking what a student already knows within a questionnaire. By asking via a questionnaire prior to the lecture, the lecturer can develop the module so that concepts that are well understood by students do not have to be repeated. This information can then be used to ensure the time is not spent on repeating concepts but on new areas of interest to students.

The lecturer developed the student questionnaire using Microsoft Forms. The anonymous questionnaire provides an automated answer to the student, with a constructive comment on the validity of the response and a pointer to further resources. These additional resources are highlighted with specific page numbers within the reading list. The questionnaire is integral to the learning process by outlining the student's progress before lectures and as a learning resource to develop a fuller appreciation of the topic.

Previous research by the author [1] has shown that students sometimes need help understanding receiver operator characteristic (ROC) curves. Most students, 66% correctly identified that the area under the curve (AUC) for a ROC curve "provides a summary of how well the model works across a range of thresholds" [1]. Students' anonymous answers were used to examine this issue more thoroughly. In 2022, the lecturer needed significant time within the online lecture to explain the example. The lecturer provided further guidance before the class in 2023. The lecturer also shortened the introduction to the problem during the online lecture with a worked example, which allowed more time for an extended discussion and adequate time to discuss the ethical implications of interpreting the data.

The lecture material was based on a hypothetical example of cancer diagnosis to identify patients for treatment based on Halligan, Altman, and Mallet [10]. The worked example was adapted for 2023 to provide an additional model for students to calculate the relevant figures for sensitivity and specificity. During the online discussion, the lecturer explained the reference diagnosis scale using Table I. The lecturer also explained the graph's axes and plotted the ROC curve, Fig. 2. A worked example for threshold 5 was outlined, assuming this is the highest accuracy of a cancer diagnosis. At this threshold, 17 patients are correctly identified with cancer (34% sensitivity), and only 1 with a false positive diagnosis (98% specificity). The lecturer then used threshold 3 as a further example. The lecturer reminded students this would include all patients for threshold 3 and above. Students calculated the total of patients diagnosed with cancer and percentage sensitivity. Students also provided the answer for the relevant specificity.

TABLE I. REFERENCE DIAGNOSIS

REFERENCE DIAGNOSIS/THRESHOLD	1	2	3	4	5	TOTAL
PATIENTS WITH CANCER (TRUE POSITIVES)	1	3	8	21	17	50
PATIENTS CLEAR OF CANCER (FALSE POSITIVES)	17	21	8	3	1	50

Adapted from S. Halligan, D.G. Altman, and S. Mallet [10]. Adapted with Permission

Students discussed which threshold to select. Students noticed that lowering the threshold would correctly identify more cancer patients, with a consequent increase in the rate of false positives. However, students also appreciated that lowering a threshold may create a moral dilemma if there are limited resources to cope with additional tests and procedures. Students recognized that many communities do not have the same access to healthcare as others. Students suggested that accuracy in diagnosis, whether in this example or AI systems in surgery [11] within their recommended reading, is essential, and this will often need the medical professional's expertise to interpret the AI systems' recommendations.

Guided instruction for students unfamiliar with topics rather than allowing students to explore unguided has shown to be effective [12], [13]. Although optional, students who retook the questionnaire achieved the

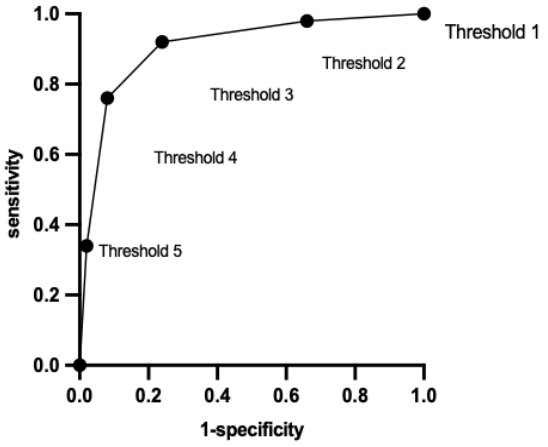


Figure 2. The ROC curve drawn during the lecture. Adapted from S. Halligan, D.G. Altman, and S. Mallett [10]. Adapted with Permission.

correct answer for the related questions. The responses suggest the guidance contributed to a better technical understanding. Students noted that the pre-reading [10] highlights that medical professionals often find it challenging to interpret these ROC curves. Many students found the discussion helpful, as they realized that professional experience is still required to make an informed decision.

C. *Questions to Understand Students' Interests*

The second section of the pre-lecture questionnaire aims to understand students' interests [2]. Within this section, one of the questions asked their views regarding the extraction of resources using a Likert scale: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree.

Question: The extraction of resources (such as cobalt and lithium) for green technologies should consider the following statements:

1. The carbon emissions of the resources
2. The impact on the environment and biodiversity
3. That technological advances in the Global North should not be at the expense of the Global South

Most students, 86%, answering this question either "strongly agreed" or "agreed" that impact on the environment and biodiversity should be considered. Many students, 71%, agreed that carbon emissions should be taken into consideration. Although there was less consensus on the last statement, whether "…advances should not be at the expense of the Global South", most students answering this statement agreed.

As students frequently mentioned healthcare bias in their online discussions, the lecturer selected bias as the introductory topic. Analyzing discussions and questionnaires determined two additional AI topics, "emotion recognition software, and AI and sustainability" [1], which were updated for 2022/23.

D. *Refinement of the Schema*

The lecturer used ATLAS.ti version 22 [14] for qualitative analysis of the reading list. The visualization helped to rapidly identify the frequency of the coverage of projects, countries, and perspectives [1]. This information helped identify repeated instances of concepts and enabled a reduction in the content for students to study, lessening the overall cognitive load. Analysis of the reading material helped develop the schema to emphasize various phenomena such as bias, "affective communities" [1], and policies or regulations that need to change, Fig. 3.

Kirschner, Sweller, and Clark [13] outline that learners must acquire the necessary schemas to scaffold their knowledge to develop expertise. The learning process is accelerated if pieces of information are processed simultaneously as a schema. Schemas can be considered akin to "patterns in our heads" [15]. Using schemas makes the learning and recall process more effective as it places less demand on working memory. According to Cognitive Load Theory [16], processing novel information causes a load on the learner's working memory. Additional working memory resources are then required to construct schemas to store in long-term memory [17]. The schema provides a framework for both the lecturer and student to explore each schema segment, helping to scaffold information and reduce the load on working memory [18].

Students have referred to the schema Fig. 3 to identify the affected communities and whether people have a say in regulations or how their resources, including data, are used. When students discussed Microsoft's research [19] to find more sustainable energy solutions for data centers by placing them on the seabed, they raised questions regarding the environmental impact on marine ecosystems, potential sea temperature increase, and the consequences to local marine ecosystems. In response to these student discussions, the relevant schema element, relating to "affective communities" Fig. 3, now encompasses the environment and ecosystems. This elaboration of the schema supports both students' learning and the lecturer to deliver the lecture content.

E. *Increased Range of Perspectives for Topics*

During 2022/23, the class discussed energy efficiency in more detail. This encompassed power usage effectiveness (PUE) measurements: the efficiency of the data center via the ratio of energy use overall compared with that used for computation. Discussions covered the industry perspective on using PUE as a standard measure and the flaws of this approach. Students readily

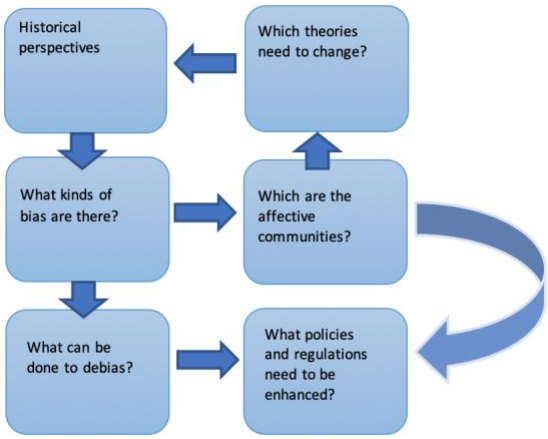


Figure 3. The schema developed for the lectures focuses on the affective communities, which theories need to change, bias, and what may have given rise to bias. The schema examines what can be done to reduce bias and what theories, policies, and regulations need to change. From G.J. Collins [1]

[2] The percentage of female students within the department for postgraduate taught modules was 32% for the year 2020/21.

appreciated that the data center would typically have a more efficient PUE value if relocated to a cooler climate. They also recognized that reliance on PUE values could hide inefficient scheduling within the data center. Students also exchanged ideas on how environmentally friendly different approaches to cooling were, such as the use of solar energy or the use of seawater.

The reading list now includes resource exploitation, particularly lithium mining in South America, from the perspective of the Indigenous Colla community and NGOs [20]. The online discussion covers the water used for lithium processing, reducing the water levels in the hypersaline lakes, and polluting water supplies for the local population. The pre-reading has also been updated to include research highlighting the impact on ecosystems [21]. Students discuss the detrimental effect on the fragile ecosystem, reducing the shrimp population within the lakes, the critical food source for flamingoes. The online lecture uses the schema Fig. 3 to discuss whether the Indigenous community has a say in what happens to these resources and whether this exploitation of resources in the Global South is ethically acceptable.

Students have expressed an interest in protecting ecosystems and climate change. A research paper within their reading outlines the use of generative adversarial networks [22] to accurately measure polar ice profiles. This research was covered in more detail during 2023 as students were interested in the effect of ice-melt on biodiversity within polar regions and climate change. As the department's research students use the university's HPC infrastructure, many are interested in the energy and mineral resources required to build and run these facilities and how AI can reduce energy consumption. The lecturer enhanced the reading list for 2023 with research papers covering AI scheduling within data centers and research outlining the application of AI to reduce energy consumption in smart cities [23].

With the increasing adoption of foundation models, such as ChatGPT, students were concerned that models would perpetuate bias across domains. Students recognized that the model being studied [24] would disproportionally affect under-represented populations in South America and Africa if used more widely. Students also explore how small training sets can improve efficacy: energy consumption, time training models, and reducing bias.

F. Key Outcomes for Online-Lectures

The reading list is now more diverse in representing authors and perspectives from across the world, Fig.1. It also includes a broader range of opinions from academics, industry, research institutes, NGOs and Indigenous communities. The AI ethics lectures have been adapted to cover topics in which students have expressed an interest: advances in AI, bias, diversity and inclusion, health, sustainability, biodiversity, and climate change.

More time was allocated to class discussions during the spring term of 2023 than compared to the previous academic year. Student discussions continued beyond the scheduled lecture time, suggesting increased engagement. The lecturer used the schema to introduce topics and refined this for future lectures emphasizing areas discussed with students, particularly bias and coverage of the environment and biodiversity.

III. DISCUSSION

Numerous research papers cover AI ethics; however, there needs to be more research conducted on how this subject should be taught [25]. Students' priorities, particularly in the environment and climate change, are not always prioritized in the suggested models of AI ethics education [26]. Including the areas of students' interests has increased engagement, as measured by the number of online posts and discussions. Although this number may not be valid, it indicates to what extent students interact with the reading lists and online lectures. Additionally, all online posts and discussions recorded were directly relevant to exploring the topics more deeply.

Providing resources before the lectures can reduce the students' cognitive load [27]. Using questionnaires to establish students' prior knowledge allows the design of the content to build on their knowledge. Having an appropriate schema aids the assimilation of new knowledge [28]. The automated feedback within the questionnaires provides further guidance and resources available. Within a crowded curriculum, it is impossible to cover every aspect of a topic; giving feedback and optional resources is one way to enhance the module.

It is essential that students are equipped to understand the societal impact of these technologies and proposed changes to the regulation of AI. Based on the schema, the discussions emphasized that communities affected by AI should have a say in AI deployment, which is a recommendation of the Ada Lovelace Institute's Policy Briefing [29]. The online lectures also provide an opportunity to discuss that researchers should have rigorous and ongoing review processes when they develop AI systems.

Although students study machine learning options, these are from a predominantly technical perspective. The AI ethics series focuses on enhancing the ethics coverage within the module and considers some examples that students have specifically mentioned they are interested in. In teaching, it is essential to provide a balance of different opinions. Students have recognized the importance of alternative perspectives and have suggested additional sources during their discussions. Including students' research paper suggestions for the reading list can increase engagement and co-create the curriculum. Within the 2022/23 questionnaire, students have indicated that there should be a more comprehensive representation of projects from different regions of the

globe. Although there is still a bias within the reading list towards authors within the Global North, the number of authors and countries represented by the Global South has substantially increased.

IV. CONCLUSION

Diversifying the reading list to include sources from students' home countries can build engagement, help students feel welcome, and create a sense of belonging. Understanding what students are interested in and including their suggestions for the reading matter helps students feel that they also have a voice in their studies. Covering topics within the curriculum that students are concerned with, health, environment, biodiversity, and climate change, demonstrates their opinions are valued.

Diversifying perspectives within engineering education instills the value of considering a more comprehensive range of opinions and technological solutions. This diversification of views enables students to realize that AI can be used to develop a greener technological future that is not dependent on exploiting the resources of the Global South.

REFERENCES

[1] G.J. Collins, "Can decolonising the curriculum provide an enhanced engineering education?" In Proc. *2022 50th Annual Conference of The European Society for Engineering Education (SEFI 22)*, Sept. 19-22, 2022, pp. 1085-1094, https://www.sefi.be/wp-content/uploads/2022/12/ebook-sefi-2022-1-1.pdf

[2] UCL Computer Science, 2023, https://www.ucl.ac.uk/computer-science/about/equity-diversity-and-inclusion/gender-equality-athena-swan/10-years-athena-swan-ucl-computer

[3] UCL Grand Challenges. https://www.ucl.ac.uk/grand-challenges/six-ucl-grand-challenges

[4] UN Sustainability Goals. https://sdgs.un.org/goals

[5] World Economic Forum, *Global Risks Report 2023*, Jan 11, 2023. https://www.weforum.org/reports/global-risks-report-2023/

[6] B. Adewumi, L.R. Bailey, E. Mires-Richards, K.M. Quinlan, E. Agyeman, A. Alabi, M. Jeyasingh, C. Konadu-Mensah, W. Lavinière, P. Mighton, T. Shortridge, D.S.P. Thomas, and N. Wassamba-Wabelua, "Cross-disciplinary, collaborative and student-led: Developing a change process for diversifying reading lists," *London Review of Education*, vol. 20, no. 1, pp. 1–17, 2022, https://doi.org/10.14324/LRE.20.1.01

[7] E. Mires-Richards, S. Field, and J. Sanderson, "See yourself on the shelf: A collaborative approach for diversifying library collections," *ALISS Quarterly*, vol. 15, no. 2, 6–, 2020.

[8] SOAS toolkit, 2018. Accessed: Feb. 03, 2023. https://blogs.soas.ac.uk/decolonisingsoas/files/2018/10/Decolonising-SOAS-Learning-and-Teaching-Toolkit-AB.pdf

[9] V.C.H., Tong, "Using asynchronous electronic surveys to help in-class revision: A case study," *British Journal of Educational Technology*, vol. 43, no. 3, pp. 465-473, 2012, https://doi.org/10.1111/j.1467-8535.2011.01207.x

[10] S. Halligan, D.G. Altman, and S. Mallett, "Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: A discussion and proposal for an alternative approach," *Eur Radiol*, vol. 25, pp. 932-939, 2015, https://doi.org/10.1007/s00330-014-3487-0

[11] G. Collins, Appendix 2 - Artificial Intelligence (AI) and Big Data. In Lester, A., Project Management Planning and Control, 8th Edition, pp. 571-589, Elsevier, 2018, https://doi.org/10.1016/B978-0-12-824339-8.15002-3

[12] J. Sweller, Instructional design in technical areas, Camberwell, Australia: ACR Press, 2003

[13] P.A., Kirschner, J. Sweller, and R.E. Clark, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," *Educational Psychologist*, vol. 41, no. 2, 75–86, 2006, https://doi.org/10.1207/s15326985ep4102_1

[14] ATLAS.ti (Version 22, 2021), ATLAS.ti Scientific Software Development GmbH. https://atlasti.com

[15] C. Kohls and K. Scheiter, "The relation between design patterns and schema theory," in Proc. *15th Conf. on Pattern Languages of Programs (PLoP '08)*, Association for Computing Machinery, New York, NY, USA, Article 15, 2008, pp. 1–16, doi: 10.1145/1753196.1753214

[16] J. Sweller, J.J.G. van Merrienboer, and F.G.W.C. Paas, "Cognitive architecture and instructional design," *Educational Psychology Review*, vol. 10, no. 3, pp. 251-296, 1998, https://doi.org/10.1023/A:1022193728205

[17] F. Krieglstein, M. Beege, G. D. Rey, C. Sanchez-Stockhammer, and S. Schneider, "Development and validation of a theory-based questionnaire to measure different types of cognitive load," *Educ Psychol Rev.*, vol. 5, no. 9, 2023, https://doi-org.libproxy.ucl.ac.uk/10.1007/s10648-023-09738-0

[18] S. Blissett, M. Goldszmidt, and M. Sibbald, "Do research findings on schema-based instruction translate to the classroom?" *Perspectives on Medical Education*, vol. 4, no. 6, pp. 334-338, 2015, https://doi.org/10.1007/s400037-015-0225-5

[19] J. Roach, "Microsoft finds underwater datacenters are reliable, practical and use energy sustainably," Sept. 14, 2020. https://news.microsoft.com/source/features/sustainability/project-natick-underwater-datacenter/

[20] N. Greenfield, Lithium mining is leaving Chile's indigenous communities high and dry (literally), NRDC, April 26, 2022, https://www.nrdc.org/stories/lithium-mining-leaving-chiles-indigenous-communities-high-and-dry-literally

[21] G. Gajardo, and S. Redón, "Andean hypersaline lakes in the Atacama Desert, northern Chile: Between lithium exploitation and unique biodiversity conservation," *Conservation Science and Practice*, 1 :e94, 2019, https://doi.org/10.1111/csp2.94

[22] M. Rahnemoonfar, J. Johnson, and J. Paden, "AI Radar Sensor: Creating radar depth sounder images based on generative adversarial network," *Sensors (Basel, Switzerland)*, vol. 19, no. 24, pp. 5479–, 2019, https://doi.org/10.3390/s19245479

[23] A. Višković, V. Franki, and D. Jevtić, Artificial intelligence as a facilitator of the energy transition, in Proc. *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO 22)*, May 23-27, 2022, pp. 494-499, https://doi.org/10.23919/MIPRO55190.2022.9803700

[24] K. Zhou, K. Ethayarajh, and D. Jurafsky, "Frequency-based distortions in contextualized word embeddings," *ArXiv* abs/2104.08465, (2021), https://arxiv.org/abs/2104.08465

[25] A. Bozkurt, A. Karadeniz, D. Baneres, A.E. Guerrero-Roldán, and M.E. Rodríguez, "Artificial Intelligence and reflections from educational landscape: A review of AI studies in half a century," *Sustainability*, vol. 13, no. 2, p. 800, 2021. https://doi.org/10.3390/su13020800

[26] E. Kazim and A.S. Koshiyama, "A high-level overview of AI ethics," *Patterns*, vol. 2, no. 9, Article 100314, 2021. https://doi.org/10.1016/j.patter.2021.10

[27] G. Sirhan, C. Gray, A.H. Johnstone, and N. Reid, "Preparing the mind of the learner," *Univeristy Chemistry Education*, vol. 3, no. 2, pp. 43-46, 1999.

[28] M.K. Seery and R. Donnelly, "The implemenatation of pre-lecture resources to reduce in class cognitive load: A case study for higher education chemistry," *British Journal of Educational Technology*" vol. 43, no. 4, pp. 667-677, 2012, https://doi.org/10.1111/j.1467-8535.2011.01237.x

[29] Ada Lovelace Institute, People, risk and the unique requirements of AI, Policy Briefing, March 31, 2022. Accessed: Feb. 03, 2023. https://www.adalovelaceinstitute.org/policy-briefing/eu-ai-act/

Virtual Electric Machines Laboratory, Requirements and Practical Realization

P.J. van Duijsen*, D.C. Zuidervliet†
THUAS, Delft, The Netherlands, *www.dc-lab.org*
* p.j.vanduijsen@hhs.nl †d.c.zuidervliet@hhs.nl

***Abstract*—A fully equipped electrical machine laboratory can become a costly factor. However the look and feel of electrical machines behavior in operation, is indispensable and very informative for students. Not only do they learn how to connect and operate the machines, but also the limitations of each type of machine, becomes visible. Not only equipment is costly, also laboratory room space, service and safety can become an issue. To reduce those laboratory costs, often simulation is used as an alternative. However, the look and feel of the real machines, is in many cases lost. In this paper a virtual electric machines laboratory is proposed, where students can be trained, using digital twins of the electric machines from the laboratory. The look and feel of the digital twins, the way the machines are connected electrically and mechanically, should represent their physical counterparts in such detail, as required in a real physical training set-up. Nameplate parameters ease the selection of machines size and power level. Adding measurement devices, such as voltage and current meters, as well as measuring torque and rotational speed, should resemble the real measurement set-up as close as possible.**

The size and power level of the electrical machine, is given by its nameplate. In the virtual electrical machine model, those parameters are entered, instead of the electrical parameters mostly encountered in simulation programs.

***Index Terms*—Virtual workbench, electric machines, simulation, animation, education, laboratory**

I. Introduction

Virtual electrical machines laboratories, will not be the replacement of a real physical laboratory. However, when teaching large numbers of students, with limited availability of costly laboratory equipment, it requires an alternative. A virtual laboratory would be the first thing that comes into mind when thinking about costs reduction. However, using such a virtual laboratory, places constraints on the way the virtual laboratory looks like. The main requirement for a practical realization of such a virtual laboratory, is simply that it should look and behave similar.

An overview and role of a laboratory for teaching electrical machines is an ongoing topic in the society meetings of the IEEE [1]. Especially in graduate courses, the use of simulation to replace the physical laboratory is a topic, that is discussed over the years. The first approaches only discuss the replacement of the laboratory by simulation [2]. The simulations are still only numerical, in the beginning of the application of computers in education [3]. The limitations of computer graphics, hampered the introduction of useful simulation environments. Also the complicated mix of block-diagram models, not representing physical component, makes simulation unattractive for undergraduate courses. Only graduate courses can profit from simulations, as these students understand the mathematics behind he electrical machines, and thus can apply the numerical models simulating the electrical machines, as for example presented in [4]. Also using equivalent electric circuit models, to model the electrical machines, is only dedicated towards graduate students, understanding the differential equations describing both the electrical machines and electric circuit components [5]. Remote labs [6], tried to fill the gap, and replace the laboratory with access to a remote laboratory. However also this remote laboratory could only be used by a limited number of students. The use of simulation programs alone, mainly remained a dedicated set of complex simulation models, requiring understanding of the models and parameters [7]. All of these simulation approaches are not fully applicable in vocational training. Vocational training requires models based rated values or nameplate parameters. Internally these models contain the complex equations, modeling all the phenomena, existing in a physical laboratory set-up. For undergraduate graduate students, the level of complexity can be increased and brushless motors and their controls are usually employed in laboratory experiments [8]. The topic of field oriented control is usually seen in graduate laboratory experiments, mostly in the form of simulation experiments [9], [10].

The main reason for implementing a virtual laboratory is given in section II. In section III the requirements on a virtual laboratory are outlined. The structure of a virtual laboratory is described in section IV. The objectives are the same as for a physical laboratory and outlined in section V. Examples of an induction machine and universal motor are presented in section VI. The next step in going more into detail, requires simulation and animation, and two typical examples, are discussed in section VII

II. Virtual laboratory

The main reason for introducing a virtual laboratory for education, are not the costs related to a real physical laboratory, but the main reason is that in a virtual laboratory, you can set-up experiments very quickly, and without any risk whatsoever. You can do any experiment, without damaging the laboratory equipment, and no matter how extraordinary, you can build

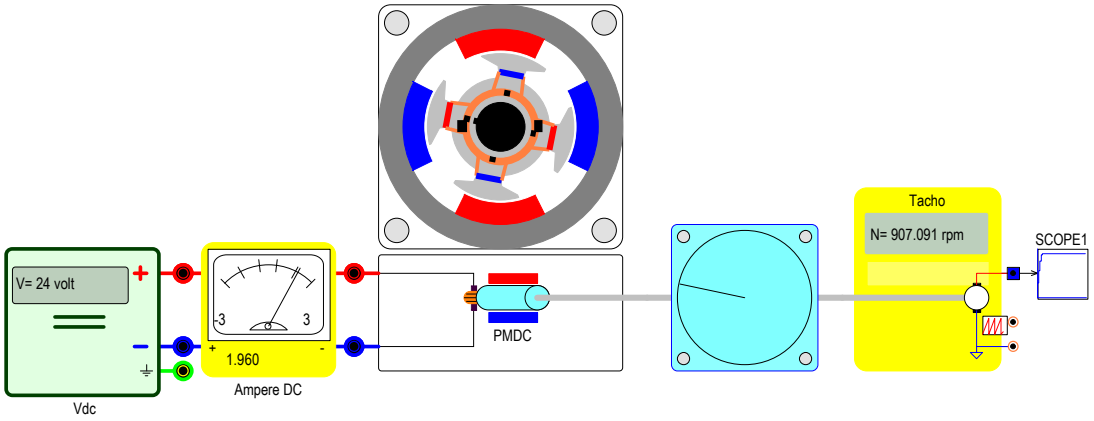


Fig. 1. Permanent Magnet DC motor driving a fan. Current is measured using an analog current meter, while a tachometer measures the speed of the rotor.

the most unrealistic experiments and learn from the mistakes. This is the main reason, since learning from mistakes is mostly what is remembered best. The background level of the students is the same as they need for a physical laboratory experiment.

A. Motivation

Students should get a low entry-level access, into laboratory experiments. The look and feel is important, but also the experience of doing things wrong, adds to the knowledge of the students. Especially the freedom to create experiments, that in real life would cause damage, contributes to the knowledge and experience level of the student. As long as the laboratory components have the same look and feel, as the real laboratory components, it contributes to the practical training of students.

B. Challenges/Risks

A virtual laboratory is undoubtedly cheaper than a physical laboratory. Students can have access to the virtual laboratory at any time and place they like and are not bounded to availability of the laboratory space. This also creates the possibility to use the virtual laboratory as demonstration during oral lectures.

The virtual laboratory experiments should be carried out in the same way, as a face-to-face traditional laboratory is organized. The tutor also has the same role during the experiments.

However, the risk is that students do not get access to a real laboratory environment, once the virtual lab is implemented. Management could cancel the physical laboratory for costs saving reasons.

C. Results

Due to the faster set-up of a laboratory experiment, without any risk of doing something wrong or destroying something, students can get results quicker. Instead of doing a single experiment, what-if scenarios are possible, not only by the students, but also by the teachers, when preparing lectures or experiments.

D. Outcomes

Application of the virtual laboratory in Delft [11], showed that students got a deeper understanding of the operation of electric machines, because they had the possibility to virtually experiment with the electric machines of any size and power level. The examples from section VI are implemented and tested, in a second year undergraduate course on electric machines in Delft [11]. The drive simulation from section VII, are to be used in a third year undergraduate or second year graduate course. The examples from section VI can also be implemented in second and third year vocational training. However, here it is advised to have pre-programmed examples, where students only have to change parameters like power level of an electric machine, or power supply output voltage.

III. Requirements

To prepare students, to do practical laboratory tests, requires them to train using simulation tools. However simulation tools do provide nice models for electrical machines, but also require parameters. This is manageable by undergraduate and graduate students, but impossible for students of vocational training and first semester undergraduate students. For them, the nameplate parameters are the first encounter, in identifying electrical machines.

A virtual laboratory [10], should mimic the real laboratory [11]. Each component in the virtual laboratory should reassemble a physical component from the real laboratory. Not only should the look and feel be the same, also connecting the components should be the same as if the students would connect the physical machines, see figure 1.

Models should also have the limitation of the real set-up, for example, measuring the angular speed using a tachometer, gives an indication of the angular speed in the range of 0 to 5 volts. The students will understand that there is a conversion factor, when going from RPM to an analog voltage, that can be measured in an electric circuit.

Figure 2 shows a typical laboratory set-up. An inverter is driving a three-phase brushless motor, and a tachometer is measuring the speed. Students can change the DC voltage of

the inverter, and this changes the amount of current supplied to the brushless motor. As this is a typical experiment for vocational training, the DC voltage to drive speed relation can be measured using DC voltage meters.

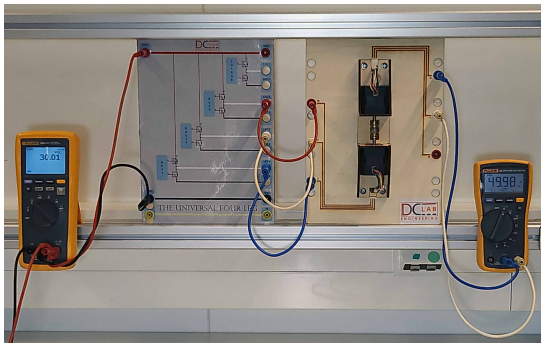


Fig. 2. Physical laboratory set-up, of three phase inverter U4L [11] with brushless DC motor and DC motor acting as tachometer.

To convert the laboratory experiment from figure 2 into a virtual laboratory experiment, the virtual experiment has to fulfill some constraints. The virtual laboratory model should have realistic parameters. The brushless motor from figure 2 has typical manufacturer data like nominal voltage, nominal speed and power level. This has to be supplied by the student in the virtual experiment, see figure 3. Also the student should carry out the measurement in the same way as in the real experiment. In simulation, graphic display of waveforms and numerical results are easy to obtain, but in this virtual experiment, the student has only access to a multimeter and eventually an oscilloscope for recording the waveforms. Finally the virtual workbench should show an animation of what would happen in the real experiment. For electrical machines, this can be done by showing the rotation of the rotor and, additionally, the rotation of electromagnetic fields in stator and/or rotor.

To mimic the laboratory experiment as realistic as possible, a virtual laboratory assignment should include:

- Connecting the components via wires with banana-plugs.
- Set parameters and/or choose the size/power rating of a component.
- Connect measuring devices like analog voltage or current meters, or digital multimeters.
- Start the experiment, instead of turning on the power supply, the simulation/animation has to be started.
- Observation of the results from the measuring devices.
- Change parameters(power/size rating) and observe their influence on the results.
- Validate the observation of the results.

IV. Structure

The structure of a virtual laboratory should be able to Conceive, Design, Implement and Operate the experiments in an engineering environment, being creative and value added according to the CDIO-Initiative [12]. Each phase of the experiment can be structured according to this initiative.

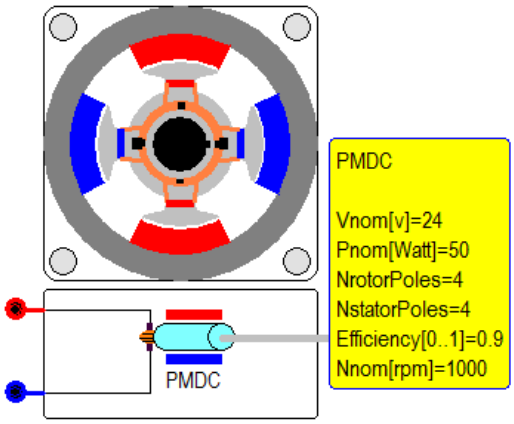


Fig. 3. Name plate parameters of the Permanent Magnet DC motor, including the number of rotor and stator poles for the animation of the motor.

- *Conceive: Theory*
  A short introduction on the theory behind the experiment from the textbook. Here the basics of the functioning should be outlined, without going into detail on the operation.
- *Design: Simulation*
  Simulation or animation, with simple model to explore working principle using simulation or online tooling [10]. This will reveal the operation as explained in the theory.
- *Implement: Set-up*
  Virtual laboratory set-up, with virtual components and animation. Select and place components and interconnect them.
- *Operate: Validation*
  Run the experiment. Validation of the virtual laboratory results, with the theoretical and/or simulation results. Comparison with expectations, and explanation of the observed laboratory results.

V. Objectives

In the virtual electric machines laboratory assignment, the objectives are the same, as for a physical laboratory. They can be extended with giving insight in phenomena's, that are only visible through animation.

A. Electromechanical

An easy start for undergraduate students is the permanent magnet DC motor, see figure 1. The only important parameter is the motor constant.

$$U = k \cdot \omega = k \cdot \frac{2\pi}{60} N \tag{1}$$

$$T = k \cdot I \tag{2}$$

where k is the motor constant describing the most important relation between electrical and mechanical quantities. However, for undergraduate students, this motor constant can be hidden as a parameter, if the student only has to enter machine nameplate parameters like, nominal voltage, nominal speed

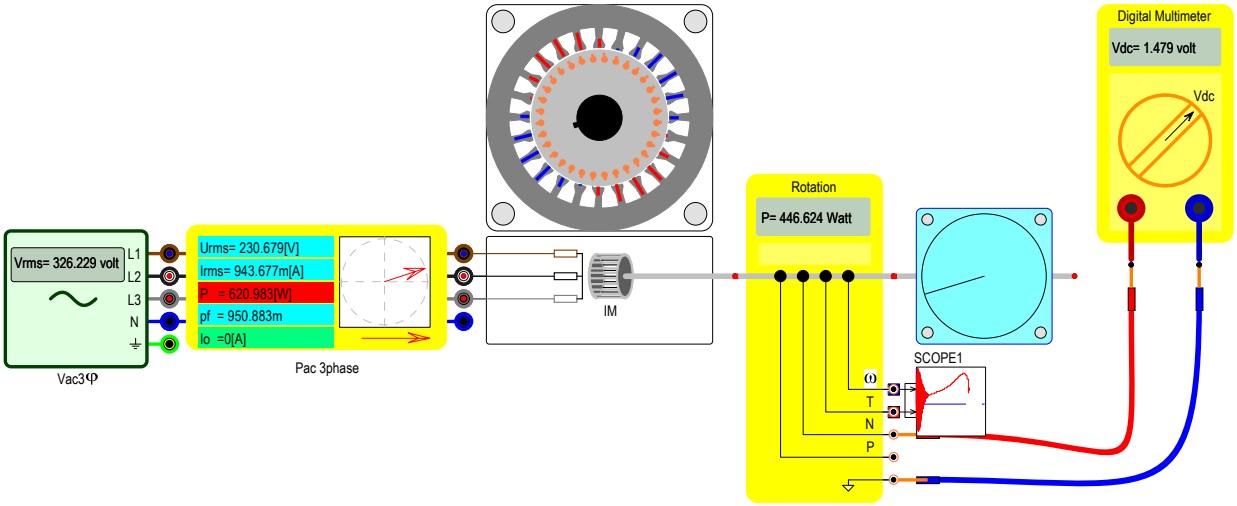


Fig. 4. Simulation of the torque-speed characteristic of an induction machine driving a fan.

and nominal power, see figure 3. From this data, the machine constant k is calculated using (1), and used internally in the model. Vocational training students, can simply choose the motor based on voltage and power level. Undergraduate students can calculate the motor constant k and verify (1).

B. Winding resistance

The winding resistance is the first parameter that defines the losses of the electrical machine. It can be measured at an early stage, using a simple multimeter. The influence of the damping effect of the winding resistance on regulation and stability, is not yet to be measured in this stage of a virtual laboratory, however it can be noticed in the dynamic behavior.

C. Winding inductance

If the machine is kept at standstill using the dynamometer, the impedance of the winding can be measured. Impedance can be measured using an AC signal generator and AC current meter. Since the winding resistance is already known, the winding inductance can be estimated. Knowing the impedance, one can set the correct parameters in a proportional-integration [PI] controlled, current regulated AC drive. Graduate students can measure the winding resistance and inductance, and use this for calculating the PI parameters of the Field Oriented Control [FOC] [9].

D. Magnetic field of the stator and rotor

Understanding what is happening inside the machine, can be visualized using the animation in the virtual laboratory. A revolving field inside the motor alongside the revolving currents in the windings, is shown as animation. Students can see the asynchronous fields in induction machines, and synchronous commutation of phase currents in brushless machines. Especially vocational training student and undergraduate students, have a need for visualization of magnetic fields, as their curriculum does not incorporate 2D/3D electromagnetic field analysis.

E. How is Torque created?

The concept of torque creation by the outer product between current and field $T = \Psi \times I$, is easily visualized in animation. It is also numerically visible in the measuring device. The linear relation between current and torque in DC motors, can be measured by the vocational training and undergraduate students, by simply adjusting the power supply. Graduate students are challenged to express and verify this relation, using the outer product of current and field.

F. How does terminal voltage arise?

Adjusting the angular speed of the motor using the dynamometer, or changing the field producing current in a DC shunt motor, shows the variation of the terminal voltage of the motor. This is a typical experiment for vocational training, as it is intuitive to see the proof of the mathematical relation from (1).

VI. Examples

A. Induction Machine

An induction machine laboratory set-up, can be used for many different experiments, see figure 4. Typically for induction motors, compared to DC and brushless motors, is the fact, that most of the induction motors are designed and thus only operate at the operating voltage they are designed for. This means that most of the smaller induction machines, with affordable power rating for a laboratory are rated for $400\text{volt}/1kW$. In vocational and undergraduate laboratory classes, this can be a problem because of safety. There are some exceptions [13] working on $23/40\text{volt}$, but especially in vocational training, measuring rated voltage and current levels, gives more insight for the student on what to expect. The $23/40\text{volt}$ motors are especially suited for control experiments using a low voltage inverter, since students can touch all wires during operation. In the virtual laboratory, the machine ratings

are like in a real drive. This means, that students can perform measurements on 100kW motor, without any safety issues.

The winding resistance of the induction machine is simply measured using a digital multimeter, see figure 5. From this the winding losses can be estimated, if the induction machine is in operation.

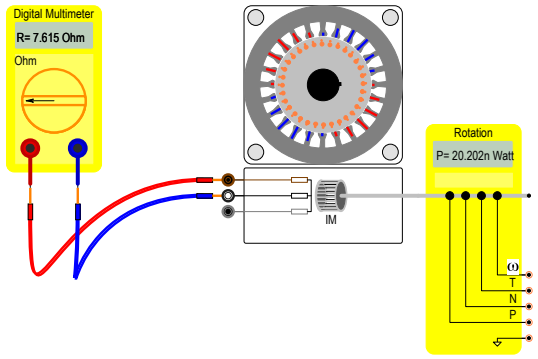


Fig. 5. Measuring the winding resistance using a multimeter.

The simulation of start-up in a virtual laboratory, allows the recording and visualization of the torque-speed characteristic of the induction machine, see figure 4. Interesting for students is many times the observation, that is looks slightly more complicated during start-up, than the ideal characteristic as seen in most textbooks, see Scope1 in figure 4.

A typical assignment in an induction machine laboratory is to measure the torque slip curve, see figure 6. Using a dynamometer, the mechanical load of the rotor can be adjusted, by specifying the load torque. The speed of the rotor can be measured and tabulated, see table I.

Torque[Nm]	Speed[rpm]	slip[%]
-5	.	.
.	.	.
.	.	.
5	.	.

TABLE I
TORQUE-SPEED CHARACTERISTIC TO BE COMPLETED BY THE STUDENT

B. Universal motor

Running an universal motor from the mains AC power grid, reveals a vibrating noise. In a virtual laboratory, using oscilloscopes, the origin of this audible noise can be visualized, see figure 7. Although the students select an universal motor and only specify nominal operating values, the underlying model is more complex. It includes the saturation of the stator and rotor iron in the internal model, however, this is not directly visible to the students as a separate parameter. The internal parameters are calculated from the rated values of the machine, simplifying the set-up for the students.

First of all, power delivered from a single AC source is pulsating at double supply frequency. Secondly the saturation of the stator and rotor iron in the universal machine, gives

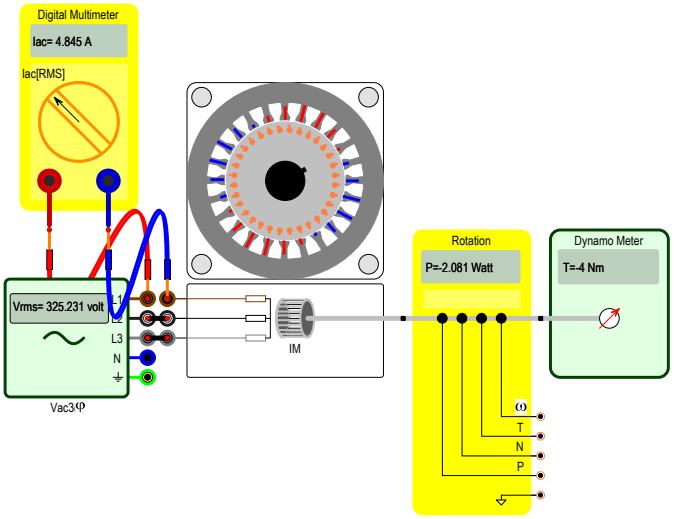


Fig. 6. Application of a dynamometer to measure the slip of the induction machine for varying load.

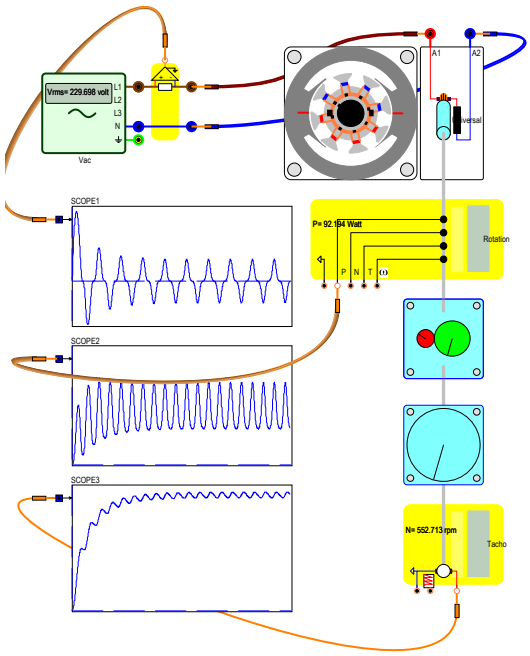


Fig. 7. Universal motor driving a fan, showing the typical torque harmonics.

rise to higher harmonics in the waveform for the current, see Scope 1 in figure 7. Observing this current waveform, makes explaining the torque pulsations easier, see Scope2 in figure 7. Finally the resulting rotor speed can be seen, see Scope3 in figure 7, explaining the audible noise.

VII. DRIVE SIMULATIONS

Once the students understand the working of machines, the next step in an undergraduate course, is to add the controller and use the machine in a drive system. From the laboratory models, students can measure voltages, speed, current as constant values. But in a drive, the dynamic behavior becomes

important. The simulation becomes more detailed and also the control parameters have to be added. Because of the many wiring details and different required simulation results, the simplified machine and drive models are better suited. Instead of measuring constant values, start up behavior or transient response becomes more important. The simulation in figure 8, shows the start up of a variable speed drive with quadratic load characteristic. Scope2 clearly shows the difference between the motor and load torque, and from this the acceleration of the motor can be explained. When both torque characteristics meet, the drive is at stationary speed. Changing parameters like load mass or inertia, driving voltage, machine size or machine inertia, can be studied using this simulation.

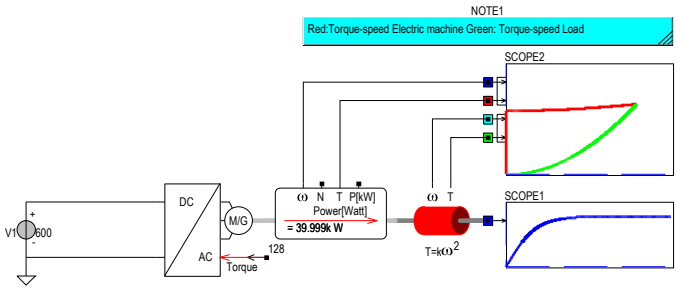


Fig. 8. Variable speed drive with quadratic load. Both motor torque and load torque are displayed in a single scope2 as function of the rotor speed, to show the remaining torque difference responsible for the acceleration of the drive.

The application of a speed or position profile in combination with a variable speed drive, only makes sense when there is adequate control. The variable speed drive blocks includes a PI controller. Setting up such a simulation experiment becomes easier when everything is integrated into a single block and students are not distracted by the many details of the control structure. They only have to set parameters of the PI controller, like they would do in an industrial environment. The simulation in figure 9 shows a given torque, speed and position profile. The speed is the integral of the acceleration, while the position is the integral of the angular speed. Therefore, defining the acceleration profile in the form of a given torque profile, feeding this in the variable speed drive, reveals the final speed of the variable speed drive.

CONCLUSIONS

For vocational training and undergraduate students, a virtual laboratory can be applied, when the components have the same look and feel as the real components. The students should be able to set-up the virtual laboratory experiment, using banana plugs and multimeters, like the student would do in a physical laboratory experiment. The models require only nameplate and rated parameters, but internally, the models are complex and the required parameters are automatically calculated from the nameplate and rated values. Non-linearity, saturation and harmonics, is included in those models. Since animation can be applied, also the rotation of magnetic fields or the commutation of currents in the machine can be visualized, being an extra educative quantity compared to a physical laboratory.

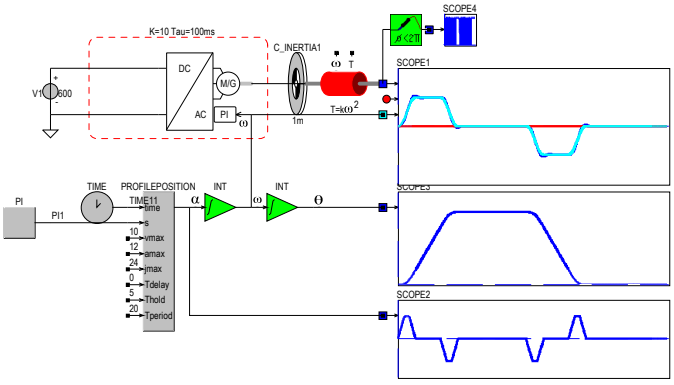


Fig. 9. Variable speed drive, where the angular speed of the machine is set by a defined speed profile. The PI controller finally regulates the speed according to the speed profile.

For last year undergraduate and graduate students, the more detailed drive simulations can be applied.

REFERENCES

[1] G. G. Karady, "Roll of laboratory education in electrical power engineering education," 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, PA, USA, 2008, pp. 1-3, doi: 10.1109/PES.2008.4596141.

[2] R. Belmans and W. Geysen, "Impact of the computer developments on the education of engineering graduates in electrical machines and drives," Proceedings. Electrotechnical Conference Integrating Research, Industry and Education in Energy and Communication Engineering',, Lisbon, Portugal, 1989, pp. 728-732, doi: 10.1109/MELCON.1989.50146.

[3] H. A. Smolleck, "Modeling and analysis of the induction machine: a computational/experimental approach," in IEEE Transactions on Power Systems, vol. 5, no. 2, pp. 482-485, May 1990, doi: 10.1109/59.54556.

[4] F. Flinders, S. Senini and W. Oghanna, "Mixed electrical and mechanical simulations using dynamic systems analysis packages," Proceedings of the 1993 IEEE/ASME Joint Railroad Conference, Pittsburgh, PA, USA, 1993, pp. 87-93, doi: 10.1109/RRCON.1993.292959.

[5] K. T. Chau and C. C. Chan, "A SPICE compatible model of permanent magnet DC motor drives," Proceedings of 1995 International Conference on Power Electronics and Drive Systems. PEDS 95, Singapore, 1995, pp. 477-482 vol.1, doi: 10.1109/PEDS.1995.404875.

[6] P. Bauer, V. Fedak and O. Rompelman, "PEMCWebLab - Distance and virtual laboratories in electrical engineering: Development and trends," 2008 13th International Power Electronics and Motion Control Conference, EPE-PEMC 2008, Poznan, Poland, 2008, pp. 2354-2359, doi: 10.1109/EPEPEMC.2008.4635615.

[7] P. Casals-Torrens, "Virtual Laboratory for Learning Asynchronous Motors in Engineering Degrees," in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, vol. 8, no. 2, pp. 71-76, May 2013, doi: 10.1109/RITA.2013.2258234.

[8] A. Üzel, D. C. Zuidervliet and P. J. v. Duijsen, "Educational Set-up for Brushless Motor Drives," 2021 17th Conference on Electrical Machines, Drives and Power Systems (ELMA), Sofia, Bulgaria, 2021, pp. 1-6, doi: 10.1109/ELMA52514.2021.9503025.

[9] P. J. v. Duijsen and D. C. Zuidervliet, "Teaching Field Oriented Control using Animation," 2021 17th Conference on Electrical Machines, Drives and Power Systems (ELMA), Sofia, Bulgaria, 2021, pp. 1-6, doi: 10.1109/ELMA52514.2021.9502966.

[10] Caspoc Simulation and Animation, Simulation Research, [Online] Available: https://www.caspoc.com,

[11] DC laboratory, The Hague University, TIS, Delft. [Online] Available: http://dc-lab.org

[12] CDIO Initiative, [Online] Available: www.cdio.org

[13] ETS-Didactic, 3-phase asynchronous motor, [Online] Available: https://ets-didactic.de/en-gb/home/drehstrom-asynchronmotor-ac3-23-40v-20w-35004-2078293/

Hierarchical Approach in Modeling and Simulation of Power Electronics for Education

P.J. van Duijsen*, D.C. Zuidervliet†
THUAS, Delft, The Netherlands, *www.dc-lab.org*
* p.j.vanduijsen@hhs.nl †d.c.zuidervliet@hhs.nl

***Abstract*—For teaching power electronics circuits, circuit simulation is well adapted, to show the working principles of the various types of power electronics converters. However, depending on the complexity of the models used, the teaching purpose might be troubled by the implementation details, required for the specific simulation program. Instead of focusing on the power electronics working principles, limitations and/or parametrization of models, becomes a disturbing factor. Depending on the goal of the simulation results, modeling should be fitted exactly to the knowledge level and availability of parameters. In this paper, a hierarchical modeling method is presented, which allows the simulation, of basic principles, up to detailed circuits. Linear models and open-loop operation, are for the understanding of the basic principles. Generic models are added to simulate closed-loop behavior. Detailed component models, are finally required, for the simulation of closed-loop feedback, EMI and thermal behavior. Finally, parasitic components are added, for detailed electromagnetic interference and component stress simulation. The basic working principles are modeled using simple models, and then, depending on the required simulation results, the models are expanded into more detailed models.**

***Index Terms*—Modeling, power electronics, switched mode power supply, simulation, generic model, current mode control**

I. Introduction

The role of simulation in the education of power electronic, and especially switched mode power supplies is nowadays not questionable any more. However, in the past, this used to be different, and simulation was more or less accepted on the fence, regarding modeling effort and the accuracy of simulation results.

With the availability of microcomputers in the early seventies, modeling and simulation for power electronics became within reach for many power electronics engineers. In integrated circuit design, simulation already was taking his place, simply because of the large size of the circuits and the inability to breadboard such large circuits. From those early developments in circuit simulation programs, Spice [1], developed at the Berkeley University, became the most used electronics simulation package. Because it was very flexible and the original source code was freely available, it was adopted by many research institutes and commercial developers. Commercial packages for microcomputers started to appear, and some of them even exists today. Since so many engineers were using Spice, it somehow became a verb [2]. Spice became the de-facto standard in circuit simulation and is also widely used for power electronics, although there were no models to tackle the switching behavior correctly, other than detailed simulation models.

The first approaches dedicated towards using simulation for power electronics, relied heavenly on the mathematical background of users. Knowledge on applying state space equations and numerical routines, were required from the users [3]. The application of state space equations, versus the Modified Nodal Analysis [4] for simulation, was extensively discussed during power electronics specialist conferences [5].

In the eighties and nineties, new simulation programs were developed, especially for switched mode power supplies [SMPS] [6], [7]. Multilevel modeling techniques were introduced, that allow a more dedicated modeling of power electronics, control and electrical machines [8], and let to the development of the program Caspoc [9]. An overview of simulation programs by the end of the nineties, is given in [10].

By the turn of the century, simulation becomes accepted for studying power electronics behavior [11]. Nowadays, simulation is so common, that it is simply integrated into undergraduate and graduate curriculum [12], [13].

Simulation is used in the popular textbooks on power electronics. Especially Spice is included in [14]–[18], while Caspoc is used in [19], [20].

The need for the hierarchical approach is first outlined in section II, and in section III, the three modeling approaches are introduced. The advantages of a high simulation speed, to first get an idea of the power electronics circuit, is described in section IV. The generic approach, which includes much more functional control and semiconductor models with parasitic components, is discussed in section V. The approach where the semiconductor models are modeled in detail, is discussed in section VI

II. Hierarchical approach

Teaching power electronics requires a broad knowledge from the students. Not only electronics, but also topics like magnetic, thermal, control and electromagnetic interference are important. The undergraduate laboratory courses, are mostly focusing on a single specific topic. Because the operation of power electronics requires knowledge from all of the above mentioned topics, also modeling and simulation of all those topics could overwhelm the students.

A. Approach

The approach shown in this paper is intended to start at the second year undergraduate level, where students do have a background in basic electronics and electrical circuits, but are not yet familiar with detailed electronics circuit analysis and behavior. However during the three year studies, they are developing their knowledge and skills on the above mentioned topics. During this three year period, the students are thus confronted with a first understanding of the basics, up to a profound and detailed analysis in their last year. It would be possible to have different software for the different knowledge levels of the students, but giving them a hierarchical software approach, where they can start with simple basic models up to detailed models, allows them to share user experience and build-up useability knowledge, over the years. The continuous use of the same modeling methods with increasing complexity, is inline with the development of their theoretical knowledge of the subject under study.

B. Multilevel

The multilevel approach [6], [8], [9] was a first attempt to the hierarchical approach, but the focus was on solving the engineering simulation issues, not the educational aspects. The hierarchical approach for the theory of power electronics, is also applied in the early textbooks [14], [17] and [18]. In the explanation of the theory, simple ideal switch models instead of detailed Mosfet models were used, to demonstrate the operation and behavior of the power electronics. The derivation of the mathematical relations is first based on the basic approach, the switch model, but finally extended to include the typical details of the Mosfets. A similar approach is used for simulation, where the students start with using a simple simulation model, but gradually improve the models with more details. The approach for the explanation of the theory in [14], [17], [18], is followed in [15] and [16] The hierarchical approach for modeling and simulation is practiced using the multilevel techniques [6], [8] and extended more generally to electric systems in [11].

C. Software

As an example of mismatch in use of software for education, are the many issues with convergence problems of the numerical integration methods used in the spice-based simulators [2], [8], [10]. If students are confronted with these kind of problems, not related to their actual learning process, time and effort is spend on solving those problems. Time and effort that could have been spend on their actual learning process of understanding and practicing the subject of power electronics. Therefore the entry level should be free of these above mentioned software problems, and dedicated to the teaching subject.

D. Level

The hierarchical approach as described in this paper, is thus applicable to the set of power electronics courses for undergraduate and graduate study. The examples in the next sections, detail how this hierarchical approach is used to structure the broad range of topics relevant to education of power electronics.

E. Implementation

To implement this hierarchical approach for modeling and simulation, the development of undergraduate and graduate courses, should address the importance of combining the theory, simulation and laboratory experiments, within the same semester, or half semester [12], [13].

The hierarchical approach was applied to the undergraduate third semester power electronics courses in Delft, Netherlands [12], [13], [22], where the entry level is only a basic understanding of electric circuits and electronics. The undergraduate students in the third semester do not yet have an understanding of numerical integration methods. The underlying numerical integration methods therefore have to be hidden to the students, and only models with basic parameters should be applied. Students thus practice the influence of those basic parameter(s), and are not affected by problems of the underlying software.

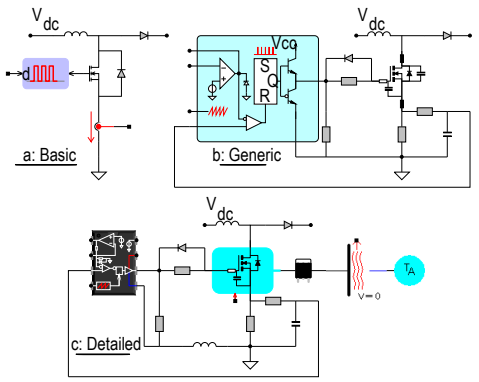


Fig. 1. Modeling approaches for SMPS. a:Basic model with constant duty-cycle and ideal current measurement, b:Generic peak-current-mode control model and inclusion of parasitic components around the semiconductor, c:Detailed semiconductor and control IC model [21].

III. Virtual Workbench

A virtual workbench is basically modeling and simulation. However the question is, if the model should be as simple and general as possible with less parameters, or if it should look as close as possible equal to the final device. The hierarchical approach is split over three types of models:

- Basic ideal model
- Generic functionality model
- Detailed component model

The followed approach, is to start with a simple basic ideal model, with a limited set of parameters. Gradually, the level of detail should be improved, using generic models, to finally have a model, based on real component specification.

It should however be emphasized that detailed control design, power loss estimation or EMI can be evaluated from each of the three model approaches. Figure 1 shows the differences between the various approaches.

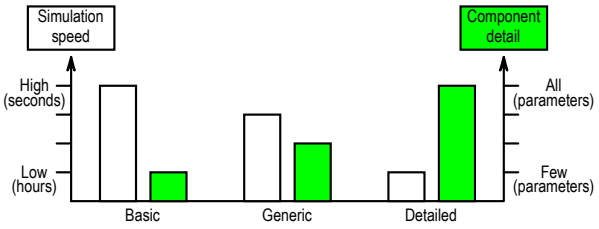


Fig. 2. Simulation speed decreases, while component detail increases, when going from a Basic towards a Detailed model.

In figure 1-a, a basic model is shown, where only an ideal switch model represents the semiconductor. It is either on or off, but has a series resistor, to model the conductance losses. If the current through the semiconductor is required in a peak-current-mode control [16], it can be measured using a simple ideal current sensor. Especially in the basic ideal model, we can extensively use the advantage of the multilevel modeling approach, in that we model the control circuitry using a block-diagram, where only signals are modeled [9]. In figure 1-a, the gate of the Mosfet circuit model, is directly controlled from a block-diagram signal [6].

In figure 1-b, the generic model is shown. Although this is still based on ideal models for semiconductors, the controllers and especially the analog controllers are modeled using a generic model. Such a generic model includes typical functionality of SMPS control IC's, such as build-in oscillator, current limitation and an operational amplifier, with correctly defined bandwidth and gain. Typical control IC's for SMPS for voltage mode and peak-current-mode, being single or dual ended, share a large common set, of functionality and behavior. The semiconductor models for the Mosfet, diode or IGBT, are in principle the basic ideal models, but now they are extended with the most important parasitic components that influence their behavior and the waveforms during switching. For example, the lead inductances are mainly responsible for transient voltage peaks [14]. The junction capacitors C_{GS}, C_{GD}(not drawn in the symbol) and C_{DS} define the charge required during switching as well as the switching losses [14].

Figure 1-c, shows the detailed component model. Here the typical component circuit models are employed, which contain the electrical specifications for that typical IC. Compared to the generic models, the oscillator is modeled inside the detailed model, in the same way it is operating in the real IC. It means that you can do the same analog circuit tricks, like using this signal for control purposes, such as Constant-On-Time control, or add external triggering or synchronization. Also the fan-out of the gate-driver is modeled, as well as the external charge pump circuit components, that have influence on the transient behavior. The semiconductor switch, is modeled using a non-linear relation between the gate-source voltage and drain-source current, to model the non-linear trans-conductance of the semiconductor.

Surprisingly, you can model most of the typical engineering questions and design issues with each of the three modeling approaches. Table I shows the three approaches and their strength per application.

TABLE I
ENGINEERING QUESTIONS AND DESIGN ISSUES DEPENDENCY ON THE MODELING APPROACH

	open-loop	closed-loop	Power Loss	EMI
Basic	++	+	-	-
Generic	+	++	0	0
Detailed	+	0	++	+

In table II, the modeling approaches are shown for typical engineering questions and design issues. Although in principle each modeling approach can be used for each topic, there are differences in the results and what is sought for. For example, the waveforms for open-loop or closed-loop control can be simulated with any of the approaches, but the basic simple model will have the shortest simulation time. For EMI, the generic model will already show the most important high frequency transients, where the basic model can only show the low frequency harmonics. The detailed model will not reveal more information compared to the generic model, when it comes to the transients, but will require a much longer simulation time. The power loss is best evaluated using the detailed model, as it includes the detailed non-linear behavior, during switching of the semiconductors. Depending on what is expected to be investigated using modeling and simulation, a choice has to be made, among the three modeling approaches.

Figure 2, shows the differences in component detail and simulation speed for the three approaches. Clearly visible is the high simulation speed for the basic model, compared to the detailed model. However, the detailed model will show more details in the waveforms, but requires more simulation time.

IV. BASIC MODELING APPROACH

The basic model for the boost converter simulation, is shown in figure 3. The purpose of this simulation, is to reveal the start up behavior and the maximum inductor current, if the boost converter is started with a constant dutycycle. Also the inductor current ripple and the influence of the equivalent series resistor R_{esr} of the capacitor C_1, can be studied using this simulation. The maximum current during start up is around three times the maximum inductor current for nominal operation. The duty cycle is set to the ideal value of $d = 0.6$

$$\frac{V_{out}}{V_{in}} = \frac{1}{1-d} \tag{1}$$

If an output voltage of 48 volt is expected, the duty cycle should be set a little bit higher, to compensate for the losses due to the voltage drops, in the boost converter. A value of $d = 0.6$ is used, and ideally we expect:

$$\frac{V_{out}}{V_{in}} = \frac{1}{1-0.6} = \frac{1}{0.4} = 2.5 \tag{2}$$

$$V_{out}^{ideal} = 2.5 \cdot 20 = 50\text{volt} \tag{3}$$

Since the output voltage in the simulation is less than 48 volts, it shows that using the basic simulation, the voltage drop over the series resistance and output diode, are already modeled.

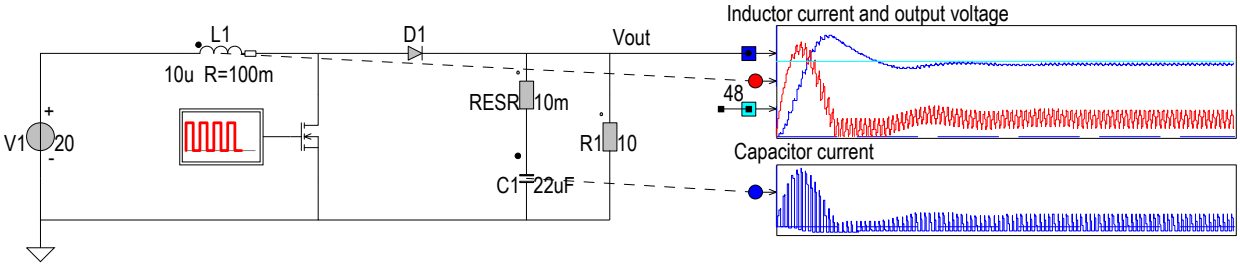


Fig. 3. Basic model for the open-loop simulation(Total simulation time $T = 1ms$, step-size $dt = 100ns$) of the Boost converter [21].

TABLE II
SIMULATION GOALS

	Basic	Generic	Detailed
Control	• Steady state Behavior, • Start up,Stability	• Control limits, • Stability	• Component non-linearity
Losses	• Conduction, • Inductor winding loss, • Capacitor ESR loss	• Conduction, • Inductor winding loss, • Capacitor ESR loss	• Conduction, Switching, • Inductor winding loss, • Core loss approximation, • Capacitor ESR loss
EMI	• Harmonics from the switching waveforms	• Harmonics from the switching waveforms and parasitic components	• EMI from the parasitic components combined with dynamics from semiconductors

In the basic model approach, these conduction losses are already modeled, by simply inserting series resistance in the circuit. The inductor has a series resistance of $100m\Omega$, the Mosfet has series resistance $R_{DS} = 100m\Omega$ modeled, as well as the forward voltage drop of the diode, being equal to around 0.7 volt. The conduction losses in the output capacitor are modeled by the series resistance $R_{esr} = 10m\Omega$.

Since the Mosfet model is ideal, the simulation speed is high, but the only parameter used for modeling the Mosfet, is the on-state resistance R_{DS}. But although we use simple models, we can already study the harmonic content in the input current, being equal to the inductor current. Also the ripple in the output capacitor, and thus in the output voltage, is simulated, showing the influence of the series ESR of the capacitor. Adding capacitors with different ESR in parallel will directly show the influence on the output voltage ripple. Although a control could be added in this simulation, the generic model approach is better suited for this as is outlined in section V.

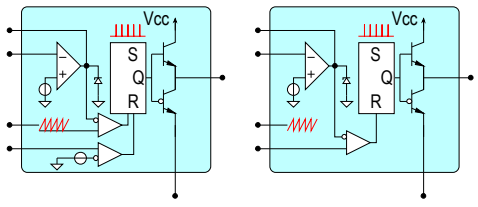


Fig. 4. Generic models for voltage mode(left) and current mode(right) controllers [21].

V. GENERIC MODELING APPROACH

Closing the loop can be done using analog or digital control. In this simulation example, we choose a peak-current-mode control IC. There are various types from different manufacturers, but internally they show the same functioning, see figure 4. Important when closing the analog control-loop, is the compensation network and bandwidth of the Opamp from the control IC. The current sense filtering and slope compensation, are added to the simulation.

Since the series resistance of the inductor and capacitor are modeled, the output voltage ripple as well as the conduction losses are included in the generic modeling approach. The poles and zeros of the control to output small signal transfer function are modeled in the generic model, and therefore, effects of the loop compensation circuit can be studied. Measuring the inductor current for the peak-current-mode control, is done using a simple sense resistor, and an additional low pass filter. These simple components in combination with the current sense input pin of the generic IC model, will give adequate information on the shape of the waveform of the measured current.

The non-linear behavior of the Mosfet has little influence on the stability and transient response of the closed-loop control, and is therefore omitted in the generic model. However, parasitic wire inductance and junction capacitance can be added, to model some of the transients during switching. Since the generic Mosfet model already includes the gate to source capacitance C_{GS}, the influence of the gating resistor can be

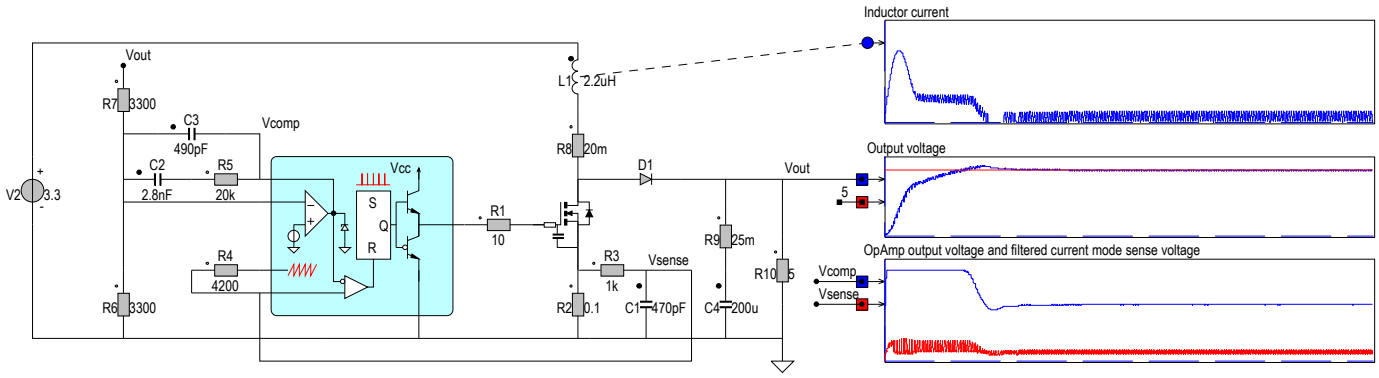


Fig. 5. Generic model for the peak-current-mode control and semiconductor in the simulation(Total simulation time $T = 1ms$, Step-size $dt = 100ns$) of the Boost converter [21].

studied to some extend. For a more detailed simulation of the influence of the gating resistance on the transient switching performance of the Mosfet, the detailed modeling approach as outlined in section VI is required. Since there are more parameters to be included in the simulation compared to the basic simulation, the model is more accurate when it comes to transient behavior, but also requires more simulation time. This is visible in figure 2, where the number of parameters of the model is increased compared to the ideal model. Because of the more refined control model, the simulation speed for the generic model, is decreased compared to the basic model.

Figure 5 shows a boost converter with peak-current-mode control. The closed-loop control is modeled using a generic peak-current-mode model.

A single ended peak-current-mode control IC, is directly driving a Mosfet via a gating resistor. The Mosfet model includes the gate-source capacitance C_{GS} and internal gate resistor $R_{Gate}^{internal}$, as well as the threshold voltage level V_{TH}. The current is measured through the sense resistor R_2. The low-pass filter, build from R_3 and C_1, filters the transient response and removes the sharp spikes, that could lead to wrong current measurements. The losses in the current sense resistor, $I_{DS}^2 \cdot R_2$ are modeled, and the delay of the low-pass filter are thus included in the simulation. To compensate for sub-harmonic oscillation, slope compensation [16] is added in the simulation by R_4. The oscillator pin from the generic peak-current-mode control IC provides a sawtooth voltage, as commonly seen in current mode IC circuits.

The closed-loop feedback circuitry, is a type II, modeled by the components R_5, C_2 and C_3 [15]–[17]. Optimizing the control and changing the pole locations, directly shows the influence on the start-up behavior of the circuit.

The generic model approach in the simulation in figure 5, is especially suited for testing the closed-loop control. It includes all components that have to be designed, and has just enough parameters, like bandwidth and gain in the generic control IC model. It is therefore ideally suited for an undergraduate course, on SMPS control.

VI. Detailed Modeling Approach

To study the influence of the parasitic inductance in the circuit, the rise and fall times for the current in the Mosfet have to be simulated. For this, the detailed Mosfet is used. Here the non-linear behavior of turn-on and turn-off, of the Mosfet is included. It depends on the non-linear relations between V_{GS}, V_{DS} and I_{DS} and the temperature of the junction. Small inductors, modeling short PCB traces of around $10mm$, in the size of typically $10nH$ to $20nH$ per lead [14], are added in the simulation. Over these parasitic inductors, over-voltages during switching are simulated.

The peak-current-mode control IC is included in the detailed simulation, which includes the details of the internal oscillator circuit and its dependency on the external components around the oscillator pin. Slope compensation and the influence of the slope compensation resistor R_4 on the oscillator voltage V_{rc}, are included in this detailed simulation.

The transient behavior as revealed by the simulation of the detailed model, would be equal to a simulation using the generic modeling approach. However, in the detailed simulation, the voltage peaks during switching are visible in the simulation results.

The thermal behavior is included, by adding a thermal model for the Mosfet package. The temperature on the junction is simulated, and influences the on-state resistance R_{DS} of the Mosfet. The detailed model is typically applied in undergraduate and graduate design courses, where the emphasis is on evaluating the influence of the various design details in the experimental results, based on simulation results.

Conclusions

Applying simulation in undergraduate and graduate courses, requires an hierarchical approach, in order to correctly match the complexity of the simulation model, to the knowledge level of the student. Starting with basic models in second year undergraduate courses, and via generic models for second and third year undergraduate, detailed models are applied in last year undergraduate and second year graduate courses.

Instead of focusing on a single model or simulation method, the idea is to use a different type of model, depending on

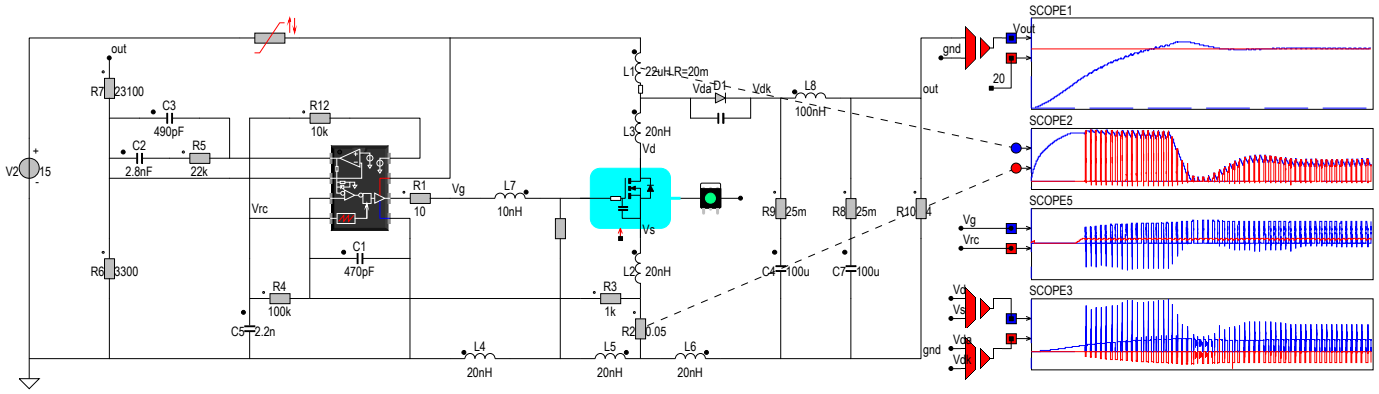


Fig. 6. Detailed model for the peak-current-mode control IC and Mosfet in the simulation(Total simulation time $T = 1ms$, Step-size $dt = 10ns$) of the Boost converter [21].

the needs for simulation results. A basic model would show just enough information regarding the voltage and current waveforms. A generic modeling approach is to be applied when the control has to be designed, while a detailed model is used when thermal issues and EMI have to be simulated.

Depending on what simulation result is required, a modeling approach has to be selected. The basic model will reveal the currents and voltage waveforms, including the influence of series resistance of the passive and active components. Only conduction losses can be simulated using the basic model. The generic model is used when simulating closed-loop behavior. Feedback compensation circuitry can be studied using the generic model. If the influence of parasitic components has to be studied in detail, a detailed semiconductor model is required, where the switching delay is model.

The main advantage for education of the hierarchical approach, is that only those effects are modeled and simulated that are of interest. Students only have to supply a minimum set of parameters, if not all simulation details are required, and will be able to correlate the simulation results due to parameters variations.

REFERENCES

[1] L. W. Nagel and D. O. Pederson, "Simulation Program with Integrated Circuit Emphasis (SPICE)," presented at 16th Midwest Symp on Circuit Theory, Ontario, Canada, April 12, 1973. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/1973/22871.html

[2] C. C. McAndrew, "How Come SPICE Is a Verb?: The Natural Course of a Diverse Engineering Tool," in IEEE Solid-State Circuits Magazine, vol. 11, no. 1, pp. 14-18, Winter 2019, doi: 10.1109/MSSC.2018.2882279.

[3] H. A. Owen, A. Capel and J. G. Ferrante, "Simulation and analysis methods for sampled power electronic systems," 1976 IEEE Power Electronics Specialists Conference, Cleveland, OH, USA, 1976, pp. 45-55, doi: 10.1109/PESC.1976.7072897.

[4] Chung-Wen Ho, A. Ruehli and P. Brennan, "The modified nodal approach to network analysis," in IEEE Transactions on Circuits and Systems, vol. 22, no. 6, pp. 504-509, June 1975, doi: 10.1109/TCS.1975.1084079.

[5] G. W. Wester, "Power electronics modeling, design and simulation session summary," 1975 IEEE Power Electronics Specialists Conference, Culver City, CA, USA, 1975, pp. 62-62, doi: 10.1109/PESC.1975.7085568.

[6] P. J. van Duijsen, "Multilevel modeling and simulation of power electronic systems," 1993 Fifth European Conference on Power Electronics and Applications, Brighton, UK, 1993, pp. 347-352 vol.4.

[7] P. Bauer and P. J. van Duijsen, "Large signal and small signal modeling techniques for AC-AC power converters," Conference Record of the Power Conversion Conference - Yokohama 1993, Yokohama, Japan, 1993, pp. 520-525, doi: 10.1109/PCCON.1993.264201.

[8] G. A. Franz, "Multilevel simulation tools for power converters," Fifth Annual Proceedings on Applied Power Electronics Conference and Exposition, Los Angeles, CA, USA, 1990, pp. 629-633, doi: 10.1109/APEC.1990.66362.

[9] P.J. van Duijsen, "Multilevel modeling and simulation of power electronic converters and drive systems.", Proceedings of the PCIM 1994, Nuremberg, Geermany, 1994

[10] O. Apeldoorn, "Simulation in power electronics," Proceedings of IEEE International Symposium on Industrial Electronics, Warsaw, Poland, 1996, pp. 590-595 vol.2, doi: 10.1109/ISIE.1996.551009.

[11] P. Bauer and P. J. Van Duijsen, "Challenges and Advances in Simulation," 2005 IEEE 36th Power Electronics Specialists Conference, Dresden, Germany, 2005, pp. 1030-1036, doi: 10.1109/PESC.2005.1581755.

[12] P. J. van Duijsen and D. C. Zuidervliet, "Structuring a Switched Mode Power Supply Course, Part I: Lectures," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 1319-1324, doi: 10.23919/MIPRO55190.2022.9803707.

[13] P. J. Van Duijsen and D. C. Zuidervliet, "Structuring a Switched Mode Power Supply Course, Part II: Laboratory," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 1325-1330, doi: 10.23919/MIPRO55190.2022.9803446.

[14] N. Mohan, T.M. Undeland, W.P. Robbins, Power Electronics: Converters, Applications, and Design 3rd Edition, Wiley 2002, ISBN:978-0471226932

[15] S. Ang, A. Oliva, Power-Switching Converters, 3rd ed. CRC Press 2011, ISBN:978-1439815335

[16] M.K. Kazimierczuk, Pulse-Width Modulated DC-DC Power Converters 2nd Edition, Wiley 2015, ISBN:978-1119009542

[17] J.G. Kassakian, M.R. Verghese, G.C. Schlecht, Principles of Power Electronics, Addison-Wesley 1991, ISBN:978-0201096897

[18] R.W. Erickson, D. Maksimovic, Fundamentals of Power Electronics 3rd ed., Springer 2020, ISBN:978-3030438791

[19] J. Pollefliet, Electronic Power Control: 1 Power Electronics, 8th Edition, Lannoo Publishers, 2019, ISBN:978-9038225258

[20] J. Pollefliet, Electronic Power Control: 2 Electronic Motor Control, 8th Edition, Lannoo Publishers, 2019, ISBN:978-9038225265

[21] Caspoc Simulation and Animation, Simulation Research, [Online] Available: https://www.caspoc.com,

[22] DC laboratory, The Hague University, TIS, Delft. [Online] Available: http://dc-lab.org

Teaching Introductory Parallel Programming using Two-Player Online Games

Sathiamoorthy Manoharan and Xinfeng Ye
School of Computer Science
University of Auckland
New Zealand

***Abstract*—Parallel programming is a powerful tool for computing tasks faster and more efficiently. Teaching an introduction to parallel programming requires explaining the fundamentals of parallelism, the various types of parallel programming constructs, and the tools and techniques used to build and debug parallel programs. Teaching introductory parallel programming can be a challenging task. This paper is an experience report, detailing how a two-player online game is built as a motivating focal point to engage students and explain various parallel programming concepts and constructs.**

***Index Terms*—teaching programming, parallel programming, asynchronous programming, two-player games**

I. Introduction

The benefits of parallel programming are readily realized in the modern computing landscape where processors have many cores. It is a powerful technique that can be used to drastically improve the performance of applications and make them more efficient. As such, it is essential for students to have a good understanding of parallel programming techniques in order to best utilize the technology.

A parallel program course would typically cover many fundamental concepts such as threads, processors, cores, distributed systems, and synchronization. It would compare single-threaded and multi-threaded programming, and explore the various types of parallel programming such as data parallelism, task parallelism, and pipeline parallelism.

For an introductory course on parallel programming, some of these concepts can be difficult to learn without hands-on practical work. We took a project-based learning [1] approach whereby the students built a two-player online game from the ground-up to to gain a deeper understanding of the concepts and apply their knowledge in a practical setting.

This paper is an experience report that discusses the benefit of developing a two-player online game using a project-based learning approach in achieving the learning outcomes of our introductory course on parallel programming in the School of Computer Science.

The rest of the paper is organised as follows: Section II reviews some of the recent work related to ours. Section III outlines the project as specified, while section IV discusses how the project was executed. Section V adds further considerations and reflections. The final section concludes with a summary.

II. Related Work

This section discusses some of the recent related work in teaching parallel programming courses. Earlier works are reviewed systematically by Carneiro Neto and colleagues [2].

Younis and colleagues [3] use a series of multi-themed group projects to introduce parallel programming to their computer organization class. They report results showing significant improvement to the students' parallel programming skills, which they attribute to their use of project-based learning.

Chen reports on the design of a parallel programming course for junior students, drawing from prior expertise in teaching senior students [4]. Nearly half of the course is devoted to programming exercises, offering students plenty of hands-on experience.

Danelutto and Torquati leverage structured parallel programming principles and several parallel programming frameworks to improve the teaching and learning process [5]. They argue that over a period of eight years, their teaching approach has shown a steady trend of increase in student performance as implied by the student grades.

Martins and colleagues use programming-contest-style challenges in parallel programming to encourage students to learn parallel programming [6]. They argue, using quantitative and qualitative scores, that such challenges engage students and motivate them to master the concepts required to solve the challenges.

Marzulo and colleagues, just like Martins and colleagues, use challenges to teach parallel programming [7]. In this case, the challenges are parallel programming marathons, which, according to the authors, kindle student interest and increase motivation.

Kurniawati presents a short paper on their experience with teaching parallel programming, and notes the difficulty students faced in designing and debugging parallel programs [8].

Conte and colleagues stress the importance of the teaching parallel programming early in the degree program so that students have the background and skills to use parallel programming concepts throughout their degree [9].

III. The Project

The project required students to design and implement a multithreaded game server suitable for two-player games (such as Backgammon, Checkers, Chess, Go, etc.). The server helps to pair up players and to coordinate the exchange of game

moves. It does not know anything about the underlying game itself. In other words, the game server does not attempt to understand the game moves or enforce any rules. The game rules are left to the two players to enforce, just like in a real board game. This allows the game server to be generic, and enables the students to focus their design and development solely on the multithreading aspects.

The server was to be programmed using the C# programming language, and needed to be constructed using a synchronous server socket from the ground up. Higher-level APIs (such as HTTPListener) were not allowed, and the students had to implement the required multi-threading and synchronization themselves.

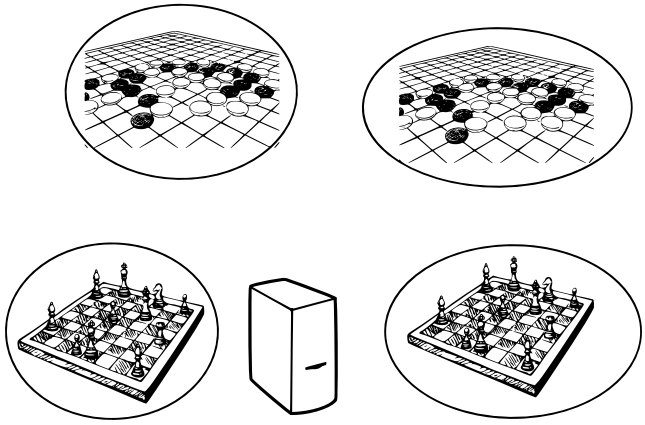


Fig. 1: An illustration of the multithreaded game server for two-player games. The game server does not attempt to understand the game moves or enforce any rules. The game rules are left to the two players to enforce, just like in a real board game. This allows the game server to be generic, and enables the students to focus their design and development solely on the multithreading aspects.

The server was to use HTTP REST as the basis for communication with the game clients. To this end, the following (GET) endpoints were suggested:

/register
: This endpoint generates a random username for a player, registers this name, and returns to the user the registered name. The player is required to pass this username in all subsequent transactions for identification.

/pairme?player={username}
: This endpoint attempts to pair the given player with another player. It returns a game record (or a suitable subset of it). The game record is a tuple containing a game ID, game state, username of the first player, user name of the second player, last move of the first player, and the last move of the second player. When there is no other player waiting, the game record will contain a newly allocated game ID (which could be a GUID), a game state indicating ”wait”, and the username of the requesting player as first player. The rest of the elements in the tuple are not defined. When there is a waiting player, the game record will be the game record first created for the waiting player, updated to add the second player and the state indicating ”progress”. The state ”progress” tells both players that the game can now begin. The endpoint can be invoked by both players as many times as they want before the commencement of the game. This helps the first player to ”poll” the state to see if a second player has been paired up.

/mymove?player={username}&id={gameId}&move={move}
: This endpoint, when used during the game’s ”progress”, will supply the user’s move to the server. The ”last move” of the player in the corresponding game record will be updated with this supplied move.

/theirmove?player={username}&id={gameId}
: This endpoint, when used during the game’s ”progress”, will collect the other player’s move from the server. The game server will supply the ”last move” of the other player from the game record corresponding to the given game ID.

/quit?player={username}&id={gameId}
: This endpoint notes to the server the intention of the player to quit the game. The server will remove the game record corresponding to the given game ID. Attempt by the players to access the game record (for example, to get a move) will fail after the record is removed.

Students were told to implement other endpoints for diagnostic or informational purposes (e.g., a /debug endpoint or a /version endpoint) as required.

They also needed to take appropriate actions when errors are encountered (e.g., invalid endpoints, malformed endpoints, invalid parameters to endpoints, etc.).

Typically, a single thread will handle a single player’s endpoint interactions. The server, for the sake of efficiency, should keep the connection alive to handle the player’s requests in the same thread. However, it is possible that the client may, from time to time, close the connection. For example, a browser will close the connection if it deems the connection inactive for a period of time. The server should be able to gracefully handle this scenario and continue to serve the client (potentially using a new thread).

Students were asked to think about what information is shared among the threads and avoid potential race conditions efficiently using appropriate concurrency controls.

In order not to complicate the server implementation more, students were asked not to persist any of the information. Meaning that, there was no database backend involved – the information is kept only in nonpersistent memory and lost if the server is restarted.

One thing that the server specification did not include is the ability to support multiple game types. This was intentional, as we wanted to see if any student picks this up (and none did).

A. Testing

The students were asked to come up with an appropriate testing strategy for verifying the correctness of the server. They

were required to build a browser-based client for a chosen two-player game to help with their testing and verification plan, and to demonstrate game play using their server.

B. Learning Outcomes

The parallel-programming-related learning outcomes of this project are the following

1) being able to create and use multiple threads
2) understanding critical sections
3) understanding and using mutexes and condition variables
4) avoiding deadlocks
5) ordering events

The marking rubric took into account these learning outcomes. A high mark in the project therefore is likely to indicate that the student met most of the learning outcomes.

IV. Project Execution

Most projects require one to have a broad knowledge of many concepts to execute and complete them successfully. In addition to parallel programming skills, the students required good classic programming skills, an understanding of network programming, in particular socket programming, basic knowledge of HTTP, and systematic testing skills. While most of these topics had been covered in prerequisite subjects the students had studied, we observed gaps in their knowledge as well as lack of practice becoming a barrier to the successful completion of the project.

The key skills a large proportion of students lacked are incremental development and effective testing strategies, both of which are prerequisite material and are well outside of the scope of this subject.

This required us to have additional help sessions to guide students on iterative development. In particular, for this project, we outlined an iterative development plan as follows.

1) Compile and run the synchronous server socket example from the documentation. You will see that the server closes the connection soon after sending the response, and telnet client will terminate since the server closed the connection.
2) Now change the server so that it does not close the connection soon after sending the response. When you have successfully done this, you can use the same telnet client to send requests to the server and receive responses. This connection is now alive, and this is what HTTP's keep-alive does (loosely speaking, since HTTP has a timeout).
3) Now that you have a server that keeps the connection alive, test with another concurrent telnet client (while your first one is still active and connected to the server). You would see that the second telnet client would connect to the server but it won't receive any response to requests it sends. The first client is keeping the server busy here.
4) Try closing the first client and see if the server now responds to the second client. You will see that it won't. This is because the server does not see that the first client has closed the connection. You need to find out how to detect this and get the server to be ready to serve the second client when the first client quits.
5) At this point you have a server that keeps the connection alive and is able to tidy itself up when clients disconnect from the server.
6) Now you are ready to get multithreading into the play to make the server service multiple clients concurrently.

Yet, a number of students did not test their implementation fully. In addition, some of the students assumed that an HTTP request would come in a single packet, an assumption that would not hold if a slow client such as telnet were to be used. This resulted in the average project mark being 48%, less than the pass mark of 50%. We therefore gave everyone a second chance to test and resubmit their project, after providing student-specific feedback in one-to-one sessions. Fig. 2 shows each student's mark for their initial submission as well as their second attempt.

Note that, the class size is small (just 28 students, as can be inferred from Fig. 2), and therefore having one-to-onse feedback sessions with students was possible.

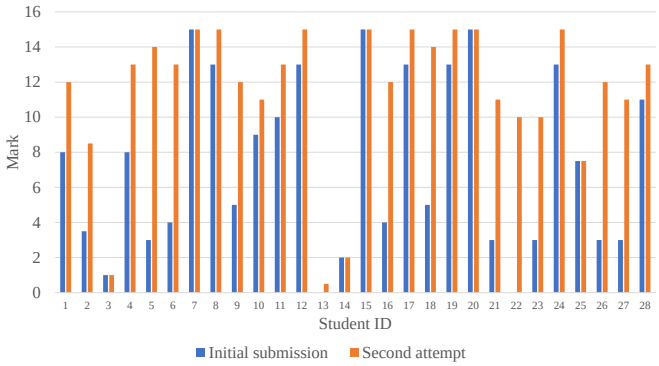


Fig. 2: Initial marks and marks for the second attempt for each student. The project was marked out of 15 and contributed to 15% of the final mark.

On average, students gained 4.2/15, or 28%, marks in their second attempt. If we exclude the three students who already scored full marks in their initial submission, then the gain becomes 4.7/15, or 31%. The average project mark after the second attempt rose to 76%. These observations lead us to believe that the feedback from the initial submission contributed to better marks in their second attempt, and consequently a better overall learning.

Seven students did not use proper synchronization or locking in their implementation. One of the students had a demonstrable deadlock in their game play.

Most students spent a lot of effort to complete the project to a high standard, and informal feedback from the students indicated that they highly appreciated the project-based learning approach and being able to develop a complete game from the ground up. Being generic, the game server allowed students to implement a game client of their choice. Students therefore implemented a variety of good, and working game clients. Fig. 3–Fig. 6 show screenshots of some of these clients.



Fig. 3: A chess client implementation.

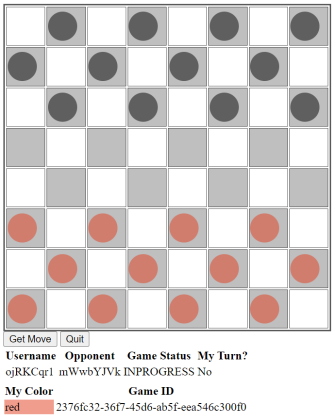


Fig. 4: A checkers client implementation.

V. Discussion

Teaching introductory parallel programming can be a difficult task, especially for those who are new to the subject. There is a lot of material to cover, and it can be difficult to decide what topics to focus on and how to best approach them.

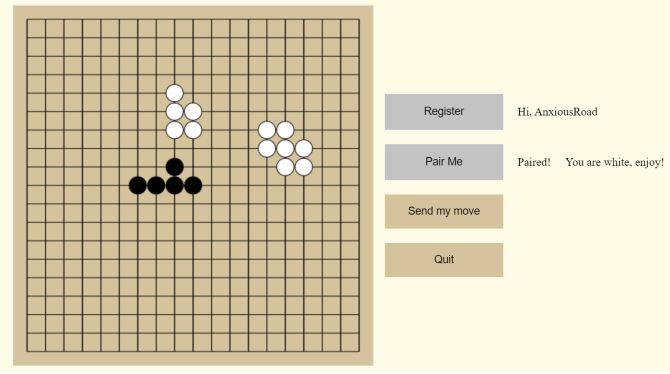


Fig. 5: A gomoku client implementation.

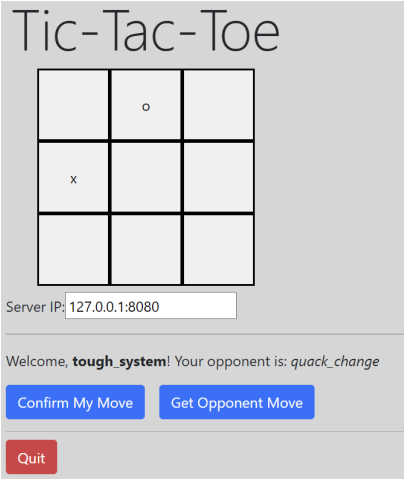


Fig. 6: A tic-tac-toe client implementation.

A practical project such as the one used in this paper provides students with opportunities to explore and extend the topics discussed in class. It helps them gain a better understanding of how parallel programming works in practice in a wide context. Formative feedback and fostering collaborative discussions during the project also help.

Most of the difficulties a number of students faced did not arise from the main focus of the course. They had difficulties with incremental development, testing, debugging, and also lacked network programming skills. While project-based learning is fun and effective, it requires students to have a broad understanding of many related subjects in order to successfully complete the project.

VI. Conclusions

Teaching introductory parallel programming can be a challenging task. However, by providing hands-on experience, and giving students opportunities to explore, it is possible to create an effective and engaging course. Giving students a chance to re-work their practical work after feedback is also a strategy to improve the overall gain in learning outcomes.

References

[1] G. E. Veselov, A. P. Pljonkin, and A. Y. Fedotova, "Project-based learning as an effective method in education," in *Proceedings of the 2019 International Conference on Modern Educational Technology*, ser. ICMET 2019. New York, NY, USA: Association for Computing Machinery, 2019, pp. 54–57.

[2] J. A. Carneiro Neto, A. J. Alves Neto, and E. D. Moreno, "A systematic review on teaching parallel programming," in *Proceedings of the 11th Euro American Conference on Telematics and Information Systems*, ser. EATIS '22. New York, NY, USA: Association for Computing Machinery, 2022.

[3] A. A. Younis, R. Sunderraman, M. Metzler, and A. G. Bourgeois, "Case study: Using project based learning to develop parallel programing and soft skills," in *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2019, pp. 304–311.

[4] X. Chen, "Designing a parallel programming course for lower-division students," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2020, pp. 1009–1011.

[5] M. Danelutto and M. Torquati, "Increasing efficiency in parallel programming teaching," in *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 2018, pp. 306–310.

[6] G. Martins, P. S. Lopes de Souza, D. Jose Conte, and S. M. Bruschi, "Learning parallel programming through programming challenges," in *2020 IEEE Frontiers in Education Conference (FIE)*, 2020, pp. 1–9.

[7] L. Marzulo, C. Bianchini, L. Santiago, V. Ferreira, B. Goldstein, and F. França, "Teaching high performance computing through parallel programming marathons," in *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2019, pp. 296–303.

[8] R. Kurniawati, "Teaching parallel programming with Java and Pyjama," in *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 2*, ser. SIGCSE 2022. New York, NY, USA: Association for Computing Machinery, 2022, p. 1109.

[9] D. J. Conte, P. S. L. de Souza, G. Martins, and S. M. Bruschi, "Teaching parallel programming for beginners in computer science," in *2020 IEEE Frontiers in Education Conference (FIE)*, 2020, pp. 1–9.

Curriculum Development for Socially Responsible and Sustainably Acting Engineers

M. Derda*, M. Wedel ** and M. Albrecht***
* Technische Universität Berlin/Department of Fluid System Dynamics, Berlin, Germany
** Technische Universität Berlin/Department Work, Technology and Partizipation/ Department Fachdidaktik Arbeitslehre, Berlin, Germany
*** Technische Universität Berlin/Department Fachdidaktik Arbeitslehre, Berlin, Germany
mareen.derda@tu-berlin.de

***Abstract* - It is the self-proclaimed ideal of university education to provide students with knowledge and skills to act responsibly and sustainably. There seems to be a university-wide consensus that ethics, sustainability, digitization, diversity, and scientific practice should be added to the subject-specific syllabi. What might be construed as a contradiction between industry demands and societal norms is in fact a response to the corporate needs of industry and business. They too require engineers to be trained to think systemically and react adequately to changing societal conditions. This also includes the on-going challenge of diversity and gender equality in industry and business, and in particular the well-established demand to attract more women to pursue a career in engineering. Based on recent experiences in the curriculum development process at the Technische Universität Berlin, the article discusses, based on first qualitative insights, how the development of the curricula can work in the given context.**

Keywords* - *curriculum development, ethics, sustainability, digitization, diversity, scientific practice, including all status groups

I. INTRODUCTION

Not all citizens are engineers, but all engineers are citizens. Given the key role engineers hold in shaping technological development and its impact on social development it is essential that they are empowered to shape their actions with an awareness of their overall social responsibility. Fostering the corresponding competencies is an educational mission. In that regard it is the university's mission to provide students (as representatives and multipliers within society) with knowledge and skills for responsible and sustainable action.

The argument to be developed shall be understood in the context of the canon of values of European-style liberal democracies as enshrined in the Charter of Fundamental Rights of the European Union [1]. Therefore what 'responsible and sustainable action' means is not argued based on ideas of ethical pluralism [2] or individual opinions as part of a democratic will-building process, but on well-established principles and core values that safeguard citizens' rights and their duties as part of constitutionally consolidated social contracts within the European Union [3].

Building on these principles there seems to be a consensus in higher education that research agendas and teaching curricula should be complemented by content and competence building around ethics, sustainability, digitization, diversity, and good scientific practice [4].

In the first part of this article, we illustrate the named topics that should be integrated into modules and curricula to educate socially responsible and sustainable engineers. The remainder of the paper explains why these choices were made. In addition to the requirements from industry and business, this section also addresses a targeted increase in the attractiveness of the courses, especially for women. Finally, it is discussed how, based on our experience, integration into and revision of the curricula can succeed.

II. WHAT TO INCLUDE IN THE CURRICULUM FOR SOCIALLY RESPONSIBLE ENGINEERS

A. Digitization and its Ethics

In 2021 the European Commission proposed a policy program "for the Digital Decade" including targets and objectives for digital skills, infrastructures, and the transformation of businesses and public services [5].

In addition, the Commission proposed a "European Declaration on Digital Rights and Principles" in which it elaborates the value-based and human-centered approach, whereas technology should serve, benefit, and empower all people through responsible and diligent action by all actors – public and private – involved in its design [6].

In that regard the importance of acquiring digital skills through training and lifelong learning is emphasized, along with the commitment to bridge the digital gender divide and support efforts to develop digital competences such as media literacy and critical thinking for active participation in society and democratic processes [6].

The idea of a declaration of digital rights and principles for the European Union continues a supranational development that is described as digital constitutionalism [7] and which aims to strengthen the digital sovereignty of European citizens in a legally binding way [8]. Consequently, the European institutions have established (or are preparing) a regulatory

framework to safeguard European values in the digital sphere [9]: General Data Protection Regulation; Digital Service Act; Digital Markets Act; Data Governance Act; Draft Machinery Regulation; Revision Product Liability Directive; Draft Artificial Intelligence Act (AIA).

In the context of this paper, the risk-based approach of the draft AIA shall be highlighted as an example for digitization as a hole. The goal of this Regulation is not to ban technologies but to avoid undesirable application scenarios [10]. The prerequisites for this approach are on the one hand, that undesirable application scenarios are clearly defined, and, on the other hand, that producers, providers, users and affected persons have the skills, knowledge and understanding that guarantees an informed deployment of respective technologies as well as a sufficient level of awareness about opportunities and risks of AI, and thereby promoting its democratic control [11].

The ethical basis that defines the framework for undesirable applications scenarios in the context of establishing a trustworthy AI ecosystem has been defined by the High-Level Expert Group on Artificial Intelligence set up by the European Commission [12]. Four ethical principles – respect of human autonomy, prevention of harm, fairness, and explicability – and seven key requirements have been proposed: Human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being; accountability. This normative approach will only be executed as intended if respective actors have the knowledge and skill to do so.

Considering our discussion, we follow the argument of [11, 13] that digital skills, knowledge, and competencies, that is Digital Literacy, must be promoted in all sectors of society and must not be limited to learning about tools and technologies but also to learn about the social dimension, that is the value and risk-based approach, of (digital) technologies. Engineers play a key role in that effort because they bear special responsibility as experts and potential innovators in their respective fields and are gatekeepers when it comes to the implementation of social values within technology and its development.

While digital competencies within the EU have been defined as part of a key competence framework for lifelong learning, interlinked with other competences in particular for citizens and educators [14, 15] we argue such a framework needs to be refined for engineers with a particular focus on fostering an ethical mind-set [12]. The goal must be to make engineers aware that they can and should participate in shaping the societal development. This includes all stakeholders, e.g. those involved in making the products (the designers and developers), the users (companies or individuals) and other impacted groups (for whom decisions are made during the research, development and deployment phase).

B. Ethics for Engineers

Ethics describes the theory of morality, which deals with the analysis, reasoning, justification and criticism of moral norms, principles, values and more [16].

In [17] the development, motives and possibilities to integrate Ethics across the curriculum are presented in detail. The article mainly refers to the integration of ethics in engineering. One origin of this initiative lies in the 80s in the United States, where as a reaction to some political and financial scandals and a perceived increasing decay of moral values as a result of declining religiousness, the teaching of ethical content was gaining more importance.

According to [18], teaching ethics has five main goals, these are: Stimulate the moral imagination; recognize ethical problems; create a sense of moral obligation; develop analytical skills; tolerate and reduce disagreements and ambiguities.

With regard to engineering training, we argue, that ethical content should not be taught separately from technical content but should be linked to it. The consequences of technology should be assessed and reflected on, for example, using case studies on ethical issues or dilemmas.

This is of particular importance since digital competencies for responsible engineering are – as deduced above – based to a great extent on ethical approaches. It might therefore be useful to link general reflections on ethical principles with applicable approaches currently discussed e.g., in the context of trustworthy AI. Having key principles and corresponding requirement (human oversight, transparency, accountability etc.) consolidated in a respective framework will be helpful when it comes to the implementation and translation of ethical principles into technology and its development.

C. Sustainability

Building on the EU's logic of competence profiles and the intertwined nature for ethics and digitization identified above it seems reasonable to assume that there is also a high degree of intersection regarding the topic of sustainability. According to the definition of the United Nations Sustainable Development Group, sustainability is the vision of a peaceful, just, socially inclusive, intra- and intergenerational coexistence of all people while at the same time respecting nature and the environment under the prerequisites of ecologically and socially compatible economic activity [19].

Accordingly, to "embody sustainability values, and embrace complex systems, in order to take or request action that restores and maintains ecosystem health and enhances justice, generating visions for sustainable futures", knowledge, skills and attitudes must be promoted by way of education and training [20].

Again, when it comes to sustainable development, that is "the many processes and pathways used to stimulate development, or achieve progress, in sustainable ways" Engineers play a key role because they bear special responsibility as experts and potential innovators in their respective fields and are gatekeepers when it comes to the implementation of sustainable approaches within technology and its development [20].

D. Diversity

The UNESCO Salamanca Declaration emphasizes that every child has a right to education, and schools should accommodate learners' individual characteristics, abilities, and learning needs [21]. Since this concept of individuality is not limited to general education in schools, this mandate must be extended to teaching in higher education as well, where the challenge lies in balancing the individual prerequisites of the students with the social and professional requirements outlined above. To achieve this, good teaching must be designed based on seven principles [22], whereas good teaching (1) promotes contact between students and lecturers, (2) cooperation between students, (3) active learning, (4) gives prompt feedback, (5) places special emphasis on study-related activities, (6) sets high standards, and (7) respects different abilities and skills.

How do these considerations relate to diversity and variety? Good teaching is inclusive teaching and therefore normal learning groups are no different from inclusive learning groups [23]. Teaching should be fundamentally designed to be diversity-sensitive, building upon students' prerequisites, while encompassing all aspects of diversity, including different talents and abilities, and enabling them to engage with the teaching content [24, 25].

Three fundamental structures that notably impact teaching quality are identified: cognitive activation, constructive support, and class management [26]. Effective class management enables efficient and goal-directed learning, while cognitive activation and constructive support provide students with opportunities to learn based on their individual prerequisites. Individual learning and processing paths should serve as the foundation of learning, which necessitates appropriately oriented teaching that considers these quality dimensions [27].

E. Scientific Practice

In 2017, the academic senate of the Technische Universität Berlin passed a statute to ensure good scientific practice [28]. Accordingly, all members of this institution see the safeguarding of good scientific practice in research and teaching as one of their top priorities. In addition, it is understood as a core task of the university to convey the basic principles of scientific work to the students. The basic principle here is honesty towards oneself and others, which is both an ethical norm and the basis of scientific professionalism. Specifically, it means working according to regulations in the respective discipline, storing primary data, documenting processes and results in a comprehensible and comprehensive manner, appropriately evaluating and involving contributions from others in publications, and respecting third-party intellectual property. The students should be introduced to the rules of good scientific practice from the beginning of their studies in order to prevent scientific misconduct in the long term, whether knowingly or unknowingly. During the first phase of the course, the emphasis is placed on respecting the intellectual property of others through the teaching of citation guidelines and rules.

III. Why should these topics be included in the engineering curriculum

The enrichment of engineering studies with "non-technical" components, such as social and cultural studies, ecological aspects and the debate with the consequences of technology is seen as a glimmer of hope to attract more female students [29]. This hope was confirmed, for example, at the Carnegie Mellon University, where a change in the curricula in the computer science department increased the proportion of women among first-year students from 7% to 42% within 5 years [30].

Not only women are attracted by a stronger integration of socially relevant topics, but also young men, who do not necessarily base their decision to start studying engineering on a mere interest in technology [29]. Since topics such as climate change and sustainability, social justice and the consequences of technological progress (AI) are increasingly becoming the focus of young people already at school, improved visibility of these topics in engineering studies promises to increase the attractiveness of engineering courses. Likewise does the clarification of the influence of an engineering profession in connection with the environment and society.

But the goal of creating forward-looking, comprehensive, sustainable and socially responsible engineering sciences also meets the needs of industry and business. Many companies owe their reputation in society to the consideration of social, ethical, and ecological factors and the encouragement of responsible innovations. Investments are made in the development of new technologies, products, services, and business models that are socially accepted and desirable. The consideration of gender and diversity aspects in product research and development is increasing [31]. Impacts, that an innovation will have, are already anticipated during the development and construction phase and considered accordingly in order to benefit society and the environment and to largely rule out bad business investments [32].

The 21st century post-modern, value-based engineer is a person who, in addition to technical knowledge, has also developed soft skills during their studies; who is aware of the responsibility of their own work, their own decisions and their effects, who is able to reflect on their decisions and their own actions from different points of view, who can perceive and evaluate risks and risk groups and who can react adequately to constantly changing social and technical conditions [33].

IV. How to include these topics in the engineering curriculum

Even if the content to be included in a program has been decided, curriculum development is a lengthy process, at least in Germany. Here we would like to briefly describe how the development started at the Technische Universität Berlin. Where we are now, and what was considered helpful.

In 2018, a mission statement for teaching was adopted at Technische Universität Berlin, which was developed during a year-long participatory process involving all

status groups. This mission statement includes educational goals and contents. For example, teaching should be modern, innovative, digital, practice-oriented and international. Students should be enabled to face technological change and its social effects with professional qualifications and a sense of responsibility, while learning to consider the ethical consequences of their actions [34]. Based on this guiding principle, the General Study and Examination Regulations (AllgStuPO) were also changed. The new version has been in force since August 2021. Here, a clear requirement for the curricula of the study programs emerges [35, §44, 3]:

[1]In the study programs, the rules of good scientific practice are taught at the earliest possible stage and continuously trained. [2]Students learn to place knowledge and action in an overarching historical, social and cultural context and to consider ethical consequences of action in order to be able to contribute to sustainable development. [3]It is to be ensured that all students have completed relevant study content amounting to at least 12 LP by the time they graduate.

The mission statement and the new AllgStuPO, constitute binding guidelines for all faculties which have to revise their courses accordingly. These obligations primarily initiate and facilitate a process of course development; the need for revision is also likely to reduce, for example, idle, fruitless discussions with professors about the indispensability and inevitability of their original content. However, there are also some potential risks.

Specifying a number of credit points to be guaranteed in each program may tempt people to offer courses that cover exactly the required topics, detached from the technical content of engineering. Then the integration of the topics might remain too superficial and conceptual. In order to build and develop competencies such as systemic thinking, critical reflection or technology assessment, the technical context dimension of the engineering modules and the competencies dimension must be integrated in a meaningful manner, e.g. room for discussion must be established and critical questions, case studies or vignettes that do not have a clear right or wrong solution, but invite the development of own position, should be foreseen.

Furthermore, the framework does not answer how the integration of topics or the required scope of the credit points can or should be ensured. This vagueness in turn leads to uncertainties. The desired integration of the topics in each module makes it difficult to prove the credit points. It remains unclear whether the collection of partial points by a percentage e.g. of sustainability and ethics in modules of construction and mechanics, is equally recognized. A central readjustment is required here, which creates clear guidelines and instructions for this.

With regard to ensuring integration into the modules, this can be demonstrated by the teaching staff via the module descriptions, on the one hand, and on the other hand, the students' perspective can also be queried with regard to the integration achieved by means of the teaching evaluation.

Regardless of the vagueness of the new AllgStuPo, the Faculty V of Traffic and Machine Systems translated its implications by first holding a closed teaching conference to inform the faculty members about the new requirements. Building on this, a Think Tank was founded which, with the participation of all status groups and fortunately a high level of student participation, is driving the process of curriculum development forward in monthly meetings.

The low-threshold discussion rounds in the Think Tank, which took place or still takes place in hybrid form both in presence and digitally, contributed in particular to identifying the needs of students in the various study programs. These will also be incorporated into further work on the study and examination regulations.

The first step conducted within the Think Tank was a curriculum mapping. It was analyzed which existing modules already contain or address topics such as digitization, sustainability, ethics and gender and diversity. Since this analysis does not create a complete picture of the courses at Faculty V, it is planned to fill the gaps through respective questions in a bi-annual teaching evaluation that takes place at the University.

The second step was dedicated to reflections about possible means to ensure the integration of these topics into the curriculum. One possibility is the establishment of a compulsory elective area, from which modules dealing with the named topics are to be selected based on a certain number of credit points. However, as mentioned above, it must be ensured that integration and reflections take place within the specialized engineering modules. To assure this, the following possible formats were developed in the Think Tank: Accentuation, empowerment, embedding, development of an integrated required foundation course at the beginning of the Bachelor's degree program [36].

Accentuation refers to subject-specific content that is reflected through a special focus on a topic. E.g. a course on *Sustainability in Space Flight*. Content on satellite operations and manned and unmanned spaceflight is reflected against the background of sustainability and both positive aspects, such as the observation of climate-related changes on our planet through satellite images, and negative aspects, such as waste in orbit, are highlighted.

Empowerment refers to enabling teachers of a module through dedicated trainings on how to enrich their own courses with content on digitization, sustainability, ethics, gender and diversity or scientific practice and to link these together.

In the case of embedding, external experts are invited to the courses to give input in a kind of mini-building block. It is then up to the module leaders to build on this input, integrated in to the on-going course and encourage the students to reflect upon it.

Work is currently in progress on an introductory compulsory course for the bachelor's degree in mechanical engineering, in which the basics of ethics, sustainability and diversity are integrated into a subject-specific introduction. There is already an Introduction to Mechanical Engineering module with a strongly practice-oriented character. This module is specifically intended

for the beginning of the degree program. It is meant to qualify the students to function as multipliers for aforementioned topics and to reduce early dropouts by enriching the quality of the studies through the integration of socially relevant topics.

In addition to these formats, certificate programs such as the Sustainability Certificate and the Berlin Ethics Certificate, which can be obtained at the Technical University of Berlin, provide students with an opportunity to explore the topics of sustainability and ethics in greater depth.

The faculty also seeks advice on curriculum development. For example, a peer-to-peer faculty advisory process is currently taking place for the mechanical engineering department under the guidance of the Higher Education Forum for Digitization (HFD). In this advisory process, the University of Bremen and the Technische Universität Berlin provide peers who support each other. In addition, they receive input and feedback from experts and the HFD on the basis of a self-report on the status of digitization at the university, the faculty, and academic degree programs.

During the HFD consultation process, facilitators from the HFD and experts from different universities who provided their expertise regarding faculty development, curriculum development and digital strategies were invited. A crucial realization was the change of perspective from a focus on subject content to competencies. The central question is: What competencies characterize our engineering graduates? The OKR (Objectives and Key Results) method [37] helped us to plan the next steps. The competence orientation confirmed the importance of integrating the above-mentioned topics into the modules, since competencies to be achieved consist of professional, methodological, social and personal competence.

The involvement of student groups and their perspective on teaching is essential for the further development of the curricula. A promising method to be implemented is the lottery method of participation, as already practiced at the Hochschule Ruhr West [38]. Here, candidates for a series of workshops were drawn at random and then personally invited by the vice president. The students involved reported that both the aspect of "winning" in the lottery and the personal approach by the president were decisive for their participation and that they probably would not have responded to a general call.

A key recommendation from the consulting process was to define achievable goals and communicate them using a step-by-step plan. Even though our original goal of the HFD consulting process was to develop a digitalization strategy, we had to rethink and define other goals and stages along the way. These are the exchange and networking of all parties involved - students, teachers, management - and the joint development of a competence profile for prospective engineers. Only then it can be determined which opportunities and methods need to be established, what framework conditions, incentive systems and support structures are required. This ultimately leads to the development of a strategy. It's important to set small goals and be bold and positive to move forward.

The next steps are the analysis and evaluation of relevant policy papers and the development of a list of requirements for prospective engineers. This will be followed by a teaching conference on the topic of graduate profiles and the development of an exemplary graduate profile. Only after this has been done and the goals have been clearly defined the curriculum process will be continued. Ultimately, we are starting a specific degree program where we are going through the entire process of redesign based on the competencies of the graduates. The extension of this redesign to all engineering programs will be implemented gradually over the next few years.

As this process of redesigning programs takes time, it must be accompanied by the sensitization and empowerment of lecturers to integrate digitization, sustainability, ethics, gender & diversity, and scientific practice into their lectures. To this end, we are planning a low-threshold, regular, subject-specific exchange between lecturers and to strengthen their networking. On the topics to be integrated, we want to implement peer-to-peer consultations, mutual shadowing, and the possibility of advice and support from experts in a timely manner. The next concrete step in this direction will be workshops at the next internal faculty conference on the topics: How can I integrate social issues into my courses? and How can I make the integration of social issues visible in the module description?

V. Conclusion

It is the self-proclaimed ideal of university education to provide students with knowledge and skills to act responsibly and sustainably. Accordingly, future engineers are to be enabled to shape their actions in awareness of their responsibility for society. To this end, the curricula must be supplemented with content and the development of competencies in the areas of ethics, sustainability, digitalization, diversity, and good scientific practice. In addition to the comprehensive qualification of engineering students for their later professions, the integration of socially relevant topics will make their studies more attractive, not least for women. On the way to developing a new curriculum, we believe that all status groups should be involved and that the needs of in particular students should be considered. Even if the integration of the topics into existing courses seems like a compromise compared to the desired fundamental paradigm shift in teaching, it is a first important step in the right direction and one that can be implemented more quickly, especially at very large universities than the complete redesign of a study program. A rapid implementation towards the integration of these topics into all academic degrees, and syllabi is of utmost importance given the rapid changes in and around our societies driven i.e. by the development of new technologies as a result of engineering activities. We argue it is time to act now, with courage, motivation, and an evidence-based willingness for necessary adaptions.

References

[1] European Union, “Charter of Fundamental Rights of the European Union”, 2012, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT

[2] H. Roberts, J. Cowls, E. Hine, F. Mazzi, A. Tsamados, M. Taddeo, L. Floridi, "Achieving a 'Good AI Society': Comparing the Aims and Progress of the EU and the US", in Sci Eng Ethics 27, 68 2021, https://doi.org/10.1007/s11948-021-00340-7

[3] Bundesverfassungsgericht, „Beschluss vom 22.10.1986 - 2 BvR 197/83", 1986, https://openjur.de/u/56233.html

[4] European University Association, "Universities without walls – A vision for 2030", 2021, https://www.eua.eu/resources/publications/957:universities-without-walls-%E2%80%93-eua%E2%80%99s-vision-for-europe%E2%80%99s-universities-in-2030.html

[5] European Commission, Directorate-General for Communications Networks, Content and Technology, "A path to the digital decade : common governance and coordinated investment for the EU's digital transformation by 2030", 2021, https://data.europa.eu/doi/10.2759/027045

[6] European Commission, "European Declaration on Digital Rights and Principles", 2022, https://digital-strategy.ec.europa.eu/en/library/european-declaration-digital-rights-and-principles

[7] E. Celeste, (2019) "Digital constitutionalism: A new systematic theorization" in International Review of Law, Computers & Technology, 2019, 33:1, 76-99, https://doi.org/10.1080/13600869.2019.1562604

[8] L. Floridi, Luciano, "The European Legislation on AI: A Brief Analysis of its Philosophical Approach", 2021, http://dx.doi.org/10.2139/ssrn.3873273

[9] M. Veale, M. Zuiderveen, F. Borgesius Frederik (2021):, "Demystifying the Draft EU Artificial Intelligence Act" in Computer Law Review International, 2021, H.22(4), p. 97-112, ssrn.com/abstract=3896852

[10] European Commission, "Proposal for a Regulation of the European Parliament and the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legsilative acts", 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

[11] European Parliament, "Opinion of the Committee on Legal Affairs for the Committee on the Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))", 2022, https://www.europarl.europa.eu/doceo/document/JURI-AD-719827_EN.pdf

[12] European Commission, Directorate-General for Communications Networks, Content and Technology, "Ethics guidelines for trustworthy AI", 2019, https://data.europa.eu/doi/10.2759/346720

[13] Gesellschaft für Informatik e.V. (2016), "Dagstuhl-Erklärung: Bildung in der digitalen vernetzten Welt", 2016, https://dagstuhl.gi.de/dagstuhl-erklaerung

[14] Vuorikari, R., Kluzer, S. and Punie, Y., "DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes", 2022, doi:10.2760/490274, JRC128415.

[15] Y. Punie, C. Redecker, "European Framework for the Digital Competence of Educators: DigCompEdu", 2017, doi: 10.2760/159770

[16] D. Hübner, " „Ethik und Moral" / „Typen ethischer Theorien" / „Aspekte von Handlungen" / „Stufen der Verbindlichkeit"" in Forschungsethik. Eine Einführung, M. Fuchs, T. Heinemann, B. Heinrichs, D. Hübner, J. Kipper, K. Rottländer, T. Runkel, T. M. Spranger, V. Vermeulen, M. Völker-Albert, Eds. Stuttgart / Weimar 2010, pp.1–39.

[17] C. Mitchal, E.E. Englehardt, "Ethics Across the Curriculum: Prospects for Broader (and Deeper) Teaching and Learning in Research and Engineering Ethics" in Sci Eng. Ethics, 2019 Dec;25(6):1735-1762. doi: 10.1007/s11948-016-9797-7, Epub 2016 Aug 22. PMID: 275498012016.

[18] D. Callahan, S. Bok, "Ethics teaching in higher education." New York: Plenum Press, 1980

[19] UNSDG, "What does the 2030 Agenda say about universal values?", 2023, https://unsdg.un.org/2030-agenda/universal-values

[20] G. Bianchi, U. Pisiotis, C. Giraldez, "GreenComp The European sustainability competence framework", 2022, doi:10.2760/13286

[21] UNESCO, "Die Salamanca Erklärung und der Aktionsrahmen zur Pädagogik für besondere Bedürfnisse", 1994, https://www.unesco.de/sites/default/files/2018-03/1994_salamanca-erklaerung.pdf

[22] A. Winteler, Profesionell lehren und lernen, 4th ed., Darmstadt, WBG, 2011

[23] U. Preuss-Lausitz, "Inklusion-Modewort oder Hoffnungsträger? Was ist neu an Inlusion und wie kann sie gelingen?," Pädagogik. Weinheim, vol. 9, pp. 41-45, 2012.

[24] H. J. Walberg, "Improving the Productivity of America's Schools," Educational Leadership, vol. 41, pp. 19-27, May 1984.

[25] J. Hattie, "Influences on Student Learning," Geoffpetty.Com. Inaugural Lecture, University of Auckland, August 2, 1999. http://geoffpetty.com/wp-content/uploads/2012/12/Influencesonstudent2C683.pdf

[26] M. Kunter, Th. Voss, "Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse," in Professionelle Kompetenzen von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV, M. Kunter, J. Baumer, U. Klusmann, S. Krauss, M. Neubrand, Eds. Münster: Waxmann, 2011, pp. 85-113.

[27] H. Spiegel, M. Walter, "Heterogenität im Mathematikunterricht der Grundschule," in Heterogenität als Chance, Vom produktiven Umgang von Gleichheit und Differenz in der Schule, K. Bräu, U. Schwerdt, Eds. Münster: LIT-Verlag, 2005, pp. 219-223.

[28] Amtliches Mitteilungsblatt. II Bekanntmachungen. Präsidium. Satzung zur Sicherung guter wissenschaftlicher Praxis an der TU Berlin. Der Präsident der TUB. 08.03.2017, https://www.static.tu.berlin/fileadmin/www/10000060/FSC/Promotion___Habilitation/Dokumente/Grundsaetze_gute_wissenschaftliche_Praxis_2017.pdf

[29] C. Wächter, "Nachhaltige Ingenieursausbildung," in C. Leicht-Scholten, Eds. Gender and Science. Perspektiven in den Natur- und Ingenieurwissenschaften. 2007, pp. 109-118, https://library.oapen.org/bitstream/handle/20.500.12657/22552/1007604.pdf?sequence=1&isAllowed=y

[30] A. Fisher and J. Margolis, "Ten Keys to Involve More Women in Academic Computing," in Gender and Science. Perspektiven in den Natur- und Ingenieurwissenschaften, C. Leicht-Scholten, Eds. 2007, pp. 109-118, https://library.oapen.org/bitstream/handle/20.500.12657/22552/1007604.pdf?sequence=1&isAllowed=y

[31] L. Schiebinger, "Gendered Innovations in Science, Health & Medicine, Engineering, and Environment." https://genderedinnovations.stanford.edu/

[32] RRI Toolkit, https://rri-tools.eu/business-and-industry

[33] A. Kamp, "Navigating the Landscape of Higher Engineering Education," 2020, https://www.4tu.nl/cee/publications/navigating-thelandscape-of-higher-engineering-education-4tu.cee-web-def.pdf

[34] Das Leitbild für die Lehre der Technischen Universität Berlin, https://www.tu.berlin/lehren/profil/leitbild-fuer-die-lehre/

[35] Amtliches Mitteilungsblatt. I. Rechts- und Verwaltungsvorschriften. Akademischer Senat. Ordnung zur Regelung des allgemeinen Studien- und Prüfungsverfahrens an der TUB. Der Präsident der TUB. 11.08.2021. https://www.static.tu.berlin/fileadmin/www/10000000/Studiengaenge/StuPOs/AllgStuPO_deu.pdf

[36] S. Ammon, "Possible formats," unpublished

[37] R. Panchadsaram, "What is an OKR? Definition and Examples", https://www.whatmatters.com/faqs/okr-meaning-definition-example

[38] I. Müller-Vogt, J. Kopper, "Partizipation durch Losverfahren," strategie digital. Magazin für Hochschulstrategien im digitalen Zeitalter, vol. 3, 2022, p.56.

Income Prospects of Engineering Graduates in Croatia

A. Tecilazić
Algebra University College, Zagreb, Croatia
ana.tecilazic@algebra.hr

***Abstract* - Research on the employability of graduates based on data collected in the Eurograduate Pilot Survey shows that the field of study is one of the three most important factors for graduates with a professional bachelor's degree in Croatia to obtain a job with a higher monthly income. The results of the analysis conducted on a sample of graduates with bachelor's degrees obtained in 2016/2017 from professional higher education institutions in Croatia show that graduates in the broad field of "Engineering, manufacturing and construction" (ISCED F) find employment in occupations with a gross monthly income that is on average EUR 676 higher than the gross monthly income of graduates with comparable qualifications in the reference educational field. Factor analysis was applied to reduce the number of measured variables indicating the same constructs of human, social or cultural capital to produce factor scores that were used as predictors of monthly income in a multiple linear regression analysis. These research findings support the core propositions of human capital theory that investing in quality and relevant education and training brings economic benefits to individuals and society.**

Keywords* - *Eurograduate, engineering education, employability, income, human capital theory

I. INTRODUCTION

The most recent developments in education and labor market policies show that human capital is becoming increasingly important for economic growth and the development of inclusive societies. Moreover, the possession of relevant skills and competences arguably increases individuals' chances of achieving better outcomes in the labor market. Building on the main principles of human capital theory, supplemented by the stipulations of cultural capital and social capital theories, as presented in Section II. Theoretical and Conceptual Framework, this paper explores the relationship between the field of study and earnings after graduation. Focusing on the relationship between the field of study and salary level, the assumption that a causal process exists between the field of study and earnings level was tested. Data on early labor market outcomes of higher education graduates collected through the first pilot European graduate tracking study - Eurograduate - were used to analyze the predictors influencing salaries in the first year after graduation. The dataset collected through Eurograduate contains responses from a total of 16,408 graduates from eight European countries. The dataset used for the analysis presented in this paper includes 464 individuals who graduated from a bachelor's degree program at professional higher education institutions in Croatia in 2016/2017. The sample is presented in more detail in Section III. Research Methods.

II. THEORETICAL AND CONCEPTUAL FRAMEWORK

The theoretical framework for the analysis presented in this paper is based on the human capital theory developed in the 1960s by economists Gary Becker and Theodore Schultz [1], [2] and complemented by the theories of cultural and social capital advanced by French sociologist Pierre Bourdieu [3], [4]. Human capital is generally and broadly defined as a person's skills and competencies. De la Fuente and Ciccone describe it as "knowledge and skills embodied in people and accumulated through schooling, training and experience that are useful in the production of goods, services and further knowledge" [5]. There is a large body of research on the private returns from investment in education and skills that demonstrates a positive relationship between investment in human capital and labor market outcomes as measured by the level of individuals' earnings [5], [6]. De la Fuente and Ciccone estimate that "one year of training increases wages by as much as 5%." [5].

However, to focus only on people's skills and knowledge in explaining labor market outcomes would ignore the broader social context that plays an important role in creating education and training opportunities for people [7], the role of social capital in creating human capital [8], and the importance of social networks in finding better paying jobs [9], [10], [11]. According to Bourdieu, cultural capital may be institutionalized in the form of parents' educational qualifications, while social capital, which is composed of family and personal social networks and connections, is defined as "the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition - or, in other words, to membership in a group" [12].

The theoretical framework as delineated by human capital theory and social and cultural capital theory formed the basis for the construction of a conceptual framework for analyzing graduate employability. In examining the historical development of the construct of employability, an operational model based on the theories of human, cultural and social capital was developed in the context of a doctoral thesis Employability of Higher

Education Graduates: Transition From Higher Education to the Labor Market [13], which was used as a conceptual framework for the research presented in this paper. Figure 1 shows an extract from the operational model, which includes the constructs of human capital, cultural capital and social capital as predictors of employability as measured by the level of monthly income. Assuming that human capital formation depends, among other things, on the field of study, this paper hypothesized that individuals with degrees in engineering, manufacturing and construction would have a higher chance of employment with higher earnings.

III. RESEARCH METHOD

The data used in this paper were collected as part of the Eurograduate - first pilot graduate tracking survey conducted in eight European countries (Austria, Croatia, Czech Republic, Germany, Greece, Lithuania, Malta, Norway) between October 2018 and February 2019 [14]. The survey included bachelor and master graduates from two cohorts, one and five years after graduation, i.e. graduates from 2012/2013 and 2016/2017. The final dataset was weighted using population data and includes responses from a total of 16408 graduates. The analysis presented in this paper is conducted on a sample of bachelor graduates from professional higher education institutions who graduated in Croatia in 2016/2017. The final dataset contains the sample of a total of 464 graduates who responded to the question about their completed study program and the questions about their earnings. The sample is distributed across nine broad fields of study according to the broad European classification ISCED-F, which includes: Education; Arts and Humanities; Social Sciences; Journalism and Information; Business, Administration and Law; Information and Communication Technologies (ICT); Engineering, Manufacturing and Construction; Agriculture, Forestry, Fisheries and Veterinary; Health and Welfare; Services. Table 1 provides a description of the sample distributed across the study areas.

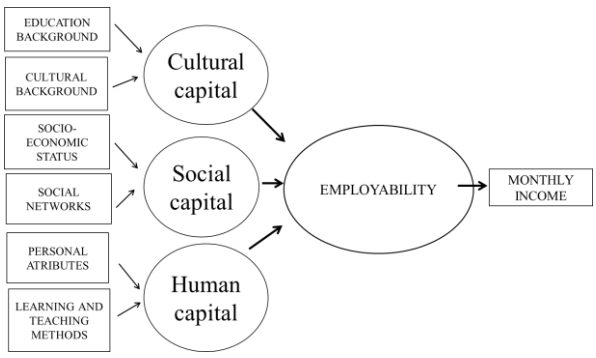


FIGURE 1. OPERATIONAL MODEL (EXTRACT)

TABLE 1. SAMPLE DISTRIBUTED WITH RESPECT TO THE FIELDS OF STUDY

ISCED-F	Frequency	Percent
Education	10	2.1
Arts and Humanities	4	0.9
Social Sciences. Journalism and Information	13	2.8
Business. Administration and Law	161	34.6
Information and Communication Technologies	40	8.7
Engineering. Manufacturing and Construction	80	17.3
Agriculture. Forestry. Fisheries and Veterinary	14	3.0
Health and Welfare	100	21.6
Services	42	9.0
Total	**464**	**100.0**

The survey instrument was a standardized online questionnaire with data centrally collected and anonymized by the Deutsche Zentrum für Hochschul und Wissenschaftsforschung (DZHW) – German Centre for Higher Education Research and Science Studies.

In analyzing the data, a descriptive and comparative statistical analysis of the sample was carried out on the one hand, and a linear regression analysis on the other hand, in order to determine the causal relationship between the field of study and the amount of gross monthly income. Descriptive statistical analysis was conducted in such a way that the respondents were divided into five categories with the following gross monthly incomes: (1) up to 1,000 EUR; (2) 1,001 to 2,000 EUR; (3) 2,001 to 3,000 EUR; (4) 3,001 to 4,000 EUR; (5) 4,001 to 5,000 EUR (6) 5,001 EUR and above. The descriptive analysis showed that 90.2% of the respondents in the whole sample receive a gross monthly income of up to 2,000 EUR (Table 2), indicating that a more sophisticated data set was needed in a linear regression analysis to test the influence of field of study on monthly income.

The comparative analysis between respondents grouped by field of study shows that all respondents with a degree in the field of education (100%) fall into the first category with the lowest monthly income, while the share of respondents with the highest monthly income (5,000 EUR and more) belongs to the group of people with a degree in engineering, manufacturing and construction (Figure 2).

TABLE 2. GROSS MONTHLY INCOME IN EUR

	Frequency	Valid Percent	Cumulative Percent
up to 1,000 EUR	269	58.0	58.0
1,001 to 2,000 EUR	149	32.2	90.2
2,001 to 3,000 EUR	31	6.8	97.0
3,001 to 4,000 EUR	8	1.8	98.8
4,001 to 5,000 EUR	0	0.0	98.8
5,001 EUR and more	6	1.2	100.0
Total	**464**	**100.0**	

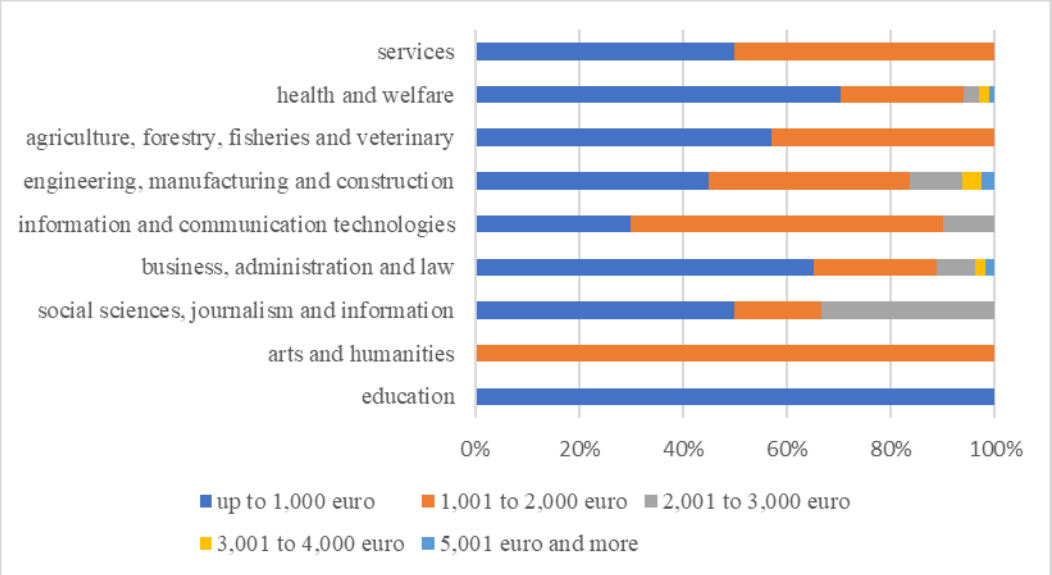


FIGURE 2. MONTHLY INCOME IN RELATION TO THE FIELD OF STUDY

Linear regression analysis was used to predict the relationship between field of study as the independent variable and gross monthly income as the dependent variable, controlling for other possible predictors identified in the operational model (Figure 1). Predictors of monthly income included measured and latent variables. Factor analysis was applied to reduce the number of measured independent variables indicating the same constructs of human, social or cultural capital to create factor scores as latent variables that were used as predictors of monthly income in the multiple linear regression analysis.

The independent variables were included in the analysis as covariates in blocks. The first block contained eight variables related to fields of study, with "Services" as the reference field of study, which was therefore omitted from the block. The second block contained factor scores related to the individual's characteristics and family cultural background, and the third block contained factor scores related to learning and teaching methods and high-impact practices (volunteering, internships, work experience, study abroad) during studies. The fourth block contained variables related to job search strategies as follows: "Finding a job through advertisements in electronic media"; " Finding a job through the Croatian Employment Service"; "Finding a job by contacting the employer on one's own initiative"; "Finding a job by being approached by an employer"; "Finding a job through family, friends and acquaintances".

IV. RESULTS

The results of linear regression (Table 3) on a sample of individuals with bachelor's degrees from professional higher education institutions in Croatia show that when controlling for other variables, the following factors predict the search for a job with higher income: the field of study, volunteering and internships during studies, while the experience of studying abroad, previous work experience and finding a job through the Croatian Employment Service have a statistically significant and negative effect on finding a job with higher monthly income.

Regression analysis showed that among graduates with a bachelor's degree obtained at professional higher education institutions, those with qualifications in the fields of "Engineering, Manufacturing and Construction" are more likely to find a job with higher income than graduates in the reference field of study "Services". After controlling for other variables, graduates with bachelor's degrees in "Engineering, Manufacturing and Construction" are employed in jobs with an average gross monthly income 676 EUR higher than graduates with comparable qualifications in the education field of "Services" ($p < 0.001$; $B = 676.195$). This result is consistent with the comparative analysis of a sample of graduates with bachelor's degrees obtained at a professional higher education institution (Figure 2).

The results of the analysis confirm the hypothesis and show that, controlling for other variables, the field of study has a significant predictive effect on obtaining higher-income employment and that individuals with degrees in engineering, manufacturing and construction have a higher probability of obtaining higher-income employment. It should be noted, however, that much of the data was anonymised or categorised into broader groups in accordance with the General Data Protection Regulation.

TABLE 3. PREDICTORS FOR FINDING EMPLOYMENT WITH HIGHER INCOME

	Unstand. Coef.		Stand. Coef.	t	Sig.	95.0% Confidence Interval for B	
	B	SE	Beta			Low.B.	Up. B.
(Constant)	892.148	185.209		4.817	0.000	527.894	1256.402
Education	-93.429	315.982	-0.016	-0.296	0.768	-714.875	528.018
Arts and Humanities	-458.387	597.933	-0.039	-0.767	0.444	-1634.351	717.577
Social Sciences, Journalism and Information	-9.294	315.040	-0.002	-0.030	0.976	-628.887	610.299
Business, Administration and Law	205.000	162.312	0.115	1.263	0.207	-114.220	524.221
Information and Communication Technologies	277.623	206.423	0.086	1.345	0.180	-128.353	683.598
Engineering, Manufacturing and Construction	676.195	175.462	0.298	3.854	0.000	331.112	1021.278
Agriculture, Forestry. Fisheries and Veterinary	133.896	270.393	0.028	0.495	0.621	-397.889	665.682
Health and Welfare	208.956	178.773	0.098	1.169	0.243	-142.640	560.552
Family cultural capital	-4.068	48.956	-0.004	-0.083	0.934	-100.350	92.213
Human capital (education outcomes)	-106.562	55.665	-0.099	-1.914	0.056	-216.038	2.915
Passive learning and teaching methods	-66.798	44.220	-0.075	-1.511	0.132	-153.767	20.171
Active learning and teaching methods	-4.353	47.978	-0.005	-0.091	0.928	-98.713	90.007
Study abroad	-165.999	68.508	-0.119	-2.423	0.016	-300.734	-31.264
Volunteering	141.891	50.105	0.147	2.832	0.005	43.348	240.434
Work experience	-103.185	43.250	-0.123	-2.386	0.018	-188.246	-18.123
Internship	136.034	56.617	0.122	2.403	0.017	24.684	247.384
Finding a job through advertisement in (online) newspaper	24.212	116.458	0.012	0.208	0.835	-204.828	253.252
Finding a job through the Croatian Employment Service	-300.282	122.675	-0.144	-2.448	0.015	-541.550	-59.015
Finding a job by contacting the employer on one's own initiative	-187.630	114.539	-0.091	-1.638	0.102	-412.896	37.636
Finding a job by being approached by an employer	170.935	129.306	0.070	1.322	0.187	-83.374	425.243
Finding a job through family, friends or acquaintances	3.421	105.193	0.002	0.033	0.974	-203.464	210.306

For example, programmes of study were grouped into broad areas of study, which limits the researcher's ability to conduct a more detailed analysis and provide evidence of comparative differences between qualifications from different education fields.

V. CONCLUSION

From the perspective of human capital theory, it pays to invest in quality and relevant higher education. Based on the theories of human capital, cultural capital and social capital, the application of the operational model for the analysis of employability to a set of data collected in the European pilot Eurograduate survey revealed a causal relationship between field of study and income. After controlling for other predictors of employability, the results of the analysis showed that an engineering education increases the chances of getting a first job.

REFERENCES

[1] G. S. Becker, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, vol. 70, no. 5, Part 2, pp. 9–49, Oct. 1962, doi: 10.1086/258724.

[2] T. W. Schultz, "Investment in Human Capital," p. 20, 1961.

[3] Bourdieu, "Cultural Reproduction and Social Reproduction," in *J. Karabel & A. H. Halsey (Eds.), Power and Ideology in Education (pp. 487-511)*, New York: Oxford University Press, 1977.

[4] P. Bourdieu and J. C. Passeron, *Reproduction in education, society, and culture*, 1990 ed. London ; Newbury Park, Calif: Sage in association with Theory, Culture & Society, Dept. of Administrative and Social Studies, Teesside Polytechnic, 1990.

[5] A. de la Fuente and A. Ciccone, *Human capital in a global and knowledge-based economy: final report*. Luxembourg: Office for Official Publications of the European Communities, 2003.

[6] M. Badescu, B. D'Hombres, E. Villalba, European Commission, Joint Research Centre, and Institute for the Protection and the Security of the Citizen, *Returns to education in European countries: evidence from the European Community Statistics on Income and Living Conditions (EU-SILC)*. Luxembourg: Publications Office, 2011. Accessed: Jan. 01, 2021. [Online]. Available: http://dx.publications.europa.eu/10.2788/31157

[7] B. Baranović, K. Doolan, I. Jugović, O. Klepač, I. Košutić, and S. Puzić, *Teorijske osnove, ciljevi i metodologija istraživanja. In: Baranović, B et al. Koji srednješkolci namjeravaju studirati? - pristup visokom obrazovanju i odabir studija*. Zagreb: Institut za društvena istraživanja, 2015.

[8] J. S. Coleman, "Social Capital in the Creation of Human Capital," *The American Journal of Sociology*, vol. Vol. 94, Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure, pp. 95–120, 1988.

[9] M. S. Granovetter, *Getting a job: a study of contacts and careers*, 2nd ed. Chicago: University of Chicago Press, 1995.

[10] N. Lin, "Social Networks and Status Attainment," *Annual Review of Sociology*, vol. 25, pp. 467–487, 1999.

[11] Y. Chen and B. Volker, "Social capital and homophily both matter for labor market outcomes – evidence from replication and extension," *Social*

Networks, vol. 45, pp. 18–31, 2016, doi: https://doi.org/10.1016/j.socnet.2015.10.003.

[12] P. Bourdieu, "The Forms of Capital u Richardson, J., Handbook of Theory and Research for the Sociology of Education;" *Westport, CT: Greenwood, pp. 241–58*, p. 15, 1986.

[13] A. Tecilazić, "Zapošljivost osoba s kvalifikacijama visokog obrazovanja: tranzicija iz visokog obrazovanja na tržište rada (Nacrt doktorske disertacije)." 2023.

[14] C. Meng, K. Wessling, K. Mühleck, and M. Unger, *Eurograduate pilot survey: design and implementation of a pilot European graduate survey.* LU: Publications Office, 2020. Accessed: May 30, 2021. [Online]. Available: https://data.europa.eu/doi/10.2766/149071

Vertikalna i horizontalna usklađenost inženjerskog obrazovanja u Hrvatskoj: rezultati istraživanja EUROGRADUATE

A. Tecilazić*, L. Mršić**
* Visoko učilište Algebra, Zagreb, Hrvatska
** Visoko učilište Algebra, Zagreb, Hrvatska
ana.tecilazic@algebra.hr, leo.mrsic@algebra.hr

***Sažetak* – Rad obuhvaća metodologiju istraživanja i izvedene rezultate o zapošljivosti osoba s kvalifikacijama visokog obrazovanja stečenima u Hrvatskoj nad podacima prikupljenim u sklopu europskog istraživanja Eurograduate. Istraživanje daje relevantne uvide o horizontalnoj i vertikalnoj usklađenosti obrazovanja inženjera sa zanimanjima u kojima se zapošljavaju nakon stjecanja kvalifikacije. Teorijski okvir istraživanja obuhvaća ekonomsku teoriju ljudskog kapitala i sociološku teoriju socijalnog i kulturnog kapitala, a konceptualni okvir proizlazi iz povijesnog razvoja konstrukta zapošljivosti. Rezultati analize napravljene na uzorku prvostupnika s kvalifikacijama stečenima na veleučilištima 2016./2017. godine, u obrazovnom području „inženjerstvo, proizvodnja, građevina" (ISCED-F), pokazuju da je omjer šansi za njihovo zapošljavanje u zanimanju koje je u skladu s razinom stečene kvalifikacije više nego četiri puta veći u odnosu na omjer šansi pojedinaca koji su kvalifikaciju preddiplomske razine na veleučilištu stekli u referentnom obrazovnom području. Također, rezultati analize pokazuju da je za istu skupinu omjer šansi za zapošljavanje u zanimanju koje je u skladu s područjem studijskog programa također više nego četiri puta veći u odnosu na omjer šansi pojedinaca referentne skupine (u konkretnom slučaju skupine „usluge"). U provedbi analize kontrolirani su ostali mogući prediktori zapošljivosti koji su faktorskom analizom reducirani na manji broj latentnih faktora indikativnih za konstrukte ljudskog, socijalnog i kulturnog kapitala.**

Ključne riječi - Eurograduate, inženjersko obrazovanje, zapošljivost, vertikalna usklađenost obrazovanja, horizontalna usklađenost obrazovanja, konkurentnost na tržištu rada

I. UVOD

Razvoj gospodarstva temeljenog na suvremenim tehnologijama i inovacijama zahtijeva kvalitetno i relevantno visoko obrazovanje temeljeno na najnovijim znanstveno-istraživačkim postignućima. Praćenjem i analizom zapošljivosti pojedinaca s kvalifikacijama stečenima kroz visoko obrazovanje moguće je utvrditi koji čimbenici doprinose zapošljavanju u zanimanju koje je u skladu s razinom i profilom stečene kvalifikacije, odnosno zapošljavanju u horizontalnoj i vertikalnoj usklađenosti obrazovanja i zanimanja.

Zapošljavanje u skladu sa stečenom kvalifikacijom mjerilo je zapošljivosti korišteno u empirijskim istraživanjima provedenim u Hrvatskoj [1][2][3]. Ono upućuje na relevantnost i odgovarajuću kompleksnost znanja i vještina stečenih kroz obrazovanje u odnosu na zanimanje koje pojedinac obavlja [4]. Ovakav kriterij zapošljivosti odgovara osnovnim postavkama teorije ljudskog kapitala prema kojoj ulaganje u stjecanje znanja i vještina doprinosi materijalnom i nematerijalnom povratu pojedincu i društvu [5][6]. Kako bi se objasnili ishodi na tržištu rada pojedinca sa završenim inženjerskim obrazovanjem primijenjen je operativni model za analizu zapošljivosti koji se temelji na teorijama ljudskog kapitala te kulturnog i socijalnog kapitala, kako je predstavljeno u poglavlju II. Teorijski i konceptualni okvir.

Nakon teorijskog i konceptualnog okvira definirana je metodologija istraživanja te ispitana hipoteza prema kojoj šire obrazovno područje studijskog programa „inženjerstvo, proizvodnja, građevina" (prema međunarodnoj klasifikaciji ISCED-F) pozitivno utječe na zapošljivost mjerenu vjerojatnošću zapošljavanja u zanimanju koje je u skladu s razinom stečene kvalifikacije te s obrazovnim područjem studija. Hipoteza je provjerena i potvrđena primjenom logističke regresije na uzorku osoba s kvalifikacijama visokog obrazovanja stečenima završetkom prijediplomskog stručnog studija 2016./2017. godine na veleučilištima u Hrvatskoj. Podatci su prikupljeni u sklopu šireg europskog istraživanja Eurograduate u razdoblju od listopada 2018. do veljače 2019. godine.

Kao podlogu za istraživanje, autori su koristili skup podataka prikupljen u sklopu istraživanja Eurograduate uključuje 16.408 sudionika ispitivanja koji su kvalifikaciju visokog obrazovanja stekli na prijediplomskoj ili na diplomskoj razini 2013./2014. ili 2016. /2017. godine u jednoj od osam zemalja sudionica ispitivanja (Austrija, Češka, Grčka, Hrvatska, Litva, Malta, Norveška, Njemačka). Karakteristike analiziranog uzorka na kojem se bazira ovaj rad opisane su u poglavlju III. Metode istraživanja.

II. TEORIJSKI I KONCEPTUALNI OKVIR

Teorijski okvir istraživanja polazi od ekonomske teorije ljudskog kapitala te je nadopunjen sociološkom teorijom kulturnog i socijalnog kapitala. Ljudski kapital najčešće se definira kao „znanja, vještine, kompetencije i obilježja pojedinca koji utječu na njegovu osobnu, društvenu i ekonomsku dobrobit" [7]. Kulturni kapital obuhvaća materijalna kulturna dobra, ali i internalizaciju

kulturnog nasljeđa u osobnosti pojedinca [8]. Socijalni kapital podrazumijeva odnose i poveznice među ljudima koji pripadaju određenim društvenim skupinama povezanima kroz zajedničke interese, vrijednosti, norme i stavove i koje olakšavaju društvenim akterima pristup različitim resursima [8][9] uključujući i pristup kvalitetnom obrazovanju, stjecanju perspektivne kvalifikacije te zapošljavanju u prestižnim zanimanjima [10][11].

Iz ovako postavljenog okvira teorije ljudskog kapitala koja naglašava važnost znanja i vještina za uspjeh na tržištu rada te sociološke teorije koja djelovanje pojedinca razmatra kroz društveno okružje, kulturno-obrazovno nasljeđe obitelji te društvene mreže koje pojedinac oblikuje kroz prostor i vrijeme, proizlazi konceptualni okvir za analizu konstrukta zapošljivosti. Konceptualni okvir primijenjen u ovoj analizi ilustriran je operativnim modelom zapošljivosti osoba s kvalifikacijama visokog obrazovanja u ranoj fazi karijere (Slika 1.) razvijenim kroz opsežnije istraživanje provedeno u okviru doktorske disertacije „Zapošljivost osoba s kvalifikacijama visokog obrazovanja: tranzicija iz visokog obrazovanja na tržište rada“ [3] i testiran na odabranom uzorku. Predstavljeni operativni model zapošljivosti kao moguće prediktore zapošljivosti obuhvaća kulturni, socijalni i ljudski kapital i njihove ključne odrednice te zapošljivost analizira kroz tri ključna pokazatelja. U ovom radu analiza je ograničena na jedan od tri pokazatelja, konkretno usklađenost razine stečene kvalifikacije i područja studijskog programa s prvim značajnijim zanimanjem koje pojedinac pronalazi u godini dana nakon diplome. Ispitana je i potvrđena hipoteza prema kojoj područje studija ima značajan učinak na omjer šansi koje pojedinci s kvalifikacijama prve razine visokog obrazovanja stečenima na veleučilištu ili visokoj školi imaju za zapošljavanje u zanimanju koje je u skladu s razinom stečene kvalifikacije i područjem studija.

III. METODE ISTRAŽIVANJA

U analizi su korišteni empirijski podatci prikupljeni u okviru pilot europskog istraživanja Eurograduate u kojem je, uz osam europskih država, sudjelovala i Hrvatska. Anketni instrument je prethodno standardizirani upitnik koji uključuje pitanja koja se odnose na demografska obilježja, akademski i radni status, iskustvo studiranja te je obuhvatio teme koje se odnose na procjenu relevantnosti stečenog obrazovanja, znanja i vještina, međunarodne mobilnosti, volonterstva te iskustva rada za vrijeme studiranja u odnosu na tržište rada te procjenu demokratskih vrijednosti i uključenosti u društvene i političke aktivnosti.

Upitnikom su obuhvaćene osobe koje su u vrijeme studija bile redoviti ili izvanredni studenti. Visoka učilišta na kojima su sudionici ispitivanja stekli kvalifikacije visokog obrazovanja uključuju privatna i javna visoka učilišta koja su podijeljena u dvije skupine, sveučilišta i ona visoka učilišta koja to nisu (ne-sveučilišta) što uključuje veleučilišta i visoke škole.

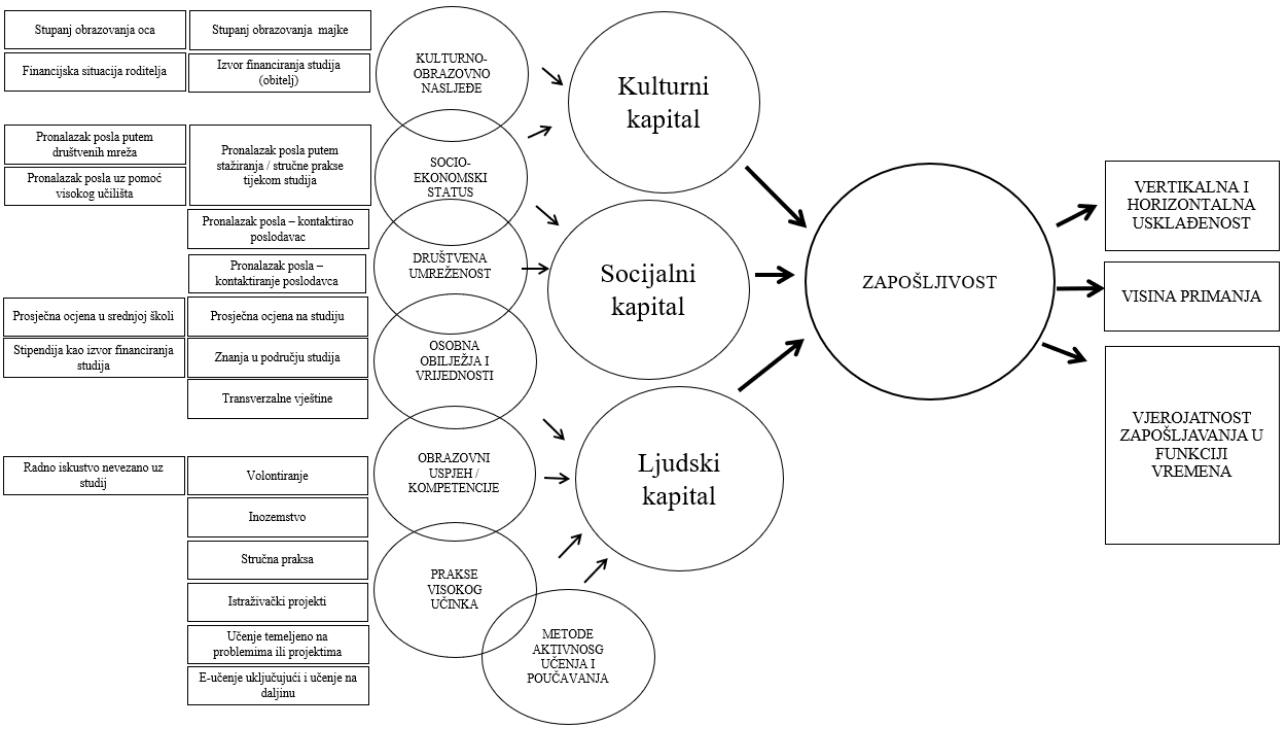
Radi provedbe istraživanja kontaktirano je svih 134 visokih učilišta u Republici Hrvatskoj, ali je tek njih 119 imalo valjane podatke za kontakt sudionika ispitivanja zbog čega su u istraživanje ušli samo osobe koje su kvalifikacije stekli na jednom od tih 119 visokih učilišta. Sudionici ispitivanja su kontaktirani preko visokog učilišta na kojem su stekli kvalifikaciju te uz potporu istraživačkog tima Studijskog centra socijalnog rada na Pravnom fakultetu Sveučilišta u Zagrebu. Upitnik se popunjavao od listopada 2018. godine do veljače 2019. godine nakon čega je slijedila obrada podataka koji su za osam zemalja sudionica istraživanja prikupljeni centralizirano.

Uzorkovanje je provedeno prema standardiziranim smjernicama koje je izradio koordinator istraživanja, Deutsche Zentrum für Hochschul und Wissenschaftsforschung (DZHW) Na prikupljenim podacima primijenjeno je utežavanje u odnosu na populacijske podatke te skup podataka obuhvaćenih ovom analizom uključuje odgovore od ukupno 930 osoba koje su u 2016./2017. godini kvalifikaciju stekle završetkom preddiplomskih stručnih studija na veleučilištima i visokim školama u Hrvatskoj (Tablica 1.). Među sudionicima ispitivanja dominanto zastupljena grupa su pojedinci u dobnoj skupini do 24 godine (54,8%) te je nešto više sudionika ženskog spola (54,0%), nego muškog. Studijski programi podijeljeni su na osam obrazovnih područja studija, klasificiranih prema međunarodnoj klasifikaciji ISCED-F te uključuju: „obrazovanje“, „društvene znanosti, novinarstvo i informacije“, „poslovanje, administracija i pravo“, „informacijske i komunikacijske tehnologije“, „inženjerstvo, proizvodnja i građevina“, „poljoprivreda, šumarstvo, ribarstvo i veterina“„ „zdravstvo i socijalna skrb“ i „usluge“. Osobe s kvalifikacijama visokog obrazovanja u obrazovnom području „inženjerstvo, proizvodnja i građevina“ čine 20% analiziranog uzorka.

Ispitivanje prediktora zapošljavanja u zanimanju odgovarajućem stečenoj kvalifikaciji provedeno je logističkom regresijskom analizom. Zavisna varijabla, zapošljavanje u zanimanju odgovarajućem stečenoj kvalifikaciji i interpretira se kroz vertikalnu i horizontalnu (ne)usklađenost koje se utvrđuju temeljem procjene sudionika ispitivanja o razini kvalifikacije i o području obrazovanja potrebnima za obavljanje posla na kojem je zaposlen.

TABLICA 1. UZORAK PREMA DEMOGRAFSKIM PODATCIMA I OBRAZOVNIM PODRUČJIMA STUDIJA

Dob	
do 24 godine	54,8%
25 – 29 godina	20,7%
30 – 34 godine	9,5%
35 – 39 godina	7,3%
40 – 49 godina	6,7%
50 i više godina	1,0%
Spol	
muškarci	46,0%
žene	54,0%
Obrazovno područje studija	
Obrazovanje	2%
Društvene znanosti, novinarstvo i informacije	3%
Poslovanje, administracija i pravo	37%
Informacijske i komunikacijske tehnologije	8%
Inženjerstvo, proizvodnja i građevina	20%
Poljoprivreda, šumarstvo, ribarstvo i veterina	4%
Zdravstvo i socijalna skrb	19%
Usluge	8%



SLIKA 1. Operativni model zapošljivost

Nezavisna varijabla koja se ispituje ovom analizom je područje studijskog programa klasificirano prema međunarodnoj klasifikaciji ISCED-F. Međutim, ograničenja koja proizlaze iz anonimizacije podataka prikupljenih kroz međunarodno istraživanje Eurograduate su takve da nije moguće utvrditi povezanost sudionika ispitivanja s konkretnim studijskim programima, odnosno kvalifikacijama koje se njima stječu.

U logističkoj regresiji kontrolirani su ostali mogući prediktori zapošljivosti osoba s kvalifikacijama visokog obrazovanja koji su prethodno faktorskom analizom reducirani na manji broj latentnih faktora indikativnih za konstrukte ljudskog, socijalnog i kulturnog kapitala, a koji su uključeni u analizu kao kovarijate u blokovima. Slijed uključivanja blokova nezavisnih varijabli određen je obzirom na vremenski slijed formiranja prediktora. U prvom bloku uključena je kontrolna varijabla „status za vrijeme studiranja“ (samo u analizi ispitivanja vertikalne usklađenosti, gdje se pokazala značajnom) uz varijablu „ISCED obrazovno područje studijskog programa“ gdje je kao referentna vrijednost određena vrijednost „usluge“. U drugom bloku uključeni su faktorski bodovi koji se odnose na obilježja pojedinca i njegova kulturnog nasljeđa, a u trećem faktorski bodovi koji se odnose na metode učenja i poučavanje te oni koji se odnose na prakse visokog učinka za vrijeme studija. U četvrtom bloku uključene su mjerene varijable koje se odnose na strategije pronalaženja posla: „putem oglasa u elektroničkim medijima“; „putem Hrvatskog zavoda za zapošljavanje“; „sam(a) kontaktirao/la poslodavca“; „obratio se poslodavac“; „preko obitelji, prijatelja i poznanika“.

IV. REZULTATI ISTRAŽIVANJA

Rezultati analiza provedenih na uzorku osoba s kvalifikacijama preddiplomske razine stečenima na veleučilištu ili visokoj školi pokazali su da na prvo značajnije zapošljavanje u zanimanju koje je u skladu s razinom stečene kvalifikacije (vertikalna usklađenost) statistički značajno utječu četiri skupine prediktora i to: status za vrijeme studiranja, područje obrazovanja, metode učenja i poučavanja, iskustvo studiranja u inozemstvu te samoinicijativno kontaktiranje poslodavca kao uspješna strategija pronalaženja zaposlenja (Tablica 2.). S druge strane, na prvo značajnije zapošljavanje u zanimanju koje je u skladu s područjem studijskog programa (horizontalna usklađenost) statistički značajno utječu tri skupine prediktora: područje obrazovanja, obrazovni uspjeh te volontiranje za vrijeme studija (Tablica 3.).

Provedena analiza potvrdila je postavljenu hipotezu te pokazuje da, uz kontroliranje varijabli koje upućuju na ljudski, socijalni i kulturni kapital, uz istaknute prakse visokog učinka i metode učenja i poučavanja za vrijeme studija, područje studija ima vrlo značajan učinak na zapošljavanje koje je vertikalno i horizontalno u skladu s kvalifikacijom. Regresijska analiza je pokazala da će posao u skladu s razinom kvalifikacije vjerojatnije naći osobe s kvalifikacijama na preddiplomskoj razini stečenima na veleučilištima i visokim školama i to u područjima obrazovanja, društvenih znanosti, novinarstva i informacija, informacijske i komunikacijske tehnologije, inženjerstva, proizvodnje i građevine te zdravstva i socijalne skrbi, u odnosu na osobe s kvalifikacijama preddiplomske razine stečenima na veleučilištu ili visokoj školi u referentnom području usluga. Vezano za zapošljavanje u horizontalnoj usklađenosti, regresijska analiza je pokazala da će posao u skladu s područjem studija vjerojatnije naći osobe s kvalifikacijama

preddiplomske razine koja se stječe na veleučilištima i visokim školama u područjima informacijske i komunikacijske tehnologije, inženjerstva, proizvodnje, građevine te zdravstva i socijalne skrbi, u odnosu na osobe s kvalifikacijama iste razine i vrste u referentnom području usluga.

Promatrajući posebno osobe s kvalifikacijama stečenima završetkom stručnih studija u području inženjerstva, proizvodnje i građevine na veleučilištima i visokim školama moguće je utvrditi da se takve osobe imaju više šansi zaposliti u zanimanjem koje je vertikalno i horizontalno u skladu s njihovom kvalifikacijom. Kontrolirajući ostale varijable, omjer šansi za zapošljavanje u zanimanju koje je u skladu s razinom

TABLICA 2. Prediktori koji utječu na zapošljavanje; horizontalna usklađenost

Prediktor	**B**	**S.E.**	**Wald**	**Df**	**Sig.**	**Exp(B)**
Status studenta (redovni)	-0,834	0,288	8,399	1	0,004	0,434
ISCED područje			27,183	8	0,001	
Obrazovanje	2,472	1,098	5,066	1	0,024	11,844
Društvene znanosti, novinarstvo i informacije	2,235	0,923	5,859	1	0,016	9,346
Poslovanje, administracija i pravo	1,173	0,434	7,316	1	0,007	3,232
Informacijske i komunikacijske tehnologije	1,782	0,551	10,463	1	0,001	5,943
Inženjerstvo, proizvodnja i građevina	1,455	0,491	8,770	1	0,003	4,283
Poljoprivreda, šumarstvo, ribarstvo i veterina	0,010	0,731	0,000	1	0,989	1,010
Zdravstvo i socijalna skrb	2,217	0,513	18,662	1	0,000	9,177
Obiteljski kulturni kapital	0,157	0,135	1,350	1	0,245	1,170
Obrazovni uspjeh (Ljudski kapital)	0,232	0,156	2,200	1	0,138	1,261
Metode pasivnog učenja i poučavanja	-0,248	0,118	4,413	1	0,036	0,780
Metode aktivnog učenja i poučavanja	-0,342	0,133	6,597	1	0,010	0,711
Studij u inozemstvu	-0,453	0,193	5,510	1	0,019	0,635
Volontiranje	0,119	0,141	0,720	1	0,396	1,127
Radno iskustvo	-0,030	0,119	0,062	1	0,804	0,971
Stručna praksa	-0,139	0,151	0,849	1	0,357	0,870
Putem oglasa u elektroničkim medijima	-0,319	0,309	1,066	1	0,302	0,727
Putem Hrvatskog zavoda za zapošljavanje	0,061	0,338	0,032	1	0,857	1,063
Samoinicijativno kontaktiranje poslodavca	0,534	0,315	2,864	1	0,091	1,705
Kontaktirao poslodavac	0,269	0,336	0,644	1	0,422	1,309
Preko obitelji, prijatelja i poznanika	0,144	0,283	0,261	1	0,609	1,155
Constant	-0,461	0,531	0,754	1	0,385	0,631

TABLICA 3. Prediktori koji utječu na zapošljavanje; vertikalna usklađenost

Prediktor	**B**	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**
ISCED područje			23,488	8	0,003	
Obrazovanje	1,709	1,257	1,847	1	0,174	5,522
Društvene znanosti, novinarstvo i informacije	-0,459	0,832	0,304	1	0,581	0,632
Poslovanje, administracija i pravo	0,647	0,429	2,268	1	0,132	1,909
Informacijske i komunikacijske tehnologije	1,213	0,606	4,007	1	0,045	3,365
Inženjerstvo, proizvodnja i građevina	1,443	0,522	7,636	1	0,006	4,232
Poljoprivreda, šumarstvo, ribarstvo i veterina	-0,962	0,731	1,732	1	0,188	0,382
Zdravstvo i socijalna skrb	1,479	0,533	7,688	1	0,006	4,390
Obiteljski kulturni kapital	0,021	0,151	0,019	1	0,891	1,021
Obrazovni uspjeh (ljudski kapital)	0,427	0,180	5,631	1	0,018	1,532
Metode pasivnog učenja i poučavanja	-0,186	0,128	2,124	1	0,145	0,830
Metode aktivnog učenja i poučavanja	0,128	0,144	0,784	1	0,376	1,136
Studij u inozemstvu	-0,203	0,201	1,025	1	0,311	0,816
Volontiranje	0,348	0,146	5,687	1	0,017	1,416
Radno iskustvo	-0,250	0,141	3,143	1	0,076	0,779
Stručna praksa	-0,202	0,170	1,423	1	0,233	0,817
Pronalazak posla putem oglasa u elektroničkim medijima	0,670	0,396	2,857	1	0,091	1,954
Pronalazak posla putem Hrvatskog zavoda za zapošljavanje	0,207	0,397	0,272	1	0,602	1,230
Pronalazak posla uz samoinicijativno kontaktiranje poslodavca	0,332	0,356	0,872	1	0,350	1,394
Pronalazak posla na način da ga je kontaktirao poslodavac	-0,355	0,368	0,930	1	0,335	0,701
Preko obitelji, prijatelja i poznanika	-0,023	0,316	0,005	1	0,941	0,977
Konstanta	0,263	0,515	0,261	1	0,609	1,301

Opis polja u tablicama: Prediktor - služi kao oznaka na temelju koje se prognozira odnosno mjeri; B - koeficijent za u nultom modelu; S.E. - predstavlja standardnu pogrešku oko koeficijenta za konstantu; Wald i Sig. - predstavljaju Waldov hi-kvadrat test koji testira nultu hipotezu da je konstanta jednaka 0. Ova se hipoteza odbacuje jer je p-vrijednost (navedena u stupcu pod nazivom "Sig.") manja od kritične p-vrijednosti od . 05 (ili .01); df - predstavlja stupnjeve slobode za Waldov hi-kvadrat test. Postoji samo jedan stupanj slobode jer u modelu postoji samo jedan prediktor, naime konstanta; Exp(B) - predstavlja potenciranje B koeficijenta, što je omjer izgleda. Ova je vrijednost dana prema zadanim postavkama jer je omjere izgleda lakše protumačiti nego koeficijent, koji je u jedinicama log-odds.

stečene kvalifikacije za pojedince koji su kvalifikaciju preddiplomske razine stekli na veleučilištu ili visokoj školi, u obrazovnom području „inženjerstvo, proizvodnja, građevina“ veći je 428% (B = 1,455; p = 0,003), u odnosu na omjer šansi pojedinaca koji su kvalifikaciju preddiplomske razine stekli na veleučilištu ili visokoj školi u obrazovanom području „usluge“ dok je njihov omjer šansi za zapošljavanje u zanimanju koje je u skladu s područjem obrazovanja veći 423% (B = 1,443; p = 0,006), u odnosu na omjer šansi pojedinaca koji su kvalifikaciju preddiplomske razine na veleučilištu ili visokoj školi stekli u obrazovanom području usluga.

V. ZAKLJUČAK

Zapošljavanje u horizontalnoj i vertikalnoj usklađenosti važan je pokazatelj strukturne usklađenosti sustava obrazovanja i tržišta rada [12] te, iz perspektive teorije ljudskog kapitala, pokazuje da učinkovito ulaganje u obrazovanje pojedincu i društvu donosi materijalni i nematerijalni povrat [5][6]. Osobe koje uspješno pronađu posao u kojem u većoj mjeri mogu primjenjivati znanja i vještine stečene kroz studij bolje će ostvariti povrat uloženog u obrazovanje od onih koji se zapošljavaju u zanimanjima u kojima su dostatne kompetencije niže razine ili znanja i vještine koje nisu nužno stekli kroz studij. Stoga se osobe koje se zapošljavaju sa višom razinom horizontalne i vertikalne usklađenosti pokazuju zapošljivijima u odnosu na one sa nižom razinom promatrane usklađenosti. Uz primjenu operativnog modela za analizu zapošljivosti razvijenim na temelju skupa podataka prikupljenih kroz europsko pilot istraživanje Eurograduate, te uz kontroliranje mogućih prediktora njihove zapošljivosti, rezultati analize su predloženim su modelom pokazali da inženjersko obrazovanje povećava šanse za pronalaženje prvog zaposlenja u skladu s razinom stečene kvalifikacije i područjem studija.

REFERENCE

[1] T. Matković, “Obrasci tranzicije iz obrazovnog sustava u svijet rada u Hrvatskoj. Doktorska disertacija. Zagreb. Pravni fakultet Sveučilišta u Zagrebu, Studijski centar socijalnog rada,” 2011. https://www.voced.edu.au/content/ngv:16446 (accessed Dec. 19, 2020).

[2] I. Rimac and J. Ogresta, Analiza mogućnosti praćenja tranzicije na tržište rada diplomiranih s veleučilišta i visokih škola kombiniranjem podataka visokih učilišta, Hrvatskog zavoda za zapošljavanje i Hrvatskog zavoda za mirovinsko osiguranje. Zagreb: Ministarstvo znanosti i obrazovanja, 2018.

[3] A. Tecilazić, “Zapošljivost osoba s kvalifikacijama visokog obrazovanja: tranzicija iz visokog obrazovanja na tržište rada (Nacrt doktorske disertacije).” 2023.

[4] L. Dacre Pool and P. Sewell, “The key to employability: developing a practical model of graduate employability,” Education + Training, vol. 49, no. 4, pp. 277–289, Jun. 2007, doi: 10.1108/00400910710754435.

[5] G. S. Becker, “Investment in Human Capital: A Theoretical Analysis,” Journal of Political Economy, vol. 70, no. 5, Part 2, pp. 9–49, Oct. 1962, doi: 10.1086/258724.

[6] T. W. Schultz, “Investment in Human Capital,”, 1961.

[7] B. Keeley, Human capital: how what you know shapes your life. Paris: OECD, 2007.

[8] P. Bourdieu, “The Forms of Capital u Richardson, J., Handbook of Theory and Research for the Sociology of Education;,” Westport, CT: Greenwood, pp. 241–58, 1986.

[9] P. Bourdieu and J. C. Passeron, Reproduction in education, society, and culture, 1990 ed. London ; Newbury Park, Calif: Sage in association with Theory, Culture & Society, Dept. of Administrative and Social Studies, Teesside Polytechnic, 1990.

[10] Antony Abraham Jack, “The Privileged Poor: How Elite Colleges are Failing disadvantaged students,”.

[11] G. C. Wolniak and M. E. Engberg, “Do ‘High-Impact’ College Experiences Affect Early Career Outcomes?,” The Review of Higher Education, vol. 42, no. 3, pp. 825–858, 2019, doi: 10.1353/rhe.2019.0021.

[12] G. C. Wolniak and E. T. Pascarella, “The effects of college major and job field congruence on job satisfaction.,” Journal of Vocational Behavior, vol. 67, no. 2, pp. 233–251, 2005, doi: 10.1016/j.jvb.2004.08.010.

SSE

International Conference on

SOFTWARE AND SYSTEMS ENGINEERING

Steering Committee

Chairs: Tihana Galinac Grbac, Juraj Dobrila University of Pula, Croatia
Darko Huljenić, Ericsson Nikola Tesla, Zagreb, Croatia

Program Committee

Members: Stipo Čelar, University of Split, Croatia
Andrej Grgurić, Ericsson Nikola Tesla, Zagreb, Croatia
Igor Ljubi, Ministry of Public Administration, Zagreb, Croatia
Mladen Sokele, Zagreb University of Applied Science, Croatia
Nikola Tanković, Juraj Dobrila University of Pula, Croatia
Mladen Tomić, University of Rijeka, Croatia
Linda Vicković, University of Split, Croatia

PAPERS

Role, Importance and Significance of Software Quality

F. Témolé*, D. Atanasova**
* Capgemini Hamburg, Germany
** University of Ruse „Angel Kanchev“, Department of IIT, Ruse, Bulgaria
ftemole@t-online.de, datanasova@uni-ruse.bg

***Abstract* - In software development processes, as in other forms within the mechanical environment, the objectives of least costs, adherence to due dates and item quality must be taken into consideration. Great software system must satisfy different already concurred quality perspectives or product characteristics. Since of the gigantic significance of quality within the advancement of computer program items, its administration ought to be coordinates into software ventures. Software quality administration ought to make and keep up program quality with the assistance of specialized and hierarchical measures. As it were when the above-mentioned goals have been satisfied can one talk of financial computer program advancement. Software quality management can ensure that the software has the required level of quality. Quality objectives should be set, and their degree of fulfilment can be determined during development. In this paper, the essential findings from the expert interviews on the role and importance as well as the significance of software quality are brought together and a reflection is made between statements from business practice (experts) and science (theory), whereby the state of both research and practice can be decisively improved.**

Keywords - component; formatting; style; styling; insert (key words)

I. INTRODUCTION

First of all, the term quality must be defined in order to subsequently be able to determine the term software quality. According to DIN EN ISO 9000, quality is understood to be a degree " to which a set of inherent characteristics of an object fulfils requirements" [3, 4]. Quality is thus reflected in how the product fulfils the existing requirements. There are numerous definitions in the literature that look at the term quality from different perspectives (e.g. objective or marketing-oriented view). Garvin [6] examined product quality, listed characteristic features and described five views (transcendent view, product-based view, user view, manufacturer view, value-based view).

TABLE I. FIVE VIEWS OF QUALITY, ACCORDING TO GARVIN [6]

View	Description
Transcendental view	No precise quality definition possible, but is absolutely and independently recognizable from other quality characteristics as a whole. Only through the highest demands, standards and performance can this quality be achieved.
Product-based view	Quality can be derived from the product characteristics and is thus measurable via the properties or ingredients.
User view	Quality is characterised by the fact that individual customer needs are met and the customer is supplied with the most suitable product for his specific needs.
Manufacturer's view	Quality is achieved when given specifications or legal requirements are met.
Value-based view	Quality is characterised by the fact that the services are provided at a cost that is acceptable to the customer (price-performance ratio of the product).

DIN EN ISO 9001 defines the requirements for a company's quality management and is designed to be industry-neutral. According to DIN EN ISO 9001 [4] a quality management concept must meet the following mini-mum requirements:

- Project-specific quality model with the relevant quality characteristics
- Determination of the parameters with the corresponding procedure
- Weighing the benefits and costs of the quality management system

Due to the special features of software products, a quality management adapted for software should be applied, which is to be regarded as part of the quality management of the company. Quality management can be carried out in a process- or product-oriented manner. According to Liggesmeyer [16] there are analytical and constructive methods that allow the development of comprehensible, reliable and easily modifiable software products and enable the early detection of errors. Software quality management can ensure that the software has the required level of quality. On the one hand, the quality of the software must be assured with appropriate measures (e.g. inspection of the requirements document), and on the other hand, the development process must be checked to ensure that defined quality characteristics are adhered to (e.g. regulations on the performance of tests). A distinction can be made between internal and external software quality.

II. DIMENSIONS AND MEASURING QUALITY

A. Dimensions of quality

Garvin [6] identified eight dimensions for analysing quality (Figure 1): The Performance dimension (1) includes the scope of performance, the primary operating characteristics of the product. The features dimension (2)

includes special or supplementary performance characteristics that go beyond the normal level. Reliability (3) reflects the likelihood that the product will behave consistently throughout its lifetime. The conformity dimension (4) indicates the extent to which the design, the expression of features corresponds to standards or previously defined specifications. The dimension durability (5) distinguishes between technical and economic durability. Technical durability is about durability until the product deteriorates to the point where repair is no longer possible. Economic durability refers to repair costs, as well as losses due to downtime. This means that the benefit of the product must be calculated up to the point where repair is no longer economical. The ease of maintenance dimension (6) refers to the speed, courtesy and competence of maintenance. The aesthetics dimension (7) assesses the external appearance and impression. Perceived quality (8) covers the way and feeling of the consumer in handling the product [6].



Figure 1. Eight dimensions of quality according to [6]

These dimensions of quality cannot simply be transferred to software products. Since there is no absolute quality, the determination of software quality is complex. Software quality can be circumscribed by numerous different properties. Quality models have therefore been developed for software products that contain quality criteria in the various dimensions. The defined quality characteristics can be used as a checklist to comprehensively ensure the quality requirements and thus provide a basis for estimating the resulting efforts and activities required during system development [9]

B. Internal and external software quality

If the focus is on how well a system is built, then the internal software quality is meant, i.e. a consideration is made from the developer's point of view. Here it must be pointed out very early on that the software architecture also has a quality requirement, which is often omitted. This internal quality is not visible to the user. Important characteristics of internal software quality are, for example, readability, comprehensibility, extensibility, changeability, maintainability and reusability, all of which refer exclusively to the source code [19]. Internal quality ensures that the software can be better maintained in the long term and thus minimises the technical risks and costs as much as possible [17]. Therefore, it primarily addresses the technical debt and is an investment in the future for the respective system.

If the focus is placed on how the system behaves with regard to the specified expectations, then we are talking about external software quality, i.e. the focus is on the customer or user perspective. In this case, the quality is focused on the needs and requirements of the end user [15]. The technical correctness, reliability, ease of use, usability and performance of the system are important characteristics of external software quality. The higher the quality of the products we deliver, the fewer reworking afterwards. The big challenge is to identify the different needs of each user group and to meet their respective expectations [1].

C. Product and process quality

Another view of software quality, as already explained, is product quality, which corresponds to the degree of quality of the product. This is often determined by how error-prone the system in question is. Its determination represents both a definable and measurable quantity and thus enables an objective assessment of quality. It is extremely important that we deliver functioning software. Software quality consists of determining the efficient and effective processes to ensure software development. A survey conducted by Javed and colleagues [12] showed that management plays a major role in quality assurance. It is the primary responsibility of team managers to support team members (e.g. through training) and provide them with a good working environment. Within the framework of quality management, measures of quality measurement, quality improvement and configuration management must also be provided [12]. Quality management in software development can be distinguished into

- Quality management focused on the software product and
- Quality management focused on software development.

Product quality is characterised, for example, by reusability, portability, reliability, manageability, flexibility and usability and depends on rigorous and systematic standards and procedures [18]. These quality characteristics determine the objective measurability of the quality of a software product [2]. Internal software quality, a characteristic of structural properties of the software product, includes the source code, the architecture and design of the software and the documentation (e.g. conception, requirements, API specification) [7]. Thus, only the requirements for the software product are considered, which can be used as criteria for product evaluation and which serve as guidelines for software development. Product-based quality assurance checks whether and how the developed software product fulfils these requirements. The disadvantage of this form of quality assurance is that errors are often discovered after the software product or

an intermediate product has been realised. Product quality can be achieved through testing and static quality assurance [19].
In order to improve software quality, the efficient and effective processes for software development need to be determined. In doing so, companies can pursue different strategies for software development [7]. Process-oriented quality assurance places requirements on the software development process that serve as guidelines for the design as well as the evaluation of the IT project. Through process-oriented quality assurance, software development is designed in such a way that as few errors as possible are committed. Various measures can be used to standardise and continuously improve software development in companies. Measuring the efficiency of the process, for example, is one of the most important items for process quality [11].

For the research design used, a sample size of 58 participants was chosen due to the relevance of different roles and stakeholders along the value chain, as well as the complexity of the software itself and the processes involved. . They have been interviewed within 4 months. The interviews were tape-recorded and subsequently transcribed. The expert interviews were scheduled for a minimum 90 minutes each. The researcher was not only an observer, but was also involved in the interview process as a moderator. Above all, their different views on software quality, the different hierarchy levels and their good and long-standing professional experience in the respective field of activity were taken into account. The focus here was on investigating several groups per subprocess at different points in time.

Some experts see an overall exposure to ever-increasing production and speed pressures. New requirements are not only being brought to developers more quickly, but are also becoming increasingly imprecise and are only clarified during implementation. This makes development and testing more difficult. Therefore, the work processes would have to be changed in order to keep up with the rapid changes, but then the entire quality assurance does not fit. The testing process and the quality process would have to be changed accordingly.
In the long run, the long development cycles will no longer be possible. Product launches must not take years as before, but must be faster. Frameworks, software, etc. need to be reviewed to enable faster feedback in the testing process. Test strategies (e.g. module testing instead of integration or regression testing) must also be reviewed.
It is therefore necessary that in the future there is no longer just reaction, but action and a stronger focus on automated and customer-oriented processes or products as added value for the customer. Accordingly, a central challenge will again be software quality, which will be characterised by even more frequent releases in the future. It is extremely important to have a quality model that takes into account the industry-specific characteristics and quality requirements. However, it should be adapted or reconsidered when the way of working in the company changes. It must be checked whether the quality model with its rigid stages can cope with the changes at all. Quality assurance must be adapted to the new structures.
Product-oriented and process-oriented quality assurance must not be considered independently of each other, as both are about software quality, but this is viewed from different perspectives. Both perspectives must therefore be taken into account.

D. Importance of software quality

In companies, software quality is often only seen as a cost factor and the importance of software quality is not recognised. A good software must fulfil different, previously agreed quality aspects or product characteristics [21]. There is a high degree of agreement between literature and empirical studies on the importance of software quality per se. During the interviews it became clear that the importance of software quality is considered very high by most of the interviewees. Functioning software must be produced that is also used. The importance is rated or regarded as high to very high in the literature and by the experts. Quality is an important factor in the software industry [12]. However, some of the experts interviewed pointed out that software quality is overrated.
The assertion that the importance of software quality is not recognised was not confirmed either by the interviews conducted or by the literature. Only individual experts see software quality as overrated. However, it was found during the interviews that the significance of software quality is partly not known or not tangible. Many of the experts only considered product quality and hardly addressed process quality. Everyone understands software quality differently, so everyone has a different idea of how software quality is created. For example, the experts made the following statements about software quality:

- Good software quality is characterised by the fact that as few technical errors as possible occur.
- Software quality not only plays a role in software development, but is also related to documentation, test processes and test management.
- The software product manufacturer's ultimate goal is to deliver valuable and high-quality software to the customer that contains functionalities that the customer needs in a stable quality and has a high and stable performance. However, in doing so, there may be trade-offs in the functional scope.
- Software quality defines what the company delivers. Especially when the software is used productively, this also has a direct impact for the customer.

It became apparent that there is still uncertainty as to what constitutes software quality. The experts interviewed often only understand software quality as the

external software quality that meets the needs of the customer. Software quality is thus predominantly viewed from the customer's and user's point of view. According to the experts, one focus is on product quality during delivery and another, in terms of the quality dimensions, on performance, which, however, is not primarily in the foreground in the coordination process.

As elaborated from the literature, good software must also have internal software quality. The internal software quality can be captured by properties such as readability or changeability of the source code. This ensures better maintainability, for example. However, some software developers prefer to live out their freedom and individuality and do not recognise the necessity of rules and specifications for software development and thus for software quality. But quality also includes making decisions and not just doing what is required, but also questioning things from time to time.

Software quality should not be seen as a hindrance, but rather as an assistance. There is still a need for improvement in the company with regard to software quality. Often the company does not live up to it. The software developers must sometimes make their own decisions in order to improve the software quality. Not all software developers have the appropriate quality awareness. This has also been addressed in the literature. Javed and colleagues [12] pointed out that the attitude of the developer plays a major role in software quality. According to Distanont and colleagues [5] they are the main responsible for the quality of the software. However, developers often do not show cooperative behaviour or interest in the problem areas pointed out by quality management [5]. Therefore, the company's developers must be given a corresponding or an appropriate attitude towards software quality.

E. Causes of poor software quality

In the literature and in the interviews conducted, several causes for a lack of software quality emerged. The most important are:

- High complexity in software development
- Rapid changes and adjustments
- Lack of knowledge and overview
- Scope of the desired functionalities
- Problems with requirements elicitation

The complexity of software development has increased considerably in recent years. Ever greater demands are being placed on software development. In addition, constant changes and adaptations (e.g. to customer requirements or legal regulations) are necessary, to which the software product manufacturer must react effectively and efficiently. This is also confirmed by the experts and the literature. Unrealistic deadlines and thus tight schedules are often agreed upon, which can be a major problem factor for software quality [12]. Quality is sometimes not sufficiently considered due to speed as well as efficiency.

Poor software quality sometimes also results from a lack of knowledge. There are fewer and fewer people who have an overview of the entire project. By breaking up into individual teams, knowledge is also decentralised. The global overview suffers due to too much focus on issues of the subject departments. As a result, errors are currently arriving in production that are caused by a lack of coordination between the developers. There are too many individual teams dealing with software quality. The interaction of individual sub-processes (e.g. requirements elicitation, software development and testing) is unbalanced and not very transparent. The global overview and the holistic view are missing. Literature and science agree on this. This can lead to product delays, rework and thus higher costs. Customer and employee satisfaction can decline. A large test coverage cannot close this gap; here too, there is agreement between literature and empirics. The author's statement was confirmed that software product manufacturers must implement the business and technically highly complex requirements in high quality in short development and change cycles while taking economic aspects into account.

Software quality must not only be seen from the perspective of the technicians, but must also take into account the perspective of the specialist/business departments. This is also the view of the departments that have to sell products or manage contracts. Functionalities must be made available to these departments efficiently and quickly in order to be able to hold one's own in the face of increased competition. Quality is therefore not an end in itself, but it must also not excessively delay or block development.

The software development of the software product manufacturer emphasises meeting the needs of different user groups. This was also evident in the responses of the different users of the software products. The software developers have to deliver a functional software, which means that functionality has a slightly higher priority. This is also the view of another interviewee. The goal in the development team is to create production-ready software more quickly, which means that quality takes a back seat. Other interviewees pointed out that it is better to make concessions in certain functional features, but that performance, stability, maintainability and security should be improved instead. Ergonomics, usability, but also maintenance and modularisation play a major role in the software development process alongside software quality.

As has already been worked out on the basis of the literature, errors in requirements elicitation cause enormous effort, longer development times and subsequent economic costs. Therefore, requirements engineering is absolutely necessary. Although this is carried out in practice, according to the experts it is not consistent and comprehensive enough, which leads to problems with the delivery of the software. The topic of software quality should already play a role in the requirements process and not only be examined afterwards during testing. Software quality therefore already begins with requirements elicitation in the business departments. However, most of the time the departments are not able to formulate the requirements

precisely, or they do not know the potential of the software products. This prevents correct technical implementation. Therefore, special attention should be paid to requirements elicitation. By clearly formulating the requirements, error chains or errors in software systems can be avoided [8]. This is also the view of the experts interviewed. However, there is usually not enough time for this due to time pressure. Only one expert stated that the current software quality in his division is very good, so that other challenges (e.g. timely elicitation of requirements) are currently in focus.

The multitude of quality characteristics allow the existing expectations and needs of different experts along the software development to be taken into account. When viewed from the perspective of the developer and the software architect, for example, the focus is on how well a system is built technologically. This aspect of quality is not visible to the user. Important characteristics of this view include readability, comprehensibility, extensibility, changeability, maintainability and reusability, all of which relate exclusively to the source code. It also serves to make it easier to maintain the software in the long term and thus to minimise the technical risks as much as possible.

F. Value of software quality

Software quality also pays off financially for companies, especially for highly complex, maintenance-intensive systems [22]. The experts were therefore asked how the value of software quality is assessed in the company. Some experts said: With software quality that is taken into account right from the start, fewer corrections have to be made afterwards, so costs can be saved. The higher the software quality, the fewer problems and errors occur and the less rework the software development and IT staff have to do.

However, the economic benefits are usually not clear to the interviewees. The experts confirmed the author's statement that savings are often made in software quality assurance measures due to permanent cost pressure. The economic perspective of software quality must be communicated more intensively to those involved. Through the interviews and observations conducted, it became apparent that mechanisms for error prevention have not been established or sufficiently taken into account in all areas of the company. This leads to problems in the timely detection of errors.

G. Importance in the company

Based on the literature, it was worked out that software quality must be integrated as an integral part of the corporate strategy. The experts were therefore asked about the importance of software quality in the company. One interviewee stated that efficiency and speed are more important in his division than ensuring quality. Market entry and thus presence are paramount. In another division, software quality also plays a rather subordinate role from the company's point of view. It is often only considered as a partial aspect in retrospect, so that quality assurance does not take place from start to finish, but only when certain incidents have occurred. However, many of the experts interviewed saw a causal connection between corporate success and high product-based software quality. As a result, expensive rework could be reduced. One interviewee stated that it is part of the company strategy to deliver high-quality software.

The image of quality assurance has changed in recent years. Overall, the company has seen a significant improvement in the quality of software. However, software quality and testing are often neglected due to time pressure or in order to save costs. In practice, software quality often has to be weighed against quality. A better error culture must be lived in the company. In the literature, company policy is also identified as an important aspect of poor software quality [12]. Within the company, specific efforts must be made to ensure that all those involved in software development recognise the importance of software quality within the company. The literature has shown that there is a causal relationship between the pursuit of high product-based software quality and the company's success. However, this is not yet practised by all divisions of the software product manufacturer.

Sufficiently high software quality allows systems to be developed more economically. Only high-quality software will be used in the long term and ensures good maintainability. Existing systems could be further developed and new systems introduced with less effort, mainly in terms of personnel, if more attention were paid to the issue of quality.

III. CONCLUSION

As the interviews showed, the experts in the company tend to have a product-based view of quality, which is expressed purely through the design of the software itself. It is also clear from the expert interviews that a view-based model - such as the quality model according to IOS/IEC 25010 - is better suited for a holistic control of quality, since here the needs of the various user groups are the focus of consideration.

With regard to software quality, therefore, it is not the favoured user view that is in the foreground in the company (despite the agreement between theory and practice) [13], but the manufacturer's view in combination with the aforementioned product-based view. Thus, in the company, quality is considered to have been achieved when the product fulfils the functionalities and thus the given specifications or design requirements have been met [15]. All in all, this is to be seen rather negatively with regard to user orientation, improvement of software quality and securing the economic future of the company. Thus, in comparison with the literature, it can be formulated that the role, importance and significance of software quality [14, 20] in terms of its effect on the success of the company is not sufficiently taken into account.

The economic advantages of software quality are usually not clear to the interviewees or, due to cost and time pressure, savings are often made in software quality assurance measures. The value of software quality is partly recognised, but not consistently implemented in the

company. Software quality is often only seen as a cost factor and savings are made at this point.

The most important approaches to reviewing software quality are improvement, delivery time, overheads, the inclusion of customer requirements and how quickly the customer's requirements are implemented. Software quality management provides a framework for software development by defining and measuring software quality in the first place. The success factor software quality is largely dependent on how well the respective company succeeds in integrating and consistently using it as a fixed component of software development.

REFERENCES

[1] ASQF. (2016). Quality from the start: 20 years in the service of quality. (ASQF, ed.) Potsdam.

[2] Broy, M., & Kuhrmann, M. (2021). *Introduction to software engineering.* Berlin: Springer.

[3] DIN EN ISO 9000. (2015). Quality management systems - Fundamentals and terms. Berlin.

[4] DIN EN ISO 9001. (2015). DIN EN ISO 9001 Quality management systems - Requirements.

[5] Distanont, A. (2015). The Test Automation Adoption-A Case Study. In: Numprasertchai, H. & Meeampol, S. (Eds.): International Journal of Business Development and Research. Volume 2, Issue 1, pp. 60-78.

[6] Garvin, D. A. (1984). What Does "Product Quality" Really Mean? *MIT Sloan Management Review, 26*(1), pp. 25-43.

[7] Hassan, M., Mubashi, M., Shabi, M., & Ullah, M. (2018). Software quality assurance techniques: A review. *In: International Journal of Information, Business and Management, Vol. 10, No.4, 2018*, pp. 214-221.

[8] Hoffmann, D. (2008). *Software Quality.* Berlin Heidelberg: Springer.

[9] ISO/IEC 25010. (2011). "Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models". ISO/IEC. Retrieved 18 March 2017 from iso.org.

[10] ISO/IEC 9126-1. (2001). Information technology-Software product quality. Geneva.

[11] Jamil, M., Arif, M., Abubakar, N., & Ahmad, A. (2016). Software Testing Techniques: A Literature Review. *In: 2016 6th International Conference on Information and Communication Technology for The Muslim World*, pp. 177-182.

[12] Javed, A., Maqsoo, M., Quazi, K., & Shah, K. (2012). How to improve software quality assurance in devoloping countries. *In: Advanced Computing: An International Journal (ACIJ), Vol.3, No.2, March 2012*, pp. 17 - 28.

[13] Kitchenham, B., & Pfleeger, S. L.. (1996). Software quality. *IEEE Software, 13*(1), pp. 12-21.

[14] Kreilkamp, E. (1987). Strategic management and marketing: market and competitive analysis. Berlin: de Gruyter.

[15] Lange, C. (2011). *Software quality models.* Munich: TUM.

[16] Liggesmeyer, P. (2009). *Software Quality.* Heidelberg: Spektrum Akademischer Verlag.

[17] Petrasch, R. (2001). Introduction to Software Quality Management. Berlin: Logos Verlag.

[18] Rashid, J., & Nisar, W. (2016). How to Improve a Software Quality Assurance In Software Development - A Survey. *In: International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 8, August 2016*, pp. 99-108.

[19] Spillner, A., Roßner, T., Winter, M., & Linz, T. (2014). *Praxiswissen Softwaretest: Testmanagement* (4th ed.). Heidelberg: dpunkt.Verlag.

[20] Venelinova N., D. Antonova and I. Kostadinova (2021). Adaptive Approach of System-engineering Project Management Skills Acquisition MIPRO 2021 - The 44th International Convention on Information, Communication and Electronic Technology, Proceedings MIPRO2021.Engineering Education, , 1778-1784, doi:

[21] Wallmüller, E. (2001). "Software Quality Management in Practice: Software Quality through Leadership and Improvement of Software Processes" (Vol. 2). Munich Vienna: Carl Hanser Verlag.

[22] Zimmermann, T. (05 March 2018). *Software quality*. From testing costs - not testing too!: https://www.business-wissen.de/artikel/software-qualitaet-testen-kostet-nicht-zu-testen-auch/ retrieved

Identification of Code Properties that Support Code Smell Analysis

S. Prokić, N. Luburić, J. Slivka and A. Kovačević
Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
{simona.prokic, nikola.luburic, slivkaje, kocha78}@uns.ac.rs

***Abstract* – Code smells are structures in code that imply potential maintainability problems and may negatively impact software quality. One of the critical challenges with code smells is that their definitions are often vague, difficult to comprehend and subjective, making them hard to reliably and consistently detect and analyze by humans and automated systems. Most existing code smell detection approaches rely heavily on human interpretation and are typically supported by structural code metrics. Unfortunately, many of these approaches are incomplete and do not cover a range of code properties that could indicate potential code smells.**

This paper analyzes code smell detection approaches to identify code properties used for code smell detection and analysis. Informed by our previous work and the literature, we define five code properties used by humans and automatic detectors to identify code smells. We demonstrate how various code properties can be mapped to the 22 code smells defined by Martin Fowler. The resulting catalog of properties can help software engineers and code maintainability researchers analyze code smells and build automated code smell detectors that examine properties beyond the traditional structural metrics.

Keywords – clean code; code smell; maintainability; code quality

I. INTRODUCTION

Code smells are structures in code that indicate maintainability issues in software design or implementation [1]. These structures may negatively impact various attributes of software quality, such as maintainability and evolvability [2]. The presence of code smells in software leads to increased maintenance costs due to the inability to understand existing code, adapt and create new features to address new requirements, and fix bugs [2][3].

Software engineering experts agree that removing harmful code smells is important for sustainable software development and high-quality code [2][4]. Identifying code smells is a challenging task for both humans and automated detection systems. Previous research has shown a significant disagreement among developers, as the definitions of code smell can vary depending on the context, developer experience, and intuition [5]. Code smell definitions tend to be ambiguous and hard to grasp, as they vary depending on the programming language, framework, and coding conventions used. It is worth noting that removing code smells can be time-consuming [6] and requires a deep understanding of the codebase and the system's architecture.

Most existing code smell detection approaches are solely based on structural metrics extracted from source code [7]. Structural metrics alone are insufficient for identifying all the smells specified by Fowler [1], as they cannot capture other aspects of the system that may indicate the presence of code smells. Historical information can offer a deeper understanding of the code's evolution and help identify patterns that may indicate the presence of certain code smells (e.g., *Shotgun Surgery*) [7]. Code property such as AST (Abstract Syntax Tree) adds more context about the system's structure and organization, enabling the identification of code smells that might not be immediately obvious from the source code [8]. The presence of specific code smells (e.g., *Feature Envy*) also needs to be evaluated by analyzing the relationship between the system's classes [9][10].

This paper investigates code smell detection approaches to identify code properties used for code smell detection and analysis. Based on our previous work and the literature review, we define five code properties used by humans and automatic detectors to identify code smells: text, AST, structural metrics, change history, and relationships. We demonstrate how different properties can be mapped to the 22 code smells defined by Martin Fowler [1], highlighting which properties are useful for each code smell analysis[1].

The resulting catalog of code properties can help software engineers and code maintainability researchers analyze code smells and build automated code smell detectors that examine properties beyond the traditional structural metrics. Software experts can use the catalog in the code smell annotation process to gain guidance on the crucial properties needed for each code smell analysis.

Section 2 presents related work for previously identified code properties. Section 3 outlines our research methodology. We offer our results in Section 4 and discuss results and limitations in Section 5. In Section 6, we make our conclusions.

This research was supported by the Science Fund of the Republic of Serbia, Grant No 6521051, AI-Clean CaDET.

[1] We use the term **analysis** to bring together activities such as detection, scoping (in terms of what code is affected) and determining severity of code smell.

II. RELATED WORK

Several earlier studies have addressed identifying various code properties for code smell analysis. To the best of our knowledge, none of these studies were concerned with mapping properties to Fowler's code smells (i.e., determining the usefulness of specific properties for smell analysis).

Walter et al. [15] identified six data sources considered useful for code smell analysis: programmer's intuition and experience, metric values, AST, history of changes made in code, dynamic behavior of code, and the existence of other smells. The authors briefly described each of these data sources but did not address the mapping of these sources to code smells. Instead of determining the usefulness of specific properties for certain smells, their study was focused on Large Class detection based on some of the listed sources.

Several recent review papers have listed data sources and approaches for code smell analysis. In [6], we found a list of metrics (primarily structural) used for the machine learning detection approaches for code smells. Authors of [6] focused exclusively on the machine learning approaches. Our goal was to identify properties useful for analyzing smells, regardless of whether performed manually or automatically.

Haque et al. [11] have presented manual and various automatic approaches used for code smell analysis (e.g., metrics-based, rule-based, history-based). Still, their aim was not to map the data used in these approaches to specific smells. Furthermore, machine learning and hybrid approaches for code smell analysis are listed in [12]. Although the mentioned approaches provide insight into the data used for code smell analysis (e.g., software metrics, change history), there is no mapping between specific code smells and data.

Authors of [13] present various tools used for code smell analysis and techniques they are based on (e.g., metrics-based, history-based). Additionally, the mapping of code smells and source code metrics used for their analysis in [13] can be useful for our paper.

We have previously conducted a systematic literature review (SLR) of the most recent studies to survey the existing machine learning approaches for smell detection [14]. Through the SLR, we have identified several data sources used in these approaches. These include structural metrics, historical features, AST, and natural language features. The data sources used for each code smell were valuable for mapping the code properties and smells in this study.

III. METHODOLOGY

Our study aims to determine code properties from the existing literature, regardless of whether humans or automated systems used them for code smell analysis. The goal is not to gather properties that produce the best smell detection results but to collect the properties frequently used, either individually or in combination. Following the identification of properties, we map these properties to Fowler's code smells [1], highlighting their use for analyzing each specific smell.

To accomplish this goal, we conducted a procedure that started with reading recently published survey papers on code smell analysis [6][11][12][13] and using data from our SLR [14]. In this step, we aimed to identify properties frequently used to analyze smells. While reviewing survey papers, we discovered a study [15] that was the foundation for creating an initial catalog of properties. They identified six distinct data sources useful for smell analysis. We made changes to the initial catalog taking into consideration the properties listed in other survey papers and our SLR. We describe these changes in the results section.

The result of the first step of our procedure was a catalog of five code properties, which we then mapped to 22 code smells specified by Fowler [1]. We used information from survey papers and the studies they cited to conduct the mapping. Using the existing literature as a reference, we identified specific properties that were examined when analyzing each smell.

Researchers do not give equal attention to each smell as they tend to focus on prevalent and impactful smells [16]. The literature we examined until this phase was insufficient to map certain smells and properties. We managed to map some of these smells to some of the properties in the previous step (most often *text* and *structural metrics*), but we needed more information for the remaining properties. We searched for additional studies that focused on analyzing these smells.

Despite extensive literature search, we were unable to find information for some of the less prioritized code smells and properties. Using the expert opinion research method [17], we completed the mapping of the properties to code smells rarely discussed in the literature. Relying on the expertise of two researchers who have been analyzing code smells for several years [18][19][20], we made predictions about whether properties could be useful for analyzing a specific smell.

The experts considered code smell definitions and property descriptions to make predictions. Each expert prediction was accompanied by an explanation of their reasoning, providing an opinion on the potential usefulness of the properties for code smell analysis. We provide a comprehensive description of these predictions and reasoning in the results of our research. Disagreements between experts were resolved in the discussion, and if there was no consensus, the opinion of a more experienced expert was considered more significant.

IV. RESULTS

This chapter presents the results of our research on code properties used for smell analysis by humans or automated systems.

A. Code properties

We used six distinct data sources useful for code smell analysis from [15] as a starting point for our catalog of code properties. These include *programmer's intuition and experience*, *metric values*, *abstract syntax tree (AST)*, *history of changes made in code*, *dynamic behavior of code*, and *existence of other smells.*

From these data sources, we retained *metric values*, referred to as **structural metrics** in our catalog of properties, **AST**, and *history of changes made in code* (referred to as **change history** in our catalog). The selection of these properties was based on data from survey papers on code smell analysis [6][11][12][13] and our SLR [14], which found that these properties were often used to analyze code smells.

Structural metrics provide a way to quantify the structural properties of the code, and most existing smell detection approaches are based only on structural metrics extracted from source code [7]. An extensive list of rules for automatic code smell detection based on the metrics and thresholds defined for each metric is presented in [23].

AST is a tree representation of the code structure where each node represents a construct in the code. AST is often used as an intermediate model for various model-based techniques for code smell analysis [23]. AST contains structural and semantic information that can be used for code smell analysis [8][24][25].

The code's **change history** provides insight into how the software evolved. This information can be used to identify trends and patterns in development, providing additional useful information for code smell analysis [26]. Several studies used change history to detect code smells [26][27][28], concluding that historical information can be helpful in the analysis of some code smells (e.g., *Divergent Change* and *Shotgun Surgery*).

We replaced *programmer's intuition and experience* with a code property named **text** (source code), bearing in mind that the manual smell analysis requires the source code where developers can rely on their intuition and experience. Analyzing the source code can reveal issues in the design or implementation that code smells imply. In [21] and [22], developers were asked how they analyze code smells. Both studies found that developers look at the source code to identify code smells and evaluate their severity.

Dynamic behavior of code as a data source is rarely discussed in the literature we reviewed while identifying properties [6][11][12][13][14]. Additionally, we did not analyze the co-existence of code smells because we focused primarily on the properties of the code being analyzed to determine the presence and impact of a code smell. Therefore, our catalog excludes *dynamic behavior of code* and *existence of other smells*.

Finally, we added the fifth code property to our catalog, representing **relationships** between system classes. The analysis of relationships is necessary to determine the presence of certain smells and has been used for code smell analysis [9][10][30][31]. Relationships can indicate highly coupled parts of the system, which can imply the presence of specific code smells (e.g., *Inappropriate Intimacy*) [29].

We use the term relationships for several relationships between system classes, such as inheritance. The important information is which classes are in the inheritance hierarchy and each class's direct descendants or ancestors. The next significant relationship represents method invocation to observe invoked methods, classes they belong to, and classes that invoked those methods. Similarly, it could be helpful to analyze the attributes that are accessed or modified, the classes to which they belong, and the classes that accessed or modified those attributes. One class could also have a member variable, local variable, method parameter, or return value of another class. These relationships also indicate highly coupled parts of the system and help analyze code smells. Some structural metrics provide insight into the relationships (e.g., CBO, DIT), but semantics are difficult to analyze through metrics.

B. Mapping code properties to code smells

After identifying the code properties, we map them to 22 code smells specified by Fowler [1], highlighting which properties are useful for analyzing each code smell. The resulting mapping is summarized in Table I where we mark "Yes" if the property was used to analyze a particular smell and list the references that discuss this analysis. If there are no references next to the value "Yes", our experts have reached these conclusions and we will describe their reasoning in more detail below. Code smells for which the value "No" is filled in Table I are those for which we found no evidence in the literature of specific properties being used for analysis, nor could we determine their usefulness through the expert opinion method [17].

Text and code smells: Several studies [21][22][32][33] conclude that developers perceive the existence and impact of some code smells by analyzing text (source code). For certain code smells (i.e., *Primitive Obsession, Alternative Classes with Different Interfaces, Divergent Change, Incomplete Library Class,* and *Comments*), we did not find similar evidence in the literature, so we used the expert opinion method to complete the mapping of the text code property. We concluded that all code smells could be analyzed through text as developers can look for the structures in the source code that suggest the possibility of refactoring (i.e., removal) of smells [1].

AST and code smells: AST has been used for code smell analysis [8][24][25][33][37][38]. *Alternative Classes with Different Interfaces* is one of the code smells not found in these studies. Since that code smell refers to classes with similar methods with different names [1][23], using the expert opinion method we concluded that similar patterns in classes could be observed through AST. Other smell not found in the literature is *Parallel Inheritance Hierarchies*, but we concluded that AST can also be used to observe hierarchies, providing information for this code smell.

Structural metrics and code smells: The metric-based approach is a widely used method among the various techniques for code smell detection [34]. Structural metrics can be used in combination with thresholds in detection rules [23], as well as in machine learning approaches [35][36].

Change history and code smells: Historical information has been used for code smell analysis in several studies [26][27][28][39][41].

TABLE I. CODE PROPERTIES MAPPED TO FOWLER'S CODE SMELLS, WHERE EACH VALUE (YES OR NO) INDICATES WHETHER THE SPECIFIC CODE PROPERTY IS USEFUL FOR ANALYZING AND DETECTING A PARTICULAR CODE SMELL

Code smell	**Code properties**				
	Text	***AST***	***Structural metrics***	***Change history***	***Relationships***
Long Method	Yes[21][22][32]	Yes[8][24][25]	Yes[23][34][35][36]	Yes[39]	No
Large Class	Yes[21][22][32]	Yes[8][25]	Yes[23][34][35][36]	Yes[26][27][39][41]	Yes[42][43]
Primitive Obsession	Yes	No	Yes[23]	Yes	No
Long Parameter List	Yes[21][22]	Yes[24][37]	Yes[23][34][35]	Yes	No
Data Clumps	Yes[33]	Yes[33][37]	Yes[34]	Yes	No
Switch Statements	Yes[32][33]	Yes[24][33][37]	Yes[23]	Yes	No
Temporary Field	Yes[32]	Yes[37]	Yes[23]	No	No
Refused Bequest	Yes[21][22][32]	Yes[24][37]	Yes[23][34][35]	Yes[39]	Yes
Alternative Classes with Different Interfaces	Yes	Yes	Yes[23]	No	No
Parallel Inheritance Hierarchies	Yes[32]	Yes	Yes[23]	Yes[26][27]	Yes
Divergent Change	Yes	No	Yes[23][34]	Yes[26][27][28]	Yes[31]
Shotgun Surgery	Yes[32]	No	Yes[23][34]	Yes[26][27][28][39]	Yes[42]
Lazy Class	Yes[21][22][32]	Yes[24][37]	Yes[23][34]	No	No
Data Class	Yes[32]	Yes[37]	Yes[23][36]	Yes[39][41]	Yes[43]
Duplicated Code	Yes[22][32]	Yes[38]	Yes[23][34]	Yes[28]	No
Speculative Generality	Yes[21][22][33]	Yes[33]	Yes[23][35]	No	No
Message Chains	Yes[32][33]	Yes[24][33]	Yes[23]	No	Yes
Middle Man	Yes[21][32][33]	Yes[33]	Yes[23][35]	No	Yes
Feature Envy	Yes[21][22][32]	Yes[8][24][25]	Yes[23][34][35][36]	Yes[26][27][39]	Yes
Inappropriate Intimacy	Yes[21][22][32]	No	Yes[35]	No	Yes[29]
Incomplete Library Class	Yes	No	No	No	No
Comments	Yes	No	Yes[23]	No	No

For some code smells (i.e., *Primitive Obsession, Long Parameter List,* and *Data Clumps*), we did not find similar research in the literature. It is pointed out in [40] that these smells seem likely to grow over time. Taking this into account and using expert opinion, we concluded that change history could provide useful information for analyzing these code smells.

We also used expert opinion on *Switch Statements* code smell, relying on the definition from [1] where it is said that the problem is essentially a duplication of code. Since *Duplicated Code* can be analyzed through change history [28], we concluded that *Switch Statements* could too.

Relationships and code smells: The visualization or information of relationships was used as an aid in the code smell analysis by the authors of several papers [29][31][42][43].

For smells not included in the mentioned papers (*Refused Bequest, Parallel Inheritance Hierarchies, Message Chains, Middle Man,* and *Feature Envy*), we used the expert opinion method to finish the mapping of relationships code property. The inheritance relationship [5] and related classes/methods should be considered when analyzing *Refused Bequest*. Examining the hierarchy in which the observed instance is located, we can determine if the hierarchy is incorrectly designed and whether the observed class is appropriately placed in it.

Several inheritance-related metrics are used in [23] for *Parallel Inheritance Hierarchies* code smell. These metrics provide numerical values that indicate certain characteristics of the hierarchy in which the observed class is located. Analyzing the relationships and related classes is essential to gain a deeper understanding of this code smell.

Message Chains code smell is reflected in a series of calls to other objects [5]. These relationships between various objects should be analyzed when detecting *Message Chains* and determining how to remove this smell. Class or method infected by *Middle Man* code smell has references to other classes as it is calling their methods [23][33]. We concluded that developers should pay attention to these relationships between system's classes when analyzing *Middle Man,* to determine whether the observed class just delegates its work to other classes or if it has a meaningful responsibility.

Methods containing *Feature Envy* code smell frequently access data from other classes [5][10]. These dependency relationships should be analyzed when assessing whether *Feature Envy* is present in code, which class is being accessed the most and how this smell should be removed.

V. DISCUSSION

This chapter will discuss the results, limitations and threats to the validity of our research. These limitations can also indicate areas for improvement and further research.

For five identified properties and 22 Fowler's code smells (a total of 110 combinations for mapping), we based 60% of the mapping on the existing literature. We completed the remaining 40% of the mapping based on the conclusions of two experts. To map text property and code smells, we used the expert opinion method in 22.7% of cases (five out of 22 smells had to be evaluated using this method). In the case of AST, 36.4% of mapping (eight out of 22 smells) was done using the expert opinion method, whereby the experts concluded that 27.3% of smells could not be analyzed by AST (six out of 22 smells). In the case of structural metrics, an expert opinion

was required for only one smell, and it was concluded that the metrics are not suitable for the analysis of that specific smell. From 59.1% of smells (13 out of 22), the experts concluded that 41% (nine out of 22) could not be analyzed through change history. Finally, the relationships property was the least represented in the literature, whereby experts made conclusions for 77.3% of smells (17 out of 22), and for as many as 54.5% (12 out of 22), it was concluded that this property would not be useful for the analysis.

Given the lack of information in the literature for mapping all properties and smells, we relied on two experts to complete the mapping. A greater number of experts in this mapping stage may have led to discovery of additional relevant insights.

Apart from the five properties we have identified, several other properties for code smell analysis can be found in the literature (i.e., semantic metrics [44], dynamic code behavior [15][45], and concern metrics [46]). Semantic metrics have been discussed more in the literature in the light of refactoring and various aspects of code quality. Still, we have not found relevant papers that analyze semantic metrics for the analysis of Fowler's code smells. Dynamic code behavior was not included because no recent studies used this property, while we found this property used only for *Data Class* [15] and *Refused Bequest* [45]. We excluded concern metrics because only several code smells were analyzed using those metrics (*Divergent Change, Shotgun Surgery*, *Large Class*) [46].

Furthermore, the existing tools for smell analysis often use properties that are easier to calculate and use (i.e., structural metrics are easier to calculate than semantic or concern metrics which require a deeper understanding of the system). Considering the literature and existing tools, it would be too challenging and time-consuming to identify all properties and map them to Fowler's code smells as part of this paper. It remains an open question of which less frequent code properties would be useful for smell analysis.

While we have mapped five properties to 22 Fowler's code smells, there are additional code smells in the literature that we have not covered [47][48][49][50][51]. With such a comprehensive list of code smells, we chose to concentrate on those specified by Fowler, which have a strong presence in the literature. Future researchers in this field may consider mapping code properties to other code smells to assess their usefulness for analysis.

We should point out that we did not conduct a SLR to identify the code properties for code smell analysis. A SLR would represent a more thorough approach for gathering information, but it would also be highly time-consuming. As indicated in the description of our methodology, we mainly relied on the data we read in recently published review papers and the papers they reference. However, we did rely on our SLR [14] of the existing machine learning approaches for code smell detection. As part of that review, we collected the data sources used in detection approaches, which we found useful for this study.

We believe that researchers should go beyond the traditional structural metrics when analyzing code smells and examine other code properties to achieve a more comprehensive understanding, and improved analysis of code smells. The resulting mapping of identified properties and code smells can guide software engineers and code maintainability researchers to build automated code smell detectors that examine properties beyond the traditional structural metrics. Additionally, our results can help researchers manually analyze code smells by examining properties highlighted for specific smells and investigate the impact of other properties sufficiently presented in the literature.

VI. CONCLUSION

The first objective of our study was to identify code properties used for code smell analysis by both human and automatic detectors. The second objective was to map the identified properties to code smells specified by Fowler.

Based on our previous work, the existing literature, and expert opinion, we have defined five code properties that are used for code smell analysis: text, AST, structural metrics, change history, and relationships. Through the mapping of these properties and 22 code smells, we emphasized the usefulness of properties for specific smell analysis. Resulting mapping could guide software engineers and researchers in their attempts to analyze code smells and build automated code smells detectors.

In addition to the five identified code properties, other properties (e.g., semantic metrics, concern metrics) have received less attention in the literature. They could be examined in the future to assess their impact and relevance for smell analysis. It should be noted that there are also other code smells besides Fowler's. These code properties and code smells have yet to be explored in the literature, and the research community should strive to identify all aspects essential for analyzing smells.

ACKNOWLEDGMENT

This research is supported by the Science Fund of the Republic of Serbia, Grant No 6521051, AI-Clean CaDET.

REFERENCES

[1] M. Fowler, et al., Refactoring: improving the design of existing code. Addison-Wesley, 1999.

[2] T. Sharma and D. Spinellis, "A survey on software smells," Journal of Systems and Software, 138, pp.158-173, 2018.

[3] A. Kaur, "A systematic literature review on empirical analysis of the relationship between code smells and software quality attributes," Archives of Computational Methods in Engineering, 27, pp.1267-1296, 2020.

[4] M. Fowler, Refactoring: Improving the Design of Existing Code. Addison-Wesley Professional, 2018.

[5] M. Hozano, et al., "Are you smelling it? Investigating how similar developers detect code smells," Information and Software Technology, 93, pp.130-146, 2018.

[6] M. I. Azeem, et al., "Machine learning techniques for code smell detection: A systematic literature review and meta-analysis," Information and Software Technology, 108, pp.115-138, 2019.

[7] F. Palomba, et al., "Detecting bad smells in source code using change history information," 28th IEEE/ACM International Conference on Automated Software Engineering, pp.268-278, 2013.

[8] M. Hadj-Kacem and N. Bouassida, "Improving the Identification of Code Smells by Combining Structural and Semantic

Information," Neural Information Processing: 26th International Conference, pp.296-304, 2019.

[9] E. Murphy-Hill, et al., "Interactive ambient visualizations for soft advice," Information Visualization 12(2), pp.107-132, 2013.

[10] R. Oliveira, et al., "Collaborative identification of code smells: A multi-case study," IEEE/ACM 39th International Conference on Software Engineering: Software Engineering in Practice Track, pp.33-42, 2017.

[11] M. S. Haque, et al., "Causes, impacts, and detection approaches of code smell: a survey," Proceedings of the ACMSE 2018 Conference, pp.1-8, 2018.

[12] A. Kaur, et al., "A review on machine-learning based code smell detection techniques in object-oriented software system(s)," Recent Advances in Electrical & Electronic Engineering, 14(3), pp.290-303, 2021.

[13] G. Rasool and Z. Arshad, "A review of code smell mining techniques," Journal of Software: Evolution and Process 27(11), pp.867-895, 2015.

[14] K. G. Grujic, et al., "Machine Learning Approaches for Code Smell Detection: A Systematic Literature Review," Available at SSRN: https://ssrn.com/abstract=4299859, 2022. Unpublished.

[15] B. Walter and B. Pietrzak, "Multi-criteria Detection of Bad Smells in Code with UTA Method," Extreme Programming and Agile Processes in Software Engineering: 6th International Conference, Proceedings 6, pp.154-161, 2005.

[16] N. Sae-Lim, et al., "An Investigative Study on How Developers Filter and Prioritize Code Smell," IEICE TRANSACTIONS on Information and Systems, 101(7), pp.1733-1742, 2018.

[17] R. Wieringa, "Empirical research methods for technology validation: Scaling up to practice," Journal of systems and software, 95, pp.19-31, 2014.

[18] A. Kovacevic, et al., "Automatic detection of Long Method and God Class code smells through neural source code embeddings," Expert Systems with Applications, 204, p.117607, 2022.

[19] N. Luburic, et al., "Towards a systematic approach to manual annotation of code smells," Preprint available at TechRxiv: https://doi.org/10.36227/techrxiv.14159183.v3,2021.Unpublished.

[20] A. Kovacevic, et al., "Automatic detection of code smells using metrics and CodeT5embeddings: a case study in C#," Preprint available at TechRxiv: https://doi.org/10.36227/techrxiv.19682754.v2,2022.Unpublished.

[21] F. Palomba, et al., "Do they really smell bad? A study on developers' perception of bad code smells," IEEE International Conference on Software Maintenance and Evolution, pp.101-110, 2014.

[22] D. Taibi, et al., "How developers perceive smells in source code: A replicated study," Information and Software Technology, 92, pp.223-235, 2017.

[23] B. Bafandeh Mayvan, et al., "Bad smell detection using quality metrics and refactoring opportunities," Journal of Software: Evolution and Process, 32(8), p.e2255, 2020.

[24] S. Slinger, "Code smell detection in Eclipse," Delft University of Technology, 2005.

[25] M. Hadj-Kacem and N. Bouassida, "Deep representation learning for code smells detection using variational auto-encoder," International Joint Conference on Neural Networks, pp.1-8, 2019.

[26] F. Palomba, et al., "Detecting bad smells in source code using change history information," 28th IEEE/ACM International Conference on Automated Software Engineering, pp.268-278., 2013.

[27] F. Palomba, et al., "Mining Version Histories for Detecting Code Smells," IEEE Transactions on Software Engineering, 41(5), pp.462-489, 2015.

[28] S. Fu and B. Shen, "Code Bad Smell Detection through Evolutionary Data Mining," ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, pp.1-9, 2015.

[29] A. Jermakovics, et al., "Visualizing software evolution with lagrein," Companion to the 23rd ACM SIGPLAN conference on Object-oriented programming systems languages and applications, pp.749-750. 2008.

[30] K. Dhambri, et al., "Visual Detection of Design Anomalies," 12th European Conference on Software Maintenance and Reengineering, pp.279-283, 2008.

[31] D. Jiang, et al., "Distance metric based divergent change bad smell detection and refactoring scheme analysis," International Journal of Innovative Computing, Information and Control, 10(4), pp.1519-1531, 2014.

[32] A. Tahir, et al., "Can you tell me if it smells? A study on how developers discuss code smells and anti-patterns in Stack Overflow," Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018, pp.68-78. 2018.

[33] T. Hall, et al., "Some Code Smells Have a Significant but Small Effect on Faults," ACM Transactions on Software Engineering and Methodology, 23(4), pp.1-39, 2014.

[34] G. Lacerda, et al., "Code smells and refactoring: A tertiary systematic review of challenges and observations," Journal of Systems and Software, 167, p.110610, 2020.

[35] F. Pecorelli, et al., "A large empirical assessment of the role of data balancing in machine-learning-based code smell detection," Journal of Systems and Software, 169, p.110693, 2020.

[36] M. Y. Mhawish and M. Gupta, "Predicting Code Smells and Analysis of Predictions: Using Machine Learning Techniques and Software Metrics," Journal of Computer Science and Technology, 35, pp.1428-1445, 2020.

[37] S. Mekruksavanich, "Design Flaws Detection in Object-Oriented Software with Analytical Learning Method," International Journal of e-Education, e-Business, e-Management and e-Learning, 1(3), p.210, 2011.

[38] I. D. Baxter, et al., "Clone detection using abstract syntax trees," Proceedings. International Conference on Software Maintenance (Cat. No. 98CB36272), pp.368-377, 1998.

[39] S. A. Vidal, et al., "An approach to prioritize code smells for refactoring," Automated Software Engineering, 23, pp.501-532, 2016.

[40] M. V. Mäntylä and C. Lassenius, "Subjective evaluation of software evolvability using code smells: An empirical study," Empirical Software Engineering, 11, pp.395-431, 2006.

[41] D. Rapu, et al., "Using history information to improve design flaws detection," Eighth European Conference on Software Maintenance and Reengineering, pp.223-232, 2004.

[42] F. A. Fontana, et al., "Towards Assessing Software Architecture Quality by Exploiting Code Smell Relations," 2015 IEEE/ACM 2nd International Workshop on Software Architecture and Metrics, pp.1-7, 2015.

[43] R. Wettel and M. Lanza, "Visually localizing design problems with disharmony maps," Proceedings of the 4th ACM Symposium on Software Visualization, pp.155-164. 2008.

[44] J. Pantiuchina, et al., "Why developers refactor source code: A mining-based study," ACM Transactions on Software Engineering and Methodology, 29(4), pp.1-30, 2020.

[45] E. Ligu, et al., "Identification of Refused Bequest Code Smells," 2013 IEEE International Conference on Software Maintenance, pp.392-395, 2013.

[46] J. Padilha, et al., "On the Effectiveness of Concern Metrics to Detect Code Smells: An Empirical Study," Advanced Information Systems Engineering: 26th International Conference, 26, pp.656-671, 2014.

[47] G. Suryanarayana, et al., "Refactoring for software design smells," ACM SIGSOFT Software Engineering Notes, 40, 2015.

[48] W. H. Brown, et al., AntiPatterns: refactoring software, architectures, and projects in crisis. John Wiley & Sons, 1998.

[49] A. J. Riel, Object-Oriented Design Heuristics. Addison-Wesley, 1996.

[50] M. Lippert and S. Roock, Refactoring in large software projects: performing complex restructurings successfully. John Wiley & Sons, 2006.

[51] T. Sharma, et al., "House of Cards: Code Smells in Open-Source C# Repositories," 2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, 2017.

Migrating to a microservice architecture: benefits and challenges

Sh. Salii, J. Ajdari and Xh. Zenuni
Southeast European University, Tetovo, North Macedonia
{ss24742, j.ajdari, xh.zenuni}@seeu.edu.mk

***Abstract*—Considering the need for scalability, flexible configuration, easier development, maintenance, and rapid agile development, many organizations are moving to a microservice architecture. This migration process is undertaken by organizations to change their software architecture and mature the benefits offered by microservice architecture. However, performing a migration towards microservices is not a trivial and structured process, and many challenges have been identified along the way.**

In this research paper, we investigate the migration process and the adoption of microservices in different organizations by conducting a survey with industry practitioners, mainly focusing on the reasons for migration, the advantages that this architecture brings, and the obstacles encountered during the migration.

The expected result of the research is to contribute to the findings reported in the scientific literature, highlighting relevant problems and identifying possible directions for future research.

Keywords—Microservices, Monolith, Microservice Architecture, Monolith to Microservices, Industry Survey

I. INTRODUCTION

Microservice architecture (MSA) is a recent architectural trend gaining tremendous interest in both, industry and academia. It can be defined as an approach to developing software applications as a subset of small, independently deployable services that communicate with each other through HTTP resource APIs [5]. MSA adoption brings several advantages [6]. MSA enables a modular and loosely coupled approach to software development, where it is possible to develop, deploy, and scale each service independently making it easier to implement changes without affecting the entire system [7] [15]. Compared to monolithic architecture, MSA is more fault-tolerant since the failure of one service will not disrupt the whole system [3]. Moreover, microservices allow for different services to be built using different technologies. These advantages and more improve resource allocation and usage, system reliability, availability, and maintenance. Major companies that are using MSAs include Netflix, Uber, and Amazon.

While MSA is suitable for many types of applications, it is not the best fit for all types of software [8] [11]. Considering the economic impact, the trade-offs should be carefully evaluated so that the benefits outweigh the added complexity. The decision to architect a system should always be based on solid information, ensuring that the expected benefits will be met [20]. Adopting MSAs is a very challenging and error-prone process, for either greenfield or brownfield systems [2] [18]. It requires a significant investment in terms of design, development, testing, and maintenance.

In this study we investigate the migration journey of industry practitioners towards MSA, analyzing their migration experience with a main focus on the migration motivations, the challenges encountered and the benefits perceived. The main aim of this study is to highlight industry migration practices toward MSA. To reach this, we designed an online survey questionnaire targeting our network of IT practitioners involved in the migration process. In total, 23 industry practitioners at different professional stages from 23 different IT companies participated in our study. We collected high-quality information on the challenges faced during migration activities and the migration impact.

The main contributions of this paper include the following:

- a survey of industry practitioners that provides insights about motivations, challenges, and barriers during migration and the perceived benefits of migration.
- analysis and interpretation of gathered data.
- discussions on the obtained results.

The audience of this paper targets both researchers and practitioners in the field of software engineering, who are interested in getting an overview of existing challenges during the design and implementation of microservices, as a guide for their organizations' decisions or future scientific contributions.

The rest of the paper is organized as follows. In Section II we summarize relevant related work on the migration toward MSA. In Section III we present the methodology used to conduct the research, including research questions and hypothesis, survey questionnaire structure, and analysis. In Section IV we discuss and interpret the obtained results. In Section V we present threats to validity and in Section VI we summarize the main findings and discuss possible future work.

II. RELATED WORK

The existing body of research on MSA adoption encompasses a wide range of topics, such as motivations [13] [1], challenges [2] [12] [14] [31], and benefits [8] [15] [4] [31]. For instance, Razzaq and Ghayyur [9] use a systematic mapping study to study and examine the transition from monolithic to microservice architectures, while

Balalaie, Heydarnoori, Jamshidi, Tamburri, and Lynn [10] explored a catalog of migration and rearchitecting patterns through a qualitative empirical approach. In addition, the comparison of monolithic and microservice architectures concerning performance and scalability was investigated by Blinowski, Ojdowska, and Przybylek [8].

Various authors have focused on specific aspects of MSA adoption. Capuano and Muccini [13] conducted a Systematic Literature Review of 58 papers to exam- ine the significance of quality attributes in the MSA adoption process. Carrasco, Bladel, and Demeyer [14] presented solutions to common pitfalls in terms of bad smells, while Dinh-Tuan and Beierle [16] analyzed the resource consumption and scalability benefits of stateless microservices-based applications through a prototype implementation using a remake of Flappy Bird. Moreover, the challenges of migrating enterprise legacy source code to microservices architecture (MSA) were explored by Furda, Fidge, Zimmermann, Kelly, and Barros [18]. Their investigation highlighted issues related to multi-tenancy, statefulness, and data consistency.

Several studies employing interviews and survey questionnaires with industry practitioners have been conducted, such as those by Francesco, Malavolta, and Lago [2] and Fritzsch, Bogner, Wagner, and Zimmermann [17], which primarily explored motivations and challenges faced during MSA adoption. In contrast, our study expands the discussion to encompass the impact and benefits of MSA adoption. Auer, Lenarduzzi, Felderer, and Taibi [3] also followed a similar approach, in the form of interviews with professionals, proposing an assessment framework based on a set of metrics to consider before rearchitecting a monolithic system.

Technical lessons learned during the migration process have been documented by Gouigoux and Tamzalit [19] and Faustino, Gonçalves, Portela, and Silva [30] in their respective case studies. Furthermore, Tymchenko, N. Pleskanka, and M. Pleskanka [20] discussed the key steps required for a successful migration, whereas Ponce, Márquez, and Astudillo [4] studied the advantages and challenges of MSA adoption, including infrastructure, team organization, and migration risks. Preti, Araújo Souza, Freiberger, and De Almeida Lacerda [21] developed a migration strategy for the Public Safety Secretariat of Mato Grosso, Brazil, considering risk factors for migration. Raj, Vanga, and Chaudhary [22] focused on event-driven architecture in their study, while also examining Kubernetes [24] and tools for designing cloud-native applications [23]. Ren et al. [25] combined static and dynamic analysis to understand the characteristics of monolithic applications when migrating to MSA. Finally, Romani, O. Tibermacine and C. Tibermacine [26] proposed a data-centric process for identifying microservices in legacy software systems.

Numerous frameworks have been introduced as a result of extensive research in this area. Rubio-Drosdov [27] developed a framework for migrating monolithic systems to MSA and evaluating the performance of the resulting systems. Tankovic, Grbac, Truong, and Dustdar [33] put forth a framework for converting web applications designed for vertical scaling into cloud applications with elastic scaling capabilities, while in a separate study, Tankovic´ and Grbac, together with Žagar, introduced the ElaClo framework [34], a development framework that optimizes topologies of service-based applications based on response-time-based service-level objectives and cloud infrastructure operating costs. Santos and Silva [28] introduced a complexity metric for MSA migration, taking into account factors such as system size, complexity, dependencies, and technical and organizational challenges. Meanwhile, Selmadji et al. [29] proposed a semi-automatic approach for MSA adoption that involves decomposing the monolithic system into a set of microservices.

In summary, the existing research on MSA adoption is vast and encompasses various aspects, including motivations, challenges, benefits, and practical experiences. Our study builds upon this literature by not only examining the motivations and challenges faced by industry practitioners during MSA adoption but also focusing on the impact and benefits after adoption.

III. Study Design

This section outlines the study's design in terms of research questions, questionnaire design, execution, and data analysis.

A. Research Questions and Hypothesis

This study investigates the following research questions (RQs) inside organizations that have adopted MSA or are currently undergoing a migration to MSA:

RQ1: What are the motivations for rearchitecting a monolith system?

RQ2: What challenges are encountered during the migration process?

RQ3: What are the perceived benefits after migration?

To conduct this research, the research questions have been synthesized into 3 main hypotheses to be tested. The following hypotheses are associated with the abovementioned research questions:

H1: Organizations adopt microservices architecture due to scalability issues, complexity, and limited flexibility.

H2: The migration process to microservices architecture poses various challenges such as analyzing the current system, designing the new system, prioritizing functionalities, and overcoming technical and organizational obstacles.

H3: The adoption of microservices architecture offers benefits such as improved scalability, flexibility, fault isolation, and faster development and deployment cycles.

B. Survey Structure and Analysis

In order to ensure the rigor and integrity of the research and survey questionnaire we followed well-established

guidelines for conducting survey studies in Software Engineering by Linåker, Sulaman, De Mello, and Höst [32]. The questionnaire consists of three sections each corresponding to the information we are aiming to collect:

- Demographic Information: general information about the participants, organization, and system.
- Migration Process: motivations, system analysis, system design, and migration execution.
- Migration Impact: impact after MSA adoption.

The survey questionnaire is a web-based questionnaire, created using Google Forms and it is mainly composed of multiple-choice questions, including a few close-ended questions as well. The first section of the questionnaire consists of questions about the participant's role, experience, the organization domain, team size and the age of the system to be rearchitected. The second section of the questionnaire consists of questions about the motivations behind taking the decision to adopt MSA, the existing system assessment and analysis, the design of the new system, and the migration execution. The last section of the questionnaire consists of questions about the migration impact and the perceived benefits.

The populated form generates summarized results and we have used descriptive statistics to better represent the data about the occurrences of each given response.

IV. Survey Results

This section provides a comprehensive discussion and overview of the obtained survey results, covering personal and organizational information, system analysis, new architecture design, migration execution, and the MSA migration impact.

A. Demographic Information.

The participants were asked general questions about their role, organization, and the system they work on. Based on their answers, the participants' roles included the following: 14 out of 23 were developers, 2 out of 23 were solution architects, 4 out of 23 were Chief Executive Officers (CEOs), 2 out of 23 were product owners, and 1 out of 23 was Software Development Engineer in Test (SDET). Out of 23 participants, the majority, 15 of them had between 5 to 10 years of experience, 5 of them had 2 years of experience and 3 of them had more than 10 years of experience. Regarding organization, 20 of the participants worked in outsourcing software houses in North Macedonia, 2 in banks, and 1 in an educational institution. The team sizes varied for less than 10 members for 13 participants, between 10 and 20 for 7 participants, and 20 to 50 for 3 participants. Twelve participants were working on systems younger than 5 years, 6 of them were working on systems aged between 5 and 10 years old, and 5 participants were working on systems older than 20 years. When asked to describe the system migrated to MSA, participants answered as illustrated in Fig. 1. Fourteen of the participants described their system as a

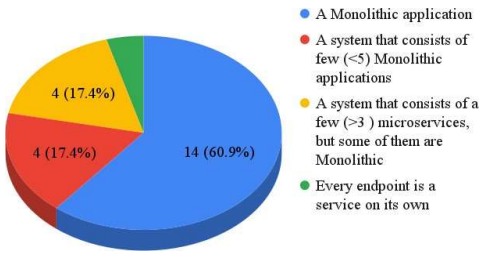


Fig. 1: System description

monolithic application, 4 of them described their system as a system that consists of few (<5) monolithic applications, 4 as a system that consists of a few (>3) microservices, but some of them are monolithic and 1 participant stated that they work on a system where every endpoint is a service on its own.

Participants were also asked about the migration execution phase as well and as illustrated in Fig. 2, we can see that 8 of the participants stated that the migration process was already completed for them, 12 of them were in the execution phase and described their migration stage as with >35% progress, and 3 of them were in the early stage of migration. The participants who were at an early stage of migration were unable to respond to questions about the benefits of MSA adoption, which we will analyze in a later phase.

In order to be able to understand more about the migration execution, we asked participants if there have been delays compared to the initial timeline estimation. As illustrated in Fig. 3, the migration process was not delayed for 14 of the participants, while 9 of them reported that they were delayed. It is important to highlight that none of the participants were highly delayed and that the number of participants who were not delayed is greater than the number of those delayed, meaning that none of the participants experienced enormous challenges.

B. Migration Motivations (RQ1)

Participants were also asked about the driver of the migration and the issues they had with their existing systems. When asked about the issues they had with their current system and what triggered MSA adoption they selected several reasons. However as illustrated in Fig. 4,

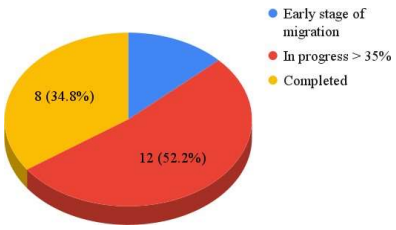


Fig. 2: Migration execution phase

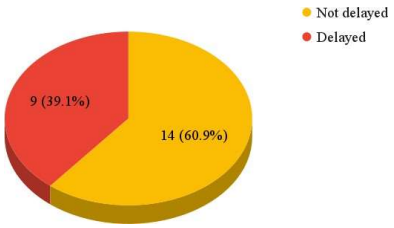


Fig. 3: Migration delay status



Fig. 4: Motivations for migrating to MSA

we can conclude that performance/scaling issues and maintainability were identified as the main drivers for migrating to MSA. The same result was reported in the literature by Fritzsch, Bogner, Wagner, and Zimmermann. [17].

Microservices allow for more flexibility and autonomy in the development and deployment process, where each service can be developed, deployed, and scaled independently, thus allowing more efficient management of the resources. On the other hand, fault tolerance, agility, and cost efficiency were also identified as important factors for deciding to rearchitect the existing systems. While in 4 cases the decision was not taken by the participants, other drivers with 3 occurrences each include the following:

- technological heterogeneity and team autonomy
- testing issues with the existing system
- incompatible system with business requirements
- technology experimentation

Based on these survey results that targeted RQ1, we can conclude that the first hypothesis is supported.

C. *Migration execution activities and challenges (RQ2).*

In order to be able to answer RQ2, the second section of the survey questionnaire consisted of questions about the analysis of the existing system, the design activities, and the migration execution. Analysis of the existing system is important in order to understand it and design the new architecture. Fig. 5 illustrates the sources used to analyze the system, where we can see that documentation and developer knowledge is considered as main resources through which the existing system was analyzed. Both documentation and developer knowledge provide valuable information about the current system, such as its structure, functionality, dependencies, and limitations.

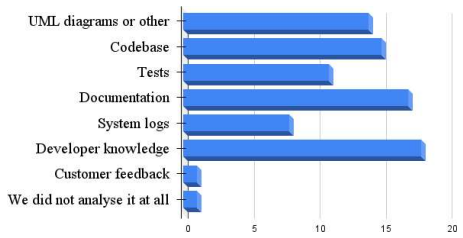


Fig. 5: Sources used to analyse the system

This information can be easily used to identify which parts of the system can be broken down into microservices, how they should be designed, and how they should interact with each other. On the other hand, 15 participants considered diagrams as a good analysis resource, while tests and codebase presented relevancy as well, with 11 and 15 occurrences each. System logs are selected by 8 participants and only one of the participants has used the client's feedback for system analysis. Finally, it is essential to mention that 1 participant did not analyze the system at all before starting the migration process, which should have caused unexpected challenges and delays.

Analyzing an existing system is not a trivial process, therefore it is important to highlight the challenges faced during system analysis. Participants were asked about the challenges they faced during system analysis, and the main obstacle they highlighted was the presence of undocumented functionalities. Fig. 6 illustrates the challenges encountered during system analysis. In addition to documentation, tests are an essential source of information, and the absence of tests and validation is a significant challenge, which occurred 14 times. Lack of database documentation and inconsistency in database documentation is listed as another challenge that occurred 10 times. At the same time, codebase and legacy technologies were also identified as relevant problems, each occurring 7 times.

As mentioned, analyzing the existing system is important in order to be able to design the new system. Survey participants were asked about the activities taken and the challenges faced when designing the new system. Fig. 7 illustrates the activities undertaken by the participants while designing the new system. As we can see, the answers are distributed across multiple options, however, the identification of dependencies is the most selected one with a total of 17 occurrences.

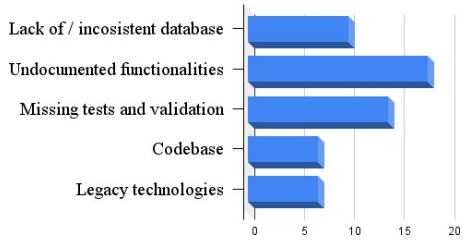


Fig. 6: Challenges encountered during system analysis

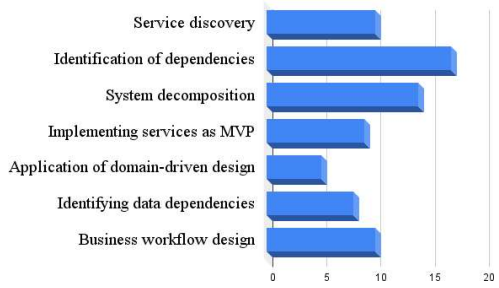


Fig. 7: Activities undertaken for designing the new system

By identifying dependencies early on in the migration process, teams can design and deploy microservices that are loosely coupled and can operate independently, which improves the overall resilience and scalability of the system. Additionally, identifying dependencies also allows teams to better manage and maintain the microservices, as they have a better understanding of how they interact with other parts of the system.

Participants were not only asked about the design of the new system but also about the impact on value delivery. In Fig. 8 we can see that 19 participants confirmed that they have implemented and delivered new functionalities during migration. Migration is not a fast process, therefore it is justifiable that new functionalities and features must be delivered while the migration is in process.

As a follow question, participants were also asked about the challenges encountered during the design phase. The results of the question are illustrated in Fig. 9. As illustrated lack of documentation of the existing system is the leading challenge with the highest number of occurrences during system design. The obtained result is consistent with the most selected challenge during the system analysis phase as well. At the same time, it is important to note that there are multiple answers referring to the identification of services, communication among them, and reducing coupling as well.

The following survey questions were focused on migration execution. Participants were asked about the processes they used for MSA adoption, where based on Fig. 10 we can see that the first-step approach was used by 8 participants, 4 of them used greenfield adoption, and 5 of them strangler pattern. It is important to note that none

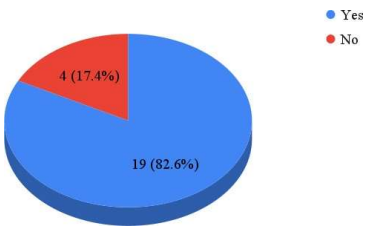


Fig. 8: New functionalities delivered during migration

of the participants used Bing Bang adoption. Migrating all services at once can lead to increased complexity and higher chances of errors. Consequently, participants were also asked about how they handled the existing data during the migration process. The results of the ways how data were handled during migration are illustrated in Fig. 11. Based on the illustrations, we can see that 18 of the participants migrated data to the new system, while 5 of them did not. Migrating data to a new system is important because it helps the organization improve its data management and storage capabilities.

To conclude the questions for this section, the survey participants were asked about the technical and organizational challenges during migration execution. As illustrated in Fig. 12 the main organizational challenge encountered is the lack of expertise/experienced resources with 16 occurrences, which challenge is encountered also during system design. Resistance to change is also a relevant challenge with 9 occurrences. We can clearly state that resistance to change is directly affected by the lack

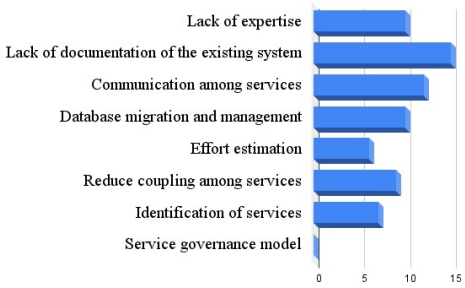


Fig. 9: Challenges encountered while designing the new system

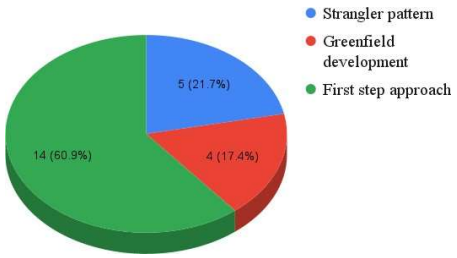


Fig. 10: Microservice adoption process

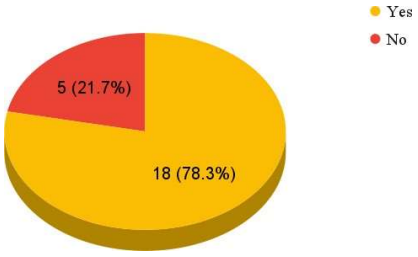


Fig. 11: Data migration

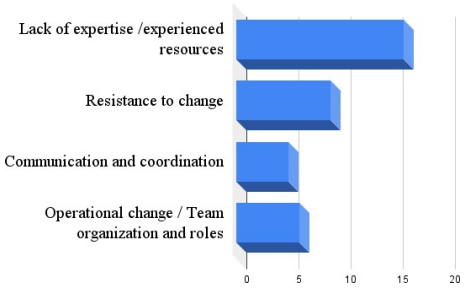


Fig. 12: Operational challenges during MSA adoption

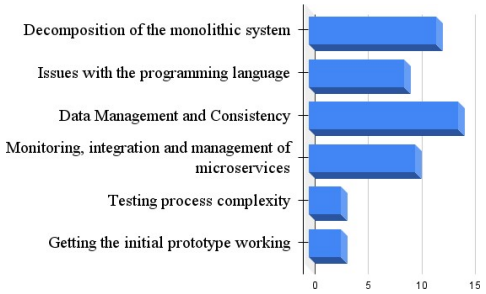


Fig. 13: Technical challenges during MSA adoption

of expertise because employees have difficulty adapting to the new system and processes.

Apart from that, other non-technical challenges that are worth mentioning include the organization of teams, role assignments, and communication/coordination between team members. Similar results are also reported in literature [17] [25] as well. On the other hand, the results of the question for technical challenges are illustrated in Fig. 13, which reveals that the most significant challenges encountered during MSA implementation were the decomposition of the monolithic system and managing data consistency and management. These challenges are consistent with those faced during the system design phase. Other challenges include monitoring, integration, and management of microservices, programming language issues, and the complexity of the testing process, with 10, 9, and 3 occurrences respectively for each challenge.

These results highlight that the migration process to MSA can indeed pose various challenges, therefore it is evident that the second hypothesis H2 is supported.

D. Migration Impact (RQ3).

We asked participants about the impact of adopting MSA and the perceived benefits of rearchitecting their systems. As seen previously, 5 of the participants had already completed the migration and 6 of them had made considerable progress and are able to evaluate the impact. The benefits of migrating to MSA are illustrated in Fig. 14. Based on the illustrations we can conclude that improved scalability is the most recognized improvement after MSA adoption with 17 occurrences, which is immediately followed by an improved performance with 15 occurrences.

Similar results are also reported in literature [31]. At the same time, it is important to mention that improved availability has 13 occurrences, while improved development velocity/team productivity has been selected 8 times. Improved maintainability and cost efficiency appear with 6 occurrences each and at the same time improved deployment velocity has 4 occurrences. Only 2 of the participants selected improved compliance while one of them stated that they are still at an early stage to recognize any benefits.

The survey results provide evidence that supports H3, indicating that the adoption of microservices architecture

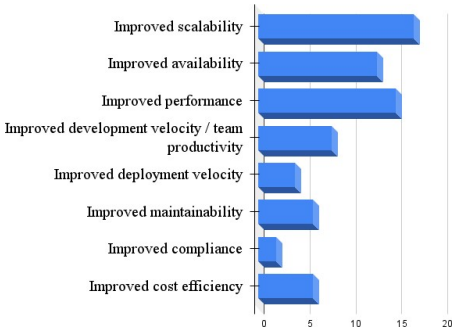


Fig. 14: Benefits of MSA adoption

offers benefits such as improved scalability, flexibility, fault isolation, and faster development and deployment cycles.

E. Correlation analysis of obtained results.

To gain a better understanding of the data collected through the survey, we used the Pearson correlation coefficient and heatmaps as analytical tools. Pearson correlation coefficient was used to measure the strength and direction of the relationships between the various factors related to migration, while heatmaps were used to visualize these correlations.

As seen in Fig. 15, from the heatmap, we are able to observe a positive correlation between the type of system to be migrated and the perceived benefits after migration. This indicates that the type of system has an impact on the perceived benefits of migration. Additionally, we found that the activities undertaken while designing the new system have a weak negative correlation with the challenges faced during migration, suggesting that these activities can influence the challenges faced during the migration process.

Regarding the correlation between migration delivery time and system type, we found a negative correlation, suggesting that the type of system to be migrated can have an impact on the estimated time it takes to complete the migration. On the other hand, the correlation coefficient between migration delivery time and challenges in designing the new system was weakly positive, indicating a positive relationship between the two variables. This

	System type	Sources to analyze current system	Activities while designing the new system	Challenges in designing the new system	Migration motivations	Migration delivery time vs Initial estimation	Technical challenges during migration	Perceived benefits after migration
System type	1.000	0.133	-0.086	-0.061	0.018	-0.178	-0.140	0.169
Sources to analyze current system	0.133	1.000	0.023	-0.046	-0.024	-0.059	-0.044	0.014
Activities while designing the new system	-0.086	0.023	1.000	0.005	-0.097	0.021	0.055	-0.012
Challenges in designing the new system	-0.061	-0.046	0.005	1.000	-0.033	-0.054	0.001	-0.008
Migration motivations	0.018	-0.024	-0.097	-0.033	1.000	-0.306	0.003	0.170
Migration delivery time vs Initial estimation	-0.178	-0.059	0.021	-0.054	-0.306	1.000	-0.129	-0.230
Technical challenges during migration	-0.140	-0.044	0.055	0.001	0.003	-0.129	1.000	-0.013
Perceived benefits after migration	0.169	0.014	-0.012	-0.008	0.170	-0.230	-0.013	1.000

Fig. 15: Correlation plot

suggests that challenges faced during the design of the new system may contribute to a longer delivery time than initially estimated.

The correlation coefficient between the motivations for migration and challenges in designing the new system is negative but weak, indicating that the motivations for migration may have an impact on the challenges faced during the design of the new system.

Finally, we found a weak positive correlation between technical challenges during migration and system type, indicating that the type of system being migrated may impact the technical challenges faced during migration. There is a weak positive correlation between technical challenges during migration and sources used to analyze the current system, suggesting that the sources used to analyze the current system may have an impact on the technical challenges faced during migration.

V. Threats to Validity

While the study contributes valuable insights into the challenges and benefits of MSA adoption, it also acknowledges several limitations that may affect the generalizability of the findings. Firstly, the small sample size of 23 industry practitioners may not fully represent the broader population of IT practitioners interested in MSA adoption. Secondly, the self-reported nature of the data may have introduced biases and may not accurately reflect the participants' actual experiences or behaviors. Thirdly, as the study only included industry practitioners, the experiences and backgrounds of the participants may not have been diverse enough to capture a broad range of perspectives on MSA adoption. While these limitations do not discredit the study's insights, they do highlight the need for further research that addresses these issues in different organizational and industry contexts.

VI. Conclusions and Future Work

Driven by the significance of microservices architecture, this study surveyed 23 industry practitioners to gain insights into their migration experiences, revealing the challenges, benefits, drivers, and issues related to the adoption process.

The study was guided by three hypotheses related to MSA adoption. First, H1 suggested that organizations adopt MSA to address issues related to scalability, complexity, and limited flexibility. Second, H2 suggested that the migration process to MSA poses various challenges, including system analysis and decomposition, new system design, prioritizing functionalities, and overcoming technical and organizational obstacles. Finally, H3 suggested that the adoption of MSA offers benefits such as improved scalability, flexibility, fault isolation, and faster development and deployment cycles. The survey results supported all three hypotheses, demonstrating that organizations are motivated to adopt MSA to overcome challenges related to system scalability, complexity, and limited flexibility and that the migration process to MSA is indeed challenging. Additionally, the results indicated that adopting MSA can offer significant benefits, such as improved scalability, flexibility, fault isolation, and faster development and deployment cycles.

Considering that decomposition of the monolithic system results to be a serious challenge during MSA adoption, as future work we aim to create a roadmap featuring timelines and milestones. This will be achieved through conducting case studies of organizations that have previously undergone MSA adoption and experienced similar challenges. By analyzing best practices and lessons learned, the roadmap will offer a clear and detailed migration plan, ensuring the project's progress remains on track and on schedule.

References

[1] D. Taibi, V. Lenarduzzi and C. Pahl, "Processes, Motivations, and Issues for Migrating to Microservices Architectures: An Empirical Investigation," in IEEE Cloud Computing, vol. 4, no. 5, pp. 22-32, September/October 2017, doi: 10.1109/MCC.2017.4250931.

[2] P. Di Francesco, P. Lago and I. Malavolta, "Migrating Towards Microservice Architectures: An Industrial Survey," 2018 IEEE International Conference on Software Architecture (ICSA), Seattle, WA, USA, 2018, pp. 29-2909, doi: 10.1109/ICSA.2018.00012.

[3] F. Auer, V. Lenarduzzi, M. Felderer, D. Taibi, "From monolithic systems to Microservices: An assessment framework," Information and Software Technology, Volume 137, 2021, 106600, ISSN 0950-5849, doi: 10.1016/j.infsof.2021.106600.

[4] F. Ponce, G. Márquez and H. Astudillo, "Migrating from monolithic architecture to microservices: A Rapid Review," 2019

38th International Conference of the Chilean Computer Science Society (SCCC), Concepcion, Chile, 2019, pp. 1-7, doi: 10.1109/SCCC49216.2019.8966423.

[5] J. Lewis and M. Fowler, “Microservices. A definition of this new architectural term.” https://martinfowler.com/articles/microservices.html.

[6] M. Cardarelli, L. Iovino, P. D. Francesco, A. D. Salle, I. Malavolta, and P. Lago, An extensible data-driven approach for evaluat- ing the quality of microservice architectures. Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, 2019, https://doi.org/10.1145/3297280.3297400

[7] P. D. Francesco, I. Malavolta and P. Lago, "Architecting with microservices: A systematic mapping study," Journal of Systems and Software, 2019, 150, 77-97, doi:10.1016/j.jss.2019.01.001

[8] G. Blinowski, A. Ojdowska and A. Przybyłek, "Monolithic vs. Microservice Architecture: A Performance and Scalability Evaluation," in IEEE Access, vol. 10, pp. 20357-20374, 2022, doi: 10.1109/ACCESS.2022.3152803.

[9] A. Razzaq, and S. A.Ghayyur, "A systematic mapping study: The new age of software architecture from monolithic to microservice architecture—awareness and challenges," Computer Applications in Engineering Education, 2022, doi:10.1002/cae.22586

[10] A. Balalaie, A. Heydarnoori, P. Jamshidi, D. A. Tamburri, and T. Lynn, "Microservices Migration Patterns. Software: Practice and Experience," 2018, doi:10.1002/spe.2608

[11] N. Bjørndal, , L. J. P. D. Araújo, A. Bucchiarone, N. Dragoni, and M. Mazzara, "Benchmarks and performance metrics for assessing the migration to microservice-based architectures". Journal of Object Technology. 10.5381/jot.

[12] M. Kalske, N. Mäkitalo, and T. Mikkonen, "Challenges when moving from Monolith to Microservice Architecture," Current Trends in Web Engineering, 2018, 32-47. doi:10.1007/978-3-319-74433-9-3

[13] A. Carrasco, B.V. Bladel, and S. Demeyer, "Migrating towards microservices: Migration and architecture smells," Proceedings of the 2nd International Workshop on Refactoring, 2018, doi:10.1145/3242163.3242164

[14] R. Capuano and H. Muccini, "A Systematic Literature Review on Migration to Microservices: a Quality Attributes perspective," 2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C), Honolulu, HI, USA, 2022, pp. 120-123, doi: 10.1109/ICSA-C54293.2022.00030.

[15] D. Faustino, N. Gonçalves, M. Portela, and A. R. Silva, "Stepwise Migration of a Monolith to a Microservices Architecture: Performance and Migration Effort Evaluation," 2022, doi:10.48550/ARXIV.2201.07226

[16] H. Dinh-Tuan and F. Beierle, "MS2M: A message-based approach for live stateful microservices migration," 2022 5th Conference on Cloud and Internet of Things (CIoT), Marrakech, Morocco, 2022, pp. 100-107, doi: 10.1109/CIoT53061.2022.9766576.

[17] J. Fritzsch, J. Bogner, S. Wagner and A. Zimmermann, "Microservices Migration in Industry: Intentions, Strategies, and Challenges," 2019 IEEE International Conference on Software Maintenance and Evolution (ICSME), Cleveland, OH, USA, 2019, pp. 481-490, doi: 10.1109/ICSME.2019.00081.

[18] A. Furda, C. Fidge, O. Zimmermann, W. Kelly and A. Barros, "Migrating Enterprise Legacy Source Code to Microservices: On Multitenancy, Statefulness, and Data Consistency," in IEEE Software, vol. 35, no. 3, pp. 63-72, May/June 2018, doi: 10.1109/MS.2017.440134612.

[19] J. -P. Gouigoux and D. Tamzalit, "From Monolith to Microservices: Lessons Learned on an Industrial Migration to a Web Oriented Architecture," 2017 IEEE International Conference on Software Architecture Workshops (ICSAW), Gothenburg, Sweden, 2017, pp. 62-65, doi: 10.1109/ICSAW.2017.35.

[20] M. Kyryk, O. Tymchenko, N. Pleskanka and M. Pleskanka, "Meth- ods and process of service migration from monolithic architecture to microservices," 2022 IEEE 16th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 2022, pp. 553-558, doi: 10.1109/TCSET55632.2022.9767055.

[21] J. P. Delgado Preti, A. N. Araújo Souza, E. C. Freiberger and T. De Almeida Lacerda, "Monolithic to Microservices Migration Strategy in Public Safety Secretariat of Mato Grosso," 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Kuala Lumpur, Malaysia, 2021, pp. 1-5, doi: 10.1109/ICECCE52056.2021.9514268.

[22] P. Raj, S. Vanga; A. Chaudhary, "Microservices Security," in Cloud- native Computing: How to Design, Develop, and Secure Microser- vices and Event-Driven Applications , IEEE, 2023, pp.289-298, doi: 10.1002/9781119814795.ch14.

[23] P. Raj, S. Vanga and A. Chaudhary, "Design, Development, and Deployment of Event-Driven Microservices Practically," in Cloud-native Computing: How to Design, Develop, and Secure Microservices and Event-Driven Applications , IEEE, 2023, pp.129-142, doi: 10.1002/9781119814795.ch7.

[24] P. Raj, S. Vanga and A. Chaudhary, "Kubernetes Architecture, Best Practices, and Patterns," in Cloud-native Computing: How to Design, Develop, and Secure Microservices and Event-Driven Applications , IEEE, 2023, pp.49-70, doi: 10.1002/9781119814795.ch3.

[25] Z. Ren, W. Wang, G. Wu, C. Gao, W. Chen, J. Wei, and T. Huang, "Migrating web applications from monolithic structure to microservices architecture", Proceedings of the Tenth Asia-Pacific Symposium on Internetware, 2018, doi:10.1145/3275219.3275230

[26] Y. Romani, O. Tibermacine and C. Tibermacine, "Towards Migrating Legacy Software Systems to Microservice-based Architectures: a Data-Centric Process for Microservice Identification," 2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C), Honolulu, HI, USA, 2022, pp. 15-19, doi: 10.1109/ICSA-C54293.2022.00010.

[27] E. Rubio-Drosdov, "A Framework for Microservice Migration and Performance Assesment", Intelligent Environments 2020 Workshop Proceedings of the 16th International Conference on Intelligent Environments, doi:10.3233/AISE200053

[28] N. Santos and A. Rito Silva, "A Complexity Metric for Microservices Architecture Migration," 2020 IEEE International Conference on Software Architecture (ICSA), Salvador, Brazil, 2020, pp. 169-178, doi: 10.1109/ICSA47634.2020.00024.

[29] A. Selmadji, A. -D. Seriai, H. L. Bouziane, R. Oumarou Mahamane, P. Zaragoza and C. Dony, "From Monolithic Architecture Style to Microservice one Based on a Semi-Automatic Approach," 2020 IEEE International Conference on Software Architecture (ICSA), Salvador, Brazil, 2020, pp. 157-168, doi: 10.1109/ICSA47634.2020.00023.

[30] H. H. S. D. Silva, G. Carneiro, and M. P. Monteiro, "An experience report from the migration of Legacy Software Systems to Microservice based architecture," 16th International Conference on Information Technology-New Generations (ITNG 2019), 183-189. doi:10.1007/978-3-030-14070-0-26

[31] P. Yugopuspito, F. Panduwinata and S. Sutrisno, "Microservices architecture: Case on the migration of reservation-based parking system," 2017 IEEE 17th International Conference on Communication Technology (ICCT), Chengdu, China, 2017, pp. 1827-1831, doi: 10.1109/ICCT.2017.8359946.

[32] J. Linåker, S. M. Sulaman, R. M. de Mello and M. Höst, "Guidelines for Conducting Surveys in Software Engineering," Department of Computer Science, Lund University, 2015.

[33] N. Tankovi ć, T. G. Grbac, H. -L. Truong and S. Dustdar, "Trans- forming Vertical Web Applications into Elastic Cloud Applica- tions," 2015 IEEE International Conference on Cloud Engineering, Tempe, AZ, USA, 2015, pp. 135-144, doi: 10.1109/IC2E.2015.15.

[34] N. Tankovi ć, T. G. Grbac, M. Žagar, "ElaClo: A framework for optimizing software application topology in the cloud environment," Expert Systems with Applications, Volume 90, 2017, Pages 62-86, ISSN 0957-4174, doi: 10.1016/j.eswa.2017.07.001.

Optimization and Parallelization of Object-Relational Mappers

Filip Jovanov*, Vladimir Zdraveski*, Marjan Gusev*, Magdalena Kostoska*
* Faculty of Computer Sciences and Engineering, Skopje, North Macedonia
filip.jovanov@students.finki.ukim.mk, {vladimir.zdraveski, marjan.gushev, magdalena.kostoska}@finki.ukim.mk

***Abstract*—Most of today's software is connected to the Internet, no matter its type/usage. The developers of these applications constantly have to make important decisions that impact the software's performance. One predominant area of these discussions is whether to use Object Relational Mappings or native SQL. In this paper, we introduce a new Java library called RQL and an algorithm to drastically optimize ORMs, specifically time and memory complexity, as well as the number of database calls that they produce. We set a hypothesis to check if this new approach performs better than the conventional one. The evaluation of conducted experiments proves that the new approach based on the division of work to separate threads, proper decision-making, partitioning, in-memory data mapping, entity pre-processing and the avoidance of Cartesian products and n+1 issues achieves better performance and reduces resource requirements.**

Keywords*—object-relational mappers, entity preprocessing, partitioning, distributed processing

I. Introduction

Our management specified a task to develop a solution based on a lot of legacy code, while also implementing new features. Around the end of development, management wanted a new functionality that was akin to GraphQL. The development team compared the remaining time to develop the new functionality with the estimated time of refactoring the whole codebase to a GraphQL implementation. We determined that it wasn't feasible.

So, we implemented a similar functionality that would meet the bare requirements and would be within the available time frame. After implementing it, we noticed that it brought a significant performance issue, especially regarding database access.

We then looked at different options to improve the performance and decided on creating a number of utility methods which we afterward turned into a library.

In this paper, we introduce said library, as well as explain the logic and decision-making algorithms related to it, along with its possibilities for parallel processing.

One real-world use case for this library is access to aggregation entities and data transfer objects (DTOs) like in a software product used by a government agency that regulates and reports on the systematization of other government agencies, or the annual salaries. These reports usually consist of many joins and a huge amount of data.

In a relational database, this can pose an issue, depending on the implemented solution like creating N+1 issues, cartesian products, technical debt, or restricting possible future functionality.

We aim to prove the following research hypothesis: *Time complexity, memory complexity, and the number of requests being sent to a database can be improved by disabling the possibility of Cartesian products and n+1 issues at run-time, as well as dividing the work into smaller jobs dealt by separate threads using a variety of different techniques.*

To prove the validity of this hypothesis we measure the response times and the counts of multiple different types of database requests, with multiple different data sets. The measurement is realized using semi-standard REST and standard GraphQL implementations, with simple JPA repositories, then, modifying the semi-standard REST implementation to use the synchronous RQL library, and finally the parallelized RQL library. Measured data are collected using Python scripts to measure the response times and Java methods to collect the number of database calls detected by P6Spy.

The rest of the paper is organized in the following structure. Section II presents similar solutions like GraphQL and Join-Monster, as well as key packages that RQL uses and an overview of similar useful papers. Section III elaborates on the Cartesian product and n+1 issues, undocumented Spring Data JPA relational rules, as well as Object Relational Mappings (ORMs) stances toward optimization. Then we dive into our library's solution, explaining how the entity pre-processing, partitioning, mapping, and parallel processing are implemented in Section IV. Analysis of all the different test cases and payloads, as well as the results and their evaluation, is presented in Section V. Finally, Section VI concludes the paper and gives directions for future works.

II. Related Work

One of the huge problems in the development of back-end services is the huge amount of effectively equivalent REST endpoints, either due to constantly changing requirements and not wanting to break previous implementations or because the problem is structured in such a way that the same entity needs to be displayed in multiple different ways. The current solution to this problem is GraphQL [1]. GraphQL allows one to "pick and choose" what is needed, by accessing a single endpoint.

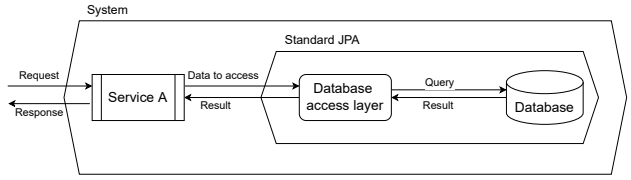


Fig. 1: General overview of system architecture.

But, GraphQL suffers from the N+1 problem [2] which can easily flood the database with unexpected requests. To solve this issue, the open-source community has created the Join-Monster [3] library, which pre-processes the GraphQL query, and creates a select to JOIN all the relations and run the query using a single database request. The issue with this approach is that this can very easily lead to a Cartesian product issue [4], which could severely tank the application's performance.

Both of these issues can be avoided by intelligently deciding the types of queries to be used based on the definitions of the entities, and, separating them into different sub-jobs, each responsible for a different segment of the entire query. RQL achieves this through the usage of multiple open-source libraries. Avoiding a restriction to the existing implementation of GraphQL, we use Cossium's EntityGraph implementation [5], which allows the dynamic generation of queries, based on the requirements that are sent to the back-end service. Another key library is MapStruct [6] because it helps us dynamically map all the attributes without risking automatic entity fetching while returning the data.

Data-parallel programming is a great paradigm that eases the parallelization of tasks and their implementation [7] which is used in RQL to map out the retrieved data to their corresponding parents. A similar idea to ours has been implemented on graph-based databases, which rewrites the queries into sub-queries, while also implementing alternate algorithms for mapping out and retrieving the data [8].

III. Database Access Optimizations and ORMs

Databases are powerful tools, but, problems arise when accessing them. Using the wrong approach for the wrong situation can mean extra seconds, minutes, or even hours during a simple lookup, in a "do anything" and "anyway" approach. Practically, they don't impose restrictions and protocols (unless they are technologically infeasible).

ORMs work in the same manner, they go by the KISS and SRP principles. They only take up the responsibility of giving database access in programming languages through objects, nothing more, nothing less. They don't try to optimize the user's database calls or try to impose any smart restrictions. This approach poses a problem because developers can make mistakes much more frequently. Two frequent types of problems due to ORMs using these principles are the N+1 issues and creating Cartesian products when they aren't wanted. Besides the principles, there is also the issue of "expected" undocumented behavior with ORMs. Such an example can be found in Spring Data JPA's "optional" property. More specifically, when an annotation has the "optional" property set to false, JPA **always** tries to access the records relating to that relation. It doesn't respect any other properties (like setting the fetch strategy as lazy), which can lead to unexpected behavior.

The new library helps to avoid this by imposing some restrictions, adding "patches" to the undocumented behavior, and making smart decisions about what actually should be accessed or not, while also giving the possibility to implement parallelization into the database access layer.

IV. Algorithms

A. General Overview

Usually, with most back-end frameworks, like Spring, the codebase is layered in such a way that there is at least a service layer that contains the business logic, and a data access layer (which is usually either an ORM or a lot of custom repositories that contain stored procedures or custom queries). In Fig. 1 we can see an example of this standard approach where Service A requests some data from Spring Data JPA (which in turn usually uses Hibernate entity manager as the data access layer) and then returns it as a result.

Fig. 3 presents a drawn-out example of this approach. Service A requests some data, but now, instead of directly calling the ORM, it first calls a divider (which also serves as an optimizer) to rebuild the query so that it runs efficiently, then, separates it into multiple sub-queries based on different criteria that are discussed later on in this same section. After that separation, it sends each of the sub-queries into different processing units that then call JPA. Afterward, when each processing unit receives the results (the parent entities as well as their relations), it sends it over to a local combiner so that any children can get re-mapped back to their parents which then gets returned as the result of each processing unit. The thread combiner collects these results and then combines them into one aggregate cohesive object that counts as the final result. Before returning the result from the RQL layer, an extra step is done which just wraps or unwraps the aggregate object so that the service layer can receive the result as if it had called JPA directly.

B. Setup

We use a specifically designed database that should break every undocumented behavior in Spring Data JPA (or in other words, designed to be as inefficient as possible), as presented in Fig. 2.

We've configured each X-to-one relation as a mandatory relation (which makes it so that even if we set the fetch type to lazy, JPA will still attempt to access these objects even if they aren't used later), for the accounts we have a many-to-many relation with itself which can easily

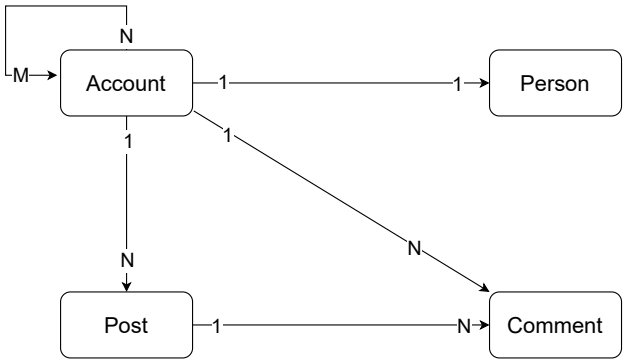


Fig. 2: ER diagrams of our sample database

throttle the performance of the application due to how JPA accesses all the relations. Another important part of the configuration is that we have defined the relation on both ends, for example in the Account and Person relation we have 2 properties, one which references the multiple posts, and one which references the accounts associated to them, which can also lead to a big throttle on performance and memory intensity due to the possibility of infinite recursion and possible depth. On a couple of the relations we also defined their "mappedBy" property which can also lead to JPA sometimes fetching the relation, even though it isn't requested or used. One more point that we'd like to stress is the circular reference between Account, Post and Comment (also keep in mind that all of these relations are 2 sided) which can lead to a heap of trouble in all aspects, performance, implementation, memory, complexity, etc. The configuration is designed in such a way that we had to override the equals and hash code methods because otherwise whenever we'd try to access any of the objects or methods, the application would simply crash.

C. Synchronous

First, we start with the synchronous aspect and optimizations of our algorithm.

Our starting point is a graph that RQL creates and uses to decide when to query specific entities. We call this graph a "reference graph". Whenever access to the database is requested, the first thing the library does is create a reference graph, which is essentially an ER diagram of the requested data with small modifications. One of those small modifications is the fact that whenever a cycle is encountered in the graph, it treats the "next in line" node as a separate node. The graph then adds the requested nodes to the tree, while also checking for bad relationship configurations (ex. a property is set with a lazy fetch strategy while also being a mandatory field). Based on the checks, it adds them in a specific way to the reference graph to avoid creating any n+1 issues by querying for them with a JOIN in their corresponding parent's batch select.

After creating the initial graph, the graph is then partitioned into special sub-graphs, based on the type of relationship in a BFS manner. If the relationship is an X-To-Many type, then it qualifies for a sub-graph. Then, it is also recursively invoked on those sub-graphs as well. While this is being processed, each sub-graph is marked with a depth level, which signifies their dependence on any previous sub-graphs. The main reason for this partitioning is to avoid Cartesian products and to add the capability for parallelism by separating one "giant" call into separate, discrete database calls.

Finally, after the partitioning, a batch query is run for each sub-graph recursively based on their depth level (each sub-graph waits for its parent so that they can receive the data to query). The query is generated through the usage of a Cossium's Entity Graph.

After all the data has been queried, it gets separated with a *HashMap* where each pair is made of the parent id as the key and a list of all its children as the value. After this separation, each parent gets their appropriate children set through reflection. Then, as the recursive functions return, the discrete data segments get combined into a final, cohesive result.

D. Parallelization

In terms of parallelism, the algorithm is implemented in such a way that it opens up multiple parallelization opportunities.

One successful attempt at parallelization was the following idea: Segment the original database call into discrete, smaller database calls where we try to gather the data as pages. This idea worked wonders. It gave the library a huge performance boost, making the calls $\sim$2-3 times faster. We implement this feature in 2 ways. The first one is by adding the capability to specify the number of threads for a particular call, and then divide the original request into that many pages. The other implementation is the inverse of the previous one, where the user would specify the maximum size of a page, and then the library would spin up as many threads as needed. With these approaches, the user can customize based on the resources they have available.

There is one downfall currently to this approach, which is a high possibility of "clogging up" the database connections since each page uses a separate connection.

Another successful attempt at parallelization was the idea of putting the HashMap segmentation and child mapping in separate threads, based on the count of child records (which can be set in spring's yml or properties configuration file). We make one big difference in the child mapping part, where we set an empty array if the property is null, and then we add the children to the list (rather than just outright setting the children). This didn't give us much performance, only about an 8-10% increase.

E. Remarks and Concerns

One point that we'd like to address is data leaks (ex. accessing a password of a user). With the algorithm model discussed thus far, this type of security is a great concern.

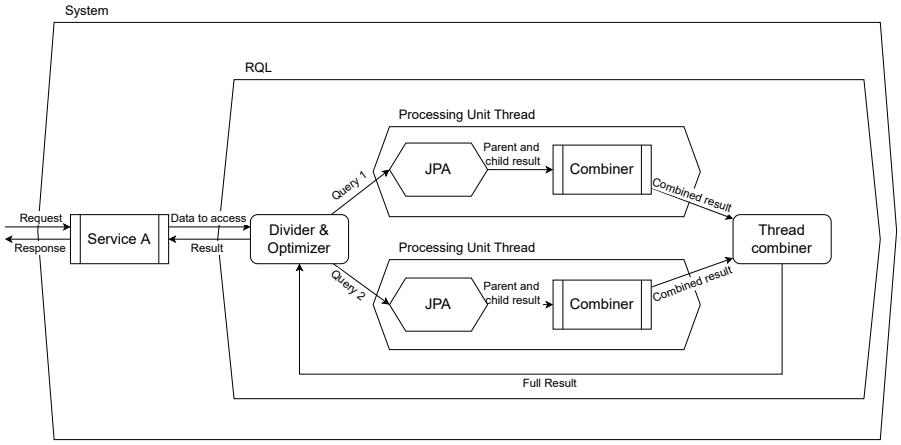


Fig. 3: General overview of proposed architecture using RQL.

We tackle this issue by adding the possibility to set an annotation on entity properties that are restricted. Then, through the usage of Spring AOP, RQL checks each database request for whether a restricted property has been requested and throws an adequate error.

Besides the possibility of data leaks, there is also the issue of over-fetching data. RQL also implements auto-generated MapStruct mappers which map the data to DTOs based on the reference graph. It can also be customized more specifically based on the scenario through the extension of those mappers, and the application of an *@InheritConfiguration* annotation [9].

V. Experimental Proof of Concept

A. Experimental methods

There are many scenarios and cases in which RQL is more or less optimized compared to its "competitors", and in this paper, we define 3 scenarios as the most important ones, each differing in the number of relations being requested, types of relations, and relational depth.

Scenario A is based on sending requests for accounts, together with all of their corresponding posts and comments. This scenario is meant to test how these libraries handle the fetching of multiple collections, which are also a part of another collection.

Scenario B will just be a simpler variant of Scenario A aiming at fetching accounts, and this time, only their posts. This is a very important scenario, precisely because it's so simple, and we'd expect that it should already be optimized.

Scenario C is dedicated to testing the depth use case, more precisely, focusing on taking accounts, their "person" relation, the account related to that person (essentially the original record), and then finally getting that account's "person".

The general basis for the testing is based on conducting tests and calculating the average of 20 requests per each number of records per scenario. The number of records per scenario is the same throughout the entire testing which is in sets of 50 records, more specifically, 50, 100, 150, 200, and 250 account records. Of course, for each account record, there will be another set of records that are fetched (but those amounts are discussed further on in this segment).

Furthermore, the database is set up in the following manner: for each account, we create 1 person and 100 posts, and then for each post, we create another 60 comments which we randomly assign to random accounts. On each initialization, we create 100 accounts, and for this test, 250 accounts are initialized (and their appropriate number of posts and comments). One more important point is that a timeout of 2 minutes is set for all cases. If a request hits this threshold, we classify it as if the result runs infinitely.

Since RQL also implements parallelization, we should also disclose the corresponding configuration. Our experiments use 5 threads, which get an equal amount of work given to them based on the algorithm.

The GraphQL queries have been defined as simple as possible since we're trying to encapsulate the performance based on the "ease of implementation" as well, or in other words, the amount of code and hours that would be needed for implementing GraphQL into a project should equal the amount it would take to implement RQL as well. The payloads in question also force the response data to be sorted based on the username of the accounts that are requested, so we can ensure that the data which is being sent as a response is the same.

B. Results

In scenario A the semi-standard JPA repository approach couldn't even operate, and while GraphQL can execute the experiment, it hits a soft cap at around 150 accounts, where the performance goes into the 2-minute mark. Other than that, RQL with our parallelization approach performs a lot faster in this case compared to the synchronous RQL approach. The semi-standard JPA failed because it created a Cartesian product that took over all

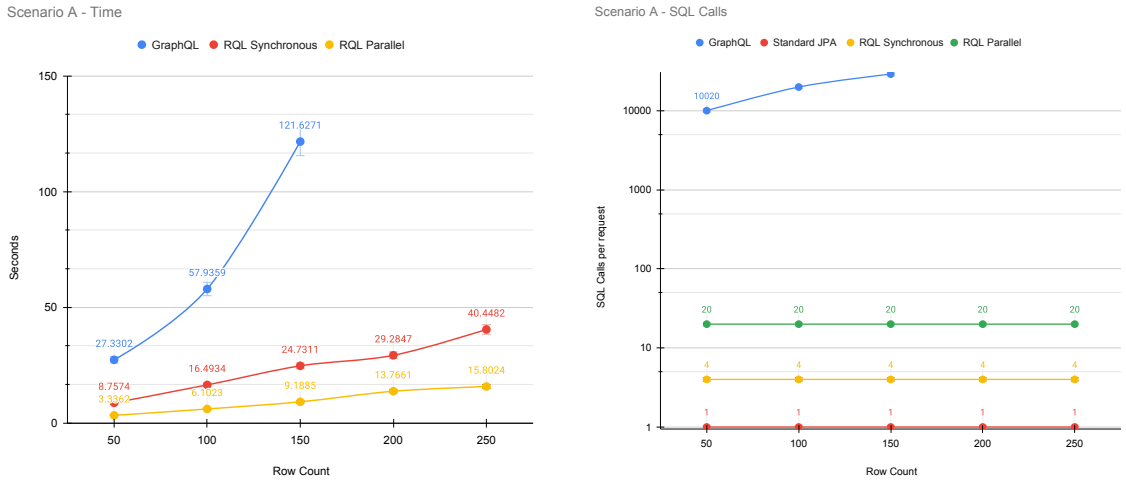


Fig. 4: Results of Scenario A.

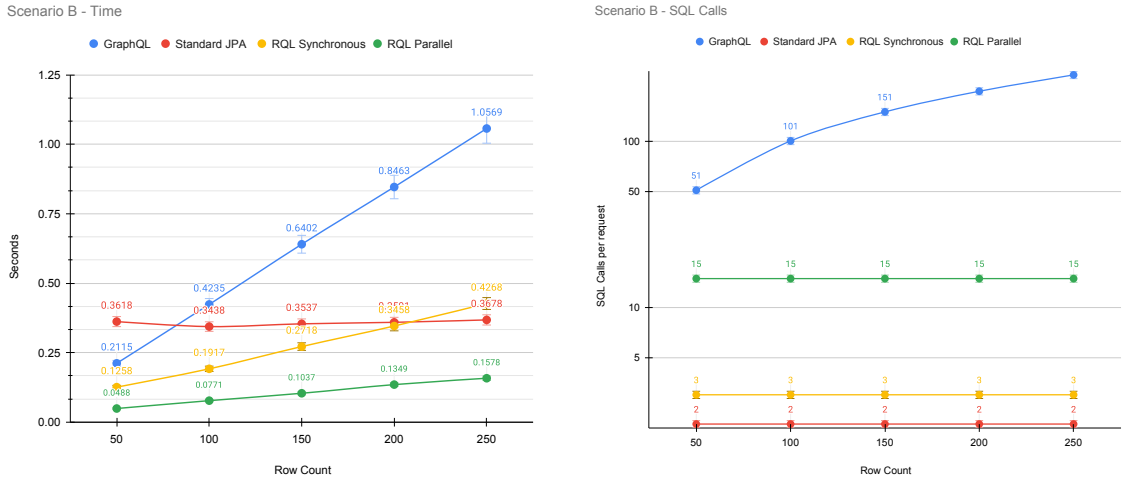


Fig. 5: Results of Scenario B.

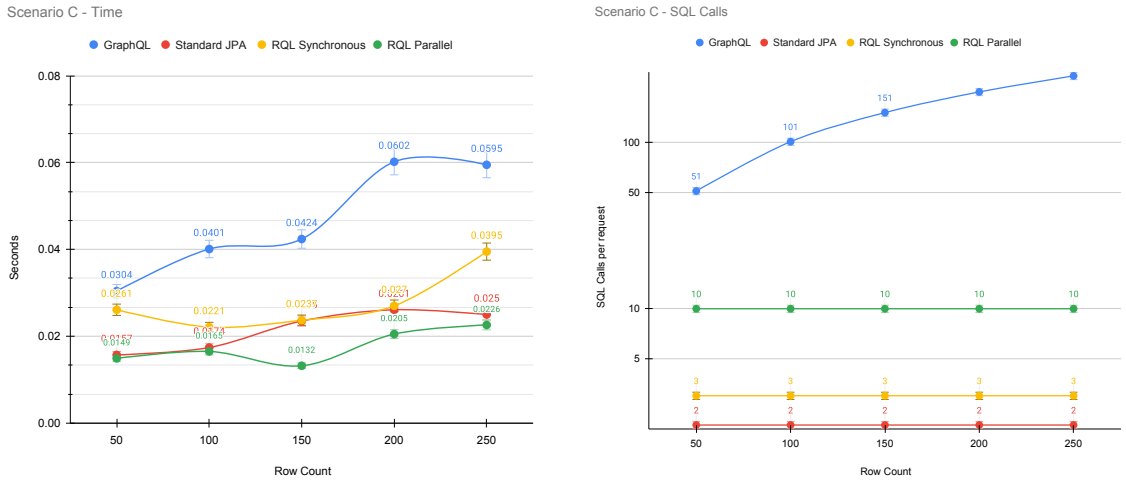


Fig. 6: Results of Scenario C.

of the system memory and hard disk memory which is dedicated to disk pagination (Fig. 4).

For scenario B the semi-standard JPA implementation managed to get to second place towards the end, and that's because RQL also values the system's memory, while JPA does not. More precisely, when running the semi-standard JPA implementation, it constantly kept taking more than the system memory made available to it (Fig. 5).

And finally, the interesting part for scenario C is the fact that the semi-standard JPA implementation was in second place (Fig. 6) throughout the entire testing.

C. Discussion

The parallelized version of RQL outperforms the synchronous one in terms of speed, while the synchronous approach uses fewer SQL calls. Other than that, both RQL approaches always outperform the GraphQL and semi-standard JPA implementations in both the number of SQL calls & speed.

Evaluating all achieved results, we conclude that indeed ORMs can improve a lot, confirming our hypothesis.

Another important fact that we can see from these results is that running a couple of extra SQL queries smartly can also drastically improve the performance and resource usage (since in the semi-standard JPA implementation we were having memory issues as discussed in Scenario A).

We also proved that some operations, currently aren't even executable with the existing technologies and their approaches.

The proof-of-concept example and insight into the introduced library are available at https://github.com/PegasusMKD/rest-graphql.

The existing implementation of the library does have a couple of edge cases, which are mainly regarding bad configurations, database & server connection throttling and deadlocks. For each of these issues, we can implement different solutions, for example, to avoid blocking server connections the library could implement a custom executor dedicated to it, and for the avoidance of database connection throttling, we can implement a producer-consumer pattern, where the library methods act as producers and we have a consumer that keeps said calls in a Queue and executes them. Finding the correct configuration has to be done with a trial-and-error approach and on a case-by-case basis. RQL will not inform the users at any point as to whether they've configured it properly or not since it can't guess the intentions, use cases, or machines that the code base will run on.

VI. Conclusion

Throughout this paper, we explained the algorithm, as well as the ideas and logic behind it. We also proved the hypothesis that by dividing the complete task into separate threads, proper decision-making, partitioning, in-memory data mapping, entity pre-processing, and the avoidance of Cartesian products and n+1 issues we can drastically improve any system's performance and resource requirements, as well as the strain it creates on the database itself.

We also discussed some issues and concerns with this algorithm like threads hoarding database connections or security issues, as well as ideas on how to fix them (ex. producer-consumer pattern for the hoarding of database connections). We've shown ways in which ORMs can improve and extend, and discussed some "hidden" behaviors within them. We also gave an alternative to GraphQL for companies which would provide some of its main features, while also drastically improving it in every aspect.

The next steps for our research would be making an official release of this library, improving on its stability, implementing a producer-consumer pattern for the parallel approach, testing it on real-world examples, and finally trying to re-implement it using different technologies and languages with more proper support for parallelization.

References

[1] The GraphQL Foundation. Graphql: A query language for your API. [Online]. Available: https://graphql.org/

[2] A. Ignjatovic. Understanding and fixing n+1 queries. [Online]. Available: https://medium.com/doctolib/understanding-and-fixing-n-1-query-30623109fe89

[3] Join-monster. [Online]. Available: https://join-monster.readthedocs.io/en/latest/

[4] A. Stankowski. (2019) Hibernate cartesian product problem. [Online]. Available: https://allaroundjava.com/hibernate-cartesian-product-problem/

[5] Cossium's entity graph. [Online]. Available: https://github.com/Cosium/spring-data-jpa-entity-graph

[6] MapStruct. Mapstruct: Java bean mappings, the easy way! [Online]. Available: https://mapstruct.org/

[7] C. Boyd, "Data-parallel computing: Data parallelism is a key concept in leveraging the power of today's manycore gpus." *Queue*, vol. 6, no. 2, pp. 30–39, 2008.

[8] F. Katsarou, N. Ntarmos, and P. Triantafillou, "Subgraph querying with parallel use of query rewritings and alternative algorithms," 2017. [Online]. Available: http://eprints.gla.ac.uk/130142/1/130142.pdf

[9] G. Morling, A. Gudian, S. Derksen, F. Hrisafov, and the MapStruct community. Inherit configuration documentation. [Online]. Available: https://mapstruct.org/documentation/stable/reference/html/#mapping-configuration-inheritance

A survey of end-to-end congestion mechanisms in the field of IoT

Dalibor Fonović*, Siniša Sovilj* and Nikola Tanković*

*Juraj Dobrila University of Pula, Pula, Croatia

dalibor.fonovic@unipu.hr

ABSTRACT

The Internet of Things (IoT) is a global network of devices capable of communicating and exchanging data with other devices and systems, mainly via the Internet, using specific communication protocols. The number of such devices that can be connected to the Internet is constantly growing due to progress in various technological areas (such as communication technologies, microelectronic circuits, sensors, embedded systems, smartphones, etc.). This leads to network congestion due to the large amount of data exchanged between devices. In addition, IoT devices are resource-constrained, further exacerbating network congestion. Network congestion leads to additional communication delays, low bandwidth, and waste of computer resources. This is one of the important aspects when some IoT applications exchange critical information (e.g., monitoring the patient's health condition in the application of IoT in smart healthcare). As the number of IoT applications increases, so does the need to modify or introduce new protocols to deal with the problems of adapting to network conditions. In the IoT layered architecture, the transport layer plays an important role in managing the end-to-end connection to the services of the upper application IoT layer. An important function of the transmission layer of the IoT layer is congestion control. This paper presents an overview of related research on the congestion control mechanism of the IoT architecture's transport layer, advantages, disadvantages, and current transport layer problems in IoT applications.

Keywords: Internet of Things; Congestion control; TCP protocol; CoAP protocol; Machine learning.

1. INTRODUCTION

IETF**1** defines the Internet of Things as a system with devices often constrained in communication and computation capabilities, becoming more commonly connected to the Internet or at least to an IP network and to various services built on top of the capabilities these devices jointly provide. Such development is expected to usher in a more machine-to-machine Internet communication with no active human mediation [1].

IoT enables the interconnection of different devices to transfer data over the network. Major applications of IoT are smart homes, smart cities, smart grids, industrial monitoring systems, health monitoring systems, environmental monitoring systems, etc.

As the number of connected IoT devices increases, Internet congestion also increases. Also, as new technologies and new communication networks develop, the complexity of network transmission has also increased. This led to new challenges in the design of network protocols. Congestion control (hereinafter referred to as CC) is one of the fundamental components of computer networks. CC plays an important role when improving the utilization of the network to achieve better IoT application performance. As future networks become increasingly complex, conventional congestion control approaches based on pre-defined rules become increasingly ineffective.

2. OVERVIEW OF THE IOT STACK APPLICATION LAYER PROTOCOLS

In the layered architecture of IoT, the application layer provides various communication protocols and acts as an interface between the application and end users or M2M communication.

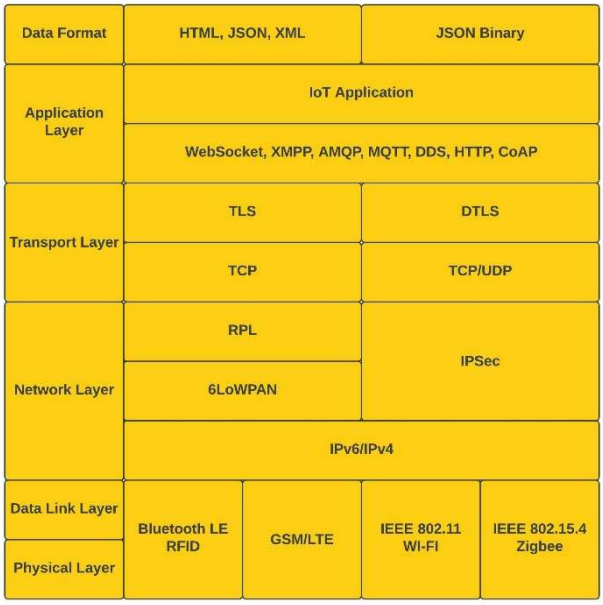


Figure 2.1: Architecture of IoT stack [21]

[1] IETF - The Internet Engineering Task Force, standards development organization for the Internet.

2.1. Request/response protocols

The request-response model is one of the basic two-way communication models on the Internet, where the calling process sends a request for data, and the responding process issues a response to the request. RESTful HTTP and CoAP protocols are often applied in IoT applications as the request/response communication pattern implementation.

2.2. MQTT

MQTT is a message transfer protocol based on publish/subscribe architecture. MQTT v3.1 was adopted by the OASIS consortium in 2013 and certified by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (ISO/IEC) in 2016. It is developed for resource-constrained devices, based on open source, reliable is also a simple protocol. Four years after MQTT 3.1 became the OASIS standard, the MQTT 5 protocol was released, bringing significant improvements and upgrades to the MQTT 3.1 protocol. In March 2019, MQTT 5.0 became the new OASIS standard.

3. REQUIREMENTS FOR NETWORK IN IOT

IoT devices have limited computing capabilities (so-called constrained devices) and limited network bandwidth. In addition, IoT devices most often use wireless networks in which (due to the nature of this type of communication) packet losses often occur (lossy networks) and have limited energy consumption (e.g., battery-powered IoT devices). Also, IoT devices have different requirements regarding communication speed, delay, and reliability according to the application method.

Another important aspect of some IoT applications is exchanging sensitive and critical information. For example, the application of IoT in smart health deals with the transfer of essential and time-critical data, which includes, for example, monitoring the health status of patients where packet loss or delay is not acceptable. That's why the goal is to reduce losses and packet delay.

TCP/IP was not originally designed to run on limited devices. The current congestion control algorithms are primarily designed for the web. They are only sometimes suitable for device-constrained environments, where nodes are limited in CPU and memory. The network is characterized by high packet loss, low bandwidth, or frequent topology change. Thus, deploying and integrating congestion control mechanisms on limited devices creates new challenges.

3.1. Congestion control

Congestion is a condition that occurs in a network when data traffic is so high that it slows down the network's response time. Congestion negatively affects the performance of IoT applications. It leads to the retransmission of packets, thereby increasing energy consumption, delay, and packet loss while reducing bandwidth and packet delivery ratio (PDR). Therefore, congestion control plays an important role in meeting the performance requirements of network communication.

Congestion control is a technique that controls the rate of sent data packets to optimize the use of network infrastructure. It is designed to avoid the degradation of network performance caused by congestion. A simplified congestion control principle is that when a packet loss is detected, the sender reduces its sending rate. Most IoT technologies use wireless (radio) connections. Wireless radio links generally have a higher transmission error rate than wired links due to their susceptibility to interference. They generally have a lower transmission speed.

An example of how congestion management can affect an IoT application's performance is a camera forest fire detection system with built-in machine learning (computer vision) models. Regardless of successful detection, sending a picture of the fire itself is still necessary. The limitation of such a system can be the communication connection (for example, we use NB-IoT technology, characterized by low bandwidth and the limited amount of traffic that the telecommunications operator allows for this type of communication). In this case, optimizing congestion management would lead to improved system performance. An example of such a system is described in the paper [2].

4. END TO END CONGESTION CONTROL IN THE IOT AREA

Traditional CC algorithms for congestion control (CC in further text) can be categorized into end-to-end CC and network-assisted CC. End-to-end approaches require only the cooperation of the sender and receiver. This category does not rely on explicit congestion information obtained from the network. Conversely, network-assisted approaches require congestion information from network devices (like routers).

Traditional end-to-end CC algorithms implemented so far rely on predefined rules. Based on these rules, the congestion window variable (CWND) and the retransmission timeout (RTO) - which determines how long the transport engine waits for an acknowledgment - are recalculated continuously. Depending on the congestion indicator, they can be classified as loss-based, delay-based, and hybrid (including both congestion indicator mechanisms). Many algorithms of one of these types have already been developed and implemented. However, modern networks are complex, diverse, and heterogeneous in their structure. Applying strict rules in a multi-parameter and dynamic environment does not perform equally well in different network conditions; rules that work nearly optimally in one environment may not work well in other environments or when network conditions change rapidly.

Considering the problem space and the variety and variability of the input parameters, new approaches based on machine learning have recently attracted much research attention. One such study was conducted by Wei et al. [3]. Unlike traditional CC algorithms, ML-based patterns rely on real-time network states to make decisions instead of predetermined rules.

Transmission Control Protocol (TCP) is the fundamental protocol of the Internet. However, IoT applications use specific communication patterns that TCP cannot effectively support. For example, some IoT devices may go into sleep mode, which causes the connection to terminate. In such cases, it is impossible to keep the connection active (keep-alive), which can be one of the

requirements in IoT applications (e.g., using the MQTT protocol). Due to situations like the above, an interest in building new reliable transport protocols on top of the unreliable UDP transport protocol has risen.

4.1. Congestion control mechanisms

Congestion management is based on the control of the size of the sending window (window size). The mechanism adjusts the window size to accommodate congestion. The window size can be considered the number of segments sent to the network, representing the number of segments and ACK[2] in transit or queued.

4.1.1. Phases of congestion control

Slow start: at the beginning, we start sending segments slowly, then exponentially increasing the speed of sending data until the congestion window (CWND) value reaches the threshold value (*ssthresh*).

Congestion Avoidance: After reaching the threshold, the sender linearly increases the size of the congestion window to avoid congestion. Each time an acknowledgment is received, the sender increases the congestion window size by 1.

Congestion Detection: there are two methods by which packet losses can be identified over the communication path: (1) by timeout (RTO) and (2) by receiving duplicate acknowledgments (*dupACK*).

4.2. Packet loss approach

When a sender has not received an ACK for a sent packet for a certain period, this usually indicates packet loss. The packet loss-based approach adjusts the sending rate when packet loss occurs. An example of an algorithm based on packet loss is TCP Vegas [4].

4.3. Delay-based approach

The delay-based approach predicts and reacts to congestion before packet loss occurs. This mechanism relies on detected network-induced transmission delays. An example of a delay-based algorithm is *TCP NewReno* [5]. Compared to loss-based approaches, delay-based approaches are more suitable for fast networks. Calculating the exact transmission delay is still a challenge. For example, a small change in packet processing time at the host can cause deviations in the measured transmission delay, which causes wrong decisions when sending. Hybrid approaches have been proposed to take advantage of both loss and delay approaches. TCP Copa is one such example [6].

4.4. ML-based approach

The complexity of today's network architectures is a significant challenge for CC. Designing a generic CC mechanism that will work on all network scenarios is challenging. The dynamic nature of even the same network can make CC operation incorrect.

Because of that, ML-based CC algorithms have been proposed to solve the above-mentioned problems. Traditional CC algorithms rely on using predetermined rules. ML-based mechanisms rely on real-time network states to make congestion control decisions. This allows better adoption of CC mechanisms to dynamic and complex network scenarios.

5. CC MECHANISMS IN THE IOT

5.1. Traditional CC mechanisms

RTO estimation. Retransmission timeout (RTO) determines how long the transport mechanism waits for confirmation (ACK) of the sent segment. The segment is considered lost if confirmation is not received within this time. An important part of the RTO calculation is determining how long it takes for the segment to go to the receiver and for the ACK to return from the receiver to the sender. This is called RTT or Round-Trip Time.

TCP uses the RTO (Retransmission Timeout) estimation algorithm defined in RFC6298 [11]. TCP adaptively determines the RTO by applying the EWMA (exponentially weighted moving average) algorithm to RTT samples.

CC mechanisms based on TCP protocol. Verma [7] et al made CC adapted to IoT. They made a comparative study comparing the new mechanism with other TCP CC mechanisms, such as TCP Cubic, TCP New Reno, and TCP BBR, simulating the IoT environment (devices with limited resources and limited network bandwidth).

The RTO estimation algorithm defined in RFC6298 is not designed to consider IoT environment scenarios. In research [8], the expected RTO algorithm defined in [9] outperformed state-of-the-art algorithms designed to enhance RFC 6298 for TCP in terms of PDR.

CoAP-based CC mechanisms. CoAP is a simple protocol, and the basic specification offers a default CoAP CC mechanism that uses an RTO with binary exponential offset (BEB) for lost packets. Since lost packets are retransmitted in an exponentially increasing time, the default CoAP CC is very simple and insensitive to dynamic network conditions. Consequently, the default CoAP CC has lower performance.

Due to UDP communication, using CoAP with existing network infrastructures (e.g., the use of NAT using firewalls) may lead to certain limitations. This is one of the reasons why there is also a CoAP over TCP specification [14].

To improve the performance of the default CoAP CC, CoAP Simple Congestion Control/Advanced CoCoA is proposed. CoCoA uses round trip time (RTT) measurements and adaptive RTO. This algorithm responds to congestion with a lower sending rate.

[2] ACK - short for "acknowledgement.". In TCP protocol an ACK packet is any packet that acknowledges receiving a sent packet.

Table 5.1: REVIEW OF RECENT RESEARCH ARTICLES ON TCP BASED CC MECHANISMS

Article	Protocol	Description
[10]	TCP Cubic	One of the most common used TCP CC mechanisms. It is supported by most of today OS.
RFC 6582[11]	TCP New Reno	Improvement of performances compared to TCP Reno
[8]	TCP uIP with CoCoA CC mechanism	Performance improvement in the IoT environment compared to RFC6298
Verma [7]	IoT based congestion control algorithm	A comparative study where the mentioned new TCP mechanism is compared in an IoT environment with other known TCP CC mechanisms. Showed the best performance, except in RTT where only TCP Cubic is better

Table 5.2: REVIEW OF RECENT RESEARCH ARTICLES ON CoAP BASED CC MECHANISMS

Article	Protocol	Description	Advantages	Disadvantages
[10]	CoCoA+	Uses RTT for RTO estimation and VBF backoff	Reduction of retransmission compared to the default CoCo model	Inconsistency in expected RTO
[12]	CoCoA++	Uses CAIA Delay Gradient for RTO estimation and PBF backoff	Low RTO with minimal latency and high packet transfer rate in various IoT scenarios	Increased end-to-end latency and computationally intensive
[13]	pCoCoA	Modification of the RTO estimation compared to the CoCoA+ algorithm	Improvement of congestion control (reduced packet loss and thus retransmission)	Increased end-to-end latency
[14]	CoAP over TCP	CoAP over TCP	Enables the use of TLS and WebSocket protocols	Weaker performance

Table 5.3: REVIEW OF RECENT RESEARCH ARTICLES ON ML BASED CC MECHANISMS

Article	Protocol	ML method	Advantages	Disadvantages
Xiao [15]	TCP - Drinc	Reinforcement learning	Increasing performance in complex and dynamic network environments	
Falahatraftar[17]	Predikcija zagušenja GRNN	GRNN model compared to SVM, decision tree, regression models	The GRNN model shows higher accuracy, reliability and stability among the forecasting methods considered	
Demir[16]	mlCoCoA	Uses SVM for dynamic RTO calculation	Performance improvement over CoCoA	Computationally intensive
Sander[18]	DeePCCI	RNN model	Performance improvement over traditional TCP CC mechanisms	
Najm[19]	C4.5 DT	Using a decision tree model for congestion control in a 5G IoT environment	Significant performance improvement over other ML machine learning algorithms	

5.2. CC mechanisms based on ML algorithms.

Xiao et al [15], proposed TCP-Drinc for smart congestion management for TCP variants. The authors use Deep Reinforcement Learning. Congestion is controlled by adjusting the window size. TCP-Drinc is compared against different versions of TCP: TCP-New Reno, TCPCubic, TCP-Hybla, TCP-Vegas and TCP-Illinois. TCP-Drinc provides maximum throughput and the second lowest round trip time (RTT).

Demir et al [16] proposed mlCoCoA, ML-based improvement in CoCoA, mlCoCoA adaptively sets CoCoA retransmission timeout (RTO) estimation parameters using an ML method.

Sander, et al [18] proposed DeePCCI for congestion detection based on packet arrival time in traffic flow.

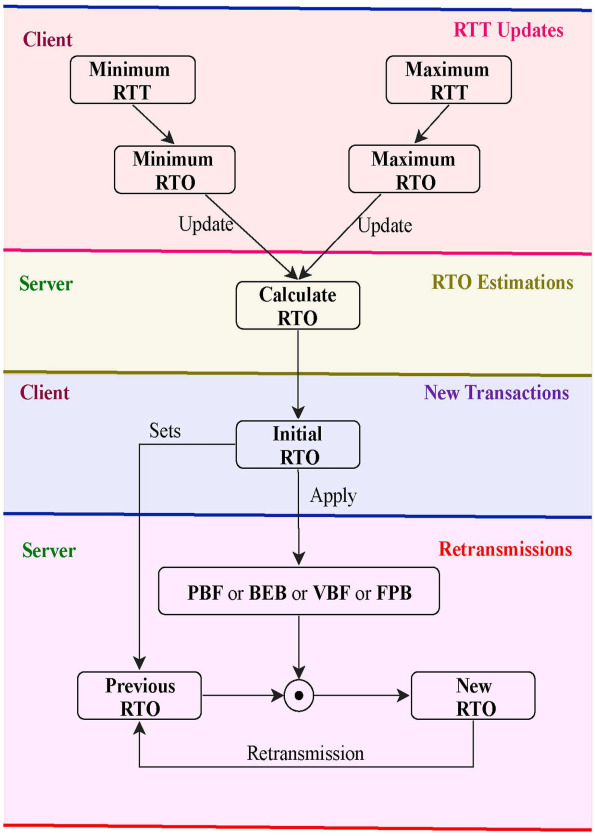


Figure 5.1: Example of RTO estimation strategy for CoAP and its improvements [21].

6. Conclusion

Network communication speed and reliability with low resource utilization in resource-constrained devices are key requirements for IoT.

Congestion negatively affects performance. Congestion leads to the retransmission of packets, thereby increasing energy consumption, delay, and packet loss while reducing the throughput and packet delivery ratio (PDR). This can have negative consequences for the end IoT application. Congestion control should be an important issue when working in IoT networks. Traditional end-to-end CC algorithms implemented so far rely on predefined rules.

There are two main ways to control congestion. In the first case, the application layer uses the CC mechanism of the TCP protocol. Examples are the HTTP and MQTT protocols that build on the TCP protocol for congestion control. The CC mechanism is implemented in the application layer in the second case. An example is the CoAP protocol which uses the CC mechanism in the application layer of the IoT model and is based on the UDP protocol.

In this paper, a literature review of existing CC mechanisms was made with an emphasis on the IoT field of application. A large number of research deal with improving the CC mechanism of the TCP protocol. There are also many works where machine learning has been used to improve the performance of the TCP CC mechanism. However, only a small amount of research is focused on applying TCP congestion control mechanisms to the IoT application area. As for the CoAP protocol, many works are investigating the improvements of the CC mechanism. A smaller number of studies use the application of machine learning for the CC mechanism of the CoAP protocol.

Acknowledgments

This paper acknowledges the support of the Croatian Science Foundation research project CSF no. HRZZ-IP-2019-04-4216 “Reliability and Safety in Complex Software Systems: From Empirical Principles towards Theoretical Models in View of Industrial Applications (RELYSOFT)” and the support of the Erasmus+ Key Action 2 (Strategic partnership for higher education) project No. 2020–1–PT01–KA203–078646: “SusTrainable - Promoting Sustainability as a Fundamental Driver in Software Development Training and Education”.

This paper has also been supported by European Regional Development Fund (ERDF) project SUNSET IoT under grant agreement number KK.01.2.1.02.0116, by Croatian Science Foundation under the project UIP-201705-9066.

References

[1] The internet of things at the IETF. Available at: https://www.ietf.org/topics/iot/ (accessed: January 22, 2023).

[2] G. Peruzzi, A. Pozzebon, and M. Van Der Meer, “Fight fire with fire: Detecting forest fires with embedded machine learning models dealing with audio and images on low power IoT devices,” Sensors, vol. 23, no. 2, 2023, doi: 10.3390/s23020783.

[3] W. Wei, H. Gu, and B. Li, “Congestion control: A renaissance with machine learning,” IEEE Network, vol. 35, no. 4, pp. 262–269, 2021, doi: 10.1109/MNET.011.2000603.

[4] L. S. Brakmo and L. L. Peterson, “TCP vegas: End to end congestion avoidance on a global internet,” IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp. 1465–1480, 1995, doi: 10.1109/49.464716.

[5] A. Gurtov, T. Henderson, S. Floyd, and Y. Nishida, “The NewReno Modification to TCP’s Fast Recovery Algorithm.” RFC 6582; RFC Editor, Apr. 2012. doi: 10.17487/RFC6582.

[6] V. Arun and H. Balakrishnan, “Copa: Practical delay-based congestion control for the internet ,” in ANRW ’18:Proceedings of the Applied Networking Research Workshop, Jul. 2018, pp. 19–19. doi: 10.1145/3232755.3232783.

[7] L. P. Verma and M. Kumar, “An IoT based congestion control algorithm,” Internet of Things, vol. 9, p. 100157, 2020, doi: https://doi.org/10.1016/j.iot.2019.100157.

[8] C. Lim, “Improving congestion control of TCP for constrained IoT networks ,” Sensors, vol. 20, no. 17, 2020, doi: 10.3390/s20174774.

[9] C. Bormann, A. Betzler, C. Gomez, and I. Demirkol, “CoAP Simple Congestion Control/Advanced,”Internet Engineering Task Force; Internet Engineering Task Force, Internet-Draft draft-ietf-core-cocoa-03, Feb. 2018. Dostupno: https://datatracker.ietf.org/doc/draft-ietf-core-cocoa/03/

[10] A. Betzler, C. Gomez, I. Demirkol, and J. Paradells, “CoCoA+: An advanced congestion control mechanism for CoAP,” Ad Hoc Networks, Apr. 2015, doi: 10.1016/j.adhoc.2015.04.007.

[11] V. Paxson, M. Allman, J. Chu, and M. Sargent, Computing TCP’s retransmission timer. RFC 6298, 2011.

[12] V. Rathod, N. Jeppu, S. Sastry, S. Singala, and M. P. Tahiliani, “CoCoA++: Delay gradient based congestion control for internet of things,” Future Generation Computer Systems, vol. 100, pp. 1053–1072, 2019, doi: https://doi.org/10.1016/j.future.2019.04.054.

[13] S. Bolettieri, G. Tanganelli, C. Vallati, and E. Mingozzi, “pCoCoA: A precise congestion control algorithm for CoAP,” Ad Hoc Networks, vol. 80, pp. 116–129, 2018, doi: https://doi.org/10.1016/j.adhoc.2018.06.015.

[14] C. Bormann, S. Lemay, H. Tschofenig, K. Hartke, B. Silverajan, and B. Raymor, “CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets.” RFC 8323; RFC Editor, Feb. 2018. doi: 10.17487/RFC8323.

[15] K. Xiao, S. Mao, and J. K. Tugnait, “TCP-drinc: Smart congestion control based on deep reinforcement learning,” IEEE Access, vol. 7, pp. 11892–11904, 2019, doi: 10.1109/ACCESS.2019.2892046.

[16] A. K. Demir and F. Abut, “mlCoCoA: a machine learning-based congestion control for CoAP,” TurkishJournal of Electrical Engineeringand Computer Sciences, vol. 28, no. 5, 2020, doi: 10.3906/ELK-2003-17.

[17] F. Falahatraftar, S. Pierre, and S. Chamberland, “An intelligent congestion avoidance mechanism based on generalized regression neural network for heterogeneous vehicular networks,” IEEE Transactions on Intelligent Vehicles, pp. 1–13, 2022, doi: 10.1109/TIV.2022.3180665.

[18] C. Sander, J. Rüth, O. Hohlfeld, and K. Wehrle, “DeePCCI: Deep learning-based passive congestion control identification,” in Proceedings of the 2019 workshop on network meets AI & ML, 2019, pp. 37–43. doi: 10.1145/3341216.3342211.

[19] I. A. Najm, A. K. Hamoud, J. Lloret, and I. Bosch, “Machine learning prediction approach to enhance congestion control in 5G IoT environment,” Electronics, vol. 8, no. 6, 2019, doi: 10.3390/electronics8060607.

[20] R. Al-Saadi, G. Armitage, J. But and P. Branch, "A Survey of Delay-Based and Hybrid TCP Congestion Control Algorithms," in IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3609-3638, Fourthquarter 2019, doi: 10.1109/COMST.2019.2904994.

[21] C. Bayılmış, M. A. Ebleme, Ü. Çavuşoğlu, K. Küçük, and A. Sevin, “A survey on communication protocols and performance evaluations for internet of things,” Digital Communications and Networks, vol. 8, no. 6, pp. 1094–1104, 2022, doi: https://doi.org/10.1016/j.dcan.2022.03.013.

Security and Privacy Concerns Associated with the Internet of Things(IoT) and the Role of Adapting Blockchain and Machine Learning - A Systematic Literature Review

Elva Leka[1,2], Enkeleda Hoxha[2], Genta Rexha[1,2]
[1]Polytechnic University of Tirana, Tirane, Albania
[2]Albanian University, Tirane, Albania
elva.leka@fgjm.edu.al; enkeleda.hoxha@student.albanianuniversity.edu.al; genta.rexha@albanianuniversity.edu.al

***Abstract* - The Internet of Things (IoT) is one of the most important technologies of our day, and it will continue to become more significant as more businesses realize the competitive advantages of having connected devices. Blockchain helps maintain correct data that is unaltered and permanent. At the same time, the Internet of Things (IoT), addresses hardware connected via the internet and may use machine learning to discover trends and produce accurate projections utilizing this data. Those three technologies have taken the place of the conventional methods that were previously employed for the design and architecture of new devices. Combining them yields a wide range of potential outcomes, some of which might result in the next great thing. We also contend that this convergence will speed up the development of autonomous models and information technology that aids companies in managing the security and privacy of their clients. We highlight and expose several challenges and upcoming research areas using machine learning algorithms and blockchain techniques to address security and privacy issues in the IoT space. In this paper, we thoroughly analyze the use of blockchain and machine learning technologies for IoT, focusing on security and privacy. Critical thoughtful writing has been heightened by examining research articles published in well-regarded magazines between 2018 and 2022. Finally, there are still challenges that can steer academics toward potential future advancements in IoT security and privacy.**

***Keywords* - Internet of Things; blockchain; machine learning; technologies; security; privacy**

I. INTRODUCTION

The internet of things (IoT), blockchain technology, and machine learning are now acknowledged as technologies that can enhance present procedures, develop innovative solutions, and disrupt numerous industries. A network of physical "things" that use software, sensors, and other technologies to connect and exchange data with other devices and systems through the internet is referred to as the "Internet of Things" [1]. IoT, which has brought about many technological advancements including voice assistants, smart homes, cars, and healthcare, has radically changed the way we live [2-4]. The recent rise of digital currency has sparked interest in blockchain, the technology that underpins it. Its advantages include decentralization, a distributed ledger, security, and information transparency [5].

Blockchain functions essentially as a list of all transactions made throughout time, much like a perpetual record book [6]. Blockchain primarily serves as a record of all prior transactions. Each participant's computer stores every piece of information as a long chain of linked data items, new items can only be added with the approval of the vast majority of users [7].

By utilizing complex algorithms to analyze vast amounts of data, machine learning can assist in demystifying the underlying patterns in IoT data. Automated systems that use statistically determined actions can supplement or entirely replace manual processes in critical activities [8]. We believe that ingesting images, video, and audio from data from IoT devices recorded on a decentralized ledger can aid machine learning in making predictions, identifying anomalies, and improving intelligence. Additionally, costs for data management, data transfer, and other essential services are decreased by assisting in the protection of data on IoT devices.

The research shows that using Blockchain to trade property between two devices, including smart contracts, offers a straightforward infrastructure [9]. Without the requirement for a centralized authority, autonomous device operation is possible with smart contracts. Blockchain can be used for interactions between people and platforms or objects in addition to IoT. On the other hand, millions of gadgets connected to the Internet of Things produce enormous amounts of data [10]. Machine learning, which extracts information from data, is fed by data.

The paper is organized as follows. In Section II, we give a background of The Internet of Things (IoT) and describe the security and privacy of IoT, further the Role of Adapting Blockchain and Machine Learning. In Section III we evaluate the Systematic Literature Review (SLR). Section IV presents the results of the SLR, and the final section describes the conclusions and future works.

II. BACKGROUND

The Internet of Things (IoT) is a network of physical devices, vehicles, home appliances, and other items that are embedded with sensors, software, and network connectivity, enabling these objects to collect and exchange data [11-12]. The IoT refers to the idea that everyday objects can be connected to the internet and can communicate with each other, without the need for human intervention. The impact of IoT can be seen in various aspects of daily life, including improved efficiency, convenience, and safety. For example, IoT-enabled devices can be used to monitor and control home appliances remotely, track the health and well-being of individuals, and optimize the use of energy in buildings [13]. IoT is also revolutionizing the way businesses operate by providing real-time data, improving decision-making, and enabling new business models. IoT is an important and rapidly growing field that is transforming the way we live, work, and interact with the world around us [14].

The security and privacy of IoT systems are critical concerns that must be addressed to ensure their widespread adoption and long-term viability. IoT devices generate and store large amounts of sensitive personal and commercial data, making them vulnerable to data breaches, hacking, and unauthorized access [15]. Some of the most common security threats to IoT systems include:

- *Data breaches*: IoT devices often collect and store sensitive personal information, making them a target for cybercriminals seeking to steal or misuse this information [16].
- *Hacking*: IoT devices can be easily hacked, especially if they are poorly secured or have outdated software. Hacked devices can be used to launch attacks on other systems, steal sensitive information, or control connected devices.
- *Unauthorized access*: IoT devices can be vulnerable to unauthorized access, which can occur when attackers gain access to devices through a weak password or by exploiting vulnerabilities in the system [17].
- *Malware*: IoT devices can be infected with malware, which can spread throughout the system and cause significant damage [18].

Securing large-scale IoT systems can be challenging, as many devices may have limited processing power, memory, and storage, making it difficult to implement traditional security solutions [19]. The need for effective and efficient security solutions is critical to ensure the security and privacy of IoT systems and the sensitive information they store and process. To address these security and privacy concerns, organizations must adopt a multi-layered approach to security, including the use of encryption, secure authentication, and regular software updates [20]. Additionally, the development of new security solutions, such as blockchain and machine learning, has the potential to enhance the security and privacy of IoT systems.

Blockchain technology has the potential to address some of the security and privacy issues associated with IoT [21]. Blockchain is a decentralized, distributed ledger that allows multiple parties to access and verify information without the need for a central authority [22]. This decentralized structure provides a number of benefits that could enhance the security and privacy of IoT systems, including:

- *Decentralization:* Blockchain technology eliminates the need for a central authority to control and manage data, reducing the risk of single points of failure and making it more difficult for attackers to compromise the system [23].
- *Immutability:* The data stored in a blockchain is tamper-proof, meaning that it cannot be altered or deleted once data is recorded on the blockchain. This feature provides a high level of trust in the data recorded on the blockchain, making it more difficult for attackers to manipulate or corrupt the information [24].
- *Transparency:* Blockchain technology provides a transparent and verifiable record of all transactions and data stored on the blockchain. This transparency helps to build trust and increases accountability, reducing the risk of data breaches or unauthorized access [25].
- *Smart Contracts:* Blockchains can support smart contracts, which are self-executing contracts with the terms of the agreement between buyer and seller being directly written into lines of code. This can increase the security and privacy of IoT systems, as smart contracts can be used to automate the enforcement of security and privacy policies [26].

The use of blockchain technology in IoT has the potential to enhance the security and privacy of IoT systems by providing a decentralized, tamper-proof, transparent, and verifiable method of storing and managing data [27]. While there is still much research to be done in this area, the potential benefits of using blockchain in IoT are significant and could play a major role in addressing security and privacy concerns in IoT.

Machine learning algorithms have the potential to improve the security and privacy of IoT systems by providing real-time threat detection and response capabilities. Machine learning algorithms can learn from historical data and identify patterns that could indicate a security threat, such as unusual traffic patterns or attempts to access sensitive information [28]. This allows machine learning algorithms to detect and respond to threats in real time, reducing the risk of data breaches or unauthorized access to sensitive information [29].

Some of the benefits of using machine learning in IoT include:

- *Real-time Threat Detection*: Machine learning algorithms can detect threats in real-time, allowing for rapid response and mitigation of security incidents [30].

- *Predictive Analysis:* Machine learning algorithms can analyze historical data and identify patterns that could indicate a security threat. This allows for proactive and predictive security measures to be put in place, reducing the risk of data breaches or unauthorized access [31].

Machine learning algorithms in IoT have the potential to real-time the security and privacy of IoT systems by providing real-time threat detection and response capabilities, predictive analysis, personalized security measures, and improved accuracy [32]. While there are still challenges to be addressed in this area, such as privacy concerns and the need for large amounts of data to train machine learning algorithms, the potential benefits of using machine learning in IoT are significant and could play a major role in improving the security and privacy of IoT systems.

III. METHODOLOGY OF RESEARCH

In light of the quick advancements in science, an SLR locates and evaluates the pertinent research to respond to a developed research issue. For the researchers tackling that problem domain, the findings are an SLR that outlines potential solution options. By categorizing the papers and attempting to offer an overview of the study to detect deficiencies, SLR aids us in providing a visual summary and a map of outcomes.

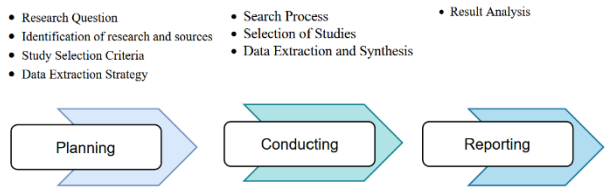


Figure 1. The Systematic Literature Review process

A. The Review's Planning

This analysis focuses on the benefits of integrating blockchain and machine learning technology into IoT devices and how they affect security and privacy. In this paper, the following research questions are especially addressed:

- **RQ1.** Could IoT devices potentially include blockchain and machine learning to further the technology?
- **RQ2.** How can we boost security by integrating blockchain and machine learning technologies into IoT devices?
- **RQ3.** What will be the effect on the cost of IoT security by combining other technologies?
- **RQ4.** Is the vulnerability assessment of IoT systems using blockchain and machine learning effective in identifying flaws?

B. Studying Techniques

A broad understanding of the literature was necessary for a thorough analysis. The relevant databases were rigorously reviewed to ensure that the information provided here is complete. We tracked down and collated relevant content. Not every exceptional work of literature met the search parameters for a variety of reasons.

TABLE I. DATABASE RESEARCH

Coverage	*Search terms*
Peer-reviewed journals, conferences, Articles	("IoT security, privacy " AND "blockchain")("IoT security, privacy " AND "machine learning")
Peer-reviewed journals, conferences, Articles	("IoT security challenges" AND "blockchain") OR ("Blockchain security" AND "machine learning privacy")
Peer-reviewed journals, conferences, Articles	("IoT privacy concerns" OR "blockchain for IoT security")
Peer-reviewed journals, conferences, Articles	("IoT privacy" AND "machine learning") ("IoT security challenges" OR "blockchain")
Peer-reviewed journals, conferences, Articles	("IoT privacy concerns" AND "blockchain for IoT security")

TABLE II. PAPER SELECTION RESULTS

Database	*Number of records retrieved*	*Number of studies selected*
Scopus	233	11
IEEE Xplore	324	10
ACM Digital Library	52	5
Science Direct	251	10
Google Scholar	124	4
Total	984	40

To accomplish this goal, a comprehensive literature search was carried out. We found 40 of the 984 papers from 5 databases to be interesting.

C. Primary Studies Chosen

The goal of the selection process is to find a group of studies that are relevant to the topic of the systematic literature review. By carefully evaluating primary publications, we seek to ensure that the review provides a strong and trustworthy interpretation of the research issue. The search string formulation was influenced by the study domain and themes, and the right content was discovered and gathered as a result.

We created the inclusion and exclusion criteria in order to perform selection on the collected material.

Inclusion Criteria:

- Peer-reviewed articles and conference papers that were published between 2018 and 2022.
- Studies that focus on the security and privacy challenges associated with the IoT.

Inclusion Criteria:

- Peer-reviewed articles and conference papers that were published between 2018 and 2022.
- Studies that focus on the security and privacy challenges associated with the IoT.
- Studies that discuss the role of blockchain and machine learning in enhancing the security and privacy of IoT systems.
- Empirical studies that provide quantitative and qualitative evidence to support their findings.

Exclusion Criteria:

- Grey literature such as thesis, dissertations, and technical reports.
- Non-English language Peer-reviewed, Journals, Conferences, and Articles.
- Studies that only focus on the technology and implementation of the IoT without addressing security and privacy concerns.
- Studies that only focus on the blockchain or machine learning without considering their application in the IoT context.

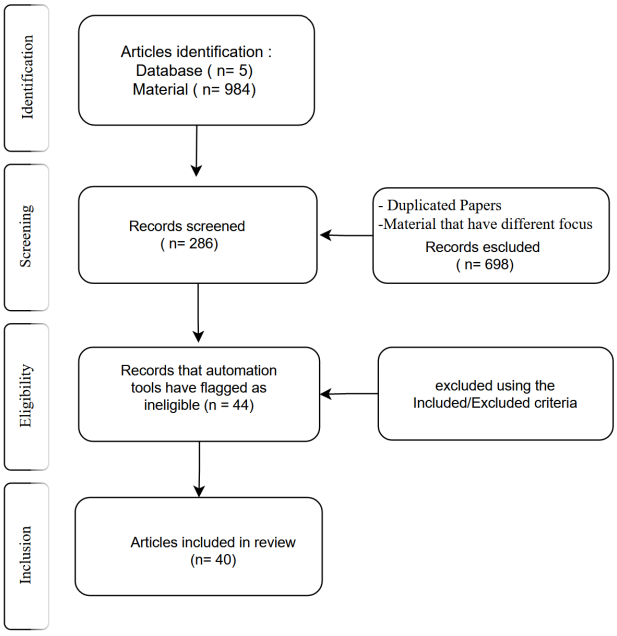


Figure 2. Diagram illustrating the research process [33]

IV. RESULTS

The number of articles generated in each year from 2018 through 2022 is shown in Figure 3. Over the past five years, the idea of incorporating blockchain technology and machines into the Internet of Things (IoT) and addressing concerns about security and privacy has developed and grown.

Table 4 below lists the words used most frequently by the authors. It should come as no surprise that "Blockchain", "Machine Learning", "Internet of Things", " Cryptography", and "Security" were the most often used terms.

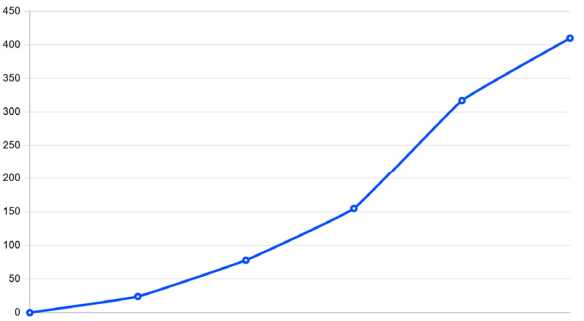


Figure 3. The number of articles published between 2018 and 2022

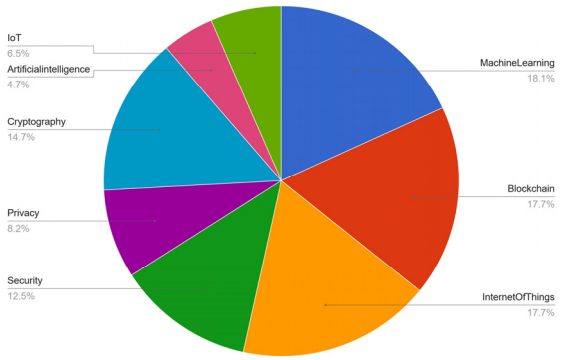


Figure 4. Frequently used subject words

The findings of our SLR with reference to each study question are presented ongoing.

RQ1. Could IoT devices potentially include blockchain and machine learning to further the technology?

To further improve the functionality and security of the IoT ecosystem, IoT devices may use blockchain and machine learning technology. Blockchain technology can provide a secure and decentralized platform for storing and managing the vast amounts of data generated by IoT devices, as stated by the authors of the article [6] [23]. We can assist in preventing data breaches and illegal access to sensitive information in this way. The massive amounts of data produced by IoT devices may be analyzed using machine learning algorithms to provide insights, discover patterns, and make predictions. This can lead to improved decision-making and automation in various industries and applications [28-29].

Incorporating these technologies presents new challenges, such as scalability, privacy, and security issues. Developers must carefully examine the benefits and drawbacks before integrating these technologies into IoT devices.

RQ2. How can we boost security by integrating blockchain and machine learning technologies into IoT devices?

Integrating blockchain and machine learning technologies into IoT devices can boost security in several ways, which have been analyzed in [34-37]:

- *Blockchain for data security:* Blockchain can provide a secure and decentralized ledger for storing and transmitting the vast amounts of data generated by IoT devices. The decentralized nature of blockchain ensures that the data is not controlled by any single entity, reducing the risk of data breaches or unauthorized access.
- *Blockchain for device authentication:* Blockchain can be used to authenticate IoT devices, ensuring that only authorized devices can access the network and the data stored on it. This helps prevent malicious actors from accessing and manipulating the network or its data.
- *Machine learning for threat detection:* Machine learning algorithms can be used to analyze the data generated by IoT devices in real-time, detecting anomalies and potential threats. This can improve the speed and accuracy of detecting and responding to security threats.
- *Machine learning for access control:* Machine learning algorithms can be used to control access to the network and its data, based on the behavior and characteristics of the devices and users. This can help prevent unauthorized access and increase the security of the network.

It is important to note that while integrating blockchain and machine learning technologies into IoT devices can significantly boost security, these solutions must be implemented with caution and proper security measures must be put in place to prevent new security risks from emerging.

RQ3. What benefits do ML and blockchain provide for the IoT privacy framework?

Machine learning (ML) and Blockchain technologies can provide several benefits for an IoT privacy framework:

- *Data protection:* ML algorithms can help in identifying and detecting unusual behaviour in the data, while blockchain can provide an immutable ledger of data transactions, thus making it difficult for unauthorized users to access the data [38].
- *Decentralization:* Blockchain allows for decentralized data storage, reducing the risk of single-point failure and enabling secure data sharing between multiple parties [23].
- *Trust and transparency*: The transparency and immutability of blockchain transactions can help to build trust between different parties involved in the IoT ecosystem, including consumers, manufacturers, and service providers [39].
- *Improved security:* ML algorithms can help in identifying security threats in real time and prevent them as analyzed in [40], while blockchain can provide an additional layer of security through cryptographic techniques [41].
- *Compliance:* The use of blockchain and ML can help organizations meet privacy regulations such as GDPR, as they enable organizations to securely manage and track personal data [42].

RQ4. Is the vulnerability assessment of IoT systems using blockchain and machine learning effective in identifying flaws?

Although it is crucial to keep in mind that no technology is impenetrable and all systems have some level of vulnerabilities, the usage of blockchain and machine learning (ML) for the security assessment of IoT systems can be successful in discovering faults. IoT devices create a lot of data, which ML algorithms may be used to analyze to find patterns and abnormalities that might point to security vulnerabilities. For example, an ML algorithm may be able to detect unusual network traffic patterns or identify unauthorized devices on a network. Similarly, blockchain can provide a secure and transparent platform for tracking and verifying data transactions, making it easier to identify and prevent unauthorized access to sensitive information, as the authors have highlighted in their paper [43]. The decentralized nature of blockchain also makes it more difficult for malicious actors to manipulate the system, which can help to increase the overall security of IoT networks.

V. CONCLUSION

This paper performed a Systematic Literature Review to provide an in-depth understanding of the current state of research on this topic. The review aims to identify the primary security and privacy issues associated with the IoT, as well as the strategies and technologies used to address these concerns.

We collected 984 papers and selected 40 papers as primary studies. The integration of blockchain and machine learning has the potential to address some of the security and privacy concerns associated with the Internet of Things (IoT). Blockchain provides a secure and transparent ledger of data transactions, which can prevent tampering and unauthorized access to sensitive information and help ensure the security of network-stored data. Machine learning algorithms can analyze large amounts of data generated by IoT devices to identify patterns and anomalies that may indicate security vulnerabilities, further as well as detect unusual network activity or identify devices that deviate from their normal behaviour.

However, there are also limitations and challenges associated with using blockchain and machine learning in IoT systems. These include the need for high computational power and data storage, the potential for data privacy breaches, and the need for a secure and reliable infrastructure. As the IoT continues to evolve and expand, it is important for organizations and individuals to be aware of these security and privacy concerns and to explore the potential benefits and limitations of integrating blockchain and machine learning into their IoT systems. In addition to ensuring the integrity and security of the IoT as a whole, this will also help protect private user information and sensitive data.

REFERENCES

[1] A. Banafa, "16 IoT, AI and Blockchain: Catalysts for Digital Transformation," in Secure and Smart Internet of Things (IoT): Using Blockchain and AI ", River Publishers, 2018, pp.113-120.

[2] A. Miglani, N. Kumar, "Blockchain Management and Machine Learning Adaptation for IoT environment in 5G and beyond Networks: A systematic review, Computer Communications, Volume 178, 2021, Pages 37-63, ISSN 0140-3664, DOI: https://doi.org/10.1016/j.com

[3] W. Choi, J. Kim, S. Lee, E. Park, "Smart home and internet of things: A bibliometric study", in Journal of Cleaner Production, Volume 301, 2021, 126908, ISSN 0959-6526, https://doi.org/10.1016/j.jclepro.2021.126908.

[4] Sh. Nazir, Y. Ali, N. Ullah, I. García-Magariño, "Internet of Things for Healthcare Using Effects of Mobile Computing: A Systematic Literature Review", Wireless Communications and Mobile Computing, vol. 2019, Article ID 5931315, 20 pages, 2019. https://doi.org/10.1155/2019/5931315

[5] M. Banerjee, J. Lee, Kim-Kwang Raymond Choo, "A blockchain future for internet of things security: a position paperDigital Communications and Networks,Volume 4, Issue 3,2018,Pages 149-160,ISSN 2352-8648, https://doi.org/10.1016/j.dcan.2017.10.006

[6] D. Shah, D. Patel, J. Adesara, "Integrating machine learning and blockchain to develop a system to veto the forgeries and provide efficient results in education sector", Vis. Comput. Ind. Biomed. Art 4, 18 (2021). https://doi.org/10.1186/s42492-021-00084-y.

[7] D. Lizcano, J. A. Lara, B. White, "Blockchain-based Approach to Create a Model of Trust in Open and Ubiquitous Higher Education" J Comput High Educ 32, 109–134 (2020). https://doi.org/10.1007/s12528-019-09209-y.

[8] Machine Learning (ML) for IoT, Available: https://www.softwareag.com/en_corporate/resources/what-is/machine-learning.html, (Accessed: Jan. 2023).

[9] G. Schmitt, A. Mladenow, Ch. Strauss, M. Schaffhauser-Linzatti, "Smart Contracts and Internet of Things: A Qualitative Content Analysis using the Technology-Organization-Environment Framework to Identify Key-Determinants", Procedia Computer Science, Volume 160, 2019, pp. 189-196, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.09.460.

[10] A. Reyna, C. Martín, J. Chen, E. Soler, M. Díaz, "On blockchain and its integration with IoT. Challenges and opportunities", Future Generation Computer Systems, Volume 88, 2018.

[11] D. Mocrii, Y. Chen, P. Musilek, "IoT-based Smart Homes: A Review of System Architecture, Software, Communications, Privacy and Security, Internet of Things, Vol. 1–2, 2018, pp. 81-98,ISSN 2542-6605, DOI: 10.1016/j.iot.2018.08.009.

[12] F. Sabrina, N. Li, S, Sohail, "Blockchain Based Secure IoT System Using Device Identity Management". *Sensors (Basel, Switzerland)*. 2022; 22(19):7535, DOI: 10.3390/s22197535.

[13] F. A. Awin, Y. M. Alginahi, E. Abdel-Raheem and K. Tepe, "Technical Issues on Cognitive Radio-Based Internet of Things Systems: A Survey," in *IEEE Access*, vol. 7, pp. 97887-97908, 2019, doi: 10.1109/ACCESS.2019.2929915.

[14] A. Banafa, "8 IoT Standardization and Implementation Challenges," in *Secure and Smart Internet of Things (IoT): Using Blockchain and AI* , River Publishers, 2018, pp. 53-58.

[15] B. K. Mohanta, D. Jena, U. Satapathy, S. Patnaik, "Survey on IoT security: Challenges and solution using machine learning, artificial Intelligence and Blockchain technology, Internet of Things, Volume 11, 2020, 100227, ISSN 2542-6605, DOI: 10.1016/j.iot.2020.100227.

[16] G. Vojković, M. Milenkovic, T. Katulić, "IoT and Smart Home Data Breach Risks from the Perspective of Data Protection and Information Security Law", *Business Systems Research*. 2020;11(3):167-185. doi:10.2478/bsrj-2020-0033.

[17] C. Hahn C, J. Kim, H. Kwon H, J. Hur, "Efficient IoT Management With Resilience to Unauthorized Access to Cloud Storage". *IEEE transactions on cloud computing*. 2022, 10(2):1008-1020. doi:10.1109/TCC.2020.2985046

[18] N. R. Nath, V. H. Nath, "Critical analysis of the layered and systematic approaches for understanding IoT security threats and challenges". Computers & Electrical Engineering. 2022; 100:107997-. doi:10.1016/j.compeleceng.2022.107997.

[19] A. Goudarzi, F. Ghayoor, M. Waseem, S. Fahad, I. Traore "A Survey on IoT-Enabled Smart Grids: Emerging, Applications, Challenges, and Outlook". Energies (Basel). 2022; 15(19):6984-. doi:10.3390/en15196984.

[20] Y. Goh, J. Yun , D. Jung, JM. Chung JM, "Secure Trust-Based Delegated Consensus for Blockchain Frameworks Using Deep Reinforcement Learning", IEEE Access. 2022; 10:118498-118511. doi:10.1109/ACCESS.2022.3220852.

[21] R. Xu and Y. Chen, "μDFL: A Secure Microchained Decentralized Federated Learning Fabric Atop IoT Networks," In IEEE Transactions on Network and Service Management, vol. 19, no. 3, pp. 2677-2688, Sept. 2022, doi: 10.1109/TNSM.2022.3179892.

[22] A. Su. Patil, R. Hamza, A. Hassan, N. Jiang, H. Yan, J. Li, Efficient Privacy-Preserving Authentication Protocol Using PUFs with Blockchain Smart Contracts, Computers & Security, Vol. 97, 2020, 101958, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2020.101958.

[23] M. Alhejaz, R. Mohammad, "Enhancing the Blockchain Voting Process in IoT using a Novel Blockchain Weighted Majority Consensus Algorithm (WMCA)", Information Security Journal. 2022;31(2):125-143. Doi:10.1080/19393555.2020.1869356.

[24] Z. Rahman, X. Yi, I. Khalil I., "Blockchain based AI-enabled Industry 4.0 CPS Protection against Advanced Persistent Threat", IEEE Internet of Things Journal. Published online 2022:1-1. Doi:10.1109/JIOT.2022.3147186.

[25] Y. Goh, J. Yun, D. Jung, J. Chung, "Secure Trust-Based Delegated Consensus for Blockchain Frameworks Using Deep Reinforcement Learning", in IEEE access. 2022; 10:118498-118511. doi:10.1109/ACCESS.2022.3220852.

[26] H. Mrabet, A. Alhomoud, A. Jemai, D. Trentesaux, "A Secured Industrial Internet-of-Things Architecture Based on Blockchain Technology and Machine Learning for Sensor Access Control Systems in Smart Manufacturing". Applied sciences. 2022;12(9):4641-. doi:10.3390/app1209464.

[27] A Dorri, C. Roulin, S. Pal S, S. Baalbaki, R. Jurdak, S. Kanhere "Device Identification in Blockchain-Based Internet of Things". In IEEE Internet of Things Journal. 2022; 9(24): 1-1. doi:10.1109/JIOT.2022.3194671.

[28] F. Hussain, R. Hussain, S. A. Hassan and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1686-1721, thirdquarter 2020, doi: 10.1109/COMST.2020.2986444.

[29] U. Farooq, N. Tariq, M. Asim, T. Baker, A. Al-Shamma'a "Machine Learning and the Internet of Things Security: Solutions and open challenges". In Journal of Parallel and Distributed Computing, Vol. 162, 2022, pp. 89-104, ISSN 0743-7315, Doi: 10.1016/j.jpdc.2022.01.015.

[30] Y. Zhao, G. Cheng, Y. Duan, Z. Gu, Y. Zhu, L. Tang, "Secure IoT edge: Threat situation awareness based on network traffic" *Computer networks (Amsterdam, Netherlands : 1999)*. 2021;201:108525-. doi:10.1016/j.comnet.2021.108525.

[31] J. Cui, L. Wang, X. Zhao, H. Zhang, “Towards Predictive Analysis of Android Vulnerability Using Statistical Codes and Machine Learning for IoT applications. *Computer communications*. 2020; 155, pp. 125-131. doi:10.1016/j.comcom.2020.02.078.

[32] M. Ferrag, O. Friha, L. Maglaras, H. Janicke H, L. Shu, “Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis”, IEEE Access. 2021, pp. 138509-138542. DOI:10.1109/ACCESS.2021.3118642.

[33] M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Online), 372, n71–n71. https://doi.org/10.1136/bmj.n71.

[34] D. Liao, H. Li, W. Wang, X. Wang, M. Zhang, X. Chen, “Achieving IoT data security based blockchain”. In *Peer-to-peer networking and applications*, 2021, 14(5), 2694-2707. doi:10.1007/s12083-020-01042-w.

[35] M. Shen, H. Liu, L. Zhu, “Blockchain-Assisted Secure Device Authentication for Cross-Domain Industrial IoT”. In *IEEE Journal on Selected Areas in Communications*. 2020, 38(5), pp. 942-954. doi:10.1109/JSAC.2020.2980916.

[36] D. Javeed D, U. Badamas, T Iqbal, A. Umar, C. Ndubuisi, “Threat Detection using Machine/Deep Learning in IOT Environments”. *International Journal of Computer Networks and Communications Security*. 2020, 8 (8), pp. 59-65.

[37] A. K. Istiaque, M. Tahir, H. M. Habaebi, S. L. Lun, A. Ahad “Machine Learning for Authentication and Authorization in IoT: Taxonomy, Challenges and Future Research Direction”. *Sensors (Basel, Switzerland)*. 2021;21(15):5122-. doi:10.3390/s21155122.

[38] M. Chanson, A. Bogner, D. Bilgeri, E. Fleisch, F. Wortmann, “Blockchain for the IoT: Privacy-preserving protection of sensor data”. In *Journal of the Association for Information Systems*. 2019, 20(9): 1271-1307, DOI:10.17705/1jais.00567.

[39] W. Liang W, N. Ji, “Privacy challenges of IoT-based blockchain: a Systematic Review”, *Cluster computing*. 2022, vol. 25, issue 3, pp. 2203-2221. doi:10.1007/s10586-021-03260-0

[40] S. Bharati, P. Podder, “Machine and Deep Learning for IoT Security and Privacy: Applications, Challenges, and Future Directions”, *Security and communication networks*. 2022, 2022:1-41. doi:10.1155/2022/8951961.

[41] N Y, P M. Radial kernelized regressive merkle–damgård cryptographic hash blockchain for secure data transmission with IoT sensor node. *Peer-to-peer networking and applications*. 2021;14(4):1998-2010. doi:10.1007/s12083-021-01135-0.

[42] A. Kounoudes, G. Kapitsaki, “A mapping of IoT user-centric privacy preserving approaches to the GDPR”. In *Internet of Things*. 2020;11:100179-. doi:10.1016/j.iot.2020.100179.

[43] C. Nartey C, E. Tchao, J. Gadze JD, “On Blockchain and IoT Integration Platforms: Current Implementation Challenges and Future Perspectives”. In *Wireless communications and mobile computing*. 2021;2021:1-25. doi:10.1155/2021/6672482.

Applying balanced scorecard in software process improvement: a case study of small software organization

Bakir Karahodža*, Elma Avdagić-Golub * and Alem Čolaković*
* University of Sarajevo, Faculty of Traffic and Communications, Sarajevo, Bosnia and Herzegovina
{bakir.karahodza, elma.avdagic, alem.colakovic}@fsk.unsa.ba

Abstract - **Software processes consist of a complex set of activities required to deliver software products within predicted quality, costs, and deadlines. To accomplish such goals, a software organization needs a quality and mature software process as a prerequisite for success. Adopting Software Capability Maturity Model Integration (CMMI) represents a well-known path in the pursuit of mature software processes. However, its implementation is a subject of a permanent effort that implies different approaches and methods, and often leads to unsuccessful or limited success, though. This is especially emphasized in small software companies given the dynamic environment influenced by different factors, including insufficient resources, changes in technology, and staff turnover. In this paper, a case study of a small software company implementing software process improvement is presented. In a tailored approach to process improvement, a specific method using the balanced scorecard as input into the IDEAL model has been designed, enabling a narrow link between business goals and specific improvement goals. The results show that the software process and selected performance indicators were improved, and suggest the potential of the proposed approach in small organizations.**

Keywords - software process improvement, CMMI, implementation, small organization, balanced scorecard, implementation method

I. INTRODUCTION

The relation between the quality of the software product and the quality of the software processes used to develop and maintain software is an ongoing challenge in software companies. One of the industry-accepted approaches for Software Process Improvement (SPI) projects is the capability maturity framework, which transforms an ad hoc process into a mature and documented process capable of producing software of predictable quality [1]. Such frameworks are the Capability Maturity Model (CMM) introduced by the Software Engineering Institute, as well as the later introduced Capability Maturity Model Integration (CMMI) [2] that contains a similar approach, yet slightly changed and tailored for wider types of software companies. The CMM-based frameworks are mostly aimed to be introduced in large software projects and companies [3]. There is a proven record of research describing various risks for success in their implementation in small and minimum companies (SMEs). They tend to have a larger possibility of failing in the improvement projects. There are a few common risks and described causes [4], [5]: high dependability of individuals and a small number of projects, a small number of employees, non-existence of the formal organization structure, small investment opportunities, insufficient resources, staff turnover, etc.

However, by using the CMMI framework and with its individual adjustment to a specific use case tailored to the organization's size and business goals, measurable effects could be reached, though [3]. There have been a few studies using balanced scorecard in SMEs, mostly providing the theoretical basis for the potential benefits of its implementation. They lack details on how to use in SPI projects or need a broader case study for validity. The objective of this research is to propose an extended approach in the initiating phase of the software process improvement project. It combines the best features of the IDEAL model and directly links the strategic business goals to internal business processes**.**

The proposed model also provides a priority-based approach to the selection of process areas using several additional factors, including the complexity of changes, commitment to improvement, and impact on customer satisfaction. Common constraints often presented in small software organizations are further taken into account by tailoring process areas and interpreting specific goals and objectives.

The rest of the paper is organized as follows: Section 2 presents the project background, including a review of previous similar research, along with the description of the software company which is the subject of process improvement, and its previous improvement efforts. In Section 3, the general method of improvement is described, including the role of the balanced scorecard analysis with other methods and defined inputs. The results of the case study are presented in Section 4, where all the necessary steps required by the IDEAL model have been performed. Finally, Section 5 presents the conclusions of the paper.

II. BACKGROUND

A. Software process improvement in small companies

The general guidelines for SPI within the CMM-based models are described by the Software Engineering Institute (SEI) in their IDEAL model [6]. It provides an

approach for continuous process improvement by describing in detail steps needed to initiate and successfully conclude the improvement program [3]: (I)nitiating: Laying the groundwork for the process improvement; (D)iagnosing: Determining the current and desired state of the maturity; (E)stablishing: Planning a specific way to achieve the desired state; (A)cting: Execution according to the previously established plan; (L)earning: Aquiring knowledge and learning from experience. However, as previously noted, the challenge for successful process improvement in SMEs contains additional issues that need to be resolved. In CMMI, there are a large number of process areas with generic and specific goals needed to achieve, having a large number of limitations and challenges [7]. To overcome the problems and limitations, opportunities presented by the specific environment of SMEs are described with emphasis on the relationship between business goals and CMMI efforts [8], [9].

Since its introduction, the Balanced Scorecard methodology has usually been used for strategic management and corporate planning. Later research questioned how it can be applied in small and medium companies given their characteristics [10]. In [11], the CMMI-based SPI is presented, where the general business goals are classified by a balanced scorecard, and later linked to the improvement criteria.

B. *Description of company*

The company described in the case study is positioned in the Bosnia and Herzegovina market with a goal to "develop and maintain business software for enterprise organizations". It is a private company with 25 employees, including managers, project managers, developers, and other employees working on their business goals.

The company previously reached the ISO/IEC 12207 [12] certificate in order to improve the company's positioning in the domestic market, but also to achieve the prerequisites for entering the international market. At the suggestion of an external consulting firm (further noted as SCC), the company's management initiated a procedure for process maturity assessment within the CMMI framework.

The formal improvement program carried out by SCC consisted of assessment activities and recommendations for the continuation of the improvement project.

TABLE I. SCC ASSESSMENT

Key Process Area / Level	**Assessment**
Repeatable (Level 2)	
Requirements management	.75
Project Planning	.5
Project Monitoring and Control	0
Supplier Agreement Management	0
Process and Product Quality Assurance	1
Configuration Management	.5

The SCC delivered its final report based on the CMM levels of maturity. Table I. presents the results of the SCC assessment, including the assessment of individual process areas, on a scale between 0 and 1. At Level 3, only three process areas with positive (0.5) assessments were recognized: Organizational Process Focus, Organizational Process Definition, and Product Integration.

III. METHOD

Based on the previously presented results, the company's management has decided to continue with the improvement activities based on measurable results that should be related to business goals.

As a direct improvement goal, it could now already be declared as satisfying Level 3, taking into account the previous maturity assessment. However, after analyzing the process areas where improvement is needed, the following issues were revealed:

- *A large number of process areas for improvement*: Having six process areas at Level 2 and seven process areas at Level 3, several dozen in total need to be implemented.
- *Lack of a standardized method for improvement*: Although the SEI offers the IDEAL model for process improvement, it's still too broad and doesn't clearly describe how to deal with each specific situation in an organization, especially in a small organization.
- *Insufficient budget for consulting services*: The management was not able to provide, nor expand the budget needed for external consulting services.

Considering these issues in the context of the research, it was necessary to create a new approach for the formal extension of the improvement program. During the analysis of the CMMI Integration model, it was found that it contains similar elements to the ISO/IEC 12207 standard. Additionally, the model provides a continuous presentation that enables the selection of process areas adapted to the direct needs of the company's business goals, without the need to satisfy all process areas before moving to a higher level of maturity.

The second question was related to the selection of process areas for improvement. The idea was to use the previous assessment given by the SCC, and additionally to align the selection and the overall improvement process with company business goals. As business goals are a subject of frequent changes in small organizations, it is necessary to first specify the business goals in more detail. The technique used for this purpose is the Balanced Scorecard table, which has been widely used and developed in practice within improvement programs.

A summary of the proposed method that combines several techniques along with the necessary inputs can be described in the following, as presented in Fig. 1:

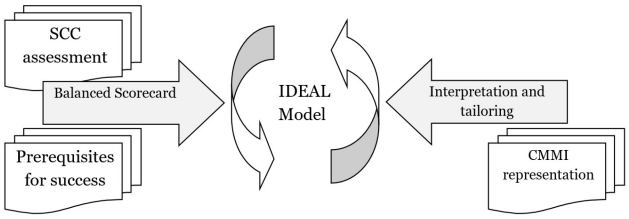


Figure 1. Proposed improvement method

1. The framework for improvement is the CMMI for Services
2. The basic model for the improvement process is the IDEAL model
3. The technique for selection, evaluation, and connection of process areas with business goals is the Balanced Scorecard table
4. Use the following input parameters:
   - Previous assessment results of SCC
   - Previously recognized prerequisites and risks of improvement in SMEs
5. Interpretation and tailoring of the CMMI model for selected business and process areas will be carried out according to SEI recommendations and practices from other cases.

There was also an additional input containing a set of prerequisites for success. This is especially emphasized in SPI projects in SMEs environment, given the previously described specific challenges.

IV. RESULTS

Following the adopted method, the IDEAL model is used to follow up the predefined phases and steps needed to accomplish the SPI project.

A. Initiating

The stimulus for change of the existing business practices is partially explained in previous chapters, and can be specified as the following:

- Changes in the company's business environment and market demand.
- Management's decision to continue with the software processes improvement project.

To establish the context in which changes should be made, the following questions were discussed and analyzed:

- What specific business goals, direct and indirect, should be realized or supported by these changes?
- What are the expected business benefits from the changes and SPI?

TABLE II. BALANCE SCORECARD TABLE

	Objectives	**Measures**	**Targets**
Financia	Increase sales revenue; Increase profitability	Income	Current Income *1.2
		Profit	Current Profit * 1.3
Customer	Acquire new customers; Retain customers; Offer standardized maintenance services	New customers	Current: 2 /year Target: 5 /year
		Customer churn	Current: 10% Target: 0%
		Maintenance services	Current: ~ 150 Target: ~ 30
Internal business process	Frequently add new features; Enable predictable maintenance costs & deadlines; Proactively control of development & maintenance	New major features	Current: 5/application/year Target: 10/application/year
		Average requests' resolution time	Current: 7 working/days Target: 3 working/days
		Coverage of tasks with automation utilities	Current: development tasks Target: support and maintenance tasks
Learning & Growth	Provide training for staff: software maintenances activities, standards & process	Training hours	Current: 0 h Target: 15 h

Given the discussion with the management, a sample of the resulting Balance Scorecard table was prepared and presented to the management. There were also other inputs taken in the preparation: analysis of quarterly and annual business and financial reports, and broad market analysis. The results are presented in Table II. Along with the strategic mapping, the key indicators for measurement and their target values were identified.

The resulting Table II shows that the critical success factors are mostly concentrated around software maintenance and customer support activities when insight is taken from the process perspective. This also leads to the conclusion about the slight, yet obvious shift in the goals taken from the management perspective: from the initial software development business, as stated by the management, towards the support and maintenance activities. The results could be explained by the current customer base: there are 20 customers with about 3,000 end-users, and 15 software applications provided and maintained. The average duration of the current maintenance contracts is 5.4 years.

The main effect of using the Balanced Scorecard is the conclusion that the software maintenance processes are the primary group of areas that should be carried out. The company's management has agreed to participate directly in those parts of the activity where their participation is necessary.

The necessary resources are allocated in terms of the following: one project manager, two senior developers, and external consultants including the authors of this paper. The formal initiation document containing previously defined goals, resource allocation and expected results of the improvement is created.

B. Diagnosing

The most critical part of the program was the Diagnosing phase, where the current and the desired - future state of the software processes should be clearly stated, as well as the development of recommendations. Having the recognized support and maintenance tasks as the pillars of internal processes, the following problems were identified:

- There is no formal department within the company responsible for support and maintenance tasks
- The activities are poorly documented
- The nature of the job is reactive, with customers dictating the entire work and deadlines, including sparse communication.

The recommendations have been completed using interpretation and tailoring of the CMMI model for selected process areas according to SEI recommendations. Recommendations from previous case studies which provide a relevant interpretation of process tailoring in SMEs were also used [9], [13], [14], and can be summarized in the following statements:

- Consider each contracted service level or business maintenance contract as an individual project
- Create maintenance projects within the Service Level Agreement (SLA)
- Enable quality assurance and control for each project.

C. Establishing

The establishing phase is required to develop a detailed work plan. The process areas presented in Table II required their prioritization due to limited resources and organizational goals defined within the Balanced Score table. The priority-based approach was developed using the following weight factors for each process area: impact on business goals, the complexity of changes, commitment to improvement, and impact on customers. The resulting priorities on a scale between 1 and 5 are presented in Table III. The process areas with an average weighting factor >= 3.5 were selected for priority in the improvement.

Before the action plan, the following actions needed for the improvement were specified:

(a) Map all internal maintenance tasks and assignments to selected process areas (bolded in Table III) and integrate them with the existing Workload system
(b) Classification of all maintenance requirements according to the ISO/IEC/IEEE 14764 standard

TABLE III. IMPROVEMENT PRIORITIES

	Process area	Weight factors				
		Impact on business goals	Complexity of change	Commitment for improvement	Impact on customers	Average
1.	**REQM - Requirements Management**	**5**	**3**	**3**	**4**	**3.75**
2.	**PP - Project Planning**	**4**	**4**	**3**	**4**	**3.75**
3.	PMC - Project Monitoring and Control	3	4	2	4	3.25
4.	**MA - Measurement and Analysis**	**4**	**2**	**3**	**5**	**4**
5.	PPQA - Process and Product Quality Assurance	3	3	2	4	3
6.	**CM - Configuration Management**	**5**	**3**	**4**	**3**	**3.75**
7.	**REQM - Requirements Management**	**5**	**4**	**4**	**4**	**4.25**
8.	SAM - Supplier Agreement Management	5	3	3	2	3.25
9.	**SD - Service Delivery**	**5**	**3**	**4**	**5**	**4.25**
10	**WMC - Work Monitoring and Control**	**5**	**3**	**3**	**4**	**3.75**
11	**WP - Work Planning**	**5**	**3**	**3**	**4**	**3.75**
12	OT - Organizational Training	3	3	2	2	2.5
13	RSKM - Risk Management	3	2	2	3	2.5

(c) Create standardized Service Level Agreements and offer them to all current customers and prospects
(d) Name a single point of contact for each customer
(e) Introduces software tools for the automation of support and maintenance tasks
(f) Perform internal training for all staff.

Based on the defined actions, a detailed plan with a schedule, activities, milestones, and responsibilities was developed.

D. Acting

According to the previously prepared action plan, the members of the process group and management were trained. The management prepared analytical reports on the implementation of existing maintenance requirements. Reports have shown that there are about 150 different maintenance and support requests from customers and end users. Requirements refer to all areas of software, system software, software, and information systems in a broader sense, hardware platforms, requests for consulting, etc. Through a detailed analysis, those 150 different requirements were reduced to about 40 standard requirements, and then the requirements were classified according to ISO/IEC/IEEE types of software maintenance.

Upon completion of the classification, mapping, and interpretation of the process according to process areas were accomplished. This activity was of crucial importance for the implementation of all other improvements within selected process areas. In Table V an extract of the interpretation of the process area Requirements management is presented.

Additionally, the entire workflow of the maintenance process was documented and presented to all stakeholders. Afterward, the standardized SLA was created and presented to customers.

TABLE V. INTERPRETATION OF PROCESS AREA REQM

Specific goal / Practice (SP)	**Interpretation**
SG 1 Manage requirements - Requirements are managed and inconsistencies with plans and work products are identified.	
SP 1.1 Understand requirements	The customer manager (single point of contact) should understand the requirements and should be committed. Present the documented requests from SP 1.2 to the customer's authorized person and ask for confirmation of understanding of the request.
SP 1.2 Obtain Commitment to Requirements	
SP 1.3 Manage Requirements Changes	The customer's authorized person is authorized to approve changes during the implementation of the request within the agreed SLA. Each person from the customer's side should have the appropriate access level to the help desk software.
SP 1.4 Maintain Bidirectional Traceability of Requirements	The SLA manager reviews requirements and checks whether they are in line with SLA and the software product, and also checks bidirectional traceability between source requirements and lower requirements.

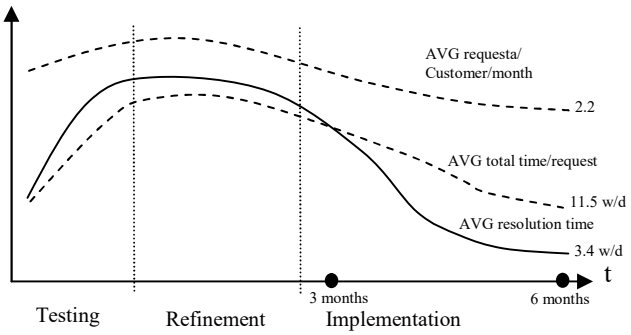


Figure 3. Monitored measures

After the completion of all the main activities from the activity plan, the solution with real examples from selected customers was tested. The testing results showed certain improvements, and a few weaknesses, though. For example, the customers demanded high urgency in resolving each documented request, explaining the requests were not classified correctly. On the maintenance side, the staff demanded an additional explanation about the term "required level of details" presented in a few interpretations of specific goals and also asked for clarification of the required amount of documentation needed to satisfy related practice. The additional training of stakeholders was completed to help with the requests prioritization and to clarify the amount of documentation.

For the further implementation of the solution, the gradual approach was selected. Along with each presented and signed SLA, the corresponding customer's training was carried out, and the single point of access with the company for each customer was named. The following implementation measures were monitored:

- Average resolution time per request (unit: working day)
- Average total time spent on the request's processing (unit: working day)
- Average number of requests per customer per month

In Fig. 3, the values per the timeline are presented. The resulting graph shows a steep growth during the testing phase for all values, following the stable values in the refinement phase, and later decline in the implementation phase. This is especially noticeable for the total time spent on each request, having the increased number of support activities introduced in the process. The staff needed additional time to acquire the necessary skills for documentation of each process, and using the implemented software tools.

E. Learning

By comparing the goals set at the beginning of the improvement cycle with the results, the following effects of the improved processes and practices can be concluded:

- Improved flexibility and readiness of the company to accept and resolve requests within a shorter resolution time with the same resources

- Improved understanding of tasks, assignments, and expected results of support staff
- Improved customer understanding and expectations regarding resources spent on each request.

Besides the positive effects, there were also several issues recognized that needed resolution during the implementation, such as the difference in levels of motivation of the staff working on the improvement. It was necessary to repeat several training sessions for customer management and the persons responsible for daily contact with customers.

Given the overall results of the improvements and the gained experience, there are a few recommendations and proposed feature actions:

- Introduce and implement software tools for configuration management (CM) and requirements management (RM)
- Automation of the workflow system and its integration with the help desk, CM and RM systems
- Prepare and launch the following improvement cycle: the goal should be to improve all process areas within Level 3 of the maturity currently not considered, e.g. PMC - Project Monitoring and Control, and PPQA - Process and Product Quality Assurance.

V. CONCLUSIONS

Small and medium software companies should provide a proportionally greater effort compared to large organizations in the SPI. There are a lot of limiting factors that prevent successful process improvement, and the pursuit of a successful approach for each case is constantly present. By using the CMMI for Services with the continuous presentation, the ground prerequisites for continuous improvement are provided in a way that can be easily applied in small software organizations with limited resources. The Balanced Score table enabled the possibility to directly link the company's business goals with the improvement efforts. Although there could be an additional discussion about whether it should be analyzed in the diagnosis phase (the description of the current state vs. desired future state), it assured its purpose. Eventually, it further enabled the selection of process areas for improvement that best match the business goals. This created a specific improvement path presented in this case study and provided the prerequisites for easier implementation of improvements according to the CMMI framework. It also ensured that efforts and resources spent on improvement would be measurable and profitable for organizations with sparse resources. As the improvement process is an ongoing and permanent project, there are a lot of further actions that need to be performed. The software company described in this case study continues with its effort and prepares its processes for the formal CMMI assessment. This is aligned with the widely accepted recommendations for continuous improvement as the primary way to meet business goals.

REFERENCES

[1] W. S. Humphrey, "Characterizing the software process: a maturity framework," *IEEE Softw.*, vol. 5, no. 2, pp. 73–79, 1988.

[2] Software Engineering Institute, "CMMI for Services, Version 1.3," Pittsburgh, Pennsylvania, 2010.

[3] B. Karahodza, E. Avdagić-Golub, and A. Čolaković, "Assessment of factors affecting the software process improvement in small organizations," *Sci. Eng. Technol.*, vol. 2, no. 1, pp. 7–15, 2022.

[4] M. Tuape and Yirsaw Ayalew, "Factors affecting development process in small software companies," in *SEiA '19: Proceedings of the 2nd Symposium on Software Engineering in Africa*, 2019, pp. 16–23.

[5] J. G. Romana Vajde Horvat, Ivan Rozman, "Managing the complexity of SPI in small companies," *Softw. Process Improv. Pract.*, vol. 5, no. 1, pp. 45–54, 2000.

[6] B. McFeeley, "IDEAL SM: A User's Guide for Software Process Improvement. CMU/SEI-96-HB-001," Carnegie Mellon University, Pittsburgh, 1996.

[7] N. Wongsai, V. Siddoo, and R. Wetprasit, "Factors of influence in software process improvement: An ISO/IEC 29110 for very-small entities," in *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2015, no. July 2020, pp. 12–17.

[8] A. F. H. Conradi, "Improving Software Process Improvement," *IEEE Softw.*, vol. 19, no. 4, pp. 92–99, 2002.

[9] G. Coleman and R. O'Connor, "Investigating software process in practice: A grounded theory perspective," *J. Syst. Softw.*, vol. 81, no. 5, pp. 772–784, 2008.

[10] H. Andersen, L. Gavin, and I. Cobbold, "Balanced Scorecard implementation in SMEs: reflection in literature and practice," in *4th SME-SME International Conference*, 2001, pp. 1–10.

[11] J. Soini, K. Harri, and T. Mäkinen, "Toward a balanced software business measurement," in *3rd Software measurement European Forum*, 2006, pp. 33–46.

[12] ISO/IEC 12207:2008, "Systems and software engineering — Software life cycle processes," 2008.

[13] P. Allen, M. Ramachandran, and H. Abushama, "PRISMS: An approach to software process improvement for small to medium enterprises," *Proc. - Int. Conf. Qual. Softw.*, vol. 2003-January, pp. 211–214, 2003.

[14] N. Paternoster, C. Giardino, M. Unterkalmsteiner, T. Gorschek, and P. Abrahamsson, "Software Development in Startup Companies : A Systematic Mapping Study," *Inf. Softw. Technol.*, vol. 56, no. 10, pp. 1200–1218, 2014.

Application of Hybrid Project Management Methodology in Development of Software Systems

Svea Urlić, Željka Car
Faculty of Electrical Engineering and Computing, Zagreb, Croatia
* svea.urlic@gmail.com ** zeljka.car@fer.hr

***Abstract* – Selection of project management methodology in development of software systems is very challenging, especially taking into consideration that every project is unique. In this paper hybrid approach as a combination of traditional plan-driven and agile methodologies of project management, with emphasis on applying Scrum in waterfall project management, in development of web-applications is analyzed. This paper provides results of seven qualitative researches done through interviews with project managers and statistical data analysis of survey respondents (n = 47). Results of qualitative research show that most projects follow a pattern in which waterfall methodology of project management is combined with agile practices - agile was used in almost all projects which were the subject of research in project phase of Development. Results of quantitative research conducted through a survey confirmed that Scrum and waterfall are the most frequently used methodologies in web application development projects and also indicated the frequency of overlapping of Development and Testing project phases. Paper provides guidelines for combining waterfall model with agile models and emphasizes the importance of planning, analyzing user requirements and defining software architecture at the beginning of the project, as well as delivering multiple minor development versions to users for testing.**

Keywords - project management; software development; waterfall; agile project methodologies; Scrum, hybrid project methodologies; hybrid model; web-application

I. INTRODUCTION

This paper provides research results on application of hybrid project management methodologies in development of web applications. The aim of this paper is to identify which project activities in web application development projects are mostly managed according to traditional waterfall model, and which activities are mostly managed according to agile methodologies, with emphasis on Scrum.

In traditional approach to project management, in accordance with the guidelines described in Project Management Body of Knowledge (PMBOK) [1], project lifecycle consists of 5 distinct phases including Initiation, Planning, Execution, Monitoring and Closure. It is characteristic that phases are sequential and that the outputs, that are the results of the work done in one phase, are the inputs for the next phase. In traditional approach, plan is the basis of the project, and therefore traditional project management methodologies are often called "plan-driven". Waterfall model was dominant software development method until the end of 20th century [2]. Waterfall model by stages of software development is shown in Figure 1.

On the other hand, agile methodologies encourage an iterative and incremental approach to development, in which all phases are completed in each iteration, and the software is delivered evolutionary [2]. The Agile Manifesto [3] consists of four value statements:

1) Individuals and interactions over processes and tools
2) Working software over comprehensive documentation
3) Customer collaboration over contract negotiation
4) Responding to change over following a plan.

According to State of agile [4] Scrum is the most used agile project methodology. Scrum life cycle is shown in Figure 2.

Hybrid project management methodologies are combination of multiple project management methodologies. Such methodologies are not strictly defined and project teams tailor them based on their experience and project needs. Results of [5] show that majority of projects that were analyzed in the survey HELENA (Hybrid dEveLopmENt Approaches in software systems development) combine traditional and agile methods.

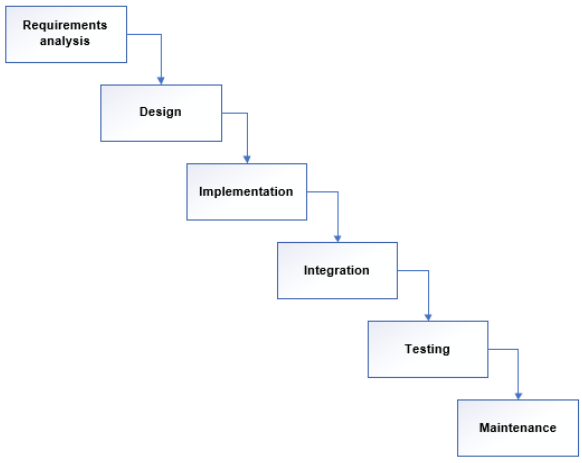


Figure 1. Waterfall model by stages of software development [2]

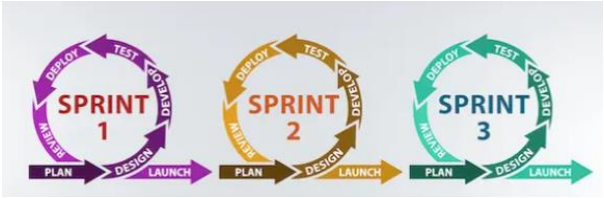


Figure 2. Scrum life cycle [6]

II. HYBRID PROJECT METHODOLOGIES

This section presents a review of results of the existing research on hybrid project methodologies and related work.

A. *Base methods for hybrid methods*

According to [7], there are eight base methods that provide the frame for hybrid development methods - Scrum, Iterative Development, Kanban, Classic Waterfall Process, DevOps, eXtreme Programming, Lean Software Development and Feature Driven Development.

B. *Core practices used in hybrid methods*

Study [8] shows that hybrid methods are composed of mostly technical practices such as Code review, Coding standards and Release planning. Within answers of participants in research [8], who combined different development approaches in one product, five practices were identified (Code review, Coding standards, Release planning, Automated unit testing and Prototyping). Of these five practices, three pairs of two composed from the five practices could be identified, which had at least 85% agreement among the participants of the study [8] who combined different development approaches in one product.

C. *Hybrid models combining traditional and agile approach*

The beauty of the hybrid project management method is that it lets the team plan before starting to work on the project, but also divides the development cycle into short-term deliveries called sprints [9].

According to [10], hybrid approaches work better in larger-scale project contexts with higher levels of requirement uncertainty. Further, the hybrid approach increases the likelihood of improving the cost-benefit ratio, compared to purely plan-driven methods [10].

Authors in [11] have shown that agility is strongly presented in development-related activities, whereas management remains on a more traditional track.

III. RESEARCH DESIGN

We conducted two researches on the topic of using hybrid project management methodologies in web application development projects.

First, we conducted qualitative research in which the respondents were only project managers.

Qualitative research was the basis for designing questionnaire for quantitative research, which included other project roles in order to determine their attitude towards project management methodologies used on web application development projects, with emphasis on hybrid models.

Overall research design is outlined in Figure 3.

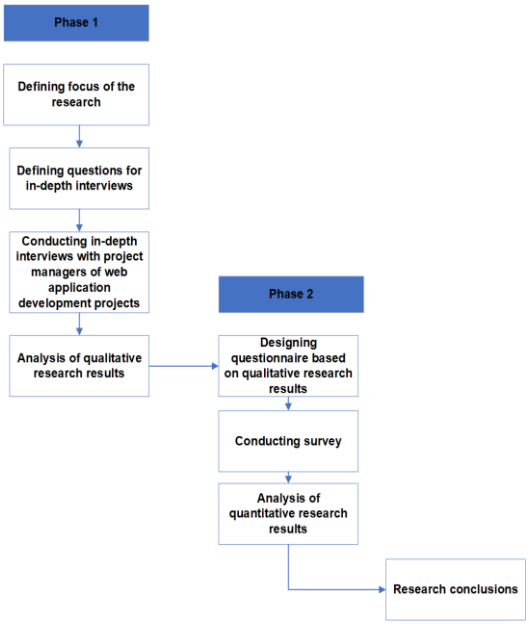


Figure 3. Research design

The aim of overall research was to determine frequency and methods of application of hybrid models in web application development projects.

IV. QUALITATIVE RESEARCH

This section presents a review of conducted qualitative research on management of web application development projects.

A. *Objective*

Qualitative research was conducted with the aim of determining which project management methodologies were used, how often different project management methodologies were combined and how successful their combination proved to be for web application development projects. The focus of the research was to determine challenges that arose on the project and the way they were overcome, from project manager's perspective.

B. *Data collection*

Data were collected from June to August 2022 through in-depth interviews. Respondents were 7 project managers from Croatia who had work experience in managing web application development projects. Projects included in the research were finished projects, respectively web applications already in usage, which respondents managed.

C. *Research questions*

Interview consisted of 24 questions, which were grouped into 8 logical units as follows:

1) Demographic questions related to project management competencies - years in profession; age; formal education in project management.
2) Characteristics of the project - project duration; size of project team; quality of communication within the project team.
3) Technical characteristics of the project - was it a new web application or improvement of existing one; complexity of integration with other systems.
4) Change management - quality of definition of user requirements; frequency of change requests.
5) Challenges that arose on the project - number of failures discovered during testing; number of failures discovered in production; largest issues in general; how were the challenges overcome.
6) Used project management methodologies - usage of hybrid models; how was initial project planning done; was there closure of each project phase or project phases overlapped.
7) Project success - in terms of meeting budget; deadlines; quality.
8) Lessons learned - the effectiveness of used project management methodology.

D. Results

In most of the projects that were subjects of qualitative research, main challenges were change requests and unclear project scope. Even projects on which planning was done at the beginning of the project encountered numerous changes during the project, so the initial plan was no longer valid. Change requests occurred near the end of a project - in project phases of Development and Testing. Communication between stakeholders was also highlighted as a common issue. Primarily there were issues in communication with the client, but also between development teams or even within development teams.

All projects that were subjects of qualitative research required integration with at least 2 external software systems. Changes of software architecture during the project occurred in half of the projects.

Almost all projects (6 out of 7) had more than 10 critical failures reported in Testing phase. Most of the projects (also 6 out of 7) were released with less than 5 critical failures.

Regarding used project methodologies, combination of traditional and agile project management methodologies, i.e. hybrid project management methodology, was used on most projects. Of traditional methodologies, waterfall model was used most often, and of agile methodologies, Scrum and Kanban. Agile methodologies were mostly used in project phases of Development and Testing. Initial project phases (Planning, Requirements analysis, Software architecture design) were mostly managed according to waterfall model. On most projects on which hybrid models were not used, only agile project methodologies were used - specifically Scrum and Kanban. Only one project was completely managed according to waterfall model.

In majority of projects, 6 out of 7, there was overlapping of project phases.

Most projects (6 out of 7) were completed within the budget, but were not completed on schedule.

In more than half projects (4 out of 7), users participated in testing during development. A clear correlation was observed between project delays and users' non-participation in testing during development, i.e. projects on which users tested only after development was completed were delayed the most because needed changes and failures were noticed quite late.

Nevertheless, most of the projects eventually met the set quality standards.

The results of qualitative research are shown in Table 1.

V. QUANTITATIVE RESEARCH

This section presents a review of conducted quantitative research on management of web application development projects.

A. Objective

Objective of this research was to detect which project methodologies and practices are used in development of web applications. The goal was to understand what hybrid project methodologies are used in web application development and how the combination of traditional and agile project methodologies aims to overcome the most common problems in projects. At last, objective was to discover which project phases are more suitable for traditional approach, specifically waterfall model, and which are more suitable for agile approach.

B. Data collection

Data were collected through survey from 29th of August 2022 to 5th of September 2022. Questionnaire consisted of 3 parts (1. Demographic data; 2. Experience from last finished web application development project; 3. General opinion on project management methodologies and practices used in development of web applications) and total of 18 questions. Questionnaire was distributed to potential respondents from 9 different companies. We collected survey responses from 47 individuals.

C. Results

Regarding age-demographics of respondents, majority of respondents, 44,7%, were in the (30-39) age bracket, followed by 21,3% (25-29 years old) and 21,3% (40-49), then 6,4% (below 25 years old) and 6,4% (above 50 years old).

The highest percentage of respondents, 34%, indicated their work experience on projects of development web applications as (5-10 years), followed by 31,9% (1-5 years), then 29,8% (over 10 years). Only 4,3% of respondents had less than 1 year of work experience on web application development projects.

In reference to duration of their last web application development project, most of respondents, 31,9%, stated that their project lasted over 2 years followed by 27,7% (6

TABLE 1. QUALITATIVE RESEARCH RESULTS

Measure/Interview No.	1	2	3	4	5	6	7
Project duration (in months)	18	6	7	18	30	22	7
Number of project team members	11	9	11	26	17	11	5
New web application/improvement of existing one	Improvement	Both	Improvement	New	New	New	New
Number of integrations	2	5	6	5	6	3	2
Change requests	Yes	Yes	Yes	No	Yes	Yes	Yes
Project phase(s) in which change requests occurred	Solution design; Development	Development; Testing	Development, Testing	-	Testing	Testing	Testing
Software architecture changes	No	No	No	Yes	Yes	Yes	Yes
Project phase(s) in which architecture changes occurred	-	-	-	Solution design	Testing	Development	Production
Number of critical failures reported during testing	>10	>10	<10	>10	>10	>10	>10
Number of critical failures reported in production	<5	<10	<5	<5	<5	<5	<5
Project methodology	Waterfall model; Scrum	Scrum	Waterfall model; Scrum; Kanban	Scrum; Lean	Waterfall model; Scrum; Lean	Kanban	Waterfall model
Project plan made at the beginning of the project	Yes	No	Yes	Yes	Yes	No	Yes
Overlapping of project phases	Yes	Yes	Yes	Yes	Yes	Yes	No
Users involved in testing during development	Yes	Yes	Yes	Yes	No	No	No
Budget met	Yes	Yes	Yes	Yes	No	No	No
Delay (in months)	6	1	-	2	12	16	1
Quality standards met	Yes	Yes	Yes	Yes	No	No	Yes

months to 1 year), then 17% (1 year to 1,5 years), then 12,8% (1,5 years to 2 years). The least, 10,6% was for project duration less than 6 months.

The most common project role among respondents was Developer (29,8%) followed by Business analyst (25,5%), Project manager (21,3%), Consultant (8,5%) and Software architect (6,4%) (Figure 4).

Regarding the size of project team, majority of respondents, 48,9%, stated that their team consisted of 5 to 10 team members, then teams of less than 5 members (29,8%), teams of 10 to 20 members (14,9%) and teams of over 20 members (6,4%).

When it comes to largest issues that occurred on their last web application development project, respondents were allowed to select multiple answers. The most common answers were the following: change requests (59,6%), unclear project scope at the beginning of the project (55,3%), inaccurate estimates of development tasks (48,9%), difficult communication with users (40,4%), too tight deadlines (36,2%) and involving users in testing too late (17%).

Even 48,9% of respondents declared that primary project methodology used on last web application development project on which they participated was Scrum. Scrum was followed by waterfall model (29,8%) (Figure 5).

However, 55,3% respondents used traditional project management methodology in project phase of Planning and requirements analysis (Figure 6).

On the contrary, only 19,1% respondents used waterfall model in project phase of Development, whereas 46,8% used Scrum (Figure 7).



Figure 4. Project role of respondents on their last web application development project

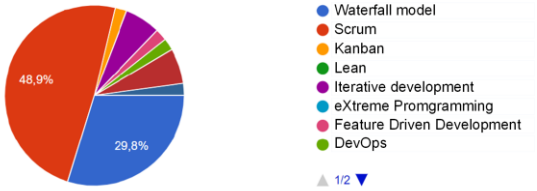


Figure 5. Primary project methodologies used on web application development projects

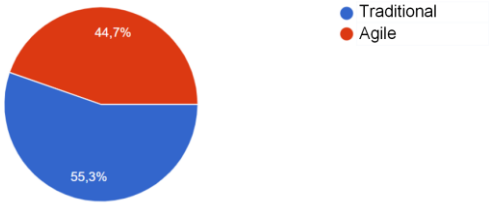


Figure 6. Project management methodology used in project phase of Planning and requirements analysis



Figure 7. Project management methodology used in project phase of Development

Moreover, answers to the question regarding used software development practices confirmed that respondents used both practices characteristic for traditional (such as Formal specification and Formal estimation) and practices characteristic for agile project management methodologies (such as Daily standup and Retrospectives). Respondents were allowed to select multiple answers to the question on used software development practices and results were the following:

- Daily standup - 80,9%
- Code review - 66%
- Formal specification - 59,6%
- Release planning - 57,4%
- Formal estimation - 53,2%
- Unit testing - 51,1%
- Retrospectives - 44,7%
- Coding standards - 38,3%
- Refactoring - 34%
- Continuous integration - 29,8%
- Prototyping - 25,5%
- Automated unit testing - 25,5%.

As regards the overlapping of project phases, on majority of projects some project phases overlapped. Project phase of Development overlapped with some of other project phases in most cases (74,5%), followed by project phase of Testing, which overlapped with some other project phase in 70,2% cases, and project phase of Requirements definition and analysis, which overlapped with some other project phase in 68,1% cases. Release phase overlapped with some other project phase in 36,2% answers, then phase of Software architecture design in 29,8% answers. But, 17% of respondents claimed there was no overlapping of project phases on their projects. Phases of Development and Testing overlapped on projects of the majority of respondents, and in a second place was overlapping of project phases of Requirements definition and analysis and Development.

More than half (53,2%) of respondents stated that users were involved in testing during project phase of Development. Over third of respondents, 38,3%, indicated that they included users in testing only after finishing project phase of Development and 8,5% of respondents have not included users in testing at all on their last web application development project.

In reference to project delay, two groups shared the first place both with 23,4% of answers - delays up to 15% of project duration and delays between 15 and 30% of project duration. Even 14,9% projects were delayed for over 50% of project duration, followed by projects delayed between 30 and 50% of project duration (6,4%). Almost third of projects, 31,9%, were not delayed but finished on time.

Regarding general opinion on project management on web application development projects, significant majority of respondents find very important (46,8%) or important (31,9%) to perform planning, requirements analysis and software architecture design in as much detail as possible at the beginning of the project. Also, 57,4% respondents completely agree and 27,7% agree that it is better to use agile project methodologies in project phase of Development. Most respondents consider very important (57,4%) and important (40,4%) to deliver multiple minor versions of application to users for testing. Even 61,7% of respondents find very important and 29,8% important to document user requirements and software architecture well at the beginning of the project to make software development and maintenance easier.

VI. GUIDELINES FOR COMBINING WATERFALL MODEL WITH AGILE MODEL

Based on the conclusions of conducted researches, guidelines for combining waterfall model and agile model are provided in Table 2.

TABLE 2. GUIDELINES FOR COMBINING WATERFALL MODEL AND AGILE MODEL

Project phase	Recommended approach	Key activities
Requirements definition and analysis	Waterfall model	Identify, clearly define and document user requirements prior to moving to next project phase.
Software architecture design	Waterfall model	Define and document software architecture prior to moving to next project phase.
Development	Agile model	Develop in iterations and deliver increments of software for testing.
Testing	Agile model	Test during development. Prepare test scenarios and acceptance criteria.
Release	Waterfall model	Formally close the project after all previous project phases are closed.

CONCLUSION

As a result of conducted research, it is evident that project teams combine different project management methodologies in development of web applications, that is, they use hybrid models in order to best respond to project challenges. Based on the results of research it is recommended to use waterfall model in project phase of Requirements definition and analysis. User requirements should be clear and well documented at the beginning of the project in order to plan the project and identify costs and achievable project deadlines. Clearly defined and documented user requirements will significantly facilitate development and maintenance of web application in the future. It is also recommended to use waterfall model in project phase of Software architecture design meaning software architecture should be defined and documented before moving on to the next project phase. Agile approach is recommended in project phases of Development and Testing. Those project phases should be overlapping, development should be done in iterations and increments of software should be delivered to users for testing. Finally, project phase of Release is recommended to be performed according to waterfall model in terms of formal closure of the project, after all previous project phases are closed.

REFERENCES

[1] „A guide to the project management book of knowledge – Seventh edition and The standard for project management“, Project management Institute, Pennsylvania, 2021.

[2] Roić, R., Ferlež, L., „Agilni razvoj softvera”, Self-published, Zagreb, 2016.

[3] Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, J., „Agile Manifesto“, available at: https://agilemanifesto.org/ (27th of June 2022)

[4] VersionOne, „11th annual State of Agile“, available at: https://www.agile247.pl/wp-content/uploads/2017/04/versionone-11th-annual-state-of-agile-report.pdf (27 th of June 2022)

[5] Noll, J., Beecam, S., „How Agile Is Hybrid Agile? An Analysis of the HELENA Data“, International Conference on Product Focused Software Development and Process Improvement (PROFES), Barcelona, Spain, 2019., pp. 341-349.

[6] Blake, S., „Your Guide To Agile Software Development Life Cycles“ dostupno na: https://www.easyagile.com/blog/agile-software-development-life-cycle/ (29th of January 2023)

[7] Tell, P., Klünder, J., Küpper, S., Raffo, D., MacDonell, S.G., Münch, J., Pfahl, D., Linssen, O., Kuhrmann, M., „Towards the statistical construction of hybrid development methods“, Journal of Software: Evolution and Process, Vol. 33, No. 1, January 2021., pp. 1-16.

[8] Tell, P., Klünder, J., Küpper, S., Raffo, D., MacDonell, S.G., Münch, J., Pfahl, D., Linssen, O., Kuhrmann, M., „What are Hybrid Development Methods Made Of? An Evidence-based Characterization“, IEEE/ACM International Conference on Software and System Processes, Montreal, Canada, 2019., pp. 105-114.

[9] Robinson, R.D., „Hybrid: A new management approach“, available at: https://www.cio.com/article/230639/hybrid-a-new-project-management-approach.html (7th of August 2022)

[10] Imani, T., Nakano, M., Anantatmula, V., „Does a Hybrid Approach of Agile and Plan-Driven Methods Work Better for IT System Development Projects?“, Int. Journal of Engineering Research and Application, Vol. 7, No. 3, March 2017., pp. 39-46.

[11] Kuhrmann, M., Münch, J.,Diebold, P., Linssen, O., Prause, C.R., „On the Use of Hybrid Development Approaches in Software and Systems Development: Construction and Test of the HELENA Survey“, Projektmanagement und Vorgehensmodelle, Paderborn, Germany, 2016., pp. 59-68

SP

International Conference on

MIPRO JUNIOR - STUDENT PAPERS

Steering Committee

Chairs: Lovro Božičević, University of Zagreb, Croatia
Dubravko Sabolić, University of Zagreb, Croatia
Pavle Ergović, University of Zagreb, Croatia

PAPERS

Analyzing Sentiment of Reddit Posts for the Russia-Ukraine War

Armin Krivičić[1], Sanda Martinčić-Ipšić[1,2]
[1]Faculty of Informatics and Digital Technologies, [2]Center for Artificial Intelligence and Cybersecurity,
University of Rijeka,
Radmile Matejčić 2, 51000 Rijeka, Croatia;
Email: armin.krivicic@student.uniri.hr, smarti@uniri.hr

***Abstract*—Currently, we are witnessing a major global crisis caused by the ongoing war in Ukraine. In this paper, we deepen the insights into people's opinions about the Russia-Ukraine war by analyzing the posts on the Reddit subreddit r/UkrainianConflict. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using the Valence Aware Dictionary and sEntiment Reasoner (VADER). We analyze the frequency of the most representative keywords in two longer periods in May and October 2022. The keyword Zelensky appeared more in positive contexts than in negative ones. Keywords Ukraine, Russia, and especially Putin appeared more frequently in negative contexts. Finally, the trend of Google searches on the war is compared with the number of daily subreddit posts.**

***Keywords*—Russia-Ukraine war, Reddit, r/UkrainianConflict NLP, tokenization, stop words, sentiment analysis, NLTK**

I. INTRODUCTION

Social media plays an important role in communication during times of crisis, such as the COVID-19 pandemic [1], [2], wars, climate change [3], economic [4] or energy crises, etc. Social media can serve as a key communication platform [5] and is potentially a source of valuable information [6]. Over the past two decades, social media have amplified the spread of information, as well as misinformation and disinformation, which can lead to an infodemic as a negative side effect [7]. Understanding how people post comments in online communication can shed light on the fundamental mechanisms by which collective thinking emerges in a social group [8].

In this light, natural language processing (NLP) [9] offers methods that contribute to better understanding and monitoring the content of crisis communication on social media. Namely, automatic keyword extraction [10], [11], topic modelling [12], [13], named entity recognition [14], text classification [15], sentiment analysis [16], fake news detection, text summarization [17], etc. are all NLP tasks employed for the social media content analysis.

Reddit is a popular social media website with about one million daily active users (as of 2015) [18]. Reddit has many subreddits – communities or user groups that focus communication on a specific topic.

The goal of the conducted research is to discover the polarity of people's sentiments regarding the war between Russia and Ukraine and how this polarity changes over time. Hence, this paper analyses the content of communication, related to the topic of the war crisis in Ukraine on the subreddit r/UkrainianConflict. We collect data from this subreddit and employ NLP techniques for sentiment analysis.

The analysis is performed on posts from the subreddit r/UkrainianConflict during the period from January 1 to October 30, 2022. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using the Valence Aware Dictionary and sEntiment Reasoner (VADER) dictionary [19]. Finally, the trend of Google searches on the war is compared to the number of daily subreddit posts.

Section II contains a short overview of the sentiment research based on Reddit. In Section III, we describe the collected dataset In Section IV, we present and discuss the results of sentiment analysis and statistics of representative keywords over different time periods along with Google Trends. In the last section, we give some concluding remarks and plans for future work.

II. RELATED WORK

Reddit has evolved in one of the most prominent social platforms on the Internet over the past decade, with 52 million daily active users, over 430 million active monthly users [20] and over 100,000 active topical communities called subreddits. Reddit is ranked as the 20th-most-visited website in the world and the 5th in the US [20], [21]. The authors of [22] performed a systematic analysis of 727 manuscripts that analyzed Reddit data. They show that the scope of Reddit research broadly spans many scientific fields, from computer science, medicine and health, natural sciences to humanities and social sciences, indicating that Reddit has gained important status as a source of information for the research and has an impact on the community [18].

Sentiment analysis (or opinion mining) is a field of research that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [23]. Reddit has been the source for the sentiment analysis of user communities in a wide spectrum of domains spanning from COVID-19 vaccination [1], [2], video games stock

prices [4], advancement of robotic technologies [24], the education of gifted children [25], public perception of electric vehicles [3], and so on.

Melton et al. in [1] studied COVID-19 vaccine-related discussions on 13 subreddits performing sentiment analysis and topic modelling. Polarity analysis suggested that Reddit communities expressed more positive sentiment than negative sentiment, and this remained consistent over time.

In [2], Yan et al. examined location-based subreddits to identify city-level variations in sentiment toward vaccine-related topics. The random forest regression model was trained to assign sentiment scores for joy, sadness, fear, and anger to comments. The authors confirm that Reddit comments can be used to better understand concerns and sentiments around COVID-19 vaccines at the local level.

The authors in [4] investigated the impact of sentiment in discussions on the r/WallStreetBets subreddit on the price dynamics of the American online retailer GameStop in 2021. They extend the VADER [19] sentiment analysis package to perform textual analysis of 10.8 million comments and confirm that sentiment has an impact on stock prices.

Savela et al. in [24] investigated sentiment in social media discussions about robotic technologies. Comments were processed using VADER and LIWC (Linguistic Inquiry and Word Count) [26] dictionaries, and logistic regression models were used to infer polarity. The results indicate that the terms "robot" and "artificial intelligence" were used less frequently in a positive context.

Ruan and Lv in [3] studied the public perception of electric vehicles from the lenses of ecology and energy to mitigate climate change and global warming. They utilized VADER dictionary for sentiment inference and concluded that issues that impact the environment (e.g., climate change) have lower sentiment scores.

The study in [27] detected emotions, specifically "hate" and "fear" in Reddit posts during the period from May to July 2022. The authors used the NRC Emotion Lexicon [28] and topic modelling and concluded that the "spikes in hope/fear, both positives and negatives, are present after important battles, but also some non-military events, such as Eurovision and football games".

With this paper, we aim to fill the gap in the analysis of the Reddit discussion on the Russia-Ukraine war in May and October 2022.

III. METHODOLOGY AND EXPERIMENTS

A. Dataset

According to Zhu et al. in [29], r/UkrainianConflict is one of the twelve most prominent subreddits that focuses on the Russia-Ukraine war. It has one of the highest numbers of active users, posts and comments. Posts from the said subreddit were collected using the PSAW (Python Pushshift.io API Wrapper) library [30] implemented in the Python 3 language. We defined a function that accepts three parameters: start date, end date and the name of the subreddit from which we want to retrieve posts. Along with the text, the date of the post was also collected in Python *datetime* format. Our function then stores the collected data in a *pandas DataFrame* format [31] and also extracts it to a comma-separated values (CSV) file. The CSV file was then separated into two parts, one for May 2022 and one for October 2022, to allow for a deeper analysis of these two time periods. These two time periods hopefully provide a stable reflection of the war-related communication, since they are not the beginning, which is characterized by excessive posting, reflecting the initial impulses of reaction (please see the frequency for February in Fig. 1). From January to October 2022, we collected a total of 101,977 posts.

B. Dataset preprocessing

In order to apply NLP techniques, the text must first be converted to lowercase. Then the text must be tokenized, that is, it must be broken down into words. Tokenization is performed using the NLTK (Natural Language Toolkit) tokenizer from the NLTK library [32]. The NLTK tokenizer uses regular expressions to determine the words in each sentence (*tokens*) and stores them in lists for each post. These lists are then stored in the new column of our *DataFrame*.

C. Dataset analysis

For each day of the period January 2022 – October 2022 we counted the posts and created a key-value dictionary. The keys are dates, while the values represent the number of posts for the corresponding date. Fig. 1 shows the created dictionary as a graph, with the dates on the x-axis and the number of daily posts on the y-axis. A peak in the number of posts can be seen in the graph, which correlates with the start of the Russian invasion of Ukraine. The invasion began on February 24 [33], while the peak was reached three days later.

One of the most informative analyses is based on the frequency of words in users' posts. To measure frequency,

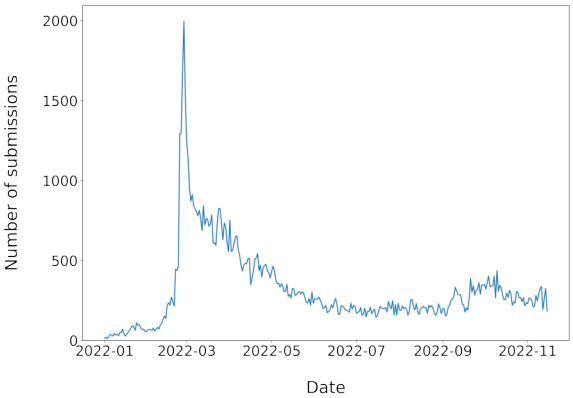


Fig. 1: Number of posts per day

we need to further preprocess the texts by removing stop words. Stop words are words that do not carry any meaning, e.g., *the, a, an, is, are, in, on*, but have a syntactical function in a text. To this end, we use a list of English stop words from the NLTK library. We create a function for comparison of the tokens in the text with a list of stop words and remove the stop words from further analysis. Then we calculate the frequency of each word in the dataset.

The top ten most frequent words (in descending order) are: **ukraine** (39,252), **russian** (32,016), **russia** (22,614), **ukrainian** (14,876), **war** (11,812), **putin** (9,470), **forces** (6,063), **military** (5,506), **says** (4,489) and **russians** (4,178). Fig. 2 shows the histogram of the top ten most frequent words.

IV. RESULTS

A. Sentiment analysis

In order to detect the polarity of Reddit posts, we use the NLTK Sentiment Intensity Analyzer [32]. The analyzer is implemented with the VADER lexicon. The VADER lexicon is a parsimonious rule-based model with over 7500 words labelled with sentiment intensity, where -4 denotes extremely negative sentiment and +4 denotes extremely positive sentiment [19]. Labelling was performed using Amazon Mechanical Turk (AMT). These words were obtained from Twitter and also include emojis and abbreviations. For this reason, the VADER lexicon is well suited for sentiment analysis in social networks.

Next, the polarity scores for each of the posts in the dataset are calculated, resulting in the *positive, negative, neutral* and *compound* scores. Positive, negative, and neutral indicate the proportion of the text that falls into each category, while the compound score is the sum of the valence scores of the individual words in the lexicon. The score is then normalized to a range of -1 to 1. The following formula is used to calculate the compound score:

$$\frac{s}{\sqrt{s^2+\alpha}}$$

where s is the sum of the valence scores while α is a constant set to 15.

As proposed by the authors of VADER lexicon [19], we retain the thresholds for the compound score as follows:

- Negative sentiment: ≤ -0.05,
- Neutral sentiment: > -0.05 and < 0.05,
- Positive sentiment: ≥ 0.05.

The initial hypothesis that the majority of posts would show negative sentiment proved to be correct, as can be seen in Fig. 3. Negative sentiment predominated with 47.2%, followed by 31.3% neutral posts and 21.5% posts with positive sentiment.

The average negative sentiment score is -0.5177, the average positive sentiment score is 0.4012, and the average neutral sentiment score is 0.0001. The average overall sentiment score is -0.1582, again confirming the hypothesis about the majority of negative posts.

It is important to examine how sentiment has changed over time. For this purpose, we need to calculate the mean sentiment for each day. If we plot these daily sentiment means, we will get a very impractical chart since there are over 300 values. Instead, we can use the rolling window calculation with a window size of 30, which is implemented in the *pandas* library. The rolling window function creates a *rolling mean* of the sentiment values, and the mean value is updated each day. The resulting graph is smoothed and plotted in Fig. 4. The most interesting part is from the end of February, when we see

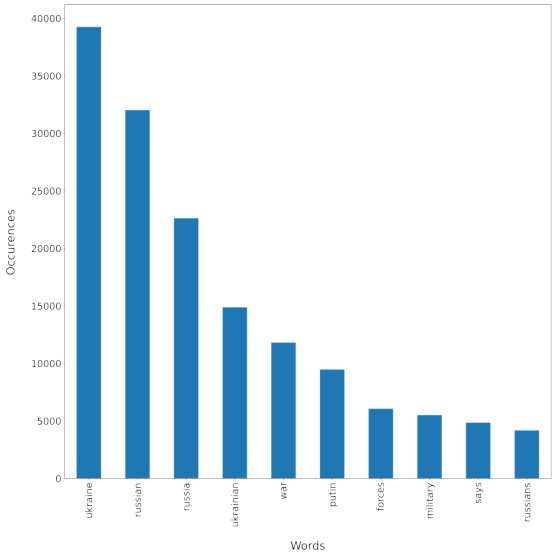


Fig. 2: Frequencies of most common words

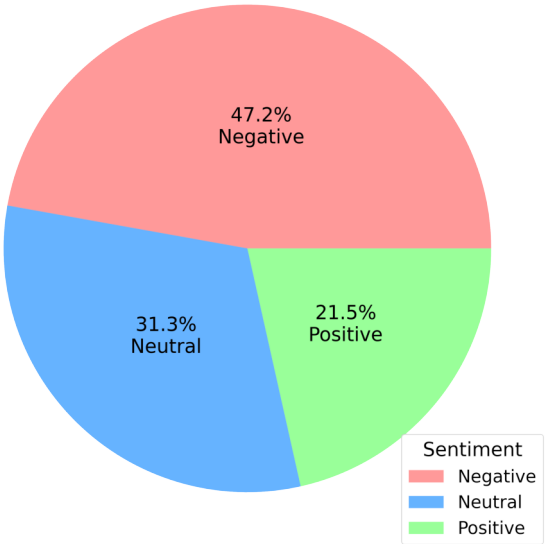


Fig. 3: Distribution of positive, negative and neutral sentiment

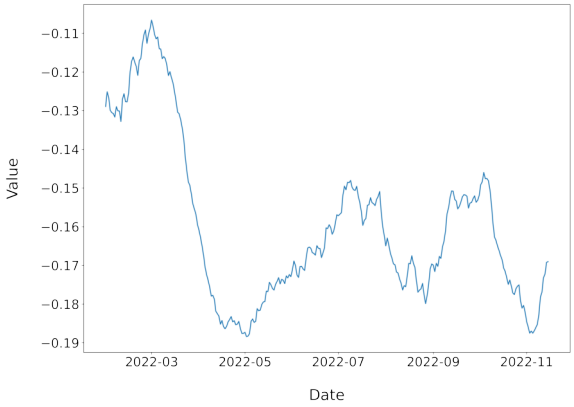


Fig. 4: Sentiment through the time with a window size of 30 days

a steep decline in the sentiment. This decline correlates strongly with the start of the Russian invasion of Ukraine.

B. Representative keywords comparison

Representative keywords used in this paper are **Zelensky, Putin, Ukraine**, and **Russia**. We divided the dataset into two parts: a positive and a negative part. In order to find out which of these representative keywords are more likely to occur in the positive context and which in the negative context, we need to calculate their percentage. By percentage we mean the ratio between the number of occurrences of the keyword (token) and the total number of tokens. The reason for this normalization is that there are many more negative posts, and the absolute numbers would be meaningless. The keyword **Zelensky** accounts for 0.25% of all tokens in positive posts. In negative posts, this percentage is 0.15%. The keyword **Putin** occurs in 0.49% of positive posts and 0.61% of negative posts. The keyword **Russia** appears in 2.63% of positive posts and 3.83% of negative posts. Finally, the keyword **Ukraine** appears almost equally in both sentiments, 2.26% in the positive and 2.33% in the negative (Tab. 1).

In addition, two longer periods are compared in terms of the most representative words. The first period is May 2022 (April 15 to June 15), and the second is October 2022 (September 15 to November 15) (Tab. 2).

As for the sentiment in these two periods, the average negative value for the month of May is -0.5268, the average positive value is 0.3965, and the average neutral value is 0. The overall average sentiment value is -0.1739,

TABLE I: Share of the most representative keywords in positive and negative sentiment

Keyword	% in positive	% in negative
Zelensky	0.25	0.15
Putin	0.49	0.61
Russia	2.63	3.83
Ukraine	2.26	2.33

TABLE II: Share of the most representative keywords in May and October 2022

Keyword	% in May	% in October
Zelensky	0.0029	0.0009
Putin	0.0132	0.0113
Russia	0.234	0.2226
Ukraine	0.0218	0.0179

slightly more negative than the overall value for 2022, which is -0.1582.

For the month of October, the average negative sentiment value is -0.525, the average positive sentiment value is 0.3938, and the average neutral sentiment value is 0.0001. The average overall sentiment value is -0.1687, which is again slightly more negative than the overall sentiment value for the year 2022. The reason for these negative-leaned sentiment values in May and October is that for the year 2022, the part up to February 24, when sentiment had higher total values around zero (i.e. was more positive in general), was also included.

C. Subreddit posts vs. Google Trends

Comparison of subreddit posts with Google Trends is performed for the same two time periods (May and October 2022). Google Trends measures the popularity of a particular term in Google searches [34]. A term can be a single word such as Ukraine or Russia or a phrase such as the Russia-Ukraine war. Google normalizes the number of searches to a range from 0 to 100, where 100 represents the day when the term was searched for the most. Additionally, we can specify a range where we want to see the popularity of the term, a region in the world, etc. The obtained data can be easily exported to a CSV file with two columns (popularity and date).

In order to compare the number of daily posts from our dataset and the popularity on Google Search, we need to normalize the values since the daily number of posts is in the range of a few hundred per day, while the popularity is at most 100. In Fig. 5 and Fig. 6, we see two graphs for May and October 2022, respectively. These graphs show the comparison, where the orange line represents the trend of Google Search for the term Russia-Ukraine war and the blue line represents the number of daily posts. In May 2022, users were more engaged on Reddit until May 9, after which activity on Reddit and Google was roughly equal. May 9 is known in Europe as Victory Day – a holiday commemorating the Soviet victory over Nazi Germany in 1945, and in Russia it is a national holiday. It was expected that the Russian president would announce an escalation of military action but that did not happen. Nevertheless, the activity on Google was very high on that day. In October 2022, activity remained the same, except for a slightly higher activity on Google in the middle of the month. The jump in activity on September 21 correlates with the day the Russian president announced a partial mobilization of military reservists [35]. The highest activity was on October 10, when Russia

launched a massive missile strike across the entire territory of Ukraine, including the capital, Kyiv [36]. The activity on Reddit follows the activity on Google, suggesting that Reddit is a legitimate source for studying war-related communication through sentiment quantification.

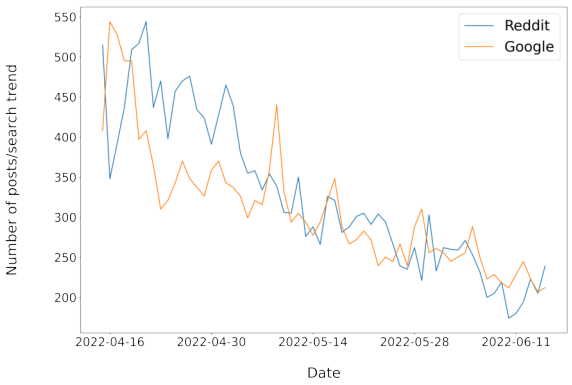


Fig. 5: Number of subreddit posts vs. Google Search trend (term Russia-Ukraine war) for May 2022

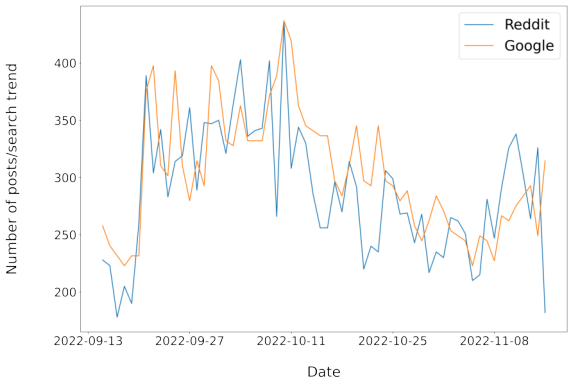


Fig. 6: Number of subreddit posts vs. Google Search trend (term Russia-Ukraine war) for October 2022

V. CONCLUSION

In this paper, we deepen the insights into people's opinions about the Russia-Ukraine war by analyzing the posts on the Reddit subreddit r/UkrainianConflict. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using VADER. We analyze the frequency of the most representative keywords in two longer periods in May and October 2022 and compared them with Google Trends.

The subreddit r/UkrainianConflict was created on 20 February 2014 [29] during Russia's occupation of the Ukrainian peninsula of Crimea [37]. A month later, Russia annexed Crimea, and to this day there are ongoing tensions between the two countries. Activity in the subreddit was low until late February 2022, which strongly correlates with the escalation of tensions and the date the Russian president ordered the invasion of Ukraine. Sentiment remained negative throughout the observed period (January 1 to October 30, 2022), dropping sharply after the invasion. Thereafter, it remained relatively stable in the range of -0.15 and -0.2.

The keyword Zelensky appeared much more frequently in positive posts than in negative ones. Keywords such as Ukraine, Russia, and especially Putin appeared more frequently in negative posts. Moreover, the keywords Zelensky and Ukraine appeared more frequently in May than in October. For keywords like Putin and Russia, the occurrence is almost the same in both months.

This research has several limitations. The first is a shorter period of the observation narrowed to May and October of 2022. In the future, we plan to expand this preliminary analysis to the whole period since February 2022. The second one is methodological. We narrowed the detection of the sentiment to the VADER dictionary. The used method showed adequate performance for the task, but for the more granular analysis of sentiments, we plan to train a multi-class classification model.

ACKNOWLEDGMENT

This work has been fully supported by the University of Rijeka under the project number uniri-drustv-18-20.

REFERENCES

[1] C. A. Melton, O. A. Olusanya, N. Ammar, and A. Shaban-Nejad, "Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence," *Journal of Infection and Public Health*, vol. 14, no. 10, pp. 1505–1512, 2021, special Issue on COVID-19 – Vaccine, Variants and New Waves. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876034121002288

[2] C. Yan, M. Law, S. Nguyen, J. Cheung, and J. Kong, "Comparing Public Sentiment Toward COVID-19 Vaccines across Canadian Cities: Analysis of Comments on Reddit," *Journal of Medical Internet Research*, vol. 23, 08 2021.

[3] T. Ruan and Q. Lv, "Public perception of electric vehicles on Reddit over the past decade," *Communications in Transportation Research*, vol. 2, p. 100070, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772424722000208

[4] S. C. Long, B. Lucey, Y. Xie, and L. Yarovaya, ""I just like the stock": The role of Reddit sentiment in the GameStop share rally," *Financial Review*, vol. 58, no. 1, pp. 19–37, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/fire.12328

[5] D. C. Glik, "Risk communication for public health emergencies," *Annual Review of Public Health*, vol. 28, no. 1, pp. 33–54, 2007, pMID: 17222081. [Online]. Available: https://doi.org/10.1146/annurev.publhealth.28.021406.144123

[6] C. Cuello-Garcia, G. Pérez-Gaxiola, and L. Amelsvoort, "Social media can have an impact on how we manage and investigate the COVID-19 pandemic," *Journal of Clinical Epidemiology*, vol. 127, 06 2020.

[7] G. Eysenbach, "Infodemiology: The epidemiology of (mis) information," *The American journal of medicine*, vol. 113, pp. 763–5, 01 2003.

[8] A. N. Medvedev, R. Lambiotte, and J. Delvenne, "The anatomy of Reddit: An overview of academic research," *CoRR*, vol. abs/1810.10881, 2018. [Online]. Available: http://arxiv.org/abs/1810.10881

[9] C. Manning, H. Schütze, and L. Lee, *Foundations of Statistical Natural Language Processing*. Cambridge, Massachusetts: The MIT Press, 08 1999. [Online]. Available: http://nlp.stanford.edu/fsnlp/

[10] S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, "An Overview of Graph-Based Keyword Extraction Methods and Approaches," *Journal of Information and Organizational Sciences*, vol. 39, no. 1, pp. 1–20, Jul. 2015. [Online]. Available: //jios.foi.hr/index.php/jios/article/view/938

[11] S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, "Selectivity-Based Keyword Extraction Method," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 12, no. 3, pp. 1–26, 2016.

[12] P. K. Bogović, A. Meštrović, and S. Martinčić-Ipšić, "Topic Modeling for Tracking COVID-19 Communication on Twitter," in *Information and Software Technologies*, A. Lopata, D. Gudonienė, and R. Butkienė, Eds. Cham: Springer International Publishing, 2022, pp. 248–258.

[13] P. K. Bogović, A. Meštrović, S. Beliga, and S. Martinčić-Ipšić, "Topic Modelling of Croatian News During COVID-19 Pandemic," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2021, pp. 1044–1051.

[14] S. Beliga, S. Martinčić-Ipšić, M. Matešić, I. Petrijevčanin Vuksanović, and A. Meštrović, "Infoveillance of the Croatian Online Media During the COVID-19 Pandemic: One-Year Longitudinal Study Using Natural Language Processing," *JMIR Public Health and Surveillance*, vol. 7, no. 12, p. 15, 2021.

[15] S. Martinčić-Ipšić, T. Miličić, and L. Todorovski, "The Influence of Feature Representation of Text on the Performance of Document Classification," *Applied Sciences*, vol. 9, no. 4, pp. 743–770, 02 2019.

[16] K. Babić, M. Petrović, S. Beliga, S. Martinčić-Ipšić, M. Matešić, and A. Meštrović, "Characterisation of COVID-19-Related Tweets in the Croatian Language: Framework Based on the Cro-CoV-cseBERT Model," *Applied Sciences*, vol. 11, no. 21, p. 10442, Nov 2021. [Online]. Available: http://dx.doi.org/10.3390/app112110442

[17] D. Aljević, L. Todorovski, and S. Martinčić-Ipšić, "Extractive Text Summarization Based on Selectivity Ranking," in *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 2021, pp. 1–6.

[18] K. E. Anderson, "Ask me anything: what is Reddit?" *Library Hi Tech News*, vol. 32, no. 5, pp. 8–11, 2015.

[19] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, vol. 8, no. 1, pp. 216–225, 2014.

[20] P. Bump. (2022, Oct) 24 Reddit Stats and Facts to Know in 2022. [Online]. Available: https://blog.hubspot.com/marketing/reddit-stats

[21] Semrush. (2022, 03) Reddit Website Traffic, Ranking, Analytics. [Online]. Available: https://www.semrush.com/website/reddit.com/overview/

[22] N. Proferes, N. Jones, S. Gilbert, C. Fiesler, and M. Zimmer, "Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics," *Social Media + Society*, vol. 7, p. 205630512110190, 04 2021.

[23] B. Liu, "Sentiment Analysis and Subjectivity," in *Handbook of Natural Language Processing*, 2010.

[24] N. Savela, D. Garcia, M. Pellert, and A. Oksanen, "Emotional talk about robotic technologies on Reddit: Sentiment analysis of life domains, motives, and temporal themes," *New Media & Society*, p. 146144482110672, 12 2021.

[25] J. Hodges, M. Simonsen, and J. Ottwein, "Gifted Education on Reddit: A Social Media Sentiment Analysis," *Gifted Child Quarterly*, vol. 66, p. 001698622210764, 02 2022.

[26] J. Pennebaker, M. Francis, and R. Booth, "Linguistic inquiry and word count (LIWC)," *Mahway: Lawrence Erlbaum Associates*, 01 1999.

[27] A. Guerra and O. Karakuş, "Sentiment Analysis for Measuring Hope and Fear from Reddit Posts During the 2022 Russo-Ukrainian Conflict," 2023. [Online]. Available: https://arxiv.org/abs/2301.08347

[28] S. M. Mohammad and P. D. Turney, "NRC Emotion Lexicon," *National Research Council, Canada*, vol. 2, p. 234, 2013.

[29] Y. Zhu, E.-u. Haq, L.-H. Lee, G. Tyson, and P. Hui, "A Reddit Dataset for the Russo-Ukrainian Conflict in 2022," 2022. [Online]. Available: https://arxiv.org/abs/2206.05107

[30] D. Marx, "Python Pushshift.io API Wrapper (for comment/submission search)," 2018. [Online]. Available: https://github.com/dmarx/psaw

[31] The pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020. [Online]. Available: https://doi.org/10.5281/zenodo.3509134

[32] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.", 2009.

[33] The Kyiv Independent news desk. (2022) Putin Declares War on Ukraine. [Online]. Available: https://kyivindependent.com/national/putin-declares-war-on-ukraine

[34] Google. Google Trends. [Online]. Available: https://trends.google.com/trends/

[35] S. McCarthy. (2022, 09) Russia announces immediate 'partial mobilization' of citizens, escalating its invasion of Ukraine. [Online]. Available: https://edition.cnn.com/2022/09/21/europe/ukraine-russian-referendums-intl-hnk/index.html

[36] (2022, 10) Timeline of missile strikes on Ukraine on October 10. [Online]. Available: https://www.ukrainianworldcongress.org/timeline-of-missile-strikes-on-ukraine-on-october-10/

[37] (2014) Russia's invasion timeline. [Online]. Available: https://war.ukraine.ua/russia-s-invasion-timeline/

Classification of crimes using machine learning techniques for national crime data

Marija Trpchevska*, Aleksandra Dedinec M.Sc.*
*Faculty of Computer Science and Engineering, University Ss. Cyril & Methodius Skopje,
Republic of North Macedonia
marija.trpchevska@students.finki.ukim.mk, aleksandra.kanevche@finki.ukim.mk

***Abstract*—The paper aims to assist in a more accurate classification of crimes provided by North Macedonia's Ministry of Internal Affairs using machine learning techniques. By means of natural language processing, data is transformed into a format suitable for term frequency-inverse document frequency (TF-IDF) vectorization. Bayesian and Support Vector Machine models are utilized and evaluated. The results show non-negligible improvement toward a successful classification of crimes, confirming the beneficial nature of machine learning techniques towards such tasks.**

***Keywords*—Natural language processing, term frequency-inverse document frequency, TF-IDF, classification, multinomial Naive Bayes, Support Vector Machine**

I. INTRODUCTION

Classification as a concept is at the core of machine learning - data exists in such overflowing and daunting quantities in today's world that one needs to be able to filter through them by automated means should he wish to get to a more complex analysis. In past decades, machine learning models have diversified and have been tailored towards solving domain-specific classification tasks with a wide variety of inputs.

In all cases, however, the output is discrete; a classifier, as the name suggests, must choose between a finite number of categories with which to bind a given input. Outputs which have continuous values fall under the umbrella of regression problems. Both classification and regression are considered as so-called supervised learning approaches [1].The implied existence of an output itself means that someone (usually a human) has gone through the unenviable task of connecting certain inputs with corresponding outputs employing some sort of prior knowledge of the valid ways in which to make connections. A subset of these pairs is then fed to a classifier model whose task is to infer the rule-set used when those pairs were labelled. To determine whether the classifier has been trained well, those inputs previously unseen by it are given for classification (the output is therefore withheld) and its guesswork is compared. A good classifier mimics the decision rules made by a human, so most guesses should in principle coincide with the outputs.

The simplest case is one in which the choice is binary - between *'yes'* or *'no'*, *'malware'* or *'software'*, *'malignant tumor'* or *'benign tumor'*. Multi-class problems, like the one which concerns this paper, involve more than two classes. However, the same principles as before apply.

II. RELATED WORK

The whole endeavour rests upon previous work by [2] in which a concerted effort is made to turn informational bulletins sourced from the Ministry of Internal Affairs of the Republic of North Macedonia's website into a usable, interactive crime map showing all criminal activities made across the country in the past thirty days. This warrants text analysis of the bulletins detailing reported crimes in a given day in order to extract relevant facts i.e. the nature of the crime, broadly classified into seven categories: **weapons** (in the Macedonian implementation referred to as пиштол) (illegal possession and distribution of firearms or causing reckless endangerment while wielding a weapon), **violence** (насилство) (physical assaults), **theft** (кражба) (aggravated and armed robberies including burglaries), **documents** (документи) (paperwork forging), **drugs** (дрога) (illegal possession and distribution of narcotic substances), **traffic** (сообраќај) (traffic violations and reckless driving) and **other** (друго) (entries that do not fit into the preceding categories). Additionally, one must know the general location of the event (city); if the report contains a street address or village, they are combined to produce a more accurate location description. Values indicating the latitude and longitude needed for map construction are generated separately from the location. The dates of publication of the bulletin and when the crime has been committed are also recorded. The resulting crime map in question can be appreciated on *http://crimemap.finki.ukim.mk/home/en*.

Information extraction is done using basic keyword tagging (drawing conclusions based on the absence or presence of certain terms). Namely, the type of crime is determined by taking into consideration the terminology used in North Macedonia's Criminal Code and matching occurrences of it in the reports themselves. This proves of immense help when the official criminal offense is given in quotes, however many entries diverge from such conventions. In those cases, further processing based on a self-concocted dictionary of common words used to explain criminal deeds is made (for example, one would look for the word *'fight'* and relate the entry to **violence** or stumble upon the word *'shot'* and assume that it describes an event belonging to the **weapons** class).

As it pertains to the general field of forensics using machine learning, many approaches have been used to tackle crime statistics from a number of sources. For

example, [3] have employed similar methods in order to classify cybercrime offenses, proposing using Naive Bayes for classification and K-means for clustering, having performed TF-IDF feature extraction beforehand, achieving 99% accuracy. Meanwhile, [4] make a comprehensive analysis of Twitter posts, trying to ascertain any criminal sentiments expressed in the tweets and classifying them as concerning or not using a text mining-based approach. With a Random forest classifier, an accuracy of 98.1% is gained. This duo of data mining followed by machine learning is preferred by [5] which make use of WEKA, an open-source data mining software in order to gauge the correlation between violent crime patterns from the Communities and Crime Unnormalized Dataset provided by the University of California-Irvine repository and actual crime statistical data for the state of Mississippi. A linear regression model provides the greatest correlation coefficient as well as the lowest error value.

III. DATA SET SOURCING AND PRE-PROCESSING

It is worth mentioning that the crime map remains an active project maintained by the authors, with entries added on a regular basis. All in all, some 40399 records have been collected dating from 21st June 2011 up until today and counting. Naturally, a narrower range has been selected for initial work. Entries belonging to the period between June 2020 and October 2021 are further isolated, having 2000 distinct data points which are then labelled (that is, a proper class for the type of crime has been given manually). This labelled data set will be taken as the ground truth upon which classifier models are evaluated for fitness.

In regard to natural language processing, a SpaCy [6] library instance is utilized due to its newly found support of the Macedonian language (with a trained *pipeline* - a step-by-step analysis of textual input on multiple linguistic levels- generously provided by the Macedonian software company Netcetera; their efforts are noted in *https://blog.netcetera.com/macedonian-spacy-f3c85484777f*). A SpaCy pipeline object boasts myriad capabilities and features. For the purposes of pre-processing, the following functionalities are used: SpaCy's tokenizer (divides sentences into *tokens* - each representing a word or collection of words that have a unique semantic value; it differs from simple white-space separation due to the injected language sensitivity when the pipeline is trained. Thus, when tokenizing the sentence *'The U.K. stands for The United Kingdom.'*, a good tokenizer would correctly identify *'U.K.'* as one token, but split off the punctuation at the end of the sentence), Parts-of-Speech tagger (identifies *'.'* as punctuation, and *'U.K.'* as a proper noun), Named Entity Recognizer (recognizes entities based on their format or real-world role, so *'U.K.'* would be recognized as a GPE or geopolitical entity, *'$1 billion'* as MONEY or monetary currency, etc.), Rule-based Matcher (looks for tokens which conform to some shape, so *'U.K.'* could be found by searching for the shape *X.X.* in the provided sentence, *X* standing for any capital letter) and lemmatizer (determines the base form of a word; *'be'* is a base form of *'was'* and *'assault'* a base form of *'assaulted'*).

Having said that, each data point goes through a text pre-processing module which renders its structure into a form recommended for more effective model training. First, each report is stripped of extra white-spaces as they are superfluous and given to the tokenizer. Of the individual tokens, those with a punctuation Parts-of-Speech tag are removed. The same applies to so-called stop words (the most frequently used words in the Macedonian language). No doubt all entries contain some sort of date, time of day, or numeric quantity, playing a similar role to the stop words. If The Named Entity Recognizer labels a token as DATE, TIME or CARDINAL accordingly, it is removed. The Rule-based Matcher steps in to remove references of people (both perpetrators and victims are referred to by initials only with their age provided in brackets, so a shape *X.X.(dd)*, *d* being a digit, is filtered for). Finally, a lower-case form of each token (this being done because for all intents and purposes, to a computer, *'Crime'* and *'crime'* are different words and will be treated as such even though it is evident that they refer to the same concept) is provided to the lemmatizer. All surviving lemmas are again joined together with white-spaces to form a "clean" report. Table I shows a sample report that has gone through said transformations.

TABLE I: EXAMPLE OF A DATA ENTRY BEFORE AND AFTER PRE-PROCESSING

Original report	**Transformed report**
СВР Охрид поднесе кривична пријава против Н.В.(18) од Скопје, поради постоење основи на сомнение за сторено кривично дело „загрозување со опасно орудие при тепачка или караница“. Тој, на 27.08.2020 во авто-кампот Љубаништа, физички го нападнал А.Ј.(20) од Охрид, при што користел и остар предмет.	свр охрид поднесе кривичен пријава скопје постоење основ сомнение сторено кривично дело загрозување опасен орудие тепач караница авто камп љубаништа физички нападнал охрид користел остар предмет

IV. TF-IDF VECTORIZATION AND MODEL SELECTION

Next, each report is vectorized using TF-IDF [7]. What intuitively comes to mind when trying to work out the main topic of some textual input is to count the number of occurrences of each word and associate the meanings carried by quantitatively prevalent words with it. When this is done to all reports, one might get a matrix with 2000 rows (one per report) and quite a large number of columns (one per unique word) representing the whole *vocabulary* of the data set. A *feature vector* is one matrix row (a report represented by word counts), *features* being

the words themselves (in subsequent text, ”feature vector”, ”data point”, ”sample” and ”report” are synonymous - the text from one report is a data point/ sample from the data set; it is later transformed into a feature vector). It is assumed that the matrix is largely sparse with many null entries. Memory efficiency aside, a more serious problem arises: longer reports will, on average, have higher counts per word and be more closely associated with a topic than shorter ones, even though both seem to describe similar events. Furthermore, if stop words are not removed beforehand, this would lead to disproportionate counts for words such as *'the'*, *'it'* or *'an'*, themselves being of no positive contribution towards determining the topic. Punctuation and commonly used named entities put an additional burden on the vocabulary with unnecessary columns.

To alleviate the issue of long reports, one switches to working with relative scaling, so the count of each word in a given report is divided by the total number of words in it i.e. their term frequency is calculated (this being the preferred method of Scikit-learn [8] - a library specifically designed with machine learning in mind and which will be used to import modules in order to implement what has been discussed). Similarly, one could calculate the logarithm of a word count (and add 1 to avoid having to take the logarithm of zero), essentially smoothing the influence of irrelevant, but frequent words.

Another refinement on top of term frequency is to down-scale weights for words that occur in many reports and are therefore less informative than those that occur only in a smaller portion and might indicate topic-specific terms. In Scikit-learn, this is done by taking the logarithm of the ratio between the total number of reports in the data set and the number of reports which contain some word w (adding a 1 to this value would make sure that words occurring in all reports are not entirely ignored).

Initially, the data set is split into a training set (80%) and a testing set (20%). The 1600 reports belonging to the training set are then fitted to a CountVectorizer followed by a TfidfTransformer (analogous to the previously described process). A vocabulary of 4912 unique words is generated. Meanwhile, their labels go through a LabelEncoder, because they are non-numerical as is and models handle numerical, categorical values better (this can be omitted, however).

Since the classification is made with discrete features, a natural choice for a model is multinomial Naive Bayes. The multinomial distribution normally requires integer feature counts, but in practice, fractional counts such as TF-IDF may also work. It is widely used in natural language processing contexts (Bayesian learning approaches have been used traditionally to filter ”spam” e-mails from ”ham” or categorize products in e-commerce settings based on description alone). The naivete of a Bayesian model stems from the fact that it presupposes conditional independence of features in a feature vector i.e. under the assumption that a given report belongs to class C, the intrinsic syntactic or semantic dependence between the

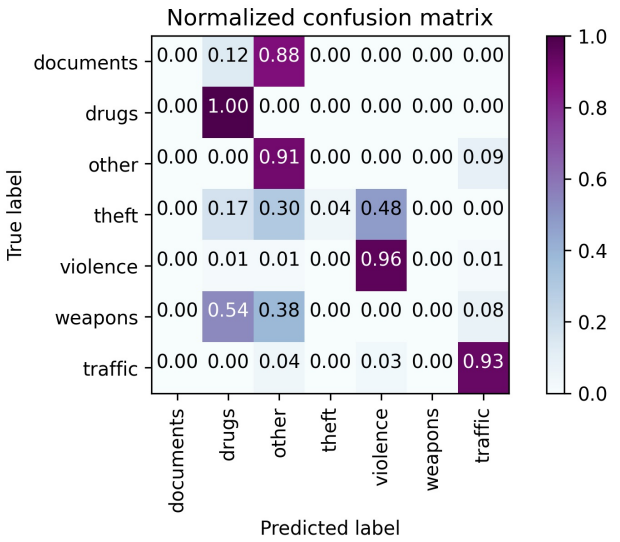


Figure 1: Confusion matrix generated by default-parameter multinomial Naive Bayes

words is ”explained away” by the class. That the word *'stabbed'* warrants occurrence of *'knife'* down the road is reduced to the probability that *'stabbed'* is present at all in a **violence** or **weapons** report (the same goes for *'knife'* and all the other words in that particular vector). These probabilities are multiplied, giving the probability that the report is part of class C. This is done for all possible values of C. Ultimately, the class with the highest probability will be chosen as the classifier's final label. While a MultinomialNB instance with default parameters is fitted to the training vectors and corresponding labels, the reports which are a part of the testing set go through CountVectorizer and are transformed only (not fitted) by TfidfTransformer. A confusion matrix in Fig. 1 shows the predicted labels and true labels accordingly. The values are normalized i.e. they closely correspond to the recall metric: if the predicted label of a data point equals that of the true label, then that point is considered a *true positive* with regard to the label examined. The others are mismatches and therefore constitute *false negatives*. The recall calculates the ratio between true positives and total predictions made (true positives + false negatives). In other words, in Fig. 1, 91% of points that actually belong to the class **other** are labelled by the classifier as such.

Unfortunately, initial results appear less than satisfactory. For classes **documents** and **weapons**, the classifier completely misses the mark and when unsure, defaults to the class **other** since the entries tend to vary most in its midst. **Theft** reports share a similar fate while all others give a surprisingly solid recall rate. The confusion between **drugs** and **weapons** can be explained by inter-sectional language (drug deals tend to be made between armed individuals); the same could be interpolated for false classifications of **theft** as **violence** (some sort of physical violence is needed when a planned theft is interrupted).

Report	True Label	Predicted Label
свр охрид поднесе кривичен пријава двајца малолетник село елшани малолетник охрид постоење основан сомнение сторени кривичен дела тешка кражба кривичен дело тешка кражба обид пријавените период градинка населба лескајца охридско употреба физичка сил насилен влегле одзема предмет	кражба	насилство
полициски службеник единица внатрешен работи свр велес подрачје велес лиши слобода велес била распишан национален потерница издржување казна затвор наредба основен кривичен суд скопје сторени кривичен дела недозволен изработување држење тргување оружје распрскувачки материја злосторничко здружување шлионажа измама прикривање уцена бил приведен полициска стан велес бил спроведен казнено поправен установа идризово издржување казната затвор	пиштол	друго
граничен премин деве баир пасош контролабило утврден извршен бришење штембил патната исправа р.албанија постоење основ сомнение сторено кривично дело фалсификување исправа бил лишен слобода предаден полициски службеник полициски стан крива паланка понатамошен постапка	документи	друго

Figure 2: Classification errors using default-parameter multinomial Naive Bayes

Fig. 2 provides examples where some discrepancy between true and predicted labels exists. A huge issue previously unmentioned is the unbalanced state of the data set. For it to be balanced, all classes need to share an equal representation. As it stands, the distribution of classes is as follows: 27.45% **other**, 18.80% **traffic**, 18.60% **drugs**, 21.25% **violence**, 7.25% **theft**, 2.05% **documents**, 4.60% **weapons**.

Taking these statistics into consideration, the confusion matrix seems to indicate a lack of data points for the three least represented classes in the data set. This will be remediated in further trials; there are no attempts to improve the classifier itself since Naive Bayes classifiers have a very limited parameter set and hyper-parameter tuning is unlikely to lead to a performance boost. TF-IDF already uses log probabilities, so there's no small-number problem. The test set itself has no zero-observation issue (all classes are present with some samples in the testing set, confirmed by looking at the support column of the classification support metric). It *does* confirm the imbalance seen in the data set, though: there are just 8 data points present for the class **documents** compared to the 118 for **other**. Pre-processing steps have already been taken in regard to the textual input itself. [9] propose some relief by smoothing out word counts based on the prominence of the class in the data set; when calculating the probability of a word occurring in a text of a given class *C*, a weighted average of the word count divided by the total number of occurrences of all words in reports of class *C* is used instead, with promising results.

Focus consequently shifts from Naive Bayes to Support Vector Machines (SVMs). SVMs have enjoyed much adoration for classification tasks because they provide robust decision boundaries between categories. They treat feature vectors as being sprawled across some vector space and then try to find a margin (linear or otherwise) which will not only separate classes but be as further from actual data points as possible, building a "buffer zone" with which to reduce possible erroneous classification on a testing set (it is better to have some physical distance between the boundary than thread the needle with one which is very close to a vector as minimal disturbances might result in the unwanted crossing of the boundary).

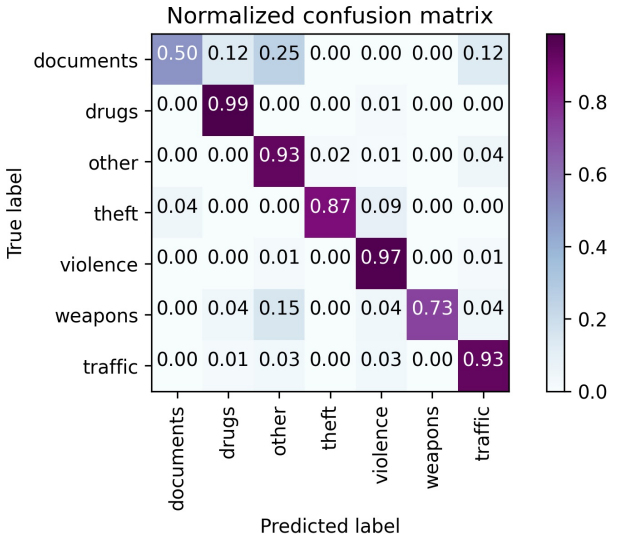


Figure 3: Confusion matrix generated by Support Vector Machine using Stochastic Gradient Descent

Those vectors closest to the margins "support" the decision of the model, hence the name. As per [10], SVMs prove to be suitable for text classification due to their handling of high-dimensional feature vectors well, especially when most features have some relevancy to the prediction (text pre-processing assures this is the case). As mentioned, even though each feature vector has a shape 1×4912, most entries are zeroes and SVMs are suited for such "dense concepts, sparse instances" formats. Finally, problems belonging to this type turn out to be linearly separable, so simpler models could account for a vast array of applications.

Returning to practical implementation, a linear SVM using Stochastic Gradient Descent is chosen (the model updates its parameter values after going through one sample at a time (or a small subset of samples - a *batch*) instead of taking the whole data set and *then* opting to tweak them. This results in a model that jumps quickly and easily through possible values for parameters while at the same time getting a correct rough estimate (order of magnitude, sign, etc.) with few samples. Therefore, an SGDClassifier is initialized with a *loss* parameter equal to *'hinge'* (it gives a linear SVM), *penalty* equal to *'l2'* (prefers simpler models when possible), *alpha* equal to *'0.001'* (how hard to penalize), a *random_state* equal to a random integer (e.g. *'42'*) so the order of the samples in the training set is reproducible across function calls because *max_iter* is set to *'5'* (how many times to call the fitting function, equivalent to the number of times to pass over the whole training data - no. of *epochs*). The resulting confusion matrix is shown in Fig. 3.

Indeed, a much better recall rate is provided for all classes compared to Naive Bayes, although **documents** and **weapons** could use more samples to differentiate them better from other types of crime. Consequently, this setup is chosen for further improvement by hyper-parameter tuning using Grid Search. This refers to the process of giving a range of values for parameters and

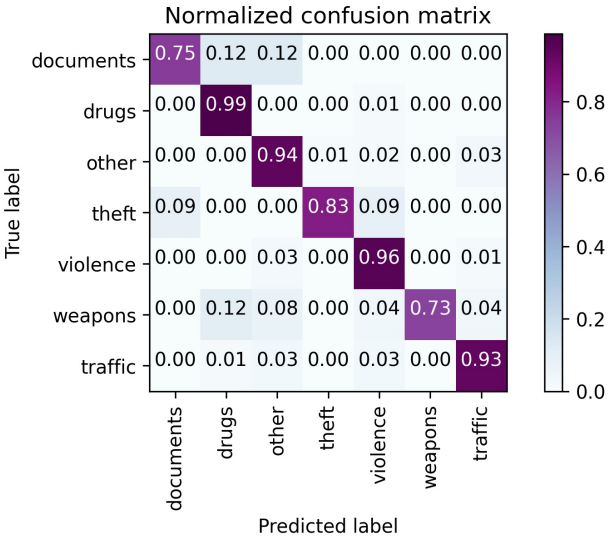


Figure 4: Confusion matrix generated by Support Vector Machine with Grid Search-tuned hyper-parameters

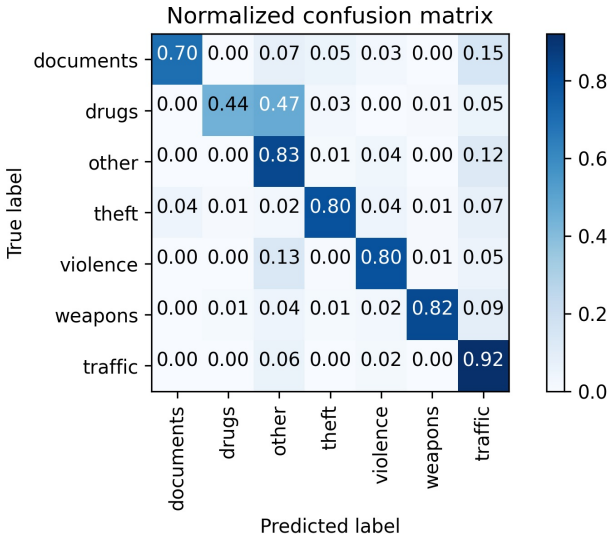


Figure 5: Confusion matrix from labels provided by crime map authors for the same data set

letting the model try out various combinations from the provided lists, eventually giving a configuration that has performed best (each configuration is tested on a subset of the provided data set using a five-fold cross-validation technique: the data set is randomly split five-ways, one of those splits (*folds*) is used as a testing set, the others are used for training; some accuracy score is achieved and the model is discarded. This is done five times so each fold gets to be a test set. The evaluation score of the model is averaged out from those five accuracy scores).

The CountVectorizer's *max_df* parameter is refined (this is a threshold value; words that have a document frequency strictly higher than the given threshold are ignored) as well as *ngram_range* (n-grams are collections of *n* consecutive words; unigrams are individual words, bigrams are formed by two words found next to each other in a report, etc. The parameter value indicates up to which *n* to build. A vocabulary is formed from all n-grams, substantially increasing its size, but hopefully, its predictive abilities to boot as words reveal more information when taken holistically and reflect their linguistic connections). The TfidfTransformer is tested upon its *use_idf* parameter (its role is self-evident) while the SGDClassifier is altered for better *loss*, *alpha* and *penalty* values. The resulting output showing the best parameters for the model is shown below:

```
{'sgdc__alpha': 0.0001, 'sgdc__loss':
'hinge', 'sgdc__penalty':
'l1', 'tfidf__use_idf':
True, 'vect__max_df': 0.75,
'vect__ngram_range': (1, 2)}.
```

New instances of the counter, vectorizer and classifier are made, initialized with the values provided by Grid Search and a new confusion matrix is constructed (Fig. 4).
As expected, the model has tuned its parameters in such a way as to provide minimal errors across classes, which has the effect of shifting some margins slightly and contributing to the major improvement of recall for **documents** (on the back of some minimal loss of recall in **theft** for example; this is a game of trade-offs after all).

V. COMPARISON OF MODEL AND CRIME MAP CONFUSION MATRICES

All that is left to do with regard to the current data set is to compare the matrix from Fig. 4 with the one made using labels for the samples made by [2] (Fig. 5) and determine whether there exists improvement in correct classification of crimes using the methods so far outlined (by comparing them with the ground truth as has been done with the classifiers).
In all classes except **weapons**, the model outperforms and gives better recall rates- substantially in the case of **drugs** (125% relative increase in recall) and marginally in others (below 20% relative increase). This is no cause for concern as the matrix itself appears quite satisfactory, so this should be considered an attempt at bettering an already solid outcome.

VI. CREATING A BALANCED DATA SET. SPLITTING CLASS 'OTHER' INTO MULTIPLE CLASSES

Since the matter of an unbalanced data set has been raised, it is of great interest to explore how dire of an issue it really is and test the Grid-Search SVM against one in which said issue has been mitigated. While at it, the class **other** is analysed and additional four types of crime extrapolated in its stead: **COVID-19** (breaking safety protocols by disobeying movement restrictions while positive for the SARS-CoV-2 virus; vaccination certificate forgeries fall under **documents**), **dog attacks** (there are overwhelming reports of individuals bitten by stray dogs), **disappearances** (cases of persons going missing or running away- not a crime per se, but such reports

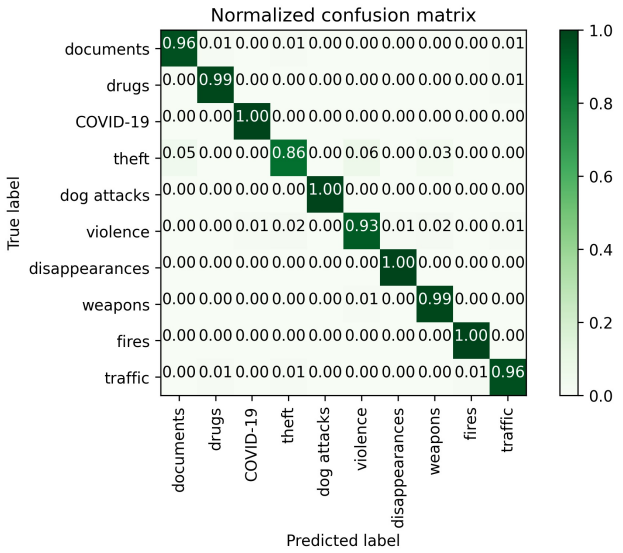


Figure 6: Confusion matrix from balanced data set by Support Vector Machine with Grid Search-tuned hyper-parameters

are common enough as to justify a separate category), **fires** (fires breaking out, intentionally or unintentionally e.g. as a result of a car crash).

Both old and new types of crimes will have around 370 different samples representing them in the new data set in order to secure proportional involvement. It might come as no surprise that a wider time period needs to be considered to achieve that. Therefore, entries spanning from June 2020 up until June 2022 are taken. If some class still lacks samples, even earlier records are added. Each new sample is manually labelled with the appropriate class and the whole data set is randomly permuted in order to break apart blocks of entries with the same label (to prevent losing the balance when splitting the set into training and testing sets, stratified splitting in which the proportion of samples from each class mirrors that of their distribution in the larger set is used).

In the end, a data set of 3694 unique data points (pre-processed into TF-IDF feature vectors) is given to the Grid-Search SVM classifier. A vocabulary composed of 32357 unigrams and bigrams is returned.

The results generated are near-perfect, as can be seen in Fig. 6; all but one of the recall rates hit the high nineties (the somewhat smaller percentage for **theft** could be due to some ambiguities regarding certain crimes. Cases where individuals have monetarily damaged institutions by deliberately withholding expenses and pocketing the difference themselves could very well be categorized as **theft** and **documents** (since the crime technically involves forging official documents). The same argument could be said for cases involving counterfeit money spending (in essence a type of forgery and apt for a **documents** label) and yet appropriate for **theft** as someone was deprived of the value of the monetary transaction. Nevertheless, trying to catch such slight variations in model predictions borders on overfitting the data- thus, no further corrections to the model itself are undertaken.

VII. CONCLUSION AND FUTURE WORK

By conforming to the best practices for handling textual information and careful analysis of model behaviour, a non-negligible improvement towards a successful classification of crimes is made, once again confirming the beneficial nature of machine learning techniques towards such tasks.

With much confidence, these same principles could be put to use in order to improve other aspects concerning the construction of a crime map. For example, more accurate location extraction could be made by semantic analysis of reports (natural language processing comes at hand immediately) so the place of birth of a suspect or victim is distinguished from the actual location of the crime in the eyes of the model. Such an undertaking might require training a Named Entity Recognizer from scratch, feeding into it a whole range of geographical locations (SpaCy's Macedonian pipeline is partially trained on Wikipedia texts written in Macedonian, so one would suppose that some geography-related articles are covered, however, it lacks much of the desired sensitivity to geolocations; this, in essence, means that further training and manual labelling of data is required, especially for street names, hospitals, banks, border crossings, etc.). Such a gargantuan undertaking nearing the informational scope of Google Maps is, of course, in need of more care and attention by a larger group of members and will be considered once those resources are secured.

REFERENCES

[1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2010.

[2] M. Jovanovik, D. Trajanov, I. Mishkovski, and D. Temelkovski, "Towards open data in macedonia: Crime map based on ministry of internal affairs' bulletins," *The 9th Conference for Informatics and Information Technology (CIIT)*, 2012.

[3] R. Ch, T. R. Gadekallu, M. H. Abidi, and A. Al-Ahmari, "Computational system to classify cyber crime offenses using machine learning," *Sustainability*, vol. 12, no. 10, p. 4087, 2020.

[4] S. Lal, L. Tiwari, R. Ranjan, A. Verma, N. Sardana, and R. Mourya, "Analysis and classification of crime tweets," *Procedia computer science*, vol. 167, pp. 1911–1919, 2020.

[5] L. McClendon and N. Meghanathan, "Using machine learning algorithms to analyze crime data," *Machine Learning and Applications: An International Journal (MLAIJ)*, vol. 2, no. 1, pp. 1–12, 2015.

[6] Y. Vasiliev, *Natural language processing with Python and spaCy: A practical introduction*. No Starch Press, 2020.

[7] A. Rajaraman and J. D. Ullman, *Data Mining*. Cambridge University Press, 2011, p. 1–17.

[8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[9] E. Frank and R. R. Bouckaert, "Naive bayes for text classification with unbalanced classes," in *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer, 2006, pp. 503–510.

[10] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," in *European conference on machine learning*. Springer, 1998, pp. 137–142.

Optimizing Vision Transformer Performance with Customizable Parameters

E. Ibrahimovic
The Second Gymnasium, Sarajevo, Bosnia and Hercegovina
emira1.ibrahimovic@gmail.com

***Abstract* – This paper experimentally examined the effects of changing the size of image patches and the number of transformer layers on the training time and accuracy of a vision transformer used for image classification. The transformer architecture was first introduced in 2017 as a new way of processing natural language and has since found applications in computer vision as well. In this experiment, we trained and tested fourteen versions of a vision transformer on the CIFAR-100 dataset using graphical processing units provided by Google Colaboratory. The results showed that increasing the number of transformer layers and decreasing the patch size both increased test accuracy and training time. However, learning curves generated by the models showed overfitting for very small patch sizes. Overall, changing patch size had a greater impact on accuracy than changing the number of transformer layers. The results also suggested that transformers are more resource-intensive to train than other models. We suppose that including a classification token could lead to shorter training times, but another experiment is needed to examine its influence on accuracy.**

Keywords* – *vision transformers; computer vision; nerual networks;

I. Introduction

Teaching machines to see and process the world around them similarly to a human is an increasingly popular task as technology is developing. Computer vision solutions are automating processes in manufacturing, surveillance, optical character recognition and robotics by reducing the amount of human error and decreasing operational costs [1], [2].

Convolutional neural networks are currently the most commonly used architecture in computer vision solutions. Even though CNNs reach accuracies as high as 91.7% on certain data sets [3], on others accuracy is limited by the amount of data and processing power available [2]. Use cases such as vehicle detection, human pose estimation and medical imaging require higher accuracies than those currently achieved [2].

New machine learning paradigms which cope with this issue are emerging and being developed at an accelerated pace [4]. The most recent development has been the introduction of the vision transformer. The transformer architecture was first introduced in [5] as a new way of processing natural language and has since become the default in the field [6].

This paper explores the effect of changing the number of self-attention layers and size of image patches on the test accuracy and training time of vision transformer models. The relationship was explored experimentally by implementing a ViT on the platform Google Colaboratory [7]. Expected results of the experiment were that test accuracy would increase as image patch size decreased. The accuracy would ideally increase as the number of self-attention layers increased. The training time was expected to increase as the number of layers increased and decrease as image patch size decreased.

II. Theoretical Background

A. Neural Networks

A neural network is a computing system which produces a set of outputs from a set of inputs by performing a complex series of computations. For classification purposes, its output is the likelihood that the input belongs to any of a number of pre-determined classes. [8]

A neural network is made up of units called nodes, which are arranged in layers. A network consists of one input layer, multiple hidden layers and one output layer. In the simplest case, layers are ordered in a row, with each layer except for the input connected to the layer before it. Each node except for those in the input layer is connected to all the nodes in the previous layer, along with one additional node named the bias unit. The connection between every pair of nodes is assigned a value called the weight. This is illustrated in Fig. 1. [9]

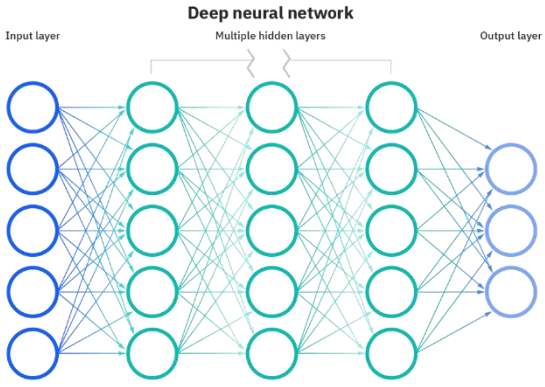


Figure 1. Each node has a weighted connection to all nodes from the previous layer [9]

The input layer contains as many nodes as there are inputs. Each input node holds a floating-point value, called the activation, corresponding to the input it represents.

The neural network performs calculations layer by layer. Each node in layer *i* takes the activation of a node from layer *i-1* and multiplies it with the weight of the connection between them. These products, along with the weight of the connection with the bias unit, are summed to get some value *S*. This value is passed to an activation function and its output becomes the current node's activation. After activations of nodes in layer *i* are all determined, the process is repeated for the following layer, stopping once the values in the output layer are computed [8].

The activations of the output nodes represent the output of the network. For classification purposes, each node in the output layer corresponds to one of the classes in the classification problem. The activation of a node in the output layer represents the neural network's certainty that the input belongs to the respective class. The class with the highest certainty is considered the network's answer to the classification problem [8].

The weights in a neural network are initially unknown and therefore assigned randomly. They are adjusted through training: a process in which a network processes a series of inputs with known outputs [8].

The type of neural network described here is a type of feed-forward neural network (FFNN) known as a Multi-layer Perceptron (MLP). It is also the basis on which all NNs work [10], [11], [9].

Convolutional neural networks have been the standard model for image classification and have shown significant success on standard computer vision benchmarks, even surpassing human-level performance on some [12]. However, there are still open problems which they face such as: difficulty to justify their successes theoretically, difficulty recognizing image manipulations designed to trick them, and difficulty giving multiple labels to images and describing the contents of an image [12].

The vision transformer (ViT), a neural network architecture inspired by the golden standard for natural language processing could prove successful in solving some of these issues. Apart from achieving higher accuracies on several public datasets than the best CNNs [13], the ViT also had a shorter training time. They have shown promise in various recognition tasks, generative modeling, multi-modal tasks, video processing, and 3D analysis [14].

B. Vision Transformers

The transformer is a neural network architecture whose job is to learn the context created by a set of inputs [15]. These inputs, called tokens, are all parts of a larger whole, e.g., words in a sentence or patches of an image. The transformer is able to use the relationships between these tokens to assign a new meaning to each of them, one which conveys the context in which each token is situated [6]. A transformer is composed of a number of identical layers called encoder blocks.

The initial input to the transformer is given to the first encoder in the stack. Each subsequent encoder receives as inputs the outputs of the encoder below it. The output of the final encoder is the output that is returned by the transformer [5]. An encoder block consists of 2 layers: a layer of self-attention, whose output becomes the input to the second layer – a feed-forward neural network. Self-attention is a transformer's primary mechanism of action.

Information about each token is embedded in a vector, where each value in the vector describes a particular aspect of the token. In the case of the vision transformer, each token embedding is a representation of a patch of an image. There are a few steps to turning an image into a sequence of token embeddings.

The input image comes in the form of a tensor: a H×W matrix, where each cell in the matrix contains C values, one for each of the primary colors, also called channels.

The image is broken down into N patches of size P×P ($N = \frac{HW}{P^2}$). The patches are then flattened into (CP^2)-dimensional vectors and made into rows of the matrix $\boldsymbol{x_p}$ of dimensions $N \times (CP^2)$. Next, these N flattened patches are transformed via a linear projection E into N D-dimensional vectors. This linear projection is learned during the training process. The results of these transformations are called patch embeddings [13].

The final piece of information that is included in the token embedding for a particular patch is its position in the image. This is done by adding a D-dimensional position embedding to the patch embedding. The position needs only reflect the patch's 1D position in the $\boldsymbol{x_p}$ matrix. The exact position embeddings are also one of the parameters that are learned through training [13].

The number of the transformer's outputs is the same as its number of inputs. Each output represents the contextualized embedding corresponding to one of the input tokens. Every output embedding will therefore contain information about the entire image, but the focus of its "knowledge" is on the specific image patch it corresponds to. It is then somewhat, but not entirely, suitable to classify the image based on one specific token [6].

For this reason, a special classification token is passed as input along with the image patch tokens. This token initially contains no information about the image, but as it is passed through the encoder stack along with other tokens, it gets imbued with context that is not localised to any specific part of the image [6]. It is the value of the output corresponding to this token that is used to classify the image at the end. The initial value of the classification token's embedding is a parameter learned through training [13].

The transformer's input can be summarized with the formula:

$$\boldsymbol{z_0} = \left[\boldsymbol{x_{class}}; \boldsymbol{x_p^1 E}; \boldsymbol{x_p^2 E}; \ldots; \boldsymbol{x_p^N E}\right] + \boldsymbol{E_{pos}} \qquad (1)$$

where: each row represents a token embedding passed as input to the transformer, $\boldsymbol{x_{class}}$ represents the initial embedding for the classification token, $\boldsymbol{x_p^i}$ represents the *i*-

th row of the matrix $\boldsymbol{x_p}$ of flattened image patches, $\boldsymbol{E} \in \mathbb{R}^{(P^2C)\times D}$ is the linear projection which transforms a flattened image patch into a D-dimensional patch embedding, $\boldsymbol{E_{pos}} \in \mathbb{R}^{(N+1)\times D}$ is the matrix of positional embeddings, where row i represents the positional embedding of token i.

Self-attention is the mechanism that allows the transformer to gain a further understanding of a token by examining the tokens which surround it. The token being contextualized in the current step is called the query token and the surrounding tokens are called key tokens.

Given a query token and a sequence of key tokens, self-attention computes a weighted average of values assigned to the key tokens. Each weight is computed during run-time and is dependent on the relationship between the query token and a corresponding key token. If the relationship is in some respect strong, the weight associated with that key will be greater, increasing the influence of the value attached to the key on the final output [6].

The process is implemented using 3 2D matrices: the key, query, and value matrices, all of which are learned during training. Each of the token embeddings received as input is multiplied by the key matrix and the query token embedding is multiplied by the query matrix. Every one of the transformed keys is then multiplied via cross-product with the transformed query to obtain an attention score, which serves as a measure of similarity between a key and the query. Then, each of the initial key token embeddings is multiplied by the values matrix and by its attention score. These final embeddings are added together to reach the final result [6].

Vision transformers use multi-headed self-attention. The sequence of operations described above is performed by every attention head, only with different parameters in the key, query, and value matrices. This way, one output embedding is calculated by each attention head. These embeddings are passed as input to a feed-forward neural network to combine them into a single embedding [6].

A vision transformer takes an image as input and segments it into N equal-sized patches. These patches are transformed via a linear projection into patch embeddings. The embeddings, along with one classification embedding, pass through a number of encoder blocks composed of multi-headed self-attention and feed-forward neural networks. The last encoder in the stack outputs N+1 contextualized embeddings corresponding to the N+1 input tokens.

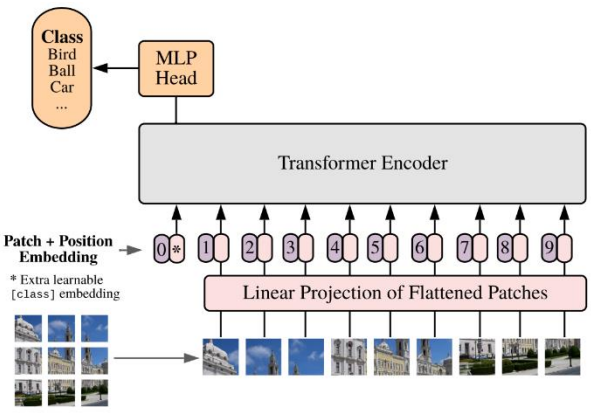


Figure 2. Overview of vision transformer architecture [13]

The embedding corresponding to the classification token is given as input to a multi-layer perceptron whose number of outputs is the same as the number of classes in the classification problem [13]. The output activations will be the desired probability distribution for the input image over all classes in the dataset. A complete overview is given in Fig. 2.

III. EXPERIMENTAL METHODOLOGY

An experiment was conducted to obtain primary data which will be used to answer the research question. 14 versions of a vision transformer were trained and tested on the platform Google Colaboratory [7]. All code was run on graphical processing units provided by Colab. The ViTs were programmed in Python 3.0 using TensorFlow's Keras, adapting code by [16].

A. *Variables*

Independent variables in models are:

- Number of transformer layers – The values chosen for the experiment were 2, 4, 8, 16. The values are in geometric progression so that changes in the dependent variables are clearly noticeable.
- Image patch size – The values chosen for the experiment were 2px, 4px, 8px, and 16px. Images in the data set are 32px × 32px.

The two independent variables were manipulated within the code. The only exceptions to the range of values chosen were pairs (P, L) ∈ {(2,8), (2,16)} because the training of these models exceeded usage limits set by Colab.

We examined the following dependent variables:

- Training time – When training a TensorFlow model, an overview of the training process is output to the terminal. Among the values in this summary is the time taken to train the model. This is the value taken for this dependent variable.
- Accuracy – Among the values in the overview of model training is the test accuracy achieved after training. This is the value taken for this dependent variable. Accuracy is the quotient of the number of images whose classes were correctly identified and the total number of images in the test set. Accuracy was chosen as an evaluation metric to be able to compare results to [3], [13].

B. *Dataset*

The dataset used for the training and testing of the models is CIFAR-100 which contains 60 000 images of dimensions 32× 32 pixels [17]. Each image is assigned to a class which describes the contents of the image. There are 100 classes with 600 images per class, examples of which can be seen in Figure 3 below.

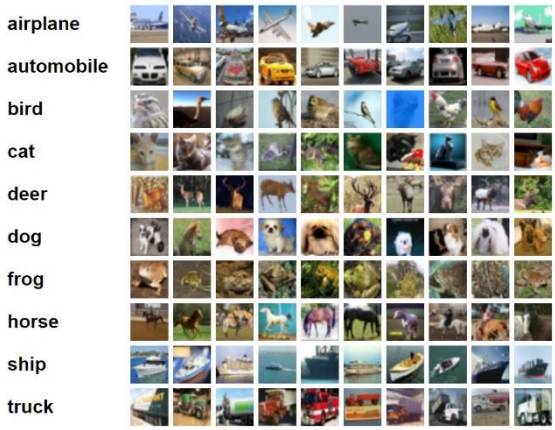


Figure 3. 10 randomly selected images from 10 classes in CIFAR-100

During each run, the dataset was split into 3 groups called the training, validation, and test sets. For each run, the test set consisted of the same set of 10 000 randomly selected images such that there are 100 images in each of the 100 classes. The validation set always consisted of 5 000 images and the test set of the remaining 45 000, though the exact images were randomly selected at the start of each run.

C. *Model Layers*

Each of the 14 models consisted of 3 parts: the input image processing layer, multi-headed self-attention, and the classification head. In the input processing layer, $\frac{32^2}{P^2}$ image patches of dimensions P×P are created and linearly projected into $\frac{32^2}{P^2}$ patch embeddings of dimension 32. The parameters of the linear projection are learned during training. This is implemented with Keras layer Conv2D [18].

Next, come L encoder blocks are implemented according to [13]. The first layer in the block is Keras Layer Normalization [19], which normalizes the activations of the layer that came before it. Doing this helps reduce training time [20]. This becomes the input to Keras Multi-Head Attention layer with 4 attention heads [21]. The outputs of this layer are added to the inputs to the block and passed to another normalization layer. Then, an MLP with one 64-node layer and one 32-node layer is attached. Finally, the outputs of the MLP are added to the outputs of the multi-head attention layer and the inputs to the block.

The classification head first normalizes the outputs of the final MHA block, then flattens the set of embedding vectors into a single vector. This is passed to a dense layer with 2048 nodes and then to another with 1024 nodes. A dense layer with 100 nodes corresponding to the 100 classes in CIFAR-100 gives the final output.

The models were trained in 200 epochs with batch size 100.

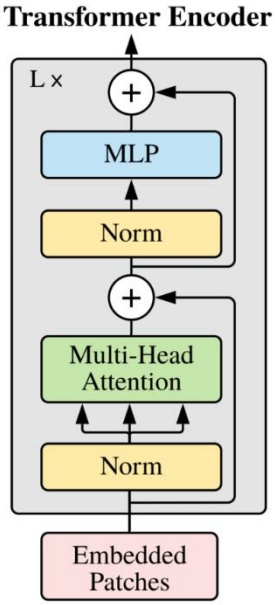


Figure 4. Multi-Headed Attention Layer Block

IV. RESULTS

A. *Model Training Time*

Training models with more transformer layers resulted in longer training times, as is shown in TABLE 1. Decreasing patch size, and therefore increasing the number of patches led to longer training times, with the exception of the transition from (P=16, L=4) to (P=8, L=4). These results have a common reason. Both increasing the number of patches and the number of self-attention layers resulted in a larger number of trainable parameters in the model (TABLE 2).

After processing each batch of input images, the transformer adjusts a large proportion (determined by the dropout rate) of parameters. Each calculation resulting from inputting an image takes one operation per parameter, and each adjustment of a parameter requires a machine to do at least one operation. Considering the fact that the number of operations a machine has to do is directly proportional to the time taken to complete them all, it follows that increasing the number of parameters would necessarily lead to an increase in training time. However, the decrease in training time seen in the transition from (P = 16, L = 4) to (P = 8, L = 4) where the increase in the parameter number is still present suggests that the relationship between parameter number and training time is not so straightforward.

TABLE 1. TRAINING TIME IN SECONDS [22]

Size of patch	Number of self-attention layers			
	2	4	8	16
16	802	1143	1171	3121
8	803	1059	1930	3303
4	1121	1814	3219	6011
2	5931	10 895	×	×

TABLE 2. NUMBER OF TRAINABLE PARAMETERS [22]

Size of patch	Number of self-attention layers			
	2	4	8	16
16	2 531 780	2 574 020	2 6585 00	2 827 460
8	3 299 780	3 342 020	3 426 500	3 595 460
4	6 440 900	6 483 140	6 567 620	6 736 580
2	19 022 660	19 064 900	×	×

Decreasing patch size from P=8 to P=4 with L=16 leads to an increase in parameter number of 3 141 120 and a near doubling of training time. At the same time, increasing the number of layers from L=8 to L=16 with P=4 leads to an increase in parameter number of just 168 960 but a near doubling in training time. This suggests that some parameters either require more operations to be adjusted or more operations are needed to compute outputs in a specific layer. The increase in parameter number resulting from a decrease in patch size is almost entirely due to the third to last layer in a model: a dense layer converting all transformer outputs into outputs. Adding another transformer layer increases parameter number by 21 120, regardless of the initial shape of the input. This result complied with a study performed by [23].

Examining the structure of these two layers gives us a possible explanation as to why one is more computationally intense than the other. During self-attention, each of the N outputs need to consider all N outputs from the previous layer, all of which vectors, and perform a series of matrix multiplications to get a result. This means that calculating all N outputs result in an order of magnitude of N^2 operations, assuming that matrix multiplication takes an order of magnitude of 1 operations. Given that the models are working with 4 attention heads each and with L layers, the final number of operations for this stack of layers is an order of magnitude of $N^2 \times L$. On the other hand, the third final dense layer in the model flattens these N vectors into a single vector of dimension $32 \times N$ which becomes its input, and returns 2048 outputs. Each of these 2048 outputs require the processing of all inputs, which is done by a single matrix multiplication, resulting in a number of operations in an order of magnitude of N. This means that as N and L increase, the time taken for the transformer encoder to process the image grows at least a whole order of magnitude faster than the time it takes for the MLP to process it.

This does not give an explanation for the decrease in training time from (P = 16, L = 4) to (P = 8, L = 4) The best answer which can be offered is either randomization steps during training, or the Google Colab platform assigning different types of GPUs to these two runs.

B. Model accuracy

Decreasing patch size always led to an increase in test accuracy, as is shown in TABLE 3. Increasing the number of self-attention layers generally led to an increase in accuracy, with the exception of the transition from (P = 16, L = 4) to (P = 16, L = 8). However, none of the accuracies obtained with these models comes close to accuracies achieved by convolutional neural networks of comparable size [3]. Searching for the reasons for this leads us to examine the learning curves produced by the models.

Learning curves [22] for models with the same patch size have nearly the same shapes, whereas learning curves for models with the same number of transformer layers differ greatly. This tells us that patch size could have a greater impact on accuracy than the number of layers. The test accuracy generally displays this pattern, despite great variation in the ranges of values obtained for one fixed variable, as shown in

TABLE 4 and TABLE 5 below. We can see that range expressed as the percentage of the mean when the number of transformer layers is kept constant is nearly always greater than the same value when patch size is kept constant. This suggests that changing patch size has a greater percentage impact on accuracy than changing the number of transformer layers does.

Observing the learning curves for P=2, P=4, and P=8, we determine they are consistent with signs of overfitting. For example, on the accuracy graph for P=4 ∧ L=8 we see that training accuracy decreases logarithmically as epochs go on, reaching a final accuracy of around 80%. The validation accuracy increases significantly within the first 25 epochs, then slows down its increase until it plateaus after the 100th epoch at around 40% accuracy. Overfitting can be solved by having a dataset with more training images, which could suggest why the ViT trained by [13] achieved such high accuracies: their models were pre-trained on datasets with 15 and 300 million images.

We can see that with increasing patch size the training and validation accuracy curves began to converge, as did the training and validation loss curves. This is consistent with signs of underfitting. Even though the training and validation accuracy curves for P=16 converge, they are both worryingly low at around 28% during the entire run. Increasing patch size while keeping the projection dimension constant led to the network trying to encode more data into the same projection vector, losing useful information in the process.

TABLE 3.TEST ACCURACY (%) [22]

Size of patch	Number of self-attention layers			
	2	4	8	16
16	28.77	29.31	28.8	29.48
8	32.76	35.12	36.17	36.94
4	35.20	39.43	41.81	44.19
2	35.82	38.88	×	×

TABLE 4. RANGE OF ACCURACIES FOR CONSTANT PATCH SIZE [22]

Size of patch	Range of accuracies	Mean accuracy	Range as % of the mean
16	0.71	29.09	2.44
8	4.18	35.25	11.86
4	8.99	40.16	22.39

TABLE 5. RANGE OF ACCURACIES FOR CONSTANT NUMBER OF TRANSFORMER LAYERS [22]

Number of layers	Range of accuracies	Mean accuracy	Range as % of the mean
2	7.05	33.14	21.27
4	10.11	35.69	28.33

V. CONCLUSION

This paper experimentally investigated the effects of changing the size of image patches and number of transformer layers on the training time and test accuracy of a vision transformer.

The results confirmed that training time increases with decreasing patch size and increasing number of transformer layers. Both of these changes correspond to an increase in the number of trainable parameters in the model, which leads to increasing the number of operations a machine has to perform, ultimately resulting in more training time. The results also showed that adding transformer layers has a greater impact on increasing training time than decreasing patch size. In particular, the change in training time per change in parameter number was generally significantly greater for parameters added by transformer layers. This suggests that transformers are inherently more resource intensive to train, confirming results by [23].

The experiment results have shown that models increasingly exhibited signs of underfitting as patch size increased, as well as that they overfit for small patch sizes. Increasing the number of transformer layers was shown to always increase the test accuracy of the models. However, this trend does not promise to continue indefinitely, since it could lead to overfitting and a corresponding fall in accuracy. Changing patch size has a greater impact on accuracy than changing the number of transformer layers.

There are a number of limitations to the chosen methodology. Firstly, GPU type was not controlled for. Google Colab is likely to have assigned different GPUs to different model runs, which could have influenced training time data, and perhaps led to the inconsistency previously observed. Secondly, because of usage limitations for free accounts on Google Colab, we were unable to run experiments for two additional values of P and L, resulting in a limited overview of the effects of the independent variables on training time and accuracy.

Though it is likely that including a classification token could lead to shorter training times, it is a question what its influence would be on the accuracy. Another experiment could be conducted to determine the exact relationship.

REFERENCES

[1] "Computer Vision Use Cases in Manufacturing | Altamira Softworks." https://www.altamira.ai/blog/computer-vision-in-manufacturing/ (accessed Nov. 06, 2022).

[2] A. A. Khan, A. A. Laghari, and S. A. Awan, "Machine Learning in Computer Vision: A Review," EAI Endorsed Transactions on Scalable Information Systems, vol. 8, no. 32, pp. e4–e4, Apr. 2021, doi: 10.4108/EAI.21-4-2021.169418.

[3] "CIFAR-100 Benchmark (Image Classification) | Papers With Code." https://paperswithcode.com/sota/image-classification-on-cifar-100 (accessed Nov. 06, 2022).

[4] R. Pugliese, S. Regondi, and R. Marini, "Machine learning-based approach: global trends, research directions, and regulatory standpoints," Data Science and Management, vol. 4, pp. 19–29, Dec. 2021, doi: 10.1016/J.DSM.2021.12.002.

[5] A. Vaswani et al., "Attention Is All You Need." 2017. doi: 10.48550/arXiv.1706.03762.

[6] I. Turc, "Transfer learning and Trasformer models (ML Tech Talks)," YouTube, Jun. 22, 2021. https://www.youtube.com/watch?v=LE3NfEULV6k

[7] "Google Colaboratory," Google. https://colab.research.google.com/

[8] "Machine Learning - Linear Regression with Multiple Variables | Coursera." https://www.coursera.org/learn/machine-learning-course/home/week/4 (accessed Nov. 26, 2022).

[9] "What are Neural Networks? | IBM." https://www.ibm.com/cloud/learn/neural-networks (accessed Nov. 26, 2022).

[10] ["Feed Forward Neural Network Definition | DeepAI." https://deepai.org/machine-learning-glossary-and-terms/feed-forward-neural-network (accessed Nov. 27, 2022).

[11] "Multilayer Perceptron Definition | DeepAI." https://deepai.org/machine-learning-glossary-and-terms/multilayer-perceptron (accessed Nov. 26, 2022).

[12] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," Neural Comput, vol. 29, no. 9, pp. 2352–2449, Sep. 2017, doi: 10.1162/NECO_A_00990.

[13] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." 2021. doi: 10.48550/ARXIV.2010.11929.

[14] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in Vision: A Survey," ACM Computing Surveys (CSUR), vol. 54, no. 10s, pp. 1–41, Sep. 2022, doi: 10.1145/3505244.

[15] Y. Kilcher, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale (Paper Explained)," YouTube, Oct. 04, 2020. https://www.youtube.com/watch?v=TrdevFK_am4

[16] K. Salama, "Image Classification with Vision Transformer," Keras.io, Jan. 18, 2021. https://keras.io/examples/vision/image_classification_with_vision_transformer/

[17] A. Krizhevsky, G. Hinton, and et al., "Learning Multiple Layers of Features from Tiny Images," Toronto, ON, Canada, 2009.

[18] "Conv2D layer." https://keras.io/api/layers/convolution_layers/convolution2d/ (accessed Nov. 25, 2022).

[19] "LayerNormalization layer." https://keras.io/api/layers/normalization_layers/layer_normalization/ (accessed Nov. 25, 2022).

[20] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer Normalization," Jul. 2016, doi: 10.48550/arxiv.1607.06450.

[21] "MultiHeadAttention layer." https://keras.io/api/layers/attention_layers/multi_head_attention/ (accessed Nov. 25, 2022).

[22] E. Ibrahimović, "Investigating the Effects of Customizable Parameters of Vision Transformers on Their Performance," Extended Essay, International Baccalaureate Organization, 2022.

[23] L. Liu, X. Liu, J. Gao, W. Chen, and J. Han, "Understanding the Difficulty of Training Transformers," EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, pp. 5747–5763, Apr. 2020, doi: 10.48550/arxiv.2004.08249.

Cardano - What Is It and How to Start Working with It

Sven Barac, Ivica Botički, Gabrijela Perković, Vjekoslav Radošević and Ivan Terzić
Faculty of Electrical Engineering and Computing, Zagreb, Croatia

Cardano is a popular blockchain platform known for its balanced approach to security and decentralization, as well as its use of smart contracts and scientific approach. In our research of the Cardano network, we analyzed its architecture and capabilities, including transactions, metadata, and wallets. Our research aimed to answer the question of the potential applications of the Cardano network in modern information systems. The result of our work is a publicly available API that enables data transfer to the Cardano network, allowing for implementation in other programs. Our research provides valuable insights into the possibilities and mechanisms of using the Cardano blockchain for modern information systems development.

Key words: Cardano, blockchain, software development

I. INTRODUCTION

Blockchain is a digital system for recording transactions and information so that it is too difficult or even impossible to alter or hack the system. It is a form of distributed ledger technology (DTL) - a protocol in which information is shared and duplicated across a network of computers in different locations, specified by its unmodifiable cryptographic signature known as hash. Its research started in 1991, but its became widely known after its first use for Bitcoin in 2009.

Blockchain is made of blocks linked together, encrypted and ran on several thousands of nodes all over the world. Blocks are compiled information stored on blockchain, with every block being duplicated and distributed across the blockchain network of computers using peer to peer network. Each block consists of its own cryptographic hash, the hash of the preceding block, timestamp of the moment of its creation, and transaction data. A hash is an immutable signature of a block generated by a mathematical function. By having stored the hash of the preceding block and the timestamps, a chain of all the blocks and its timeline can be made. Since blocks are formed in a chain, they are basically irreversible. Any alterations of one block would lead to alterations of all the other blocks.

Node is a core network stakeholder. It is a physical device that maintains the copies of the blocks [1]. Nodes keep all the copies in sync to prevent any malicious actions, store them onto own local disk, encrypt the data, take on new blocks and check their validity. They are the source of trust for this type of a system. They are characterized by redundancy, storing copies of the same information multiple times, and fidelity. When a node receives a block, it needs to verify it. Two of the most famous methods are proof of work and proof of stake [2]. In the proof of work method clients go through a mining process, which requires solving a complex math problem faster than other miners to add a block and get a reward. It is extremely energy consuming, and it could lead to centralization and manipulation due to miners gathering into mining pools. With the proof of stake method, a validator is chosen based on a specific algorithm. The validator gives a stake which they lose in case they try to manipulate the transaction data. This method consumes a lot less energy, ensures decentralization and invites more individuals, rather than "pools" to participate.

Decentralization is one of the most significant elements of blockchain. A common, but very flawed approach to storing data is one in which a single person, or a company owns a database, and the database is stored in one place. There is space for human malicious management over the database or an error in management. Data could also be mistakenly erased or rewritten. Even without taking into account the human impact, a simple power outage might lead to the entire system collapsing, while a natural disaster or a war could destroy it completely. In blockchain there are multiple copies of data all over the world, making such an outcome highly unlikely.

Hacking blockchain would require a different approach to database hacking. Modification of a block happens through consensus of the most. If one was to make an alteration of a single node, the blocks would cross-reference each other due to being synchronized, meaning that copies of the same block must have identical content at all times, and it would become apparent which node has incorrect information. To validate the alteration, hacker would have to control more than half of the copies of the block and simultaneously make changes of all of them [3]. That would be immensely expensive and most likely not feasible due to complexity. Even if the hackers managed to perform such a task, the other network members would notice such a big change in the system and react accordingly.

Blockchain has a variety of uses, including banking, healthcare, property monitoring, voting, and food production tracking for certificates of authenticity. It is used to manage more than 10 000 cryptocurrencies, including the most dominant Bitcoin and Ethereum. It works as decentralized banking system where each transaction over the currencies is stored in several blocks and verified. Transactions are transparent, meaning that they are traceable to everybody, but they are encrypted and anonymous. Moreover, one of the applications of blockchain are smart contract. They allow all participants to digitally send their portions of the deal, and upon gathering all the portions,

the contract would automatically be executed or the transaction would be stopped if its deadline expired or some other inconvenience occurred.

Despite blockchain being a promising technology, it does come with challenges. Firstly, it requires power for all the computers to run, especially when it comes to mining Bitcoin. Annual consumption of energy for mining Bitcoin amounts to somewhat more than the annual consumption of energy of Argentina. Secondly, the use of blockchain is limited due to low transactions per second rate and data storage capacity, so it could not be used as a general database. This system also provides a basis for illicit activities of dark web with not regulation currently. Moreover, since there is no single authority, regulation is even more uncertain. This leaves space for manipulation, as is the case with the mining pools mentioned previously.

II. CARDANO BLOCKCHAIN NETWORK

Cardano is an open source proof-of-stake (PoS) blockchain project that aims to provide a more secure and sustainable ecosystem for the development and execution of smart contracts and decentralized applications (dapps) [4]. Cardano uses Ouroboros, a unique proof-of-stake algorithm, which is the first proof-of-stake algorithm to be developed through peer-reviewed research [4].

Cardano network also supports the development and execution of smart contracts which were introduced in 2021. Smart contracts in the Cardano network are written in Plutus programming language/platform which is based on Haskell and Marlowe [5]; a domain-specific language (DSL) for writing and executing financial contracts [6]. In addition, Marlowe is embedded in both JavaScript and Haskell.

One of the main properties of the Cardano network is its use of multi-layer architecture. This design allows for a more flexible and modular system, as well as greater separation of concerns between different types of transactions and activities on the network. Figure 1 depicts the elements of the architecture such as the node, CLI, Daedalus wallet, GraphQL API server (queryable API for applications that use REST APIs to talk to other services [8]), and the database synchronizer (the component which is used to query historical information from the Cardano blockchain through the use of a Structured Query Language (SQL) relational database [9]). Nodes communicate via the Node-to-Node IPC protocol, which enables the exchange of blocks and transactions. On the other hand, applications on a single computer such as the wallet backends communicate with the network via the nodes using the node-to-client IPC protocol.

In the Cardano blockchain, addresses are used to identify the location where funds can be sent. They are derived from the public key of a user's wallet and are typically represented as a string of letters and numbers. Different kinds of wallets can be used on the Cardano network, some of which are software-based and others hardware-based. Software wallets can be accessed as mobile or desktop applications, providing users with access to their resources. Hardware wallets, on the other hand, are considered the most secure form of data storage and can be stored as a USB device, for example.

Transactions on the Cardano blockchain are used to transfer funds from one address to another. They consist of the sender's address, the recipient's address, and the amount of funds being transferred, along with some additional metadata such as the transaction fee. Transactions on the Cardano network are processed by nodes, which are responsible for validating the transaction and updating the blockchain's global state.

Metadata is an additional parameter or data that is part of every transaction which enables the creation of a clearer and better context for both the receiver and sender. (e.g. object with attributes in JSON file). Metadata is optional and can be added during the transaction submission, it can include information such as a message by the sender or a description of the purpose of the transaction.

Ada is the original cryptocurrency of the Cardano network with which users can receive and send their currencies and funds through transactions. In this way, users participate in the proof-of-stake consensus and contribute to greater security and validation of transactions and the network. When creating transactions, the fee for the transaction will be calculated and the term Lovelace will

TABLE I. CARDANO NETWORKS DIFFERENCES [14]

MAINNET	TESTNET
• „main network“ – live production version • Real transactions involving Adacryptocurrency and execution of smart contracts	• „Preview testnet” - get familiarised and play with cardano-node • “Pre-Production Testnet” - for those who are ready to run the mainnet but need to test it before running it • Uses „test Ada “ – have no real value (only for testing)
• The common multi-layer architecture • Plutus programming language for writing smart contracts • Open-source and actively maintained by the Cardano community	

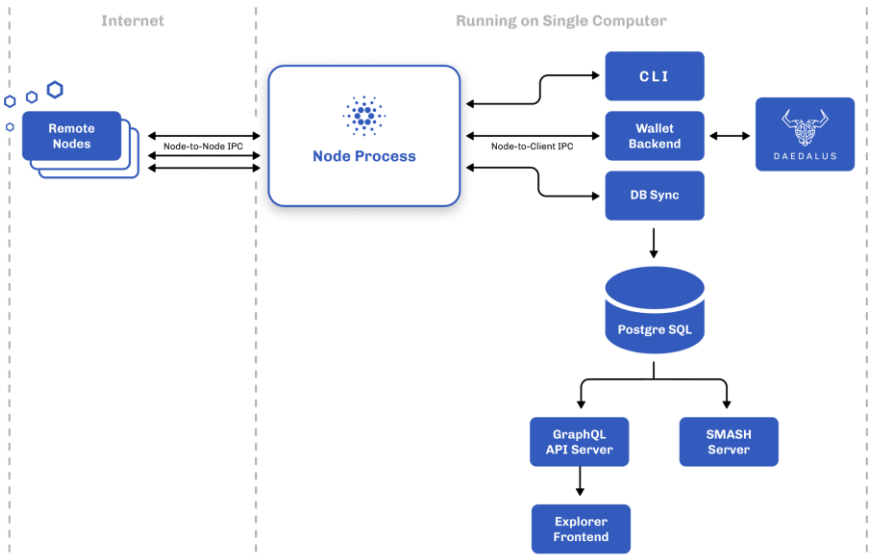


Figure 1 Cardano Arhitecture Overview [7]

be utilized. This may seem confusing, as the currency used was test Ada, and now there is a new term used. However, Ada can be expressed in smaller units named Lovelace, where one Ada contains a million of Lovelace units (similar to Satoshi and Bitcoin). In order to encourage users to actively contribute to the network and participate in the creation of new blocks, users can stake a share of their Ada currency and in return earn a reward on their stake. Staking is a way to support the network and help secure it, but also allows holders of Ada to earn a return on their investment. The total supply of Ada is limited and currently stands around 34 billion coins [10], which is the upper limit and cannot be exceeded. The mainnet and the testnet of the Cardano network refer to two different instances of the Cardano blockchain network.

In the Cardano network, the term epoch is the period of time in which new blocks are created and put on the network. Each epoch is divided into slots, where in each slot, there is a possibility of the creation of a new block. The duration of an epoch is approximately 5 days and contains approximately 432,000 slots, thereby ensuring the stability of the network [11]. At the end of the duration of each epoch, a slot leader is chosen to create a new block through the process of "epoch transition". The slot leader is elected by the stakeholders of the network proportionally to the amount of Ada they have staked. The more Ada they have staked, the more likely they are to be selected as a slot leader. This is to guarantee that the leaders are well distributed and also incentivize users to hold and stake their Ada. After a slot leader places a block on the blockchain, the process is repeated for each slot in the epoch.

III. Installing and Configuring a Cardano Node

When installing and running an instance of the Cardano node, the first thing that needs to be taken into consideration are the hardware specifications of the machine on which the node will be run. Cardano node installation turorial [12] suggests an AMD x86 or processor with at least two cores running at a minimum of 1.6GHz and on top of that, a minimum of 16GB of RAM along with 75GB of free disk space as prerequisites, and, in the case of using the mainnet network, these requirements increase. In practice, a machine with lower hardware specifications (e.g., 8GB of RAM) will struggle with compiling required binaries from source code but will still be able to run the node with sensible management of computer resources applied (it would be advisable to kill processes that are not needed and require a lot of memory to run, such as web browsers). While a machine with less than 16GB of RAM could run the node, it might not be able to run some tools and additional software associated with the node or the Cardano network, such as Daedalus Wallet [13]. Furthermore, while the recommended 75GB of disk memory might at first remain underutilized, the blocks can be filled in a short time interval. It is highly recommended to start with one of the testnet networks before switching to the mainnet and perform testing on the testnet before deploying on the mainnet.

A Cardano node can be run on the Windows operating system, various Linux distributions, and MacOS. In practice, running a node seems to be the easiest on a Linux distribution or on a Linux virtual machine. MacOS presents a solid option as it is based on Unix as well as Linux distributions, along with Ubuntu which is simple to set up and use. The official Cardano Developers site [12] did not have an installation guide for Windows at the time of preparing this report and finding Windows equivalents to Linux commands specified in the guide has proven to be a big challenge, so using a virtual machine, e.g. WSL2, is the best option should Windows be set as a target platform.

The installation process of a node is best covered in the official documentation [12]. As an alternative to going through the entire process of compiling the node binaries or should resources be limited, already compiled files can be directly downloaded. When compiling and further developing for Cardano, tools like Cabal, GHC, and Git are necessary. The process of compiling the binaries from the source code needs to be carried out with care by dully following the instructions as even the slightest mistake can lead to problems. Focus needs to be given to the required versions of the tools needed such as GHC or Cabal. The process also installs cardano-cli (command line interface), a key tool for interacting with the node allowing for a range of actions such as querying for a block at the top of the chain, creating key and address files, creating templates and signing the transaction.

Once installed the node needs to be run. The tutorial [14] on the Cardano Developers website contains almost all the instructions required to run the node. The node will need some configuration files which will give the cardano-node executable information on the kind of node the user wants to run [15], these files are necessary to run a node. Starting a node is basically just running a command in the terminal. As part of the startup, it is necessary to pick a folder for the database, decide on the port and IP address on which the node will be available, and submit the path to the configuration files. Writing a simple script that will start the node with the chosen arguments is also an option. Running the node in the terminal will display the node log, which will print out everything that the node is doing in the terminal.

The node will save new blocks to the database in the folder provided as an argument to the run command and this will be in a format similar to "Chain extended, new tip: 1e64e74bd7ac76d6806480a28017deb0aedd356fb61844e c95c429ae2f30c7c3 at slot 0". In a different terminal, the state of the blockchain, including the latest block in the node database and synchronization percentage, can be queried with cardano-cli commands (e.g., "cardano-cli query tip --testnet-magic 2" [14]). The testnet-magic flag signifies the testnet network (preview, preprod) on which the node is started (–mainnet flag is used for mainnet) [14]. Blocks, transactions, epochs, etc. for preview testnet can be explored via the Cardanoscan webapp [16]. To be able to further work with the transactions and submit them to the network, the node will have to be 100% synchronized with the network it is running on.

In addition to informing the user of successfully downloaded blocks, the node's log may also show some errors and quite a large number of notifications about peers. This is an integral part of how the Cardano network works using the Peer-to-peer (P2P) networking. Each node maintains its own set of peers classified into 3 different categories: cold, warm and hot peers. A newly discovered peer is initially added to a set of cold peers; they are peers without established network connection. A warm peer is a peer with an established connection, but no node-to-node protocols implemented, and only used for network measurements. The hot peer also has an established connection with the node, like the warm peer, and node-to-node protocols can use the connection between peers. Peers are managed by a P2P governor who maintains a certain number of peers of all types and promotes and demotes them depending on certain criteria so that there is always a certain number of peers available. A peer error can occur due to a node exception (e.g., an error in a node's file system), a network error (e.g., a failed TCP connection), or unexpected, adversarial behavior of the node [17].

Upon successfully starting the node, it is recommended to continue following the Cardano Developers documentation [18], [19] and install the Cardano Wallet. It should be downloaded and installed in a similar way to the node itself. While downloading or installing wallets, it is critical to download files from reliable and authorized sources in order to avoid theft of funds (much more important for the mainnet network). Although the Cardano Wallet is used for checking balance, sending transactions, getting transaction history details and more, it is not specifically needed to make a transaction on the Cardano network, or create a wallet for that matter.

The simplest way to create a wallet is using the cardano-cli executable [19]. Files required to build and sign a transaction are generated using commands address key-gen and address build. Running those commands with the required arguments will create three important files: a .vkey file, used to build an address, a .skey file, which is a private key used to sign transactions and approve them, both of them created by the address key-gen command. The third, an .addr file, is created from the address build command and is built from the .vkey file and contains only the wallet address, which is of this format on the preview testnet: "addr_test1vz95zjvtwm9u9mc83uzsfj55tzwf99fgeyt3gm wm9gdw2xgwrvsa5". With this address, test Ada can be claimed for free at [20].

IV. Developing software using Cardano-related frameworks

Upon successfully installing the Cardano node, the process of creating, signing and storing a transaction on an active Cardano block could be automated. Apart from using the comand line tools and scripts, one of the options for doing so would be to utilize a programming module that has all the functionalities the cardano-cli tool has. Cardanocli-js [21] is a JavaScript module that wraps every command of the cardano-cli for efficient interactions with the Cardano node. Installation of the cardano-cli-js is simple: the only necessary step is to download the package from the NPM repository using a Node package manager.

Since Cardano is not used for storing standalone data, every piece of information has to be integrated into a separate transaction as metadata to store it on the blockchain. Metadata is written in JSON format, with some restrictions including integer range 0 to 264 - 1 for values and $-(2^{64}-1)$ to $2^{64} - 1$ for keys. Floating point numbers, null values and boolean values are not supported and for that reason, they need to be transformed into another data type. Detailed description of metadata and transaction structures can be found in the official Cardano Docs guide [22].

Storing a transaction on the Cardano blockchain requires paying a fee which can be calculated using the cardano-cli tool. Using the cardanocli-js module, transaction fee can be calculated with a single function call while building the transaction. Following that, the transaction is ready to be signed which is done by using the peer's secret key. The secret key ensures that everybody on the network knows nobody tampered with the transaction after it was created. Formed and signed transaction is now ready to be stored on the platform. Storing the transaction is done by calling another cardanocli-js function which returns a unique transaction hash. That hash will be used later when querying the blockchain to find that specific transaction.

An important part of the interaction with the Cardano blockchain is fetching information from the network. These functionalities cannot be done using the cardano-cli-js, as they cannot be done in the cardano-cli command tool either. A solution to that comes with JavaScript support for Blockforst, an API that provides specific endpoints for querying the Cardano blockchain over HTTP [23]. In order to use the API, registered users have to request a Blockfrost token called "project_id" which is used as part of every HTTP request that will be sent to Blockfrost servers. The JavaScript module used for sending HTTP requests to Blockfrost servers is blockfrost-js [24]. Cardano mainnet, Cardano preprod and Cardano preview (testnet-magic-2) are the only supported Cardano networks from which Blockfrost can fetch the data. For displaying the remaining Cardano tokens, Blockfrost uses Lovelace currency instead of tAda (test Ada).

A great deal of information can be accessed using blockfrost-js module. The most important Blockfrost API endpoint covers fetching transaction data that also includes transaction metadata. To fetch the transaction data, an endpoint that accepts transaction hash as a unique value tied to every transaction (a cardanocli-js function return value) is used. Alternatively, metadata information can be fetched by targeting an endpoint that requires a metadata label identifying a metadata object, but its value is not globally unique so multiple JSON objects could be returned as a response.

Creating and storing transactions on the Cardano blockchain using the cardanocli-js module, as well as fetching transaction data using Blockfrost API are simple functionalities for interacting with the installed node and newly accessible network. Examples can be found as part of the CardanoPreviewTestnetUtils class which is packaged and uploaded to Npm repository as @tim1/cardano-testnet-tim1 [25]. The package can be downloaded and used on the systems that meet these requirements:

- installed Node.js version >= 12.19.0
- installed and configured Cardano node with version >= 1.26.1

V. CONCLUSIONS

Since its launch in 2017, Cardano has been one of the more widely used blockchain platforms available. Its unique mission is to establish an overall balance in the crypto and blockchain ecosystem by reworking the principles of security and decentralization, using capabilities of smart contracts, building a modular system through multilayer architecture and a scientific approach developed by IOHK (Input Output Hong Kong) for the Cardano network.

As an open-source program, its community is constantly expanding and working to improve the network. This community-driven development highlights the decentralized decision-making process and accelerates its innovation.

Decentralized application development (WEB3) is becoming increasingly utilized and since it is based on blockchain technologies, understanding basic concepts such as Peer-to-peer networking, proof of stake, proof of work, blocks and transactions in addition to having basic knowledge in development on blockchain network could prove to be beneficial. Cardano platform is a great way to introduce a developer to blockchain. Even if a programmer decides to use a different platform, the knowledge acquired while working with Cardano will be very valuable. Cardano has a big perspective and promising future since proof-of-stake is more environmentally friendly than proof-of-work because of the lower energy consumption which is becoming increasingly important in modern times. IOHK's blockchain platform has many uses and the mentioned code and usage is just the top of the iceberg.

REFERENCES

[1] B. Becher, "What Are Blockchain Nodes and How Do They Work?", Sep, 2022 [Online]. Available : https://builtin.com/blockchain/blockchain-node. [Accessed Jan 10, 2023].

[2] A. Rosen, "Proof of Work vs. Proof of Stake: The Biggest Differences", Oct, 2022 [Online]. Available: https://www.nerdwallet.com/article/investing/proof-of-work-vs-proof-of-stake. [Accessed Jan 10, 2023].

[3] A. Hayes, "Blockchain Facts: What Is It, How It Works, and How It Can Be Used", Sep, 2022 [Online]. Available: https://www.investopedia.com/terms/b/blockchain.asp#toc-how-does-a-blockchain-work. [Accessed Jan 10, 2023].

[4] IOHK, "Why use Cardano?". [Online]. Available: https://docs.cardano.org/new-to-cardano/why-use-cardano. [Accessed Jan. 9, 2023].

[5] IOHK, “Plutus”. [Online]. Available: https://developers.cardano.org/docs/smart-contracts/plutus. [Accessed Jan. 9, 2023].

[6] IOHK, “Marlowe”. [Online]. Available: https://developers.cardano.org/docs/smart-contracts/marlowe. [Accessed Jan. 9, 2023].

[7] IOHK, “Cardano architecture”. [Online]. Available: https://docs.cardano.org/explore-cardano/cardano-architecture. [Accessed Jan. 9, 2023].

[8] IOHK, “Cardano GraphQL”. [Online]. Available: https://docs.cardano.org/cardano-components/cardano-graphql. [Accessed Jan. 9, 2023].

[9] IOHK, “About DB Sync”. [Online]. Available: https://docs.cardano.org/cardano-components/cardano-db-sync/about-db-sync. [Accessed Jan. 9, 2023].

[10] Binance, “Cardano Price”. [Online]. Available: https://www.binance.com/en/price/cardano. [Accessed Jan. 9, 2023].

[11] IOHK, “Slots and Epochs”. [Online]. Available: https://developers.cardano.org/docs/stake-pool-course/introduction-to-cardano/#slots-and-epochs. [Accessed Jan. 9, 2023].

[12] R. Phair, “Installing cardano-node and cardano-cli from source,” Dec, 2022. [Online]. Available: https://developers.cardano.org/docs/get-started/installing-cardano-node. [Accessed Jan. 9, 2023].

[13] IOHK, “Daedalus - Cryptocurrency wallet”. [Online]. Available: https://daedaluswallet.io/en/. [Accessed Jan. 9, 2023].

[14] F. Bormann, “How to run cardano-node,” Nov, 2022. [Online]. Available: https://developers.cardano.org/docs/get-started/running-cardano/. [Accessed Jan. 9, 2023].

[15] M. Szamotulski, “Understanding your configuration files and how to use them,” Nov, 2022. [Online]. Available: https://github.com/input-output-hk/cardano-node/blob/master/doc/getting-started/understanding-config-files.md. [Accessed Jan. 9, 2023].

[16] Cardanoscan, “Preview Cardanoscan,” 2022. [Online]. Available: https://preview.cardanoscan.io/. [Accessed Jan. 8, 2023].

[17] IOHK, “Peer-to-peer (P2P) networking”. [Online]. Available: https://docs.cardano.org/explore-cardano/cardano-network/p2p-networking. [Accessed Jan. 9, 2023].

[18] T. Kammerer, “Installing cardano-wallet,” Sep, 2022. [Online]. Available: https://developers.cardano.org/docs/get-started/installing-cardano-wallet/. [Accessed Jan. 9, 2023].

[19] T. Kammerer, “Exploring cardano wallets,” Sep, 2022. [Online]. Available: https://developers.cardano.org/docs/integrate-cardano/creating-wallet-faucet/. [Accessed Jan. 9, 2023].

[20] IOHK, “Testnets faucet”. [Online]. Available: https://docs.cardano.org/cardano-testnet/tools/faucet. [Accessed Jan. 10, 2023].

[21] alessandrokonrad, “Cardanocli-js”, May, 2021. [Online]. Available: https://github.com/shareslake/cardanocli-js/blob/main/API.md. [Accessed Jan 10, 2023].

[22] T. Kammerer, “Build with transaction metadata”, Sep, 2022. [Online]. Available: https://developers.cardano.org/docs/transaction-metadata/. [Accessed Jan 10, 2023].

[23] Blockfrost, “Blockfrost.io API Documentation (0.1.50)”, Dec, 2020. [Online]. Available: https://docs.blockfrost.io. [Accessed Jan 10, 2023].

[24] blockfrost.io, “blockfrost-js”, Dec, 2022. [Online]. Available: https://www.npmjs.com/package/@blockfrost/blockfrost-js. [Accessed Jan 10, 2023].

[25] tim1, “blockfrost-js @tim1/cardano-testnet-tim1”, Dec, 2022. [Online]. Available: https://www.npmjs.com/package/@tim1/cardano-testnet-tim1?activeTab=readme. [Accessed Jan 10, 2023].

Keeping Drivers Alert: A Solution for Monitoring Driver Attention in Assisted-Driving Vehicles

Robert Jutreša*, Peter Peer* Žiga Emeršič*, Jihun Kim†
* Faculty of Computer and Information Science/Computer Vision Laboratory, Ljubljana, Slovenia
† School of Electronic and Electrical Engineering, Kyungpook National University, Daegu, South Korea
rj7149@student.uni-lj.si, peter.peer@fri.uni-lj.si, ziga.emersic@fri.uni-lj.si, soji@knu.ac.kr

***Abstract*—The advent of the new industrial revolution has led to a surge in the number of self-driving and assisted-driving vehicles on the roads. Although this development enhances the overall quality of life, it also poses new challenges and risks. One such difficulty is the need for driver attention in the event of system errors. To overcome this issue, a driver monitoring system is necessary to ensure that the driver remains focused and can take control if necessary, while also sending alerts in case of driver distraction. To address this challenge, we propose a method that employs a convolutional neural network based on the ResNet architecture. This system can recognize selected types of driver distractions in captured infrared images. We experimented with different depths of the ResNet architecture to determine the most optimal results. The best model achieved an accuracy rate of 94% in detecting various forms of driver distraction.**

***Keywords*—drowsy driving, distraction detection, CNN, ResNet**

I. INTRODUCTION

The topic of Highly Automated Driving (HAD) has been widely discussed in recent times, with leading companies such as Tesla and Waymo spearheading this technological revolution. However, there have been reports of accidents resulting from driver distraction or failure to intervene promptly if necessary. This research seeks to address this issue through a feasibility study of deploying a Convolutional Neural Network (CNN) model in a Driver State Monitoring System (DSMS) to achieve accurate classification outcomes. To this end, a dataset of driver images engaged in distraction tasks was collected, and a model was chosen based on related work, the implementation of which was then optimized. Our objective is to develop a model that can accurately identify driver distraction in real-world driving scenarios.

II. RELATED WORK

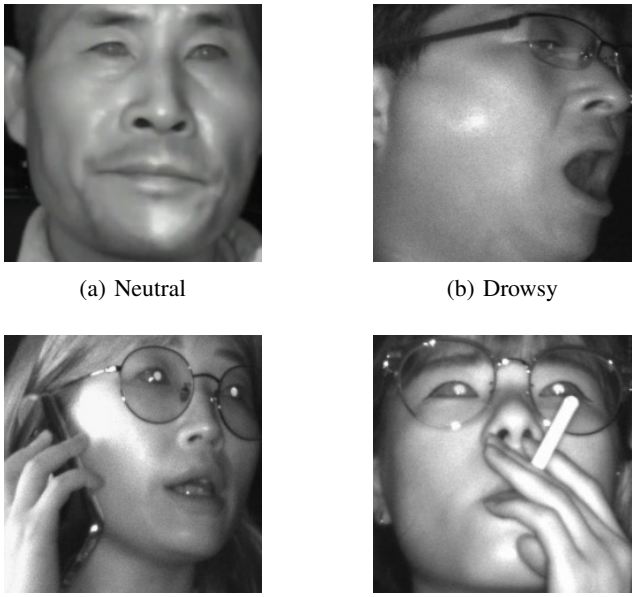
Gonçalves et al. [1] defined the important monitoring characteristics of DSMS for HAD. They said this introduces a new paradigm where driving performance metrics are no longer viable, and approaches like detecting engagement in nondriving tasks or fatigue-countering behaviors are becoming more important. They defined several distraction and fatigue indicators, based on eye and body behavior. Among these yawning was the one we focused on the most.

In 2021 Kashevnik et al. published a literature review paper [2], in which they outlined the process of developing a distraction detection system from sensor data acquisition to data processing and so forth. They also reviewed multiple papers that had neural network (NN) based solutions, which were using either self-created datasets or the AUC Distracted Driver Dataset [3], [4], the purpose of which is detecting driver distraction using driver posture from a side view. Most of the reviewed papers achieved an accuracy of 92% and above, which we set as our target benchmark.
Mase et al. [5] also performed their tests on the AUC Distracted Driver Dataset. They employed the InceptionV3 CNN architecture achieving accuracy in line with accuracies reported in papers reviewed by Kashevik et al.
Craye et al. [6] based their approach on using RGB-D cameras to capture driver state and classifying into one of five distinct classes including making a phone call, drinking, etc. They used the AdaBoost classifier and Hidden Markov Model achieving an accuracy of 90% for distraction detection. This approach is most in line with our own among the reviewed literature and thus represents the secondary benchmark for our metrics. The big difference from our proposed approach is the focus on the position of the driver's arms through a 3D point cloud and gaze estimation through iris localization. In contrast, we decided to focus primarily on the usage of convolution filters.
Liu et al. [7] approached the detection of driver detection problems, using their own developed system and dataset, where they used Support Vector Machines (SVM) and other methods of Semi-Supervised Machine Learning, achieving results as high as 97% accuracy. While not related to our work, as they used head and gaze positions and rotations as measures for their classification, it is important to acknowledge this different approach as it can also produce valuable results.
Fernandez et al. [8] made a study on the success of different approaches and techniques when approaching the problem of driver distraction detection (including the above mentioned [6], [7]), where they outlined some of the important areas to consider when developing such a system. Among these, they drew attention to the importance of different factors that modulate distractions, including but not limited to factors outside of the vehicle (traffic,

road accidents) and the driver's emotional state.

III. METHODOLOGY

A. Data Overview



(a) Neutral (b) Drowsy

(c) Phone (d) Cigarette

Fig. 1: Examples of images from each class in the train set

For our research, we used a dataset provided by the National Information Society Agency (NIA) of South Korea. The Driver Status Information Images for Preventing Drowsy Driving dataset [9] was developed for use with AI technology and the development of service applications that monitor the driver's condition through changes in facial expressions and feature points. The dataset is composed of over $250,000$ images of frontal views of $1,000$ subjects using an infrared camera with corresponding labels captured in three different settings, those being real-world driving images, images from a semi-controlled environment, and images from a controlled environment. The labeling provided for the images consists of bounding box locations and locations of facial key points as well as driver identification (gender, age, etc.) and object information. The object information includes data about the presence or absence of various occlusions (sunglasses, masks, etc.) and objects such as phones and cigarettes, in addition to data about the subjects' eyes and mouth (visibility, etc.).

B. Model Selection

As our problem was at its core a multi-label problem, we decided it would be a good choice to take it on using a CNN-based approach [10]. Kashevnik et al. [2] reported that most reviewed papers had achieved the best quality results using a CNN with a ResNet [11] architecture. According to the available data we decided on testing various depths of this architecture.

IV. EXPERIMENTS

A. Image Selection and Preprocessing

Due to the size of the original dataset, the number of images had to be limited due to time and hardware limitations. This subset was selected from both real-world driving images and images from a controlled environment. The set of images from a semi-controlled environment was omitted. From these two groups, a random sample of $26,000$ images was selected without replacement. Among these, $6,500$ belong to each of the 4 classes that were defined using the provided labeling. These 4 classes included 3 that classified a form of distraction and one neutral class. The classification schema was as follows:

- Image was classified as "subject is distracted by a phone" if there is a phone present in the image (e.g. Fig. 1c).
- Image was classified as "subject is distracted by a cigarette" if there is a cigarette present in the image (e.g. Fig. 1d).
- If the mouth of the subject isn't visible, the image was classified as "subject is drowsy" if they had their eyes closed (assuming eyes aren't obscured), otherwise, the image was classified as "subject is drowsy" if they were yawning (mouth is opened) (e.g. Fig. 1b).
- If none of the above were applicable to an image it was classified as "neutral" (e.g. Fig. 1a).

The images were then cropped based in such a way that they included the face of the subject, and potentially present phones or cigarettes. Lastly, these images were morphed to the shape of 224×224 which is the standard input size of the ResNet architecture.

TABLE I: Values of Applied Perturbations

Feature	**Range**
brightness	$[0.8, 1.2]$
rotation [°]	$[-5, 5]$
width shift	0.1 of image width
height shift	0.1 of image height
shear [°]	$[0, 0.1]$
zoom	$[0.9, 1.1]$

B. Model set up and Parametrization

All models were implemented using the open-source Python library TensorFlow [12] and following their implementation instructions. The exception to this were the ResNet-18 and ResNet-34 implementations, where the publicly available image-classifiers pip package was used. All tested models used the same set of selected images split into the train and test sets with the latter consisting of 20% of all images. The train set was then additionally split into the validation set, again with a 20% ratio, resulting in 16640 images in the train set, 4160 images in the validation set, and 5200 images in the test set. The ratio of images between classes was preserved. When loading the images into the programs additional prepossessing was applied in accordance with the documentation and

perturbation to images in the test set. The ranges and values of applied perturbations can be seen in Table I and were selected based on what changes could be present in a real-world setting. The horizontal flip attribute is excluded from this table as it doesn't have a value to list (parametrization is boolean).

TABLE II: Batch Size and Parameter Count of Models

ResNet	Batch Size	Number of parameters	
		Trainable	Nontrainable
18	60	11,708,359	7,942
34	60	21,816,519	15,366
50	40	25,636,868	53,120
101	20	44,655,108	105,344
152	20	60,321,796	151,424

All of the tested models shared this configuration when loading the images as well as the previously defined input shape. In addition to this they initialized with weights trained on ImageNet [13], to take advantage of transfer learning, frozen on top of the model since we were working with a closed set evaluation problem and a 2D average pooling layer at the end of the model. To this was attached a dense layer utilizing the rectified linear activation function (ReLu) with an output size of 1024 nodes, and lastly a dense layer utilizing the softmax function to give out the required predictions. All models used the Adam optimizer [14] and a universal early stopping system configured to reduce the learning rate by a factor of 0.5 when necessary (i.e. after 20 epochs without improvement), and stop the learning process when it reaches $1e{-}10$. Categorical cross-entropy and categorical accuracy were also universally used as measures during training. The information about the batch size used for each model and the trainable/nontrainable parameter count of the models can be seen in Table II.

V. RESULTS AND DISCUSSION

The first section below presents the numerical results of the experiments, while the second section provides commentary on those results. It is important to take into context how the experiments were run. The models were trained and tested on a personal computer equipped with an AMD Ryzen 9 5950X 16-Core processor, 16 GiB$\times$4 3200 MHz RAM, NVIDIA GeForce RTX 3070 graphics card with 8 GiB of video memory, and a 1 TB Samsung SSD 970 EVO NVMe drive. Used software included the Tensorflow 2.11 library for GPU and CUDA 11.7, the latest version of the image-classifiers package as well as necessary accompanying packages. All the times were averaged over multiple runs.

A. Results

Several statistical and empirical data points were collected during training and testing. The first set of data, seen in Table III, represents both the training and evaluation time of each model depth. The times listed are the execution times of the model training code blocks. The inconsistencies here are attributed to the varied batch size for different models (set as high as allowed by the hardware limitations for each model) and the configured early stopping preventing further training of some models at an earlier stage than others.

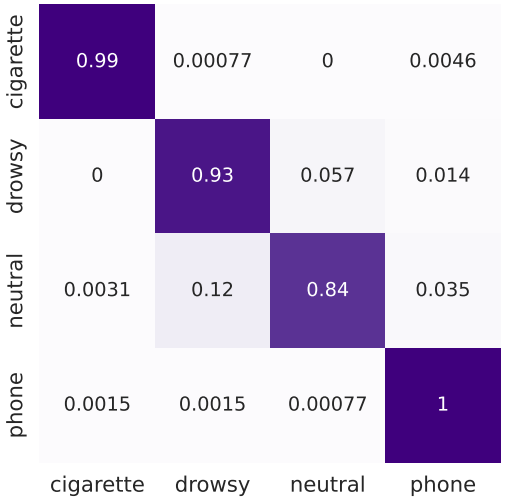


Fig. 2: ResNet-34 confusion matrix.

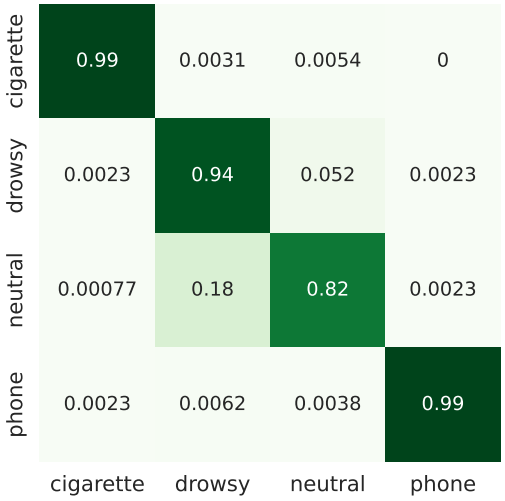


Fig. 3: ResNet-50 confusion matrix.

The results of statistical measurements done for model evaluation can be seen in the following figures. Fig. 2 and Fig. 3 show computed confusion matrices, showing the class-specific accuracy, of the two best-performing models according to statistical measurements done during model evaluation. Full results of this can be seen in Fig. 4. In the Figure the y-axis is limited, to clarify the differences between model implementations.

B. Discussion

When comparing the data it was important to consider the nature of the application of such a model. As the model is aimed toward HAD applications, it must be fast as well as accurate. With this in mind, it is quite easy to determine

TABLE III: Training and Evaluation time of various model depths

ResNet	Training time [min]	Evaluation time [s]	Batch Size	Epochs Training
18	60	7.1	60	26
34	120	7.2	60	51
50	77	8.7	40	34
101	163	16	20	55
152	109	22	20	28

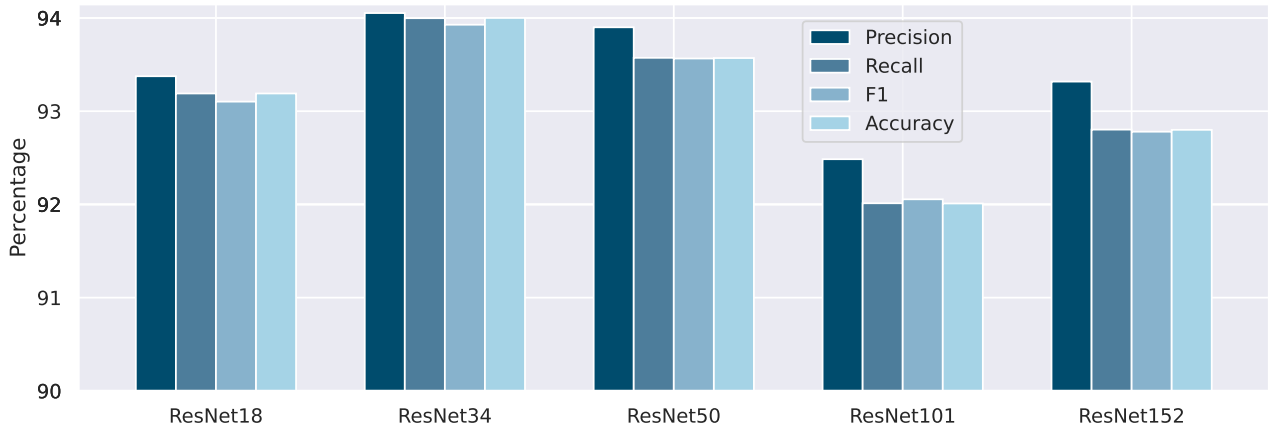


Fig. 4: Comparison of different model depths (y axis limited for clarity).

the optimal depth among the ones tested for the ResNet architecture.

Firstly looking at Table II, we can see that the ResNet-18 and ResNet-34 implementations would take up the least amount of resources based on their parameter count (size of the model), with the ResNet-50 implementation not far behind. The ResNet-101 and ResNet-152 implementations could be deemed too big for a lightweight computer installed in a car. When looking at Table III, we can see that all of the models would perform well in a real-world situation when regarding their evaluation speed of a provided image. According to Zhuk et. al. [15] the response time of a driver is between 0.773 and 2.43 seconds, meaning that even the slowest model (ResNet-152 which takes 4ms to evaluate a single image) would be sufficient to classify the current state of the driver in negligible time. The training speed is in this case, not an important factor as models would be installed in the HAD system with already trained weights.

Next, we should look at the metric data seen in Fig. 4 (note that the y axis is limited for clarity). We can see that again the lower-depth versions return better results when it comes to classification. In particular, the ResNet-34 implementation outperforms the others across all metrics, followed by the ResNet-50 implementation.

Lastly, to compare these two we look at the produced confusion matrices seen in Fig. 2 and Fig. 3. Even tho they are mostly similar, we can see that ResNet-34 has slightly higher rates of classifying the images correctly, as well as a slightly lower rate of miss-classifying the neutral as the drowsy class.

This was the most common problem when it came to the classification of every implemented depth of the ResNet architecture. As the images in the phone and cigarette classes had the items in question present in the image, and the models were loaded with pre-trained weights on the ImageNet dataset, the classification of these posed no issues. On the other hand, the drowsy and neutral classes had to be fine-tuned, and due to the nature of the dataset or our process of sampling, the images from the two classes might not have been distinct enough.

To better understand the classification described in Section IV-A, an example of the provided labels can be seen in Fig. 5. The labels presented locations and annotations for selected facial features and objects, and thus some generalization had to be done when classifying images based on those labels. While the "phone" and "cigarette" classes were straightforward, the labels didn't allow for much variability when choosing the features for the "drowsy" class. Thus we looked at the visibility and openness of the mouth and eyes in order to classify images as "drowsy". For this, we made assumptions that the subject was alone in the vehicle, and thus not talking to anyone, as well as that the potential miss-classification of a non-drowsy state as drowsy in a real-world scenario is better than vice versa. Because of these assumptions, no images, where the mouth was labeled as opened, were classified as "neutral" only "drowsy".

VI. CONCLUSION

The results of this research are promising. Based on available information a fitting CNN model was selected, and experiments were run to find the optimal implementation of the selected model. In the end, this proved to be the ResNet-34 implementation, proving to not only produce the most accurate classification

```
{
    "FileInfo": {
    ...
  },
  "UserInfo": {
    ...
  },
  "Accessory": {
    "Mask": false,
    "Glasses": false,
    "Cap": false
  },
  "Annotation": 1,
  "ObjectInfo": {
        "KeyPoints": {
            ...
        },
        "BoundingBox": {
      "Face": {
        "isVisible": true,
        "Position": [ ... ]
      },
      "Leye": {
        "isVisible": true,
        "Opened": true,
        "Position": [ ... ]
      },
      "Reye": {
        "isVisible": true,
        "Opened": true,
        "Position": [ ... ]
      },
      "Mouth": {
        "isVisible": true,
        "Opened": true,
        "Position": [ ... ]
      },
      "Cigar": {
        "isVisible": false,
        "Position": [ ... ]
      },
      "Phone": {
        "isVisible": false,
        "Position": [ ... ]
      }
    }
    }
}
```


Fig. 5: Data Labeling Schema

while remaining lightweight enough for it to be feasibly implemented in a HAD application. In addition to this we also matched and surpassed the set benchmarks outlined in Section II.

Possible improvements and future work could include further tests on already mentioned implementations, in an attempt to improve both the accuracy and portability of the model. This could include expanding the number of samples taken from the database, testing it on other possible databases, introducing new classes, freezing layers of the model, or in other ways altering the training to lower the number of parameters. Furthermore, an actual prototype setup involving both an infrared camera and a computer to test the appropriateness of the model in a real-world application could be made, together with long-term tracking of both – motion and emotion.

REFERENCES

[1] J. Goncalves and K. Bengler, "Driver state monitoring systems–transferable knowledge manual driving to had," *Procedia Manufacturing*, vol. 3, pp. 3011–3016, 2015.

[2] A. Kashevnik, R. Shchedrin, C. Kaiser, and A. Stocker, "Driver distraction detection methods: A literature review and framework," *IEEE Access*, vol. 9, pp. 60 063–60 076, 2021.

[3] Y. Abouelnaga, H. M. Eraqi, and M. N. Moustafa, "Real-time distracted driver posture classification," *CoRR*, vol. abs/1706.09498, 2017. [Online]. Available: http://arxiv.org/abs/1706.09498

[4] H. M. Eraqi, Y. Abouelnaga, M. H. Saad, and M. N. Moustafa, "Driver distraction identification with an ensemble of convolutional neural networks," *CoRR*, vol. abs/1901.09097, 2019. [Online]. Available: http://arxiv.org/abs/1901.09097

[5] J. M. Mase, P. Chapman, G. P. Figueredo, and M. T. Torres, "A hybrid deep learning approach for driver distraction detection," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2020, pp. 1–6.

[6] C. Craye and F. Karray, "Driver distraction detection and recognition using RGB-D sensor," *CoRR*, vol. abs/1502.00250, 2015. [Online]. Available: http://arxiv.org/abs/1502.00250

[7] T. Liu, Y. Yang, G.-B. Huang, Y. K. Yeo, and Z. Lin, "Driver distraction detection using semi-supervised machine learning," *IEEE transactions on intelligent transportation systems*, vol. 17, no. 4, pp. 1108–1120, 2015.

[8] A. Fernández, R. Usamentiaga, J. L. Carús, and R. Casado, "Driver distraction using visual-based sensors and algorithms," *Sensors*, vol. 16, no. 11, p. 1805, 2016.

[9] N. I. S. Agency, "Driver condition information video to prevent drowsy driving," 2021, accessed on 9. 12. 2022. [Online]. Available: https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=173

[10] J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang, and W. Xu, "Cnn-rnn: A unified framework for multi-label image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[12] T. Developers, "Tensorflow," 2022. [Online]. Available: https://doi.org/10.5281/zenodo.7604251

[13] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[15] M. Zhuk, V. Kovalyshyn, Y. Royko, and K. Barvinska, "Research on drivers' reaction time in different conditions," *EasternEuropean Journal of Enterprise Technologies*, vol. 2, pp. 24–31, 04 2017.

Teaching Programming and Microcontrollers with an Arduino Remote Laboratory Application

Antonio Bukovac, Eva Pleše, Una Maravić, Petra Petrović, Tomislav Jagušt

University of Zagreb
Faculty of Electrical engineering and Computing
Unska 3, 10000 Zagreb
{antonio.bukovac, eva.plese, una.maravic, petra.petrovic, tomislav.jagust}@fer.hr

***Abstract*. This paper presents our work on an educational web system that teaches programming and microcontroller basics using the Arduino board in remote laboratory settings. In response to the shift towards remote and distance learning brought by the COVID-19 pandemic, and in an effort to alleviate the issue of continuous equipment shortage in many Croatian schools, we developed and tested a client-server architecture-based interactive remote laboratory system called RemoLab.**

The system consists of a web-based frontend application, a backend server with an accompanying database, and additional tools for connectivity with an Arduino microcontroller and a webcam. The system allows students to remotely write code for real hardware and instantly observe the results of their actions through a connected web camera.

The RemoLab system has the potential to enable teachers in enriching their online courses or homework assignments, but also back up students in independent learning and experimentation with Arduino or microcontrollers. We believe that this application has the potential to improve the way students learn microcontroller programming and can be used as an effective tool for distance learning.

Keywords. remote laboratory, arduino, programming, technology-enhanced education, sensors, iterative learning

I. INTRODUCTION

The standard method for conducting lab exercises is to set up a laboratory on campus with all necessary equipment, allowing students to apply their theoretical knowledge from lectures to practical work. Because online education is growing in popularity due to its flexibility, practical exercises are also being adapted for online learning through remote laboratories.

Remote laboratories, which allow students to conduct experiments online without physical access to equipment, are making practical components of education possible and more accessible through remote learning [1]. In traditional labs, students interact directly with equipment through physical actions (e.g. hand gestures, pressing buttons, turning knobs) and receive feedback through their senses. In remote labs, the same interactions occur remotely through a user interface and remote infrastructure.

In this article, we will provide an overview of several remote laboratories that currently exist. Additionally, we will develop our own remote laboratory, which will be designed for teaching the programming of an Arduino board. Using a web application, users will have the ability to write code, which will then be executed on the Arduino board. As feedback, they will receive a live video feed, allowing them to see the results of their code in real-time.

II. RELATED WORK

Given that Arduino boards are increasingly being used in education and that online learning has become necessary and popular, there is an increasing number of papers on the subject of remote laboratories for programming and using Arduino boards as well as for other areas of knowledge.

Fernández-Pacheco et al. [2] highlighted the advantages and necessary components of remote laboratories. They designed their own remote laboratory using an Arduino MKR1000 Wi-Fi board, Raspberry Pi as the application server and Web camera for streaming the impact of code, providing students with an opportunity to experiment and explore the Internet of Things (IoT) using any platform, hardware, and software without the need to download additional extensions.

In their paper [3], Martin et al. described some of the many experiments that can be carried out using the

previously described remote laboratory, which are used by students in the Electronics Engineering Degree in a Distance Learning University as part of the mandatory practices of a class. The work with an LED cube, various Arduino sensors, and an LCD screen is presented.

Kalúz et al. [4] presented a low-cost remote laboratory for control education. Their remote laboratory is based on Raspberry Pi and Arduino controllers for education in the field of process control and allows efficient energy usage and problem management without a supervisor. The goal of their project is to familiarize students with specific experiments, where they cannot modify or test the programming code for the Arduino, but only change certain values that affect the outcome of these precise experiments.

In their paper [5], Carro et al. described the use of a remote laboratory as an additional tool for teaching. The paper does not relate to learning programming or the use of Arduino, but rather presents a remote smart device and a remote laboratory with an RGB LED as a tool for learning about the physics of color, various perceptions of color, etc., and for improving interaction between students and teachers.

Lima et al. [6] described and designed a remote Arduino laboratory used for learning that can be accessed through two types of applications: one uses a visual and the other a textual programming language for learning Arduino programming. The user can see the code they input being executed on a remote Arduino board.

III. REMOLAB ARCHITECTURE

This chapter explains the architectural elements of the RemoLab project. The central component is a microcontroller board Arduino Uno that acts as the main control unit. Arduino is an open source electronic kit specifically designed as a controller that regulates the workings of other electronic circuits [7]. Multiple components are connected to the board: a temperature sensor, a servo motor, RGB LED and 16x2 LCD screen which are used to provide feedback to the user and can be manipulated by the program code uploaded to the Arduino board. The connection between the Arduino board and a computer is done through the USB port (Fig. 1.).

The WebCam is used to capture images and send them to the user's screen every few seconds. Currently, all components are connected to a computer. However, for better system flexibility, components could be connected to a Raspberry Pi, which would act as the main hub for data processing and communication.

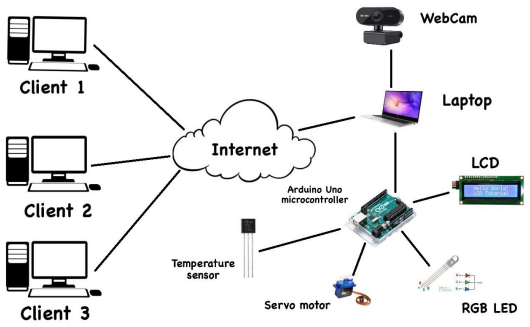


Fig. 1. RemoLab system architecture

The RemoLab system setup contains two servers, web server for the end-user application and the camera image server:

- The web server for the backend of the application runs on .NET and SQL Server database and is responsible for user authentication, showing tutorials in the application and uploading code to the Arduino microcontroller.
- The camera server runs on Node.js in combination with SocketIO and node-cam package. It is responsible for capturing and sending images to the clients.

IV. EXPERIMENTS

Several lessons were created that cover different fields of programming and electronics. Temperature sensor was chosen because it is simple and it gives students a basic understanding of sensor readings [8] and displaying results in a suitable way using LCD [9]. Servo motor is used as a component because it is an important part in dynamic projects like robotic arms [10] or remote controlled cars [11].

Students can easily access this laboratory by visiting the RemoLab website. On the website, students are presented with different categories, each containing simple tutorials and questions.

These educational resources are designed to help students in learning the fundamental concepts and commands of the Arduino device, and provide them with hands-on experience by allowing students to apply what they have learned through the tutorials on the actual Arduino device.

The code input page shows the camera image next to the code input field, allowing students to see the results of their actions in real-time (Fig. 2.). After entering the code, students can submit it to the server for compilation and execution on the Arduino. The camera

display provides feedback on the execution of the code and helps students determine whether the code is functioning as intended. This setup allows students to visualize the results of their code and make adjustments as necessary.

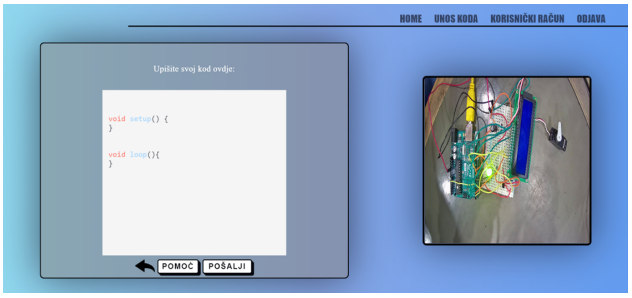


Fig 2. GUI of the experiment

There are four devices connected on the Arduino that students can program and control (Fig. 3.). The benefit of having all four devices on the same board is that students can easily experiment with multiple devices at once. This allows for a greater understanding of how different devices can interact with one another and opens up opportunities for more complex projects. For instance, the color of a RGB LED can be changed based on the measured temperature, demonstrating the relationship between two components.

Three devices, with accompanying exercises, are:

- *RGB LED*

  Goal of the RGB LED experiment is to teach basic functions/commands such as setting pins to OUTPUT, INPUT, and using the *delay* function. This will introduce students to the most important functions they will use in further experiments, in a fun and engaging way. Additionally, students will be introduced to the concept of RGB color model and have hands-on experience with it. Ability to change colors provides additional possibilities compared to a regular LED and makes the learning experience more enjoyable for the students.

- *Temperature sensor*

  The lesson with temperature sensor aims is to familiarize students with the basics of reading values from sensors. Here we use a temperature sensor as one of the basic sensors.

  In this exercise, students will also learn about converting analog inputs into meaningful values. In this example additional operations are required to obtain the temperature in e.g. degrees Celsius and this knowledge can be applied to other sensors such as humidity, light, etc.

- *Servo motor*

  In this exercise, students will learn how to control a servo motor using Arduino. Furthermore, this lesson teaches how to add "libraries" to the Arduino project and how to use them. In this particular case, students have to add the “Servo” library. This will help them to further seek and apply existing solutions that they can use to simplify certain parts of their projects.

  In this experiment students will perform a simple movement of a servo motor. Using the *delay* function and loops, they can create some interesting solutions. The aim is to familiarize students with the basic control of a servo motor and how libraries can be added and used in their Arduino code.

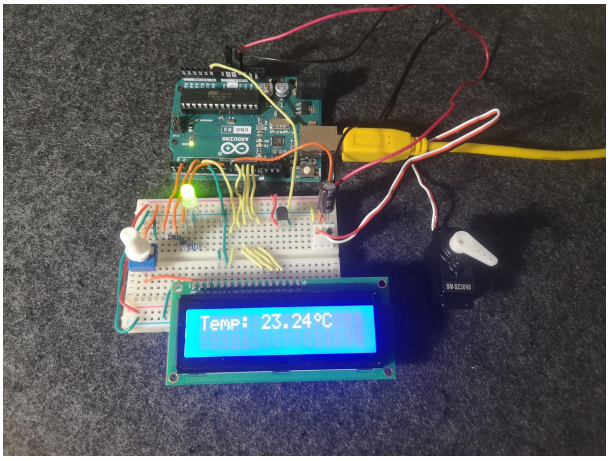


Fig. 3. Arduino board with LED, servo motor, temperature sensor and LCD screen

V. CONCLUSION

This paper describes the architecture and implementation of a remote laboratory for teaching microcontroller programming. The RemoLab system includes several components, such as a temperature sensor, servo motor, RGB LED, and a camera for live feedback. The system allows users to write code through a web application, which is then sent on server, compiled and uploaded to the Arduino board. Program execution results can be observed through the embedded web camera.This allows us to have an iterative learning experience, where we test something, observe the results, and make adjustments to try again.

The RemoLab remote laboratory provides students with a simple way to learn about Arduino programming and experiment with different components, all from the comfort of their own devices. This type of remote laboratory can be a valuable addition to any education curriculum and can make it more accessible for students to learn about technology and programming.

REFERENCES

[1] J. E. Corter, J. V. Nickerson, S. K. Esche and C. Chassapis, "Remote versus hands-on labs: a comparative study," 34th Annual Frontiers in Education, 2004. FIE 2004., Savannah, GA, USA, 2004, pp. F1G-17, doi: 10.1109/FIE.2004.1408586.

[2] A. Fernández-Pacheco, S. Martin, and M. Castro, "Implementation of an Arduino remote laboratory with raspberry Pi," in Proc. IEEE Global Eng. Educ. Conf. (EDUCON), Apr. 2019, pp. 1415–1418.

[3] S. Martin, A. Fernandez-Pacheco, G. Carro, J.A. Ruipérez-Valiente and M. Castro, "Remote Experimentation Through Arduino-Based Remote Laboratories" in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, vol. 16, no. 2, May 2021, pp. 180.-186.

[4] M. Kalúz, L. Cirka, R. Valo, and M. Fikar, "ArPi lab: A low-cost remote laboratory for control education," IFAC Proc. Volumes, vol. 47, no. 3, pp. 9057–9062, 2014.

[5] G. Carro et al., "The color of the light: A remote laboratory that uses a smart device that connects teachers and students," in Proc. IEEE Global Eng. Educ. Conf. (EDUCON), Apr. 2014, pp. 854–860.

[6] J. P. C. de Lima, L. M. Carlos, J. P. S. Simão, J. Pereira, P. M. Mafra, and J. B. da Silva, "Design and implementation of a remote lab for teaching programming and robotics," IFAC-PapersOnLine, vol. 49, no. 30, pp. 86–91, 2016.

[7] Darmawan, Aan & Ratnadewi, Ratnadewi & Sartika, Erwani & Pasaribu, Novie & Arlando, Riko. (2017). Basic Arduino Programming Training For High School Students. THE SPIRIT OF SOCIETY JOURNAL. 1. 10.29138/scj.v1i1.456.

[8] M. Kusriyanto and A. A. Putra, "Weather Station Design Using IoT Platform Based On Arduino Mega," 2018 International Symposium on Electronics and Smart Devices (ISESD), Bandung, Indonesia, 2018, pp. 1-4, doi: 10.1109/ISESD.2018.8605456.

[9] G. M. Debele and X. Qian, "Automatic Room Temperature Control System Using Arduino UNO R3 and DHT11 Sensor," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2020, pp. 428-432, doi: 10.1109/ICCWAMTIP51612.2020.9317307.

[10] V. Varshney, A. Kumar, T. Saxena, P. Jha, M. N. Tiwari and Savita, "Design and Development of Intellectual Robotic Arm," 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2019, pp. 984-987.

[11] Yuxin Jing, Letian Zhang, I. Arce and A. Farajidavar, "AndroRC: An Android remote control car unit for search missions," IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014, Farmingdale, NY, USA, 2014, pp. 1-5, doi: 10.1109/LISAT.2014.6845227.

Awareness of Croatian Citizens about the Advantages and Disadvantages of IoT Devices and Their Safety

P. Petrošanec, S. Čelan, R. Nikolić, M. Milenković
University of Zagreb, Faculty of Transport and Traffic Sciences, Zagreb, Croatia
paula.petrosanec44@gmail.com
0135265994@fpz.hr
robertina.zg@gmail.com
mmilenkovic@fpz.unizg.hr

Abstract - How knowledgeable are we about the Internet of Things? What do IoT devices mean to us in everyday life? The awareness of citizens of the Republic of Croatia will be discussed in this article on issues related to knowledge of IoT devices and their capabilities. For this paper, a survey has been conducted among the population between the ages of 19 and older on issues of familiarity with IoT devices, and an analysis of the responses was made. This paper shows how knowledgeable the citizens of the Republic of Croatia are about IoT devices, their capabilities, their security, and the possibility of omissions, data leaks, and violations of the GDPR, i.e., omissions with the sharing of their personal data. Considering that the global pandemic in the last two and a half years has stimulated computerization and the sudden expansion and ever-faster spread of IoT devices, the familiarity of EU citizens and knowledge about their capabilities has been called into question. There is also a viable need for educating and familiarizing the citizens of the Republic of Croatia with the capabilities of IoT devices and their advantages and disadvantages. This paper discusses the theoretical and practical implications of the research and gives recommendations for further research.

The key words: Internet of things, IoT devices, security, digital transformation, data privacy

I. INTRODUCTION

IoT or Internet of Things stands for a network of connected devices that has constant (24/7) access to the Internet, to exchange data with other smart devices, and holds software support that enables usage without human help.

The idea of connecting and implementing sensors to various devices appeared already in the eighties (of the 20th century), although such ideas existed in different forms much earlier [1]. John Romkey developed the first IoT device in 1990. It was a toaster that was turned on and off via the Internet [2].

The IoT primarily includes smart mobile phones, smart TV, tablets, smart watches, smart electric meters, smart lighting, smart monitoring systems, smart refrigerators, smart ovens, smart virtual personal assistants, smart heating and cooling systems, and many other different smart devices. The purpose of IoT devices is to collect data from end users and share their information to facilitate and simplify their daily life processes, i.e., their daily activities. IoT devices make it easier for users to perform everyday tasks. This paper will address the numerous advantages and disadvantages.

Users share their personal data with various third parties without knowing with whom their personal data have been shared. It is important to note; when purchasing an IoT device, end users do not sign the consent for sharing of their data, which is prescribed by the GDPR[1].

According to Ericsson research [3], there will be 22 billion devices of IoT devices by the end of 2023, and 24 billion by 2050, which indicates a significant growth of IoT devices in the next 25 years.

While Business Insider experts expect that figure to grow to 30.9 billion by 2025. [4] As the number of IoT devices increases, so does the attack surface of the cybersecurity vulnerabilities they present. More than one billion IoT attacks occurred just during 2021, of which almost 900 million were phishing attacks [5].

II. METHODOLOGY

For this paper, a quantitative questionnaire was conducted on a sample of 504 citizens/respondents from the Republic of Croatia (further: Croatia). The aim was to analyze their familiarity with the capabilities of IoT devices.

For this research, it was used a questionnaire as a data collection tool. Web-Assisted Interviewing (WAI collected the data) [6] since WAIs are easier, more efficient, and usually conducted without expenses with

[1] The basic requirements for the effectiveness of valid legal consent are defined in Article 7 and additionally stated in the introductory statement 32 of the GDPR. Consent must be freely given, concrete, transparent, informed, and unambiguous. To obtain freely given consent, it must be given on a voluntary basis. Inherently, it implies a real choice of the data subject. GDPR:EU, https://gdpr.eu/gdpr-consent-requirements/, (Accessed: 15th of February 2023)

less error occurrence, than paper questionnaires and some other types of surveys [6]. The authors have opted for ‘’Mixed-mode design with a choice of completion method’’ [6], which is used among different methods, such as web, phone, or mail, within the groups of respondents who have a connection with the University (mostly students and professors). Due to the predicted large number of respondents and its cost-effectiveness, this type of method was chosen. Data were collected from the 12th of December 2022 until the 15th of January 2023, and the questionnaire was provided in an Internet-only version via Google Forms.

The questionnaire was provided to participants using social media, such as Viber, WhatsApp, Facebook, Signal, and Telegram. The questionnaire was conducted anonymously, and the Croatian language was used.

It contained of ten (10) questions. The first question was concerning the age of the respondents. The authors decided not to ask any personal data of respondents except the age, because information is irrelevant for the meaning of the answers. The second question asked the respondents if they know of IoT devices. There were only two (2) answers available: „Yes’’ and „No’’. The third and the fourth questions were about the knowledge of respondents on IoT devices to selected from the multiple answers. Thus, respondents could also add their own answer/s. The questions also offered incorrect answers to reveal the real knowledge of the respondents.

With the fifth and sixth questions, it was attempted to examine the awareness of the security of IoT devices and the possibility of cyber-attacks via them. The seventh and eight questions included awareness and accessible education that the citizens already have in Croatia (in everyday life, examples on TV, radio, Internet, schooling system of the Republic of Croatia, etc.). The final two (2) questions examined the personal opinions of respondents about the improvement of their life by using IoT devices and whether they plan to acquire such a device soon.

All the questions from the questionnaire made it possible to check the sufficient familiarity of the citizens of Croatia with the capabilities of this type of device. Therefore, the paper offers suggestions and clarifications on different forms of education for citizens of all ages.

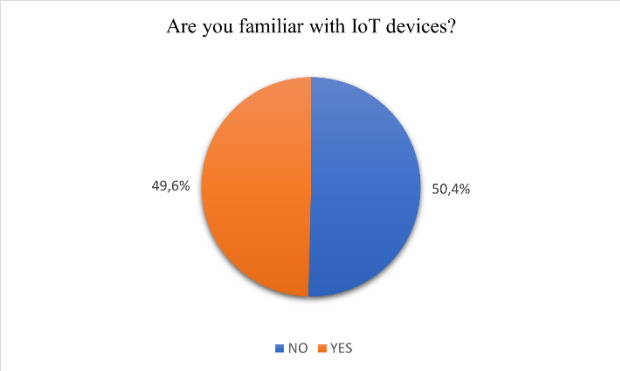


Figure 1: Question: Are you familiar with IoT devices?

III. ANALYSIS OF THE RESULTS

The questionnaire was completed by a total of 504 respondents. The first question showed the age of the respondents. 65.8% are 19 to 25 years old; 5.4% are 26 to 31 years old; 4.8% are 32 to 37 years old; 4.4% are 38 to 43 years old; 7.2% are 44 to 49 years old, and there are 12.5% of the respondents older than 50 years of age.

The second question gave an insight into the percentage of respondents’ familiarity with IoT devices. According to *Figure 1*, 49.6% of respondents are acquainted with IoT devices, and 50.4% of respondents were not acquainted with IoT devices. It is worrying to know that a large number of respondents are not familiar with IoT devices, especially when the largest percentage of respondents are young people under 25 years old.

The third question included 482 respondents. 411 respondents (85.3%) know that a TV can be a smart IoT device. 264 respondents (54.8%) are familiar with the fact that the heating system can be organized as a smart IoT system, while 224 respondents (46.5%) believe that an electric meter could also be a smart IoT device. In the same question, devices that are not IoT devices were specifically listed, and a certain number of respondents also included these devices as IoT devices, see *Figure 2*. As an example, a mixer was given, and as many as 87 respondents (15%) concluded that it could also be an IoT device.

From the answer to question 4, it is evident that 374 respondents (74.9%) own a smart TV, whereas 86 respondents (17.2%) do not have any aforementioned IoT device in their households. Under "other", the respondents added some of the smart devices they own, such as: a smart vacuum cleaner, a power socket, a smartwatch, a smart bracelet etc. The question arises whether the respondents can recognize which IoT devices they have in their household.

The fifth question, 243 respondents (49.8%) answered that IoT devices are unreliable, whereas 245 respondents (50.2%) replayed that they are reliable. Furthermore, out of 491 respondents to the sixth question, only 14 respondents (2.9%) replayed that their data is not exposed to potential threats on the Internet, while 277 respondents (56.4%) answered that their data is somewhat exposed to threats, and 200 respondents (40.7%) are aware of the dangers that can happen with their personal data on the Internet (See *Figure 5*).

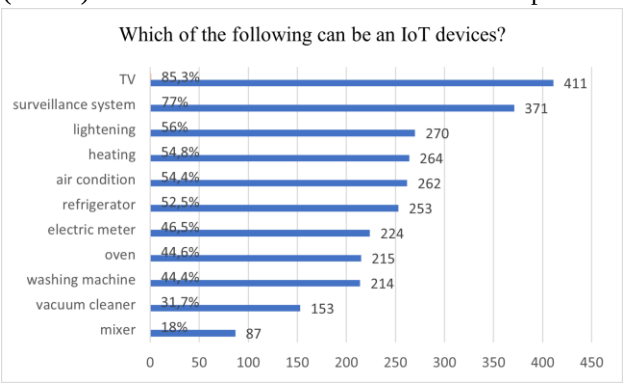


Figure 2: Question: Which of the following can be an IoT device?

In the seventh question: "Do you think that you are sufficiently knowledgeable about the capabilities of IoT devices?", 148 respondents (30.3%) answered that they are not at all knowledgeable about the possibilities that IoT devices bring.

As many as 253 respondents (51.7%) believe that they are not sufficiently knowledgeable about the operation of IoT devices and their capabilities. Only 88 respondents (18%) believe that they have sufficient knowledge about the capabilities of IoT devices. From the given answers, it can be concluded that over 80% of the respondents believe that they are not sufficiently educated about the possibilities of IoT devices.

On the 8th question, as many as 187 respondents (42.3%) answered that adequate education on IoT devices is not provided, and only six (6) respondents (1.4%) believe that education in Croatia provides sufficient education on IoT devices. All other respondents are not satisfied with education on IoT devices in the Croatian educational system (See *Figure 3*).

Out of the 476 responses we received from the answers in question nine (9), the majority of respondents believe that IoT devices and their use do not contribute to improvements to the quality of their life. In the last question, only 485 respondents answered this question while 292 respondents (60.2%) do not plan to buy an IoT device in the next year, whereas 193 respondents (9.8%) concluded that they plan to buy a smart IoT device in the next year. (See *Figure 4*)

On the basis of the questionnaire, a relatively insufficient percentage of respondents who stated that are familiar with the technology of IoT devices was revealed, and the above indicates, or may indicate, one or more of the following:

a) according to the answers to the seventh question, only 18% of respondents consider that they are sufficiently knowledgeable about IoT devices, which is insufficient for the entire Croatian system;

b) The Croatian executive state and local authorities have not regulated the introduction of IoT devices into the Croatian education system for pupils and students, or for adults through forms of lifelong learning, and they have no plans of implementing such solutions;

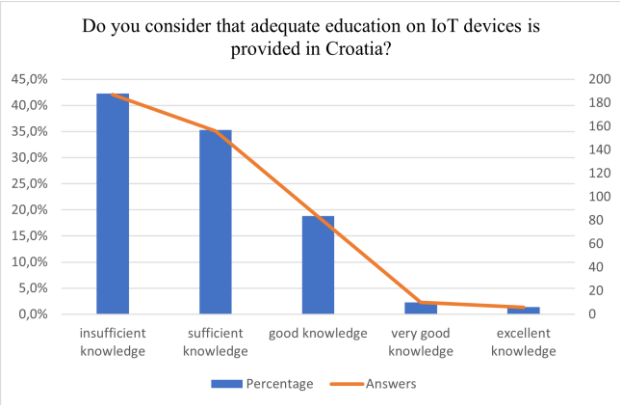


Figure 3: Question: Do you consider that adequate education on IoT devices is provided in Croatia?

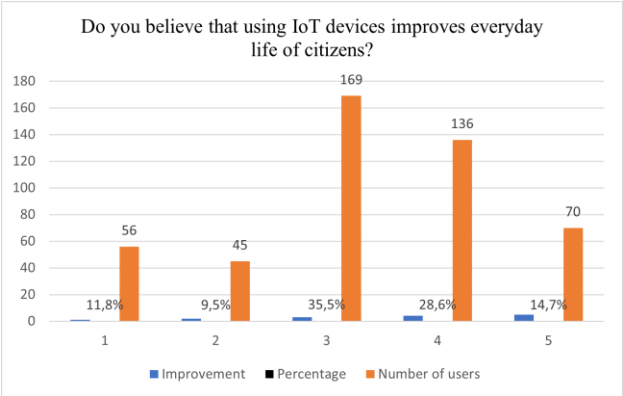


Figure 4: Question: Do you believe that using IoT devices improves the everyday life of citizens?

c) the Croatian education system does not contain the necessary education about IoT devices, i.e., about today's leading technology.

IV. LEGAL REGULATION OF IOT DEVICES IN EU AND CROATIA AND SUGGESTIONS FOR AMENDMENTS IN PRACTICE

A. Legal aspects of the security of IoT devices

The European general data protection framework, the General Data Protection Regulation (further: GDPR), does not explicitly mention IoT devices, as well as Croatian national laws, such as the Croatian Act on the Implementation of the General Data Protection Regulation [7], which contains a series of introductory statements and provisions applicable to data collection through smart home devices and the Internet of Things.

IoT products and services enable the collection of large amounts of data, some of which may be potentially sensitive. It is necessary to take protective measures, but it's important not to jeopardize the privacy of users.

The GDPR [8] requires that the service provider ensures the safety and security of the processing. It considers this further in the introductory statement 39, elaborating the principles of data protection, especially the principles of purpose limitation, storage limitation, and data minimization: "Natural persons should be aware of the risks, rules, protective measures, and rights related to the processing of personal data and how to exercise their rights related to such processing. Personal data should be adequate, relevant, and limited to what is necessary for

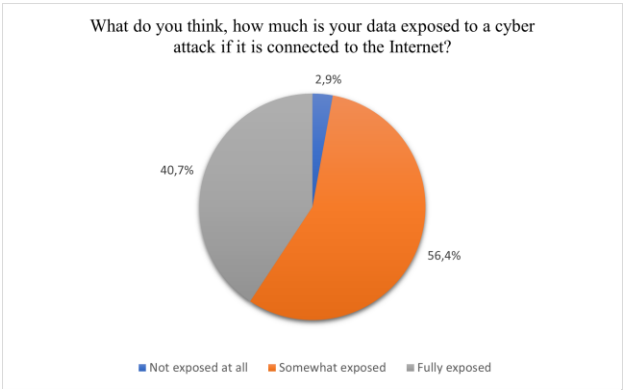


Figure 5: Question: What do you think, how much is your data exposed to a cyber-attack if it is connected to the Internet?

the purposes for which they are processed. In particular, this requires ensuring that the period for which personal data is stored is limited to a strict minimum. Personal data should only be processed if the purpose of the processing cannot reasonably be fulfilled by other means."

Hence, the results in *Figure 5* show that Croatian citizens are not familiar with the risks of trading with their personal data and the dangers they bring. As Vojković et alli mentioned: "Users should be aware of what data is being collected about them. The problem is that many IoT devices do not have high-quality protection. Consumers must be aware of this, which is not within the scope of GDPR, but consumer protection regulations." [9] The security threats posed by IoT devices have already been described in the literature. [10]

In parallel with the exponential growth of technology, the number of various directives, regulations, and standards in the field of security is growing.

The European Union Agency for Cybersecurity (ENISA) in its tasks regarding IoT and security highlights:

• Promoting harmonization of IoT security initiatives and regulations,

• Raising awareness of the need for cyber security in IoT,

• Defining the security guidelines of the software and hardware development life cycle for IoT,

• Achieving consensus for interoperability in the IoT ecosystem,

• Encouraging economic and administrative incentives for IoT security,

• Establishing secure life cycle management of IoT products/services,

• Clarification of responsibilities among IoT stakeholders.[11] [12]

According to Vojković et alli, [13]; the general legal obligation for the cybersecurity of IoT products in the EU does not exist in the EU, hence, it is incompatible with the obligation of the EU to provide for a higher level of consumer protection, as mentioned in Article 169 [14] of the Treaty of the Functioning of the European Union.

New functionality brings new data security and privacy challenges; experts noticed that the weakest links of IoT can be driven by the low-cost devices with low-security and cryptography technologies.[15]

B. Practical methods of citizen education

Education about the security of IoT devices should start from the preschool age in such a way as to create and conduct workshops. Specifically, education about IoT devices through practical classes would be introduced into the preschool curriculum.

The classes would include knowledge about threats that appear during the use of devices (e.g., mobile phones, tablets, and computers). By using this approach, it is important to raise the awareness of potential dangers that children do not understand due to their age.

Adequate lectures on IoT devices and their security should be added to the elementary and secondary school curriculum as part of computer science classes.

To ensure the education of all students in this area, it is necessary to conduct a lecture by an expert (e.g., a professor of Computer Science) at least once a year during the class community hour.

It is also important to conduct workshops and lectures on the presented topic with the faculties.

For the elderly, the local self-government should engage experts who would provide education through workshops on the dangers of their personal data, as well as their rights. In addition to workshops, it is necessary to apply education in the media through advertisements, newspapers, and magazines financed by the Government, and to obey the EU regulations and directive articles.

Additionally, ENISA has regulated the basic guaranteed level with the European cybersecurity certificate or EU declaration of conformity, which provides a guarantee that ICT products, ICT services, and processes for which that certificate has been issued, meet the appropriate security requirements. [14]

The same includes security functionalities, testing whether the products have undergone evaluation, the purpose of which is to minimize known basic risks for incidents and cyberattacks, but it is not currently implemented in practice because manufacturers do not follow the recommendations made.

Therefore, the evaluation activities that would necessarily be undertaken should include the following: review to prove the absence of publicly known vulnerabilities and testing to prove that ICT products, ICT services, or ICT processes correctly apply the necessary security functionalities.[16]

The lack of foresight is not the only IoT security issue facing newly digital industries. Another major issue related to IoT security is the technical resource limitations of these devices.[17]

IoT technology has posed new security challenges that go beyond traditional data security due to the unique multi-layered network and nature of the IoT architecture used. IoT security is challenging due to its interaction between physical and digital components. Additionally, the environment of the IoT business model, with both its regularity and contractual obligations and the liability requirements themselves, represent one of the key areas that calls for answers.[18]

Therefore, this paper shows how IoT devices pose a greater risk to personal data than previous surveillance technologies.

V. CONCLUSION

This paper aims to analyze citizens' familiarity with the capabilities of IoT devices.

It also discusses the possibility of introducing education into the entire schooling system in Croatia, starting from an early age. It is necessary to inform all respondents i.e., all users of IoT devices, regardless of age, about their existence, capabilities, possibilities, advantages, and

disadvantages, as well as about the dangers of using their personal data without their knowledge.

This paper aims to raise awareness of the need to apply and improve the legal framework in the the 27 Member States.

The idea behind this paper was to test the research hypotheses and to use the results of this questionnaire in future practice.

From the answers of the respondents can be concluded that the citizens are not familiar with the technology of IoT devices. However, they know that the system in Croatia does not provide adequate education and that the manufacturers do not inform users about the potential risks of devices of this type. This indicates necessity to introduce additional stricter regulations at the level of the European Union so that manufacturers must meet certain standards, while not setting them according to their interests.

Starting with harmonizing and complying with current regulations and recommendations, as well as obliging manufacturers to apply the same in practice.

It is necessary to raise the citizens' awareness that the use of such devices makes the life of an individual easier, but at the cost of taking away their privacy.

According to the questionnaire, respondents have no problem providing personal information about their lives, but they want to be sure that this information will not be used for any other purpose.

The security of the Internet of Things is mostly the responsibility of the manufacturers themselves, who should bear the possible consequences caused by the compromise of systems. [19]

In sum, Croatia needs a fundamental shift when we think of education concerning the IoT devices. We highly recommend the adoption of a national-level strategy for better citizen inclusion in joint decision-making processes regarding the implementation of IoT technology.

Certainly, it is expected that with the further rapid growth of IoT development, adequate and regular and up-to-date education of all citizens will be needed. It will grow human consciousness, and for this very reason, it is necessary to decrease incorrect maintenance.

Ultimately, with the growing technology, the number of security vulnerabilities and data breaches will also increase; therefore, this research continues through another paper on the education and protection of the security of IoT devices.

REFERENCES

[1] Omegasoftware, '' Internet of Things – Begginers Guide'', 2021, https://www.omega-software.hr/internet-of-things-vodic-za-pocetnike/, [Accessed 1st of February 2023]

[2] D. Šipek, DuplicoIT, ''What is IoT or Internet of Things?'', 2022, https://duplico.io/sto-je-iot-ili-internet-of-things/, [Accessed 1st of February 2023]

[3] Ericsson.com, IoT's potential, https://www.ericsson.com/en/internet-of-things, [Accessed 20th of February 2023]

[4] A. Husar, IoT Security: 5 cyber-attacks caused by IoT security vulnerabilities, 2022 https://www.cm-alliance.com/cybersecurity-blog/iot-security-5-cyber-attacks-caused-by-iot-security-vulnerabilities, [Accessed 20th of February 2023]

[5] N. Liebermann, SAM Seamless Network, '' 2021 IoT Security Landscape'', 2022, [Accessed 1st of February 2023]

[6] N. Bradburn, S. Sudman, B. Wiansink, Asking questions : the definitive guide to questionnaire design—for market research, political polls, and social and health questionnaires, John Wiley& Sons, 2004, pp. 303-304.

[7] Croatian Act on implementation of the General Regulation on Data Protection, Narodne Novine, NN 42/2018, (Narodne Novine are an Official Gazette of Republic of Croatia).

[8] Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, preamble 39, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=HR, [Accessed: 20th of February 2023]

[9] G. Vojković, M. Milenković, T. Katulić, IoT and Smart Home Data Breach Risks from the Perspective of Data Protection and Information Security Law, Business Systems Research | Vol. 11 No. 3 |2020, p. 168.

[10] Jurcut, A., Niculcea, T., Ranaweera, P., & LeKhac, A. (2020), "Security considerations for Internet of Things: A survey". SN Computer Science, Vol. 1, Article 193.

[11] ENISA, GUIDELINES FOR SECURING THE INTERNET OF THINGS, Secure supply chain for IoT, 2020, https://www.enisa.europa.eu/publications/guidelines-for-securing-the-internet-of-things (Accessed: 20th of February 2023)

[12] Versoaltima.com, IoT and Cyber Security Challenges, 2022, https://www.versoaltima.com/wp-content/uploads/2019/11/IOT-IGOR-GREGUREC.pdf, (Accessed: 21st of February 2023)

[13] Vojković, G. & Milenković, M. (2021) Liability for Early Safety Obsolescence of IoT Devices due to Information Security Reasons. U: Skala, K. (ur.)MIPRO 2021 44th International Convention Proceedings doi:10.23919/MIPRO52101.2021.9596674

[14] Consolidated version of the Treaty on the Functioning of the European Union,Article 169, *OJ C 202*, In force, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12016E169, (Accessed 20th of February 2023

[15] Vongsingthong S., Smanchat, S. (2015), "Review of data management in Internet of Things", Asia-Pacific Journal of Science and Technology, Vol. 20 No. 2, pp. 215-240

[16] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (Cybersecurity Act), SL L 151, Art. 5 and 6, In force, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.151.01.0015.01.ENG, (Accessed: 21st of February 2023)

[17] S. Shea, Tech Target, IoT Agenda, 2022, https://www.techtarget.com/iotagenda/definition/IoT-security-Internet-of-Things-security, (Accessed 21st of February 2023)

[18] A. M. Awadelkarim Mohamed and Y. Abdallah M. Hamad, "IoT Security: Review and Future Directions for Protection Models," *2020 International Conference on Computing and Information Technology (ICCIT-1441)*, Tabuk, Saudi Arabia, 2020, pp. 1-4, doi: 10.1109/ICCIT-144147971.2020.9213715.

[19] Cert.hr, Security of Internet of Things, 2023, https://www.cert.hr/sigurnost-interneta-stvari-iot/, (Accessed 20th of February 2023.)

APPENDIX - Questionaire

1. U koju dobnu skupinu pripadate?
2. Jeste li čuli za IoT uređaje?
3. Koji od navedenih mogu biti IoT uređaji?
4. Posjedujete li koji od navedenih pametnih uređaja?

5. Smatrate li da su IoT uređaji pouzdani za sigurnost vaših osobnih podataka?

6. Što mislite, koliko su Vaši podaci izloženi cyber napadu ukoliko su povezani na internet?

7. Smatrate li da ste dovoljno upućeni u sposobnost IoT uređaja?

8. Smatrate li da se u RH pruža adekvatna edukacija o IoT uređajima? 1 označava najmanju upućenost, a 5 najveću.

9. Smatrate li da korištenje IoT uređaja unapređuje svakodnevni život korisnika? 1 označava najmanji napredak, a 5 najveći.

10. Planirate li u idućih godinu dana kupiti IoT uređaj?

Device for Testing the Windings of Electrical Machines

Manuel Lonić, Vatroslav Zuppa Bakša, Goran Vujisić
Zagreb University of Applied Sciences / Electrical Engineering, Zagreb, Hrvatska
manuel.lonic@tvz.hr, vatroslav.zuppa.baksa@tvz.hr, goran.vujisic@tvz.hr

***Abstract* - The topic of the paper is the digitization of the process of testing the windings of electrical machines. Windings of electrical machines create a magnetic field in their vicinity. This field should have specific strength allocation in the surrounding area to ensure regular electrical machine operation. By mapping the field strength, it is possible to create a graph of the field strength, which should follow the specified distribution. By checking for deviation, it is possible to detect winding defects. The field is mapped with induction coils, which generate voltage when inserted in the changing electric field. This voltage is then used as an indication of the magnetic field strength. The problem is that the mapping of the field by hand requires many measurements to be made, which is both a time-consuming and an error-prone process. The paper describes the device for testing the windings of electric fields. This device makes the process overall faster and more accurate. Both software and hardware workings are explained in the paper.**

Keywords* – *winding defects; electrical machines; testing; field strength

I. INTRODUCTION

During generator and motor production, various methods are used to ensure quality and industrial standards. Therefore, it is necessary to perform serial testing as efficiently as possible to achieve the highest possible time saving and thus increase earnings. This paper is the result of a collaboration with Končar-Generators and Motors d.o.o. The company played a crucial role in facilitating the testing of the device. An opportunity occurred to save time and ease the work of engineers when performing electrical machine winding testing. This test is performed to prevent an error in the early stages of production. An impulse generator is connected to the ends of the winding, which creates voltage pulses on its clamps, thus loading the winding, which in these conditions should create an electromagnet field of a specific shape and symmetry (Figure 1).

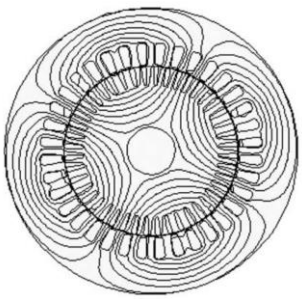


Figure 1. Shape of EMF inside induction rotor [1]

The magnetic field is sampled by a coil in which the signal used for calculations is induced.

Considering the profitability and scalability, the microcontroller was chosen for the A/D conversion of the induced signal. Further, it was necessary to design a computer application to communicate with the microcontroller and process the digital signal sent by the microcontroller. C / C ++ is used within the Visual Studio development environment for that purpose. Necessary conditions were the use of open-source technology, simplicity, and reliability.

II. DETERMINING THE HEALTH OF WINDING

A. *Winding structure*

Winding elements are inserted between laminated core teeth according to the winding diagram in Figure 2.

Figure 2 is the stator winding of 2 pole motor with 24 slots. Pole step measured by the number of slots is 12.

According to Ampere's law, a magnetic field will be created in the space around the coil. The shape can be adapted to requirements using various winding methods [2].

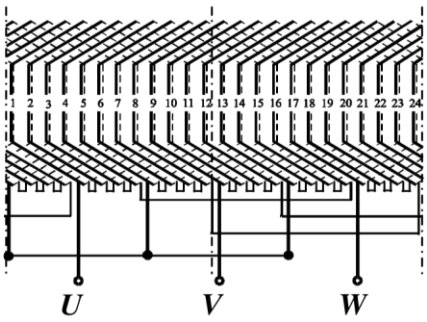


Figure 2. winding diagram [2]

B. *Determining the shape of the magnetic field*

The EMF (electromagnetic field) is measured using a coil that is placed between each of the teeth on the laminated core in such a way that it covers the slot. Faraday's law, written in expression (1), describes how an induced voltage will be created in the coil proportional to the speed of change of the magnetic flux and the number of loops.

$$\varepsilon = -N\frac{\Delta\Phi}{\Delta t} \quad (1)$$

In which:

- ε - induced voltage
- N - number of loops
- ΔΦ - change in magnetic flux
- Δt - change in time

The voltage induced in the coil serves as an indicator of an EMF intensity inside the laminated core. A low-power voltage pulse generator is used to create EMF inside the core. The device loads the winding with voltage pulses that, depending on the design of the winding, creates EMF of a certain shape and intensity. Impulses are supplied to the phase outputs (U1 – U2, W2 – W1, V1 – V2) of the winding. Each phase (U, W, V) is tested separately so that the impulse generator is connected to each end of a phase that is currently being measured. EMF intensity of each slot is then sampled by placing a coil over a laminated core slot and measuring induced voltage. The process is repeated for each slot and each phase. The data acquired this way can be used to determine pole step, phase shift, and isolation compliancy [3].

C. *Determining the form of compliant winding*

A graphical display of collected data gives a clear picture of the state of the stator and is commonly used in the analysis of results. Figure 3 is a compliant example of a 16-pole rotor with 96 slots. Pole step is 6 slots.

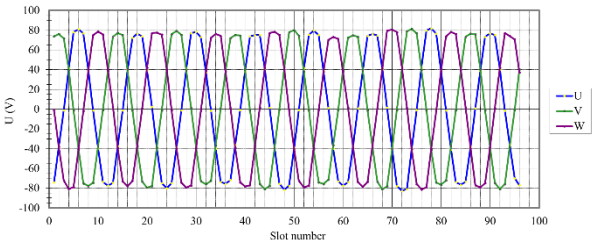


Figure 3. Compliant winding example [4]

Figure 4 shows the non-compliant measurement of incorrectly wound winding.

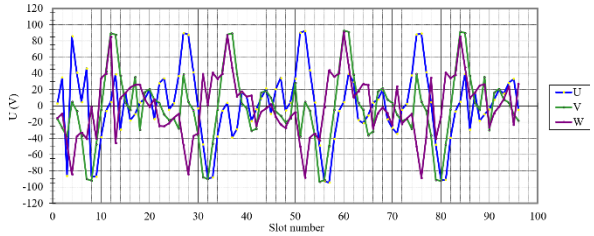


Figure 4. Non-compliant winding [4]

III. HARDWARE WORKINGS OF THE DEVICE

A. *Measurement coil*

Figure 5 shows the characteristic of different types of magnetic field sensors.

There are several types of magnetic field sensors, such as measuring coil and Hall effect sensors. From Figure 5, accuracy and sensitivity characteristics can be obtained.

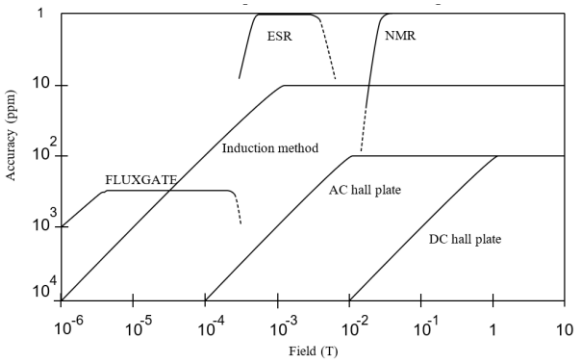


Figure 5. Measurement methods: accuracies and ranges [5]

The pulse-generating device does not supply much energy to the winding, resulting in a very weak field. Therefore, choosing a sensor that can measure a low magnetic flux density is necessary. Measuring coils (Figure 6) are used when testing windings due to their high sensitivity to the magnetic field. Their dimensions must correspond to the size of the slot [5].

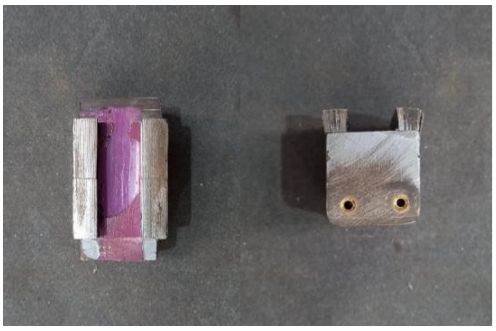


Figure 6. Measurement coil

B. *Signal processing device*

The signal conversion device is shown in Figure 7. It contains a microcontroller and a signal adjustment circuit.



Figure 7. Signal processing device

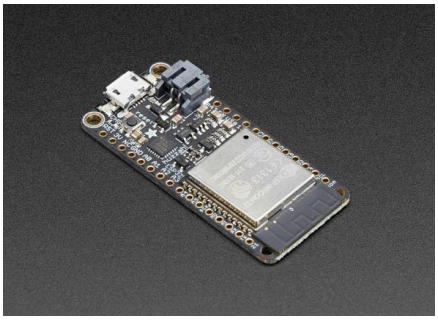


Figure 8. ESP32-based development board [6]

The microcontroller is an ESP32 model built into the Adafruit Huzzah development board (Figure 8).

It was chosen because of its 12-bit A/D converter and built-in Wi-Fi and Bluetooth capabilities, which are necessary if it is decided to modernize the way of connecting to a computer. For example, Bluetooth data transfer. Such a feature would allow the test engineer to move quickly around large stators or rotors.

The development board is insufficient for A/D signal conversion because it requires a maximum 3,3 VDC input signal while the coil's output is in the range of -300 to 300 VAC. Therefore, it is necessary to reduce the value of the signal proportionally and convert it into a DC signal.

For this purpose, a signal adjustment circuit was designed (Figure 9). Voltage reduction is performed via a voltage divider. Rectification of the signal is achieved by directing the positive part of the signal to one input, and the negative portion of the signal to another input of the microcontroller, i.e., to another A/D converter which perceives it as a positively induced signal which is later easily recognized by the software as a negative part of the signal. In this way, the resolution is increased two times compared to the resolution that can be obtained in the case of using the DC offset method.

The biggest challenge of such a circuit is determining the gain and time constant of the system. Since the device needs to work with different winding types, it must be able to cope with a wide range of EMF intensity, which can drive a signal under or over the saturation value of 0-3,3V. Therefore, it is necessary to carefully determine the gain so the device can be used on a wide range of test objects.

Formula (2) calculates the amplification of the voltage divider of the positive part of the signal. Formula (3) calculates the amplification of the voltage divider of the negative part of the signal.

$$K_{10} = R_8 / (R_6 + R_8) \quad (2)$$

$$K_2 = R_9 / (R_7 + R_9) \quad (3)$$

In which:

- K_1 and K_2 - gain of voltage dividers.
- R_8 and R_9 - resistors from which the signal is supplied to the A/D converter.
- $(R_6 + R_8)$ and $(R_7 + R_9)$ - total resistance of the divider.

K_1 and K_2 must be equal; in this case, they amount to 0,0099. However, due to the resistance error problem of non-uniformity between voltage divider R6 - R8 and the voltage divider R_7 - R_9 appears. This results in the circuit amplifying the positive and negative parts of the input signal unequally, which can lead to asymmetric results. Special resistors with a precision of 1% were used to solve this problem. Furthermore, gains of voltage resistors can be compensated in digital signal processing.

The time constant should not be too small due to the accumulation of interference but also not too large because the signal itself would not completely follow the field intensity changes. Expressions (4) and (5) calculate the time constant, which must be the same for both voltage dividers, as in the case of amplification. In this case, the time constants are 0,101 s.

$$\tau_1 = (R_6 + R_8) \cdot C_3 \quad (4)$$

$$\tau_2 = (R_7 + R_9) \cdot C_4 \quad (5)$$

In which:

- τ_1 and τ_2 - time constant.
- $(R_6 + R_8)$ - and $(R_7 + R_9)$ - total resistance of the divider.
- C_3 and C_4 - capacitor capacitance.

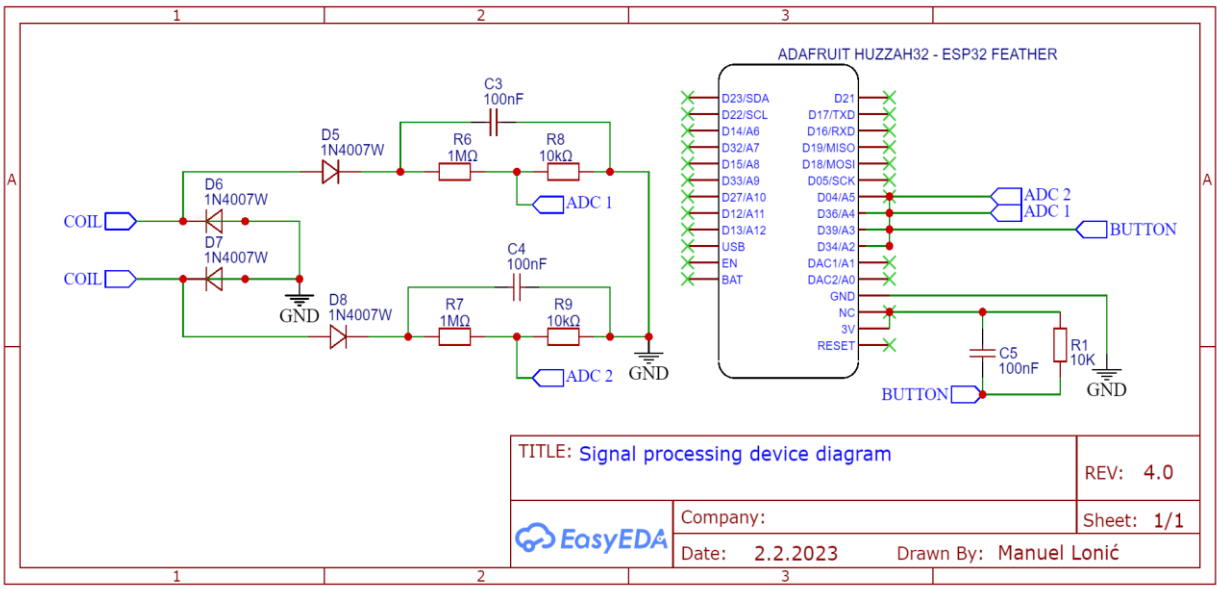


Figure 9. Schematics of signal processing device

Due to the need for system portability, the circuit was redesigned on PCB (Printed Circuit Board) technology (Figure 10).

PCB boards are carriers for electronic components on which conductors are printed, and slots are installed. The components installed on the printed circuit boards have unique housing dimensions adapted for installation on the PCB boards. The most significant advantage is the small dimensions of the final product. The PCB shown in Figure 10 is designed to be compatible with the Adafruit huzzah dimensions, meaning it can be directly installed on the development board itself.

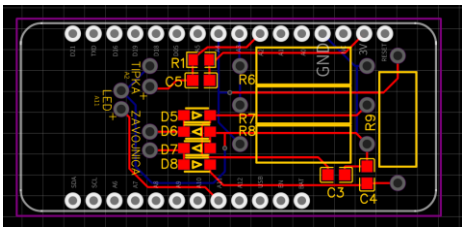


Figure 10. PCB diagram of the signal processing circuit

To achieve the system's resistance to external effects, a case (Figure 11) was designed, the dimensions of which correspond to the Adafruit development board.



Figure 11. Inner workings of signal processing device

IV. SOFTWARE WORKINGS OF THE DEVICE

A. PC application

After the obtained digital signal, it is necessary to process it and obtain values from the magnetic field of the slot, such as in figure 3. The goal is a simple computer application based on available Windows 10 technology that prints data in graphic and textual form. The application consists of a text interface (Figure 12.) for monitoring the measurement progress and a visual graph (Figure 13). The graph shows the values of the induced voltage inside the measuring coil when measuring the magnetic field around the slot.

```
D:\Završni rad\MIPRO\WindingTester.exe
Measurement name: example
Number of slots: 9
Choose the measurement type to be conducted:
Type in a number 0, 1 or 2

0 = phase measurement (W, V, U)
1 = Symmetry measurement (W, V, U)
2 = Single slot check
```

Figure 12. Textual graphics user interface

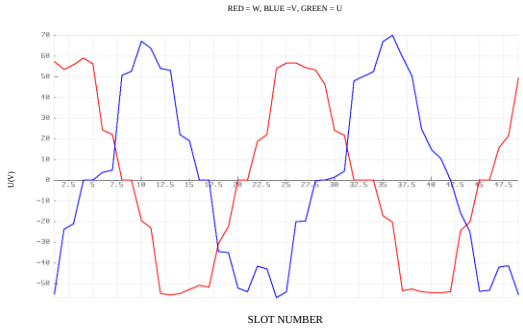


Figure 12. Graphical user interface

B. Code implementation

The software implementation conditions were to build an application on the Windows 10 platform and use open-source technologies listed in table 1. During the program's writing, problems were the creation of a communication channel between the microcontroller and the application, data processing, and display.

TABLE 1. USED LIBRARIES

Microcontroller communication	Graphing tool	Graphical user interface
SerialPort.cpp	pbPlots.cpp	opencv_world454.lib
SerialPort.h	pbPlots.hpp	-
-	supportLib.cpp	-
-	supportLib.hpp	-

Communication with the microcontroller is carried out using the RS 232 protocol, controlled by functions from the SerialPort.cpp library. In this way, the computer application communicates directly with the microcontroller. The library is available under the MIT license.

The tool used to plot the graphs from the data is pbPlots. PbPlots is an open-source plotting library that can be used with many program languages, such as C, C++, Java, JavaScript, and Python. The library is available under the MIT license.

OpenCV is used for the graphical display of results. OpenCV is an open-source library of programming functions mainly focused on real-time computer vision. The software is available under the Apache license.

C. Arduino application

The programming code for the microcontroller is written within the Arduino development environment. The Arduino code digitizes and formats the input signal into a format recognized by the application.

V. DEVICE TESTING

The measurement described below was conducted to prove the device's functionality (Figure 13).

The example below is the excitation system rotor of a hydro generator. This rotor consists of 8 poles and 72 slots.

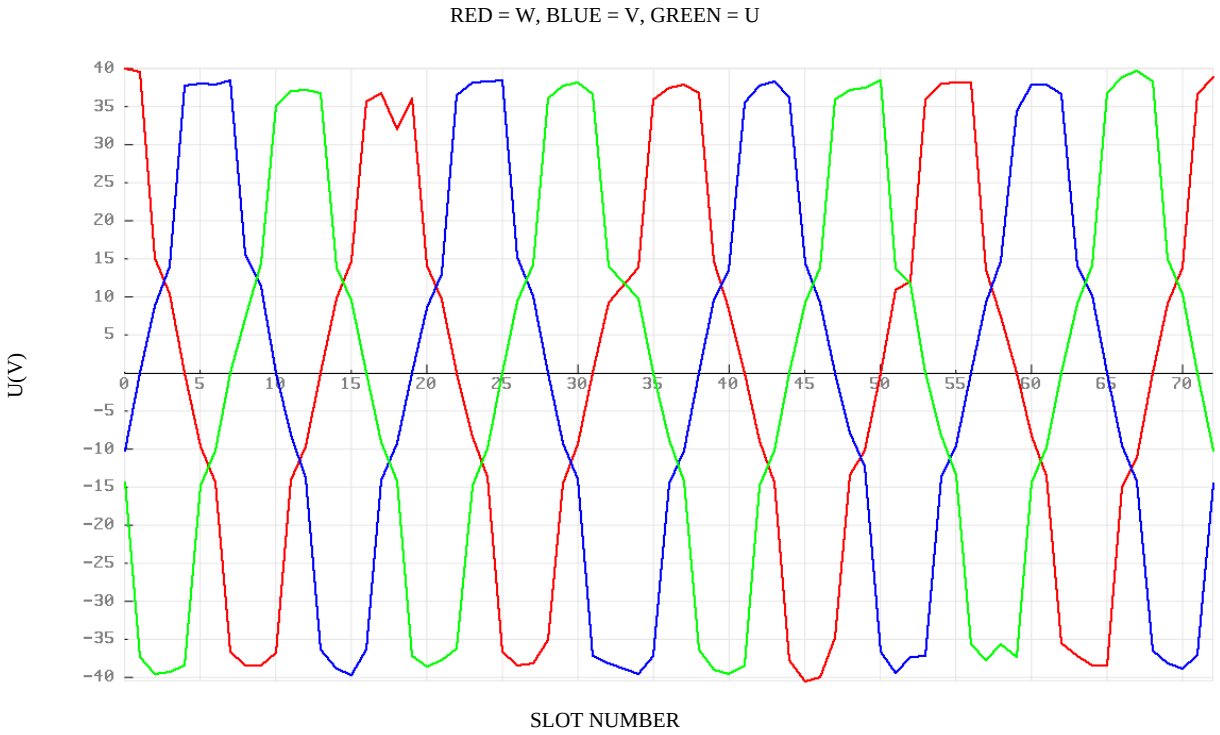


Figure 13. Measurement of the excitation system rotor of a hydro generator

Expected characteristics are a phase shift of 120° and a pole step of 9 slots per pole. Figure 13 can prove the compliancy of the winding. Furthermore, the measured results can be compared with the results obtained by the classic method of measuring each slot's EMF by hand and method, including the testing device, and thus prove the device's validity. Figure 3 shows the measurement results of a similar winding obtained by the classical method of measuring the voltage drop on the coil for each slot. From the comparison of the obtained results, it can be concluded that the results obtained by the winding testing device are credible.

The signal converter's accuracy depends on the voltage divider resistor error and the quantization error of the A/D converter. The error of the voltage divider resistor leads to a systematic error that can be compensated by software.

When calculating the voltage value of a certain point, the method of multiple sampling is applied, and the arithmetic mean of the sampled samples is calculated. The quantization error does not significantly affect the result. Only errors caused by human factors can cause incorrect results.

It is important to note that the high accuracy of the reading of the induced voltage inside the measuring coil is not overly critical because when assessing the validity of the measured object, attention is paid to the symmetry of the poles, the correct arrangement of the phases and the pole step, the case of an inter-winding short circuit results in a significant drop in the intensity of the magnetic field at only 0-20% of the expected value.

VI. CONCLUSION

Higher production output and increased employee productivity are two of the biggest reasons for implementing automation. Most tasks that are repeated in the same way can usually be automated. In addition to industrial processes, office work like creating reports and protocols can also be the subject of automation. As a result, automation can help engineers do their jobs faster and easier.

The goals of this project are to save time and eliminate errors from the routine of testing electrical machines. This solution removes any copying of data from forms to excel sheets or monotonous data processing. For example, it was estimated that using the classical method by hand when measuring windings with 72 slots takes 10 to 12 minutes more than with the winding testing device. Nevertheless, the most important thing is that there is no further processing of the data, which adds to time-saving, and eliminates the risk of mistakes during filing forms or copying data. While planning the development of the device, the shortcomings of the existing process were considered, and the good sides were retained. Furthermore, the built device is based on open-source technology, which ensures the possibility of constant upgrading and adaptation of the system to all further changes in the process.

REFERENCES

[1] J. F. A. T. Bashir Ebrahimi, "Losses Calculation in Line-Start and Inverter-Fed Induction Motors Under Broken Bar Fault," *IEEE Transactions on Instrumentation and Measurement*, p. 14, 2013.

[2] I. Mandić, V. Tomljenović i M. Pužar, Sikroni i asnikroni električni strojevi, Zagreb: Tehničko veleučilište u zagrebu, 2012.

[3] C. P, "Fault detection in three-phase rotating machines," *IEEE Transactions on Power Apparatus and System*, 1972.

[4] KONČAR – Generators and motors Inc., *Check of t rotor winding throughput with low impulse voltage.*, Zagreb, 2021.

[5] K. N. Henrichsen, "Overview of magnet measurement methods," 1998.

[6] Adafruit Industries, "Adafruit HUZZAH32 - ESP32 Feather," 2 1 2023. [Online]. Available: https://cdn-learn.adafruit.com/downloads/pdf/adafruit-huzzah32-esp32-feather.pdf. [Accessed 2 2 2023].

Electrical Scheme Digitization Using Deep Learning Methods

Marko Putak, Vatroslav Zuppa Bakša and Andrea Bednjanec
Zagreb University of Applied Sciences/Automation and Process Control Engineering, Zagreb, Croatia
mputak98@gmail.com, vzuppabak@tvz.hr, andrea.bednjanec@tvz.hr

***Abstract* - The use of software tools and applications progressively became a standard in both education and industry. A solution for hand-drawn electrical scheme digitization has been proposed to match the fast-paced dynamic of the modern world in the field of electrical engineering. The aim is to notably reduce time-consuming and error-prone electrical scheme tracing from hand-drawn to simulating software. The means have been achieved through the usage of state-of-the-art deep learning model YOLOv5 for electrical elements detection along with Python and OpenCV library for data processing. The user's input is an image of a hand-drawn circuit, and the end result is an LTspice digitized electrical scheme ready for simulation.**

Keywords - electrical scheme; digitization; deep learning; object detection; YOLOv5; LTspice

I. INTRODUCTION

Electrical circuits are a crucial aspect of electrical engineering. The conventional method of creating an electrical circuit involves manually drawing it on paper, followed by redrawing it in a simulation program during the later stages of analysis. This redrawing process is time-consuming, reduces efficiency, and can result in human error.

The intention to automate the process seems to be of great use for the purpose of removing the need for redrawing and setting the scheme in a state ready for simulation. A branch of computer science that offers a solution of the mentioned problem is artificial intelligence.

The proposed solution consists of using an electrical scheme image dataset to train a deep learning model for electrical component detection. Henceforth, detected components are processed to produce a fully functional digitized electrical scheme.

II. DATASET

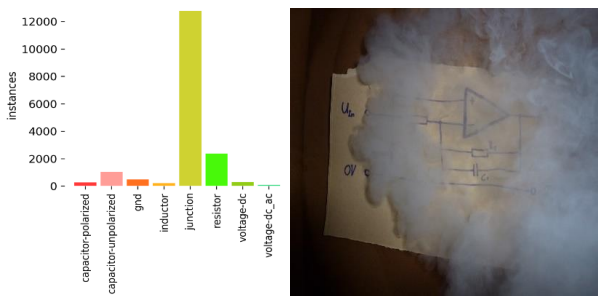
Datasets are an integral part of contemporary object recognition research [1]. Generally, the weak and unattended dataset is a bottleneck to the model's performance, leading to poor results altogether. It is of uttermost importance to squeeze in those extra hours to find a good match, as it will be significant for the rest of the project.

For this project, the CGHD1152 (Circuit Graph Hand Drawn 1152) dataset was chosen as the best available option [2]. CGHD1152 consists of 1152 *.jpg* images derived from 144 unique circuits accompanied by 48563 object annotations. Despite the dataset is divided into 45 unique classes (junction, resistor, speaker, switch, etc.), 7 classes are being used for the sake of efficiency and accuracy.

Further deconstruction and visualization of the dataset provide a clear picture of its unbalanced class distribution seen in Fig. 1 (a), and unrepresentative instances in Fig. 1 (b).

The CGHD1152 dataset was purged of outliers and unannotated images using Roboflow, a computer vision platform designed for dataset manipulation, export, and deployment [3]. Preprocessing steps, such as auto-orientation, resizing, and conversion to grayscale, were applied through Roboflow. Image augmentation was not included in the preprocessing since YOLOv5 handles it internally.

The images were divided into three standardized subsets: training, testing, and validation, comprising 70% (813 images), 15% (175 images), and 15% (174 images) of the data, respectively. In addition, roughly 1% of background images (null images) were added to reduce False Positive (FP) classification, following Ultralytics' recommendations [4].



(a) (b)

Figure 1. Dataset visualization.
(a) Class distribution, (b) unrepresentative instance

III. YOLOv5

YOLOv5 is the state-of-the-art (SOTA) model specialized for object detection. The abbreviation YOLO is derived from "You Only Look Once", meaning that the model uses single-shot detection (SSD) to predict the presence of an object and its corresponding location [4].

What makes YOLO a superior model and the perfect candidate is the usage of a grid-like division during inference that allows the model to see the input image as a whole while maintaining record-level speed and high mean average precision (mAP) [5].

A. Model Architecture

- Backbone: CSP-DarkNet53
- Neck: CSP-Path aggregation network (PAN)
- Head: YOLOv3 Head

This particular Backbone structure ensures high inference speed and accuracy while maintaining relatively low memory cost by solving the issue of repetitive gradient information passage [6].

In the Neck, the Path Aggregation Network is used to boost the information flow of the network by propagating low-level features in a feature pyramid [7].

The model Head is mainly used to perform the final detection part. It applies anchor boxes on features and generates final output vectors with class probabilities, objectness scores, and bounding boxes [8].

Considering the scarcity of data provided in the dataset[1], as well as possible applications of the model in real-time detection scenarios, the YOLOv5m version is selected based on its ideal balance of inference speed and prediction accuracy. YOLOv5m consists of 291 layers, 20.9 million parameters, and uses 48.3 GFLOPs.

B. Initial training

The initial model training was accomplished with a Jupyter notebook script running on Google Colab [9][10]. The script begins with installing the dependencies and importing the dataset. Afterward, the training procedure is initialized using pre-trained weights to accelerate the training process. The simplicity provided by Ultralytics makes the training process intuitive and easy to debug where necessary. The initial weights, yolov5m.pt, were trained on the COCO dataset [11]. To understand the general behavior of the model, most hyperparameters were set to their default values, with notable exceptions being the number of epochs (150), batch size (32), image size (640), and optimizer (Stochastic Gradient Descent).

The first training took 1 hour and 3 minutes using the Tesla T4 graphics card (GPU) and yielded satisfactory results.

The best results were at epoch 136 with mean average precision at 0.5 threshold of intersection over union (mAP@0.5) equating to 0.9888 and mAP [0.5:0.95] of 0.7353, where [0.5:0.95] represents different IoU thresholds, from 0.5 to 0.95 with a step of 0.05.

C. Hyperparameter Evolution

Having a well-defined base case of the model and its metrics, the usual procedure would involve tweaking and experimenting with different hyperparameter values used to train the model repeatedly in the hope of improvement. This is also known as hyperparameter tuning [12].

In order to avoid such time-consuming, exhaustive and iterative process, hyperparameter evolution offers an interesting solution [13].

The evolutionary algorithm uses a single gene to encode each hyperparameter that needs to be optimized for each individual. A range and resolution are specified for each gene to prevent searching irrelevant regions of the hyperparameter space. The initial population is generated by randomly choosing each gene from a uniform distribution, after which the fitness of each individual is assessed. The individuals with the highest fitness from the previous generation are used to form subsequent generations through selection, crossover, and mutation [14].

Ultralytics YOLOv5 repository has a built-in hyperparameter evolution function that uses modified mutation to find the best set of hyperparameter values. For the evolve parameters a number of epochs is 10 per default settings, and a number of generations is set to 150. Evolution lasted for 4 hours in total using RTX 3070Ti GPU, and the outcome was a fine-tuned set of hyperparameters where some of which are depicted in Fig. 2.

Marginal improvements were made by retraining the model on the new set of hyperparameters, as is depicted in the F1 curve in Fig. 3. The junction's F1 score is low compared to other classes due to its small and overrepresented nature, hence this significant disparity in the results.

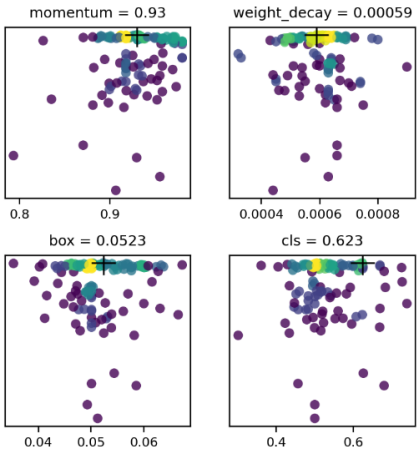


Figure 2. Hyperparameter evolution scatter plot

[1] A decent dataset contains more than 1500 images per class!

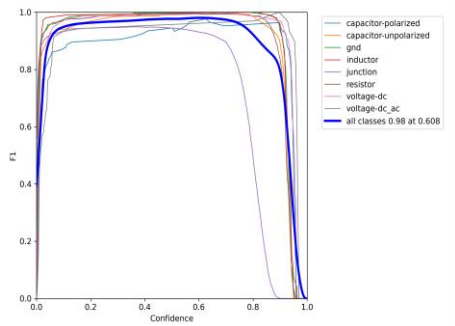


Figure 3. F1 curve of the model with tuned hyperparameters

Newly found metrics are:

- mAP@0.5 = 0.9895
- mAP@[0.5:0.95] = 0.7496
- Precision = 0.9794
- Recall = 0.9820

D. Final Model

Further results and dataset analysis discovered that increasing the image size from 640 pixels to 1280 pixels has a positive impact on the training results since the majority of the bounding boxes, especially junctions, are quite small and easily misplaced. Consequentially, changing the image size to 1280 pixels and retraining the model with evolved hyperparameters gave the best results.

Training took 2 hours and 56 minutes to complete at epoch 167, and was logged using wandb, an experimental tracking tool for machine learning. This provided several valuable metrics, such as the normalized confusion matrix shown in Fig. 4 [15].

Examining the confusion matrix, it can be seen that the model's classification is highly accurate, except for background images where 76% of the time the model confused them with junctions.

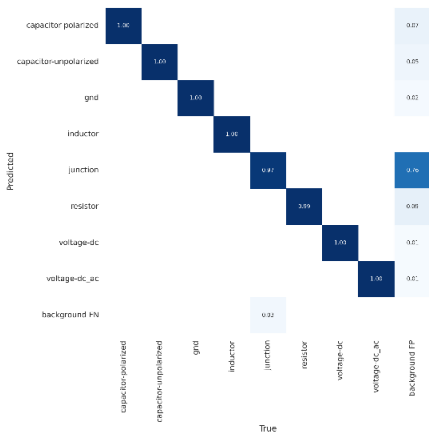


Figure 4. Confusion matrix of the final model

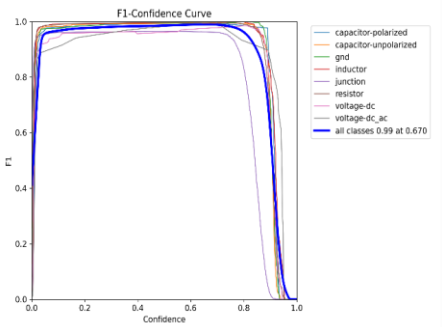


Figure 5. F1 curve of the final model

A comparison of the final model's F1 score in Fig. 5 and the evolved model's F1 score in Fig. 3 revealed that the final model had a better curvature, especially for the junction class.. This is reflected in the mAP metric, which combines precision and recall in its calculation, resulting in:

- mAP@0.5 = 0.9907
- mAP@[0.5:0.95] = 0.7614
- Precision = 0.9883
- Recall = 0.9909

making this model a perfect candidate for later use.

Overfitting occurs when a model does not generalize well from observed data to unobserved data [16]. YOLOv5 addresses this issue by using regularization techniques such as weight decay and early stopping [17]. Early stopping is a regularization method that aims to stop the training process before the model begins to overfit, specifically at the inflection point of the validation error function [18]. This is why the training ended at 167th epoch - the model stopped at the optimal time and saved the best weights to a file named *best.pt*.

The increase in image size from 640 pixels to 1280 pixels did result in improved overall accuracy for the model, however it also led to a decrease in both training and inference speed. This trade-off is commonly encountered in the field, and the choice is determined based on the specific requirements of the model application. In the case of electrical scheme digitization, the inference speed of 39.8 milliseconds is deemed appropriate for the task and therefore the model is ready to be implemented.

IV. DIGITIZATION

With the model thoroughly evaluated and ready, the next step is to begin data processing using Python 3.10 programming language and PyTorch as the primary framework [19][20].

A. LTspice

LTspice XVII, a high-performing computer program for electrical circuit simulation, has been chosen as the desired output. It is based on the SPICE simulator (Simulation Program with Integrated Circuit Emphasis), open-source computer software used to analyze and predict

the behavior of electronic circuits [21]. LTspice is widely used and accepted as the most prevalent SPICE software in the industry, making it a valuable choice due to its high user base.

Although ease of use, clarity, low computational complexity and accuracy are significant benefits of the LTspice program, the main benefit of using LTspice is its ability to render a textual type of data and create a circuit. This is achieved through a simple syntax that represents a specific circuit, making digitization possible. Fig. 6 (a) illustrates an electrical circuit in LTspice and Fig. 6 (b) its matching syntax created in the text file.

To get a better understanding of how the syntax should be written, several electrical circuits were manually drawn and inspected from their correlating text format. The conclusion was deduced for the format in which the text file should be written generally:

- <Name of the element> <coordinates>

For instance, a random wire element is written as:

- WIRE x_s y_s x_e y_e

Where x_s, y_s represents the starting point, and x_e y_e represents the ending point of the wire.

Similarly, electrical components are written:

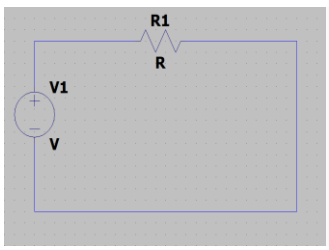
- SYMBOL res x_c y_c Rz

Where x_c, y_c represents the center point and z $\in$ {0, 360} represents the rotation of the resistor.

B. Code Implementation

The model is loaded through PyTorch's built-in function *torch.hub.load()* with custom weights as an argument. OpenCV's *.imread()* method is used to load a new image that the user wants to digitize. With the new image, the model makes predictions and visualizes the results using *.show()* method depicted in Fig. 7. Each element is surrounded by a bounding box with a confidence score.

The interpretation of the model's results is made convenient through the use of pandas, an open-source Python library for data analysis and manipulation [22]. Calling the *.pandas().xywh[0]* function on the result, a table shown in Fig. 8 is generated containing center coordinates, width, height, and other useful information regarding the predicted bounding boxes.




```
Version 4
SHEET 1 880 680
WIRE 272 80 112 80
WIRE 528 80 352 80
WIRE 112 144 112 80
WIRE 112 320 112 224
WIRE 528 320 528 80
WIRE 528 320 112 320
SYMBOL res 368 64 R90
SYMATTR InstName R1
SYMBOL voltage 112 128 R0
SYMATTR InstName V1
```

(a) (b)

Figure 6. LTspice electrical circuit.
(a) Generated circuit, (b) circuit's text format

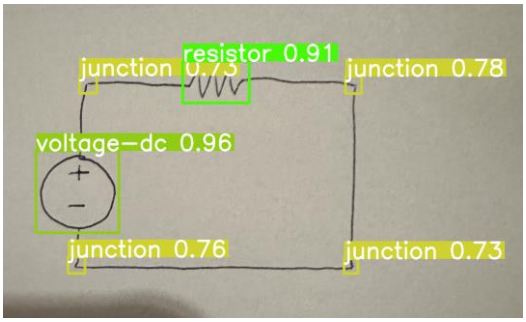


Figure 7. Visualization of model inference

As the WIRE elements in LTspice dictate the connections of the circuit and the placement of the elements, the first challenge faced was the finding of the start and end points of the wires. Upon examination of the junction class, it was deduced that the junction class was composed of two or more lines that were horizontally and vertically connected, making it a perfect match for the endpoints of the wires. Hence, by iterating over the junction class in the dataframe, the wire endpoints were derived and stored in a list.

However, simply connecting every combination of junctions is not correct, as there may not be a wire between every junction. A solution was proposed to only connect junctions that were horizontally or vertically aligned. This was achieved by finding the angle of slope formed by two given junctions.

The slope of any straight line on the *X*-*Y* plane is given by *tan* θ, where θ is the angle that the straight line makes with the positive direction of *x*-axis. Hence the equation for the slope is

$$m = \tan\theta = \frac{y2-y1}{x2-x1}. \quad (1)$$

Where m is the slope of that straight line which is inclined at an angle θ with the positive direction of x-axis, and (x1, y1), (x2, y2) are coordinates of the first and second junction, respectively.

Therefore, the angle of slope is

$$\theta = \tan^{-1}(m) = \tan^{-1}\left(\frac{y2-y1}{x2-x1}\right). \quad (2)$$

The wire is classified as horizontal if the angle of slope falls within 180 ± 10% or 360 ± 10%. Conversely, the wire is classified as vertical if the angle of slope falls within 90 ± 10% or 270 ± 10%. The elimination of non-orthogonal wires leaves the initial list comprised of only valid wires.

	xcenter	ycenter	width	height	confidence	class	name
0	750.46	1928.72	403.03	398.10	0.96	6	voltage-dc
1	1422.68	1395.59	323.33	202.92	0.91	5	resistor
2	2091.57	1409.49	77.22	83.85	0.78	4	junction
3	745.30	2285.22	82.79	78.64	0.76	4	junction
4	801.75	1407.23	83.94	86.41	0.73	4	junction
5	2079.28	2289.68	66.88	71.54	0.73	4	junction

Figure 8. The model inference dataframe

For loop was included to iterate over the list and write the coordinates to a file, which was later opened in LTspice for analysis. It was discovered that the wire validation method was effective, but the wires were not perfectly horizontal or vertical as a result of the user's hand-drawing or image capturing not being properly aligned. As a result, the user was unable to interact with the wires or elements, as LTspice only allows for manipulation of vertical and horizontal straight wires.

To overcome this challenge, a wire alignment function was implemented. To align the vertical wires on the *x*-axis, a function looped through all the points' *x*-coordinates, comparing their relative positions, and making them equal if they satisfied a proximity condition given by a threshold value. A similar procedure was implemented for the alignment of the horizontal wires using the *y*-coordinates of the points.

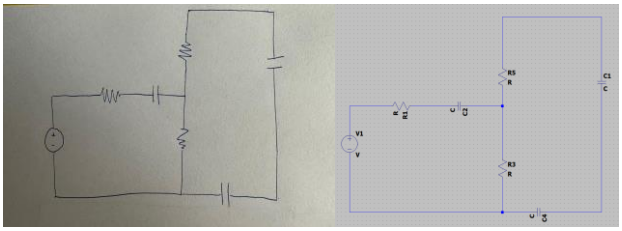
However, the aligned wires were out of reach and could not be manipulated by the user as they can only interact with wires on the edges of 16 by 16-sized grids. To address this bug and have the user be able to add, remove or modify the wires and elements, a simple wire offset method was implemented as shown in the following pseudocode:

```
for wire in all_wires_list:
    if wire[x_coordinate] % 16 == 0:
        continue
    else:
        offset_x = wire[x_coordinate] % 16
        wire[x_coordinate] -= offset_x
```

The *y*-coordinate offset correction was also applied with the same logic.

Similar to wires, electrical elements were also added to the output file using the *.write()* built-in Python function. If the element is placed on a wire, LTspice will automatically connect the element in a circuit and remove the excess wire, therefore the main challenge was determining the appropriate rotation for these elements based on whether they needed to lie on a horizontal or vertical wire. This problem was resolved by splitting the list of all valid wires into two smaller lists - *horizontal_wires* and *vertical_wires*. By comparing whether the center of the element was located between two junctions of a horizontal or a vertical wire, the appropriate rotation of the elements was assigned and included in the output file.

Upon program execution, a *.asc* file is generated in the current directory containing all the wires and elements depicted in a hand-drawn scheme and can be opened using LTspice. Fig. 9 shows an example of the input image (a) and its corresponding end product (b).



(a) (b)

Figure 9. A before-and-after example.
(a) Input image, (b) end result

C. *Limitations*

Although providing an efficient and convenient way of digitizing electrical schemes, there are some limitations to the process.

One of the main limitations is the possibility of using only six different types of elements that occur most frequently. Excluding other elements could be problematic for users who require a wide range of elements. On the other hand, the computational expenses of providing more elements are marginal, therefore, including more elements in the future is realizable.

Another constraint is the inability to use non-orthogonal wires, such as diagonal wires, due to the method used for finding valid wires. Additionally, the current dataset is insufficient and lacks the necessary data to train the model accurately.

D. *Results*

A solution has been developed to meet the challenge of creating a digitized and fully operational electrical circuit from hand-drawn designs. By using the highly accurate deep learning model YOLOv5, which was trained on a custom dataset, and processing the output from the model, a fully functional LTspice electrical circuit was created. The digitization process involves the user inputting an image, the model making predictions, and processing the output to produce a file named "output.asc" that can be opened in LTspice for simulation or measurement purposes.

The end result was further validated by using a subset of data from the dataset along with newly-drawn circuits. Although the validation procedure produced positive results, it also identified areas requiring improvement, specifically in the heuristics of digitization, which tend to produce errors when the drawings are complex or cluttered.

The digitization process is nearly independent of the drawing's complexity, making it a more efficient alternative to manual redrawing. In most cases, the entire process can be completed in less than a minute.

V. Conclusion

The proposed program for hand-drawn electrical scheme digitization is a promising solution that can greatly improve efficiency and reduce errors in the field of electrical engineering. By leveraging state-of-the-art deep learning techniques for object detection, the program successfully automates the time-consuming and error-prone process of redrawing hand-drawn circuits for simulation. While further improvements and validations are required, the program has demonstrated its potential to offer an accessible and reliable solution for a wide range of users.

Acknowledgment

I extend my heartfelt gratitude to my co-authors for their unwavering support and assistance throughout the entire project. I'd like to recognize the assistance of my fellow colleagues, and lastly, it would be remiss not to mention the support I received from Miss Bruna Duspara.

REFERENCES

[1] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," CVPR 2011, Colorado Springs, CO, USA, 2011, pp. 1521-1528, doi: 10.1109/CVPR.2011.5995347.

[2] Thoma, Felix, Johannes Bayer, and Yakun Li. "CircuitGraphHandDrawn." OSF, 8 Aug. 2022. Web.

[3] Dwyer, B., Nelson, J. (2022), Solawetz, J., et. al. Roboflow (Version 1.0) [Software]. Available from https://roboflow.com. computer vision.

[4] Jocher, G. (2020). YOLOv5 by Ultralytics (Version 7.0) [Computer software]. https://doi.org/10.5281/zenodo.3908559

[5] Thuan, Do. "Evolution of Yolo algorithm and Yolov5: The State-of-the-Art object detention algorithm." (2021).

[6] Xu, Renjie & Lin, Haifeng & Lu, Kangjie & Cao, Lin & Liu, Yunfei. (2021). A Forest Fire Detection System Based on Ensemble Learning. Forests. 12. 217. 10.3390/f12020217.

[7] Liu, Shu, et al. "Path aggregation network for instance segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[8] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[9] Kluyver, T., Ragan-Kelley, B., Fernando P'erez, Granger, B., Bussonnier, M., Frederic, J., … Willing, C. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt (Eds.), Positioning and Power in Academic Publishing: Players, Agents and Agendas (pp. 87–90).

[10] Bisong, E. (2019). Google Colaboratory. In: Building Machine Learning and Deep Learning Models on Google Cloud Platform. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4470-8_7

[11] Tsung-Yi Lin, Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., … Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. CoRR, abs/1405.0312. Retrieved from http://arxiv.org/abs/1405.0312

[12] Bardenet, Rémi, et al. "Collaborative hyperparameter tuning." International conference on machine learning. PMLR, 2013.

[13] Schmidt, Mischa, et al. "On the performance of differential evolution for hyperparameter tuning." 2019 international joint conference on neural networks (IJCNN). IEEE, 2019.

[14] Young, Steven R., et al. "Optimizing deep learning hyper-parameters through an evolutionary algorithm." Proceedings of the workshop on machine learning in high-performance computing environments. 2015.

[15] L. Biewald, "Experiment Tracking with Weights and Biases," Weights & Biases. [Online]. Available: http://wandb.com/. [Accessed: 02.02.2023.]. Software available from wandb.com

[16] Ying, Xue. "An overview of overfitting and its solutions." Journal of physics: Conference series. Vol. 1168. IOP Publishing, 2019.

[17] Loshchilov, Ilya, and Frank Hutter. "Decoupled weight decay regularization." arXiv preprint arXiv:1711.05101 (2017).

[18] Yao, Yuan, Lorenzo Rosasco, and Andrea Caponnetto. "On early stopping in gradient descent learning." Constructive Approximation 26.2 (2007): 289-315.

[19] Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.

[20] Paszke, Adam, et al. "Automatic differentiation in pytorch." (2017).

[21] Quarles, Thomas L., Analysis of Performance and Convergence Issues for Circuit Simulation, Memorandum No. UCB/ERL M89/42, University of California, Berkeley, April 1989.

Body fat percentage calculation: A Linear Regression Model on Croatian tennis players morphology

Matea Vasilj*, Vlatko Vučetić‡, Marko Sukreški‡, Dario Bojanjac*
* Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
† Faculty of Kinesiology, University of Zagreb, Croatia
‡ Ultrax technologies, Zagreb, Croatia
matea.vasilj@fer.hr

***Abstract*—Body composition is an important aspect of the fitness and performance of tennis players. Measuring body composition accurately can be a challenging task due to its complex nature and various factors that can affect the results. This study aims to examine the relationship between tennis player morphology and body fat through mathematical models, reducing the reliance on costly measurement equipment. Morphology of Croatian tennis players dataset was used, and the results of statistical analysis showed a strong correlation between player morphology, age, sex and body fat percentage. Furthermore, a PCA analysis was carried out which revealed that male and female populations become distinct when they hit puberty and allows us to explain the differences in body composition between the sexes. Using these results, a set of linear regression models were developed that can be used to predict body fat percentage. This research adds to the existing literature by highlights the significance of factors affecting body fat percentage in tennis players, which are skin fold thickness, age, and sex. The newly developed models offer a cost-effective solution for coaches to assess and monitor body composition in tennis players with skin fold measurements instead of more expensive equipment.**

Keywords*—body fat percentage, tennis player morphology, tennis, data analysis, machine learning

I. INTRODUCTION

Body composition has long been of interest to researchers in the fields of kinesiology and sports science. It is an important factor in determining athletes performance, and it is widely recognized that higher levels of body fat are associated with lower levels of performance in various sports. The Faculty of Kinesiology has been performing periodic measurements on athletes since the early 2000s, and was able to provide us with the Croatian tennis player morphology data needed for this research. In tennis, body composition is especially important as it can affect factors such as speed, agility, and endurance. However, measuring body composition accurately and reliably can be challenging due to the cost and practical limitations of traditional methods such as dual-energy x-ray absorptiometry (DXA) and underwater weighing, especially for large groups of athletes [1].

In this study, we aimed to develop a more cost-effective and practical solution for measuring body composition in tennis players. We used a dataset containing morphology data of tennis players, and applied statistical analysis, principal component analysis (PCA), and linear regression to identify the most important factors that influence body fat percentages. The results of our analysis allowed us to develop formulas for estimating body fat levels in total and for specific body parts. Furthermore, based on the same data, the critical age at which boys and girls start to differ in their morphological profile was also identified in order to help coaches in planning and modeling the training process.

This study is significant as it provides a practical and cost-effective solution for measuring body composition in tennis players. It also provides new insights into the factors that influence body fat levels in this population. By providing coaches and trainers with this information, we hope to improve understanding of body composition in tennis players and to inform future research in this area and hopefully transfer them to other sports as well.

II. RELATED WORK

Body composition is a well-studied topic in the field of sports science, and there is still a growing body of research on this topic. Previous studies have investigated body composition in a variety of sports, including wrestling [2], soccer [3], and basketball [4]. These studies have shown that body composition is an important factor in determining athletic performance, and that higher levels of body fat are associated with lower levels of performance.

A number of methods have been proposed to estimate body fat, including skin fold thickness measurements, circumference measurements, and DXA but each of these methods have their own limitations. One of the biggest problems is the requirement of costly equipment and time, which can be an obstacle for sports organizations and athletes. DXA is considered the gold standard for body composition assessment, but it is also expensive and exposes the athletes to low levels of ionizing radiation. Moreover, the Tanita scale is a widely used device for body composition analysis and measurement of body fat percentage. It uses bioelectrical impedance analysis (BIA) technology to estimate body fat percentage by sending a

low-level electrical current through the body and measuring the resistance [5]. On the other hand, skin fold thickness measurements have seen widespread utilization as a cost-effective and convenient method to estimate body fat. References [6], [7] provide evidence for the validity and reliability of skin fold thickness measurements as a method for estimating body fat, and highlight the simplicity and low cost of this method compared to other body fat assessment techniques.

There is also a growing interest in using statistical methods to analyze body composition data. For example, some studies have used regression analysis to predict body fat percentages based on skin fold thickness and circumference measurements, respectively [8], [9]. PCA was used to identify the most important factors that influence body fat levels [10]. However, most of these studies have focused on specific sports or age groups, and there is a lack of information about body composition in tennis players.

This study aimed to estimate total and regional body fat percentages using skin fold thickness, age, and gender in regression analysis. Our contribution is a comprehensive analysis of body composition in tennis players, identifying key factors that impact body fat. The insights gained have potential to guide future research in the field.

III. Dataset description

A. Structure and Types of Data

The tennis player morphology measurements were collected over a period of more than twenty years at the Faculty of Kinesiology's diagnostic center. The dataset includes measurements from 828 athletes, 370 of them being female and 458 being male. The athletes range in age from 7 to 57 years old and come from different sports clubs and have undergone multiple measurements over the years. It's worth mentioning that 718 of the participants are less than 18 years old and that the majority of the measurements were taken as part of the athletes' schooling requirements. This provides a unique opportunity to analyze the development of body composition and physical performance in young athletes.

The data used in the study consists of 73 features, including age, sex, and various morphology data. These features can be grouped into different measurement types, such as widths, lengths, circumferences, diameters, and most importantly, skin fold measurements. Additionally, the data includes body fat percentage measurements collected using the Tanita scale. Skin fold thickness measurements were taken using calipers at specific anatomical sites on the body, including the tricep, bicep, subscapular, suprailiac, abdominal, thigh, and calf, to estimate body fat percentage. It is important to mention that all used features are on the same scale. This helps to ensure that all the variables have equal importance and contribution in the regression models, and it also helps to mitigate any issues related to numerical stability or computational efficiency during the optimization process. Additionally, having all the features on the same scale allows for easy comparison and interpretation of the results, and it ensures that the variables do not dominate each other, leading to more robust and accurate predictions.

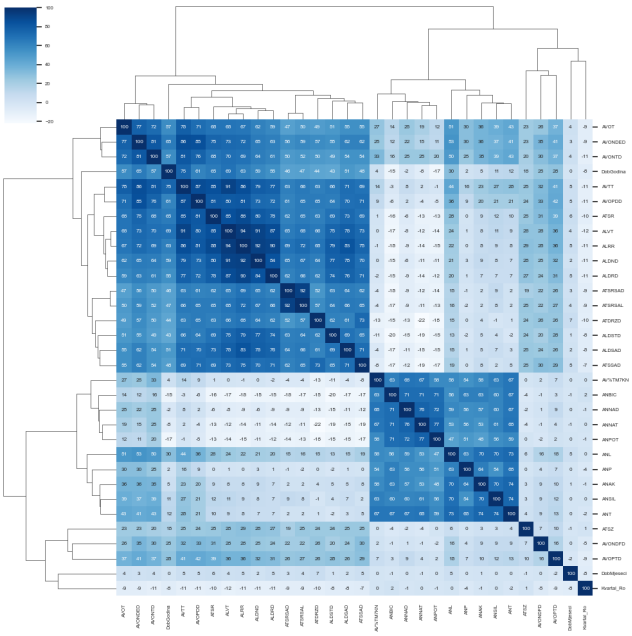


Fig. 1: Clustermap of Data Correlations

B. Data scarcity and high correlations

Having such a high dimensional dataset, we noticed a high degree of correlation among many of the features. A visual representation of this was seen through the cluster map shown in Fig. 1, which showed a clear grouping of features into two consecutive groups. This high degree of correlation is not uncommon in large datasets and can make interpretation of the data more challenging. To address this issue and improve the understanding of our data, we employed dimensionality reduction techniques. These techniques help to reduce the high-dimensional data into a lower-dimensional space, allowing for more clear and meaningful analysis. The use of dimensionality reduction will enable us to better understand the relationships between the features and how they impact the overall performance of the athletes.

Another challenge regarding the dataset are the missing values. Considering that the measurements were performed by different experts over many years, there was no guarantee that everyone would perform all the measurements found in our feature set. To handle this problem, we decided to replace the missing values with the column means if more than half of the data in that column was non-empty. If a column had less than half of the data filled, it was discarded.

IV. Methods

A. PCA

One of the research questions was related to identifying the differences between male and female players. The

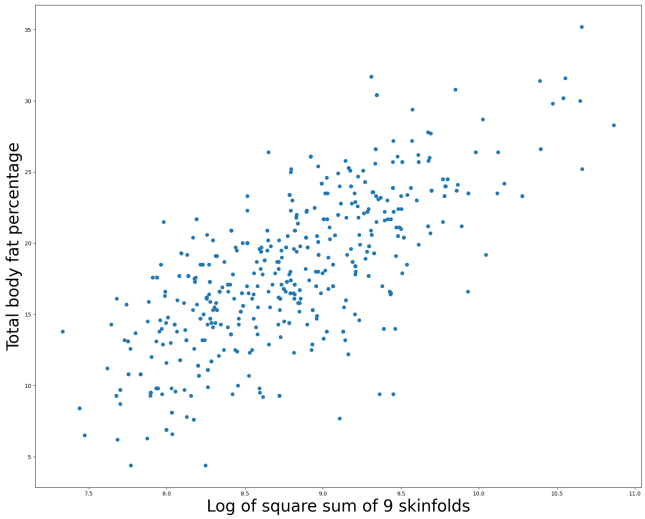
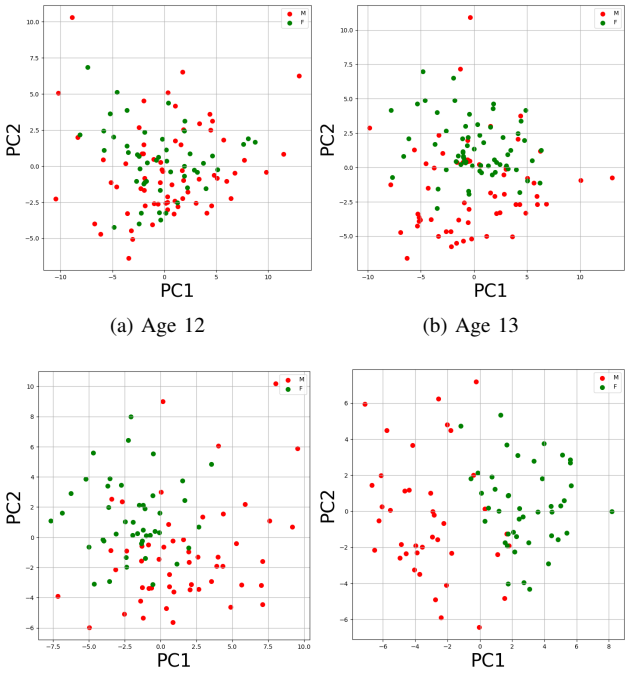


Fig. 2: Relationship between skin folds and body fat percentage



(a) Age 12 (b) Age 13

(c) Age 14 (d) Age 15

Fig. 3: PCA Analysis of Age-Related Differences between Boys and Girls

first step was to conduct a PCA analysis on all athletes using their sexes as targets. Furthermore, considering the heterogeneity of the age groups, we decided to additionally perform a PCA analysis over each age group.

The knee analysis was conducted to determine the number of components necessary to cover the most variance in each age group. The results showed that the first two components alone covered around 60 percent of the variance for each age group and additional components had a minimal contribution in variance explanation.

B. Linear regression

Linear regression was used to determine the ability to calculate the total body fat percentage and of each body part using the Tanita scale measurements. The implementation of the linear regression was performed using the statsmodels library. The primary aim of this analysis was to examine the relationship between the Tanita scale measured body fat percentages and determine whether a prediction model could be developed.

In order to determine the best set of independent features, feature selection methods from the scikit-learn library were used to identify the most important predictors for body fat percentage calculation. Specifically, the SelectKBest, Recursive Feature Elimination (RFE) and SelectFromModel methods were applied. They were used to determine which features make the biggest impact in the linear regression model for body fat percentage prediction. Number of features as a hyperparameter ranged from 2 to 10.

After observing that most of feature selection methods choose skin fold thickness measurements along side with age and sex as the most important features, we decided to take a closer look at the impact of each skin fold and its relationship to the target variable and relationship between each other. Given the still present high correlations between skin fold thickness measurements, and taking into consideration the findings from Jackson and Pollock's research, we decided to aggregate the skinfold measurements into a single feature for our linear regression model [11].

Initially, we observed a logarithmic relationship between the sum of the nine skin folds and body fat percentage. After performing a log transformation on the sum, we noticed a slight quadratic relationship. However, after squaring the log of the sum, we finally obtained a linear relationship. Linear relationship between mentioned final feature and body fat percentage in total can be seen in the Fig. 2. After repeating the research for other body fat percentages, we obtained similar results. Final formula for combined feature for body part is shown in (1).

$$cmb_ft = \log\left(\sum_{i=1}^{9} skinfold_i\right)^2 \tag{1}$$

V. Results

A. PCA

The PCA analysis revealed that female and male players start to differ at the age of 13. This is particularly important as the majority of the players in the dataset were below the age of 18. The PCA plots in Fig. 3 clearly showed the age-related differences between the two groups, highlighting the need for further research to understand these differences in greater detail.

B. Linear regression

The comprehensive search for the optimal hyperparameters was conducted by trying various combinations and evaluation of the linear regression models results. Evaluation was performed using measures from statsmodels, including adjusted R-squared, F-statistic and p-values among others, to assess the goodness of fit and statistical significance of the model.

The F-statistic is a measure of the overall significance of the regression model. It is calculated as the ratio of the explained variance to the residual variance in the model, and it is used to determine if the regression coefficients are significantly different from zero. The F-statistic is used in combination with the p-value to determine whether the model is statistically significant. The adjusted R-squared indicates how well the independent variables in a regression model predict the dependent variable, taking into account the number of independent variables in the model. It adjusts the R-squared value by penalizing the model for the addition of unnecessary variables, allowing for a better comparison of models with different numbers of predictors.

Before presenting the resulting formulas, we introduce binary indicator variable for gender, M, with a value of 1 indicating that the person is male and a value of 0 indicating that the person is female.

The results of the best linear regression models used to predict body fat percentage are summarized in Table I. The table provides the formulas and the adjusted R-squared values for each body part: Whole body, Abdominal, Leg, and Hand.

All four of our final models have high F-statistic indicating that the model as a whole is a good fit, and that the independent variables are important predictors of the dependent variable. It can be used to compare the fit of different regression models, with higher F-statistics indicating a better overall fit. Furthermore, low p-value (less than 0.05) for our predictors suggests that they are significantly related to the response and provide strong evidence against the null hypothesis that the predictor is not related to the response.

The factor of "M" (indicator for male gender) also plays a crucial role in all provided formulas, as it modifies the intercept term. This finding is consistent with the established knowledge that males generally have a lower body fat percentage [12].

The formulas developed in this study can be compared to the widely used Jackson and Pollock equations for estimating body fat percentage. These equations have been validated and widely used in the field, but have limitations as they use a small set of body circumferences and skin fold thickness measurements, and have limitations in their accuracy for certain populations and they only estimate total body density. On the other hand, the formulas developed in this study utilize a larger set of data and a linear regression approach, potentially leading to more accurate results for the specific population and data set used. A statistical test was performed in order to determine the difference between Jackson and Pollock's equations and our model. A random sample of 5 entries was taken from the models and a t-test was conducted. The results showed a significant difference between the models, with a p-value of 0.024, indicating that the models produce significantly different outputs. On the other hand, when comparing the Jackson and Pollock's results with actual measurements, a significant difference was found at a p-value of 0.0001. However, it is important to note that the validity and accuracy of the formulas developed in this study should also be thoroughly tested and validated in independent samples before widespread use.

It is commonly accepted that age is positively correlated with body fat percentage. As people age, their body fat percentage tends to increase, which is often attributed to a decline in muscle mass, hormonal changes, and a decrease in physical activity. This positive correlation between age and body fat percentage is reflected in our linear regression models, where age is consistently associated with higher body fat percentage across all four body parts. The results are in line with previous studies that have found similar trends.

Finally, it is reasonable to expect that body fat percentage would have a positive correlation with the tailored feature representing the sum of all skin folds, as the skin fold thickness is a commonly used measure of subcutaneous fat. The more body fat a person has, the thicker the skin fold measurement will be. This relationship between skin fold thickness and body fat percentage has been well established in the scientific literature, and the use of skin fold thickness as a predictor of body fat percentage is widely accepted in the field of human anatomy and physiology.

VI. Discussion

The results of our study provide valuable new insights into body composition in tennis players. Our findings suggest that body fat levels are influenced by a number of factors, including age, sex, and skin folds, and that these factors differ between boys and girls. By developing model for measuring body fat levels in total and for specific body parts, we have provided a practical and cost-effective solution for measuring body composition in tennis players.

Our study has some limitations that should be noted. The sample size of our study was relatively small, and our findings may not be representative of all tennis players. Additionally, our data were collected using a single measurement technique (the Tanita scale), and it is possible that other measurement techniques would yield different results.

Despite these limitations, our study provides valuable new information about body composition in tennis players, and it has the potential to inform future research in this area. By providing coaches and trainers with practical and cost-effective solutions for measuring body composition,

TABLE I: Body fat percentage calculation formulas

Body part	Formula	Adjusted R-squared
Whole body	$y_1 = cmb_ft * 0.1634 + age * 0.3742 - 2.563 * M$	95.6%
Abdominal	$y_2 = cmb_ft * 0.1180 + age * 0.3352 - 1.694 * M$	94.5%
Leg	$y_3 = cmb_ft * 0.2061 + age * 0.5429 - 4.256 * M$	94.6%
Hand	$y_4 = cmb_ft * 0.2117 + age * 0.5251 - 4.1068 * M$	94.7%

we hope to improve our understanding of this important factor in athletic performance.

Future studies could build upon our findings by collecting data from larger and more diverse groups of tennis players, and by using different measurement techniques to validate our results. Additionally, it would be interesting to investigate the relationships between body composition and other physical and performance measures, such as speed, agility, and endurance, in tennis players. These studies could provide additional insights into the factors that influence athletic performance in this population. Additionally, expanding the data collection to include various sports could lead to the discovery of generalizations across disciplines. This information could be utilized to develop formulas that can be applied to multiple sports.

VII. Conclusion

Our study provides valuable new insights into body composition in tennis players. We have developed practical and cost-effective solutions for estimating body fat levels in total and for specific body parts, using formulas provided in the text. We compared our models to the Jackson and Pollock's, widely considered state of the art, to evaluate our results. Our findings showed significant differences not only between our estimations and theirs, but also between their estimations and obtained measurements. The most important factors that influence body fat percentages were identified. We provided a comprehensive analysis of body composition in tennis players and reveals that boys and girls start differing in their morphology traits at the age of thirteen, which can inform training modeling for coaches.

This research has the potential to shape future studies on body composition in tennis players and to deepen our understanding in this area. Providing affordable and efficient models for determining body composition, this work aims to support the development and improvement of athletic performance in tennis players and inform coaches and trainers.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Faculty of Kinesiology's diagnostic center for providing the data used in this study and to Ultrax technologies for their expertise in the field of kinesiology, which was invaluable to the successful completion of the research.

References

[1] R. B. Mazess, H. S. Barden, J. P. Bisek, and J. Hanson, "Dual-energy x-ray absorptiometry for total-body and regional bone-mineral and soft-tissue composition," *The American journal of clinical nutrition*, vol. 51, no. 6, pp. 1106–1112, 1990.

[2] E. Franchini, C. J. Brito, and G. G. Artioli, "Weight loss in combat sports: physiological, psychological and performance effects," *Journal of the international society of sports nutrition*, vol. 9, no. 1, p. 52, 2012.

[3] P. Aagaard, J. L. Andersen, and P. Dyhre-Poulsen, "Training-related changes in body composition and muscle strength in young soccer players," *Journal of Strength and Conditioning Research*, vol. 24, no. 12, pp. 3268–3275, 2010.

[4] B. Vincent, G. Tolhurst, Z. Crowley-McHattan, and M. Batterham, "Body composition and physical performance of adolescent basketball players," *Journal of Strength and Conditioning Research*, vol. 24, no. 7, pp. 1824–1831, 2010.

[5] D. L. Simmons, M. C. Gonzalez, C. Bouchard, E. Ravussin, S. Lillioja, W. C. Knowler, P. A. Tataranni, and A. D. Salbe, "Validity of bioelectrical impedance analysis and skinfold thickness measurements for estimation of percent body fat," vol. 66, pp. 2085–2089. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/15239930

[6] F. Katch, W. McArdle, and V. Katch, "Body composition and physical performance," *Medicine and Science in Sports*, vol. 13, no. 1, pp. 15–19, 1981.

[7] T. G. Lohman, A. F. Roche, and R. Martorell, "Prediction of body density of males from skinfolds and circumferences," *Human biology*, vol. 60, no. 2, pp. 709–723, 1988.

[8] F. I. Katch, W. D. McArdle, and V. L. Katch, "Body composition and physical performance," *Medicine and Science in Sports and Exercise*, vol. 13, no. 1, pp. 15–19, 1981.

[9] T. G. Lohman, M. H. Slaughter, R. A. Boileau, J. H. Himes, A. P. Rocchini, and S. Weller, "Prediction of body density of males from skinfolds and circumferences," *Human Biology*, vol. 60, no. 2, pp. 709–723, 1988.

[10] C. Bouchard, A. Tremblay, J. P. Després, A. Nadeau, P. J. Lupien, G. Theriault, J. Dussault, S. Moorjani, S. Pinault, and G. Fournier, "Genetic and nongenetic determinants of regional fat distribution," *Metabolism*, vol. 39, no. 3, pp. 324–329, 1990.

[11] A. S. Jackson and M. L. Pollock, "Generalized equations for predicting body density of men," *British Journal of Nutrition*, vol. 40, no. 03, pp. 497–504, 1978.

[12] A. Furnham and K. Cheng, "Gender differences in body fatness and fat distribution," *International Journal of Obesity*, vol. 26, no. 3, pp. 381–390, 2002.

Visualization of Three-Dimensional Ultrasound Data

H. Hrženjak, Ž. Mihajlović
Faculty of Electrical Engineering and Computing, University of Zagreb, Republic of Croatia
helena.hrzenjak@fer.hr, zeljka.mihajlovic@fer.hr

***Abstract* – This paper presents visualization of an ultrasonic probe beam from a series of cross-section images. The goal was to simulate scanning with an ultrasonic probe as it would be done in real-life. Cross-section slices that represent the detection sensitivity of the probe were obtained by simulating a sound pressure field in the metal sample and were provided by INETEC. Different techniques for rendering the three-dimensional model of a sequence of images used for beam visualization in the making of the software are presented. The beam and its inner structure were rendered with the Volumetric Ray Marching algorithm while the polygonal mesh approximation was constructed with the Marching cubes algorithm. The final software was developed using the Unity engine. The results and visualization possibilities were presented and analyzed along with the possible improvements for the future work. The developed simulation software can be used as a helping tool in minimizing the number of redundant scans in search of defects inside an object structure during the quality control phase.**

Keywords - volume rendering, volumetric ray casting, ultrasonic probe, Unity, shaders, three-dimensional dataset

I. INTRODUCTION

Volumetric rendering is a collection of important techniques in the field of computer graphics. They are a powerful tool for data visualization as they offer a way to display data visually, offering easier and more intuitive ways to understand raw data.

Volume rendering is most notably used in medical imaging, for helping medical staff to see inside a body more accurately, using medical images obtained from procedures like computerized tomography (CT) or magnetic resonance imaging (MRI) three-dimensional models can be constructed. Besides medicine, volume rendering is finding more uses in other industries such as those that produce 3D data sets for analysis, e.g., physics for fluids, flood preparation simulations and more.

In this work, volume rendering and other techniques are used to achieve ultrasonic beam simulation. Ultrasonic beam scanners are used for non-destructive material testing of tools and metal parts for nuclear power plant inspections. The final 3D model of the beam was recreated from a series of 2D cross-section images using the ray marching technique and the marching cubes algorithm.

The main goal was to create a simulation that will visualize defect detection on a certain trajectory and make probe positioning in real testing quicker and more efficient. The models and data used in the implementation were provided by INETEC and collected using their instruments. The goal was to simulate movement of a UV probe during detection. Observing the best trajectory in the simulation could reduce the time to position the probe and scan for defects in practice.

II. ULTRASONIC PROBES

A. Representation of the pressure field

Ultrasonic probes are devices that are used for investigating the internal structure of a target object. The detection sensitivity of an ultrasonic probe can be closely estimated by the strength of the sound pressure field it generated in the specimen. The sensitivity field visualized as a heat map can be seen on the Fig 1.

The rainbow colormap was used for coloring the sound pressure field, the dark blue color represents lower pressure field density values and thus lower detection sensitivity while the dark red represents high values meaning high sensitivity. The series of this kind of images, generated with different settings - multiple UT probe angles and different focal depths, will be used as an input data for the visualization.

The formula for calculating pressure field is complex and therefore using it to calculate the pressure field multiple times for each setting would take some time. That is why visualization from images is a good option, there is no need for complex formula calculations to get good results. However, the downside of this approach is that it needs pre-generated images.

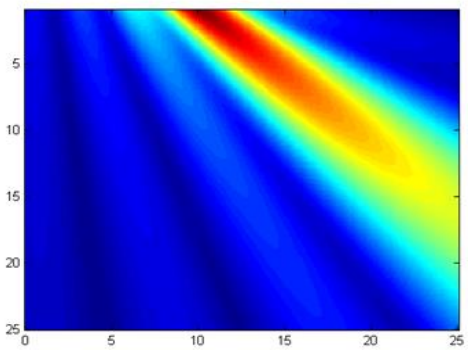


Fig. 1. Representation of a sound pressure field of the probe at a 45-degree angle

B. Ultrasonic sensors

Ultrasonic sensors are instruments that use ultrasonic waves for inspection of subsurface structures and detection of anomalies. Ultrasonic sensors work by emitting sound waves at frequencies in the ultrasonic frequency range - range greater than twenty kilohertz, higher than the audible range of human hearing. Ultrasonic sensors emit ultrasonic waves toward a target and if there is an obstacle or an object in the way, the wave will bounce back to the sensor's receiver. The distance of an obstacle is then calculated using the time the reflected wave took to return to the receiving sensor after being emitted.

Ultrasonic inspection is a family of methods most widely used for nondestructive testing, it offers a great accuracy, high sensitivity but, most importantly, it does not cause damage to an object being tested, as such it finds its application in many fields. The most familiar usage being in the field of prenatal medicine for medical screening of a fetus during pregnancy. Besides that, ultrasonic inspection is also widely used for quality control and material inspection across all major industries, including electrical and electronic component manufacturing, production of metallic and composite materials, and fabrication of structures such as airframes, engines, machinery and many others [1].

III. VISUALIZATION TECHNIQUES

Displaying data in a visual form provides clear and easily understandable representation of raw data. The main problem in visualizing volume data is displaying three-dimensional data as a two-dimensional image while not losing too much information and not losing valuable information [2]. Volume rendering and surface rendering represent the most important techniques for three-dimensional dataset visualization.

A. Surface rendering

Surface rendering method involves constructing polygonal surfaces in a dataset and rendering those surfaces. The assumption is that the original volume can be faithfully represented as a collection of polygonal surfaces. Surface rendering algorithms have a necessary preprocessing step, also called image segmentation, which involves determining a surface by extracting features from the volume data. More precisely - every pixel in an image is assigned a label such that pixels with the same label share certain visual characteristics such as color, intensity, or texture. These pixels are a part of one isosurface. The geometric primitives are then fit to the data to form a 3D surface based on the values extracted from the data in preprocessing step and rendered for display using conventional geometric rendering techniques [3].

The process of determining surface is not perfect, as it must be determined for every voxel how the surface passes through every cube of voxels. This issue is especially apparent for datasets surrounding objects that are very small, blend into their surroundings or describe poorly defined features [4]. A frequent problem that can occur is calculated surface containing voxels that don't really belong to the original object we are trying to visualize or skipping the voxels that do belong. Because of the mentioned issues, when surfaces are transparent or semi-transparent, geometric rendering techniques may not be the best choice, that is where the volume rendering comes in.

B. Volume rendering

Volume rendering is a powerful technique for the representation, manipulation and rendering of volume data. Traditional graphics techniques represent 3D objects as geometric surfaces and edges approximated by polygons and lines, but volumetric datasets can contain data and information not only on the surface, but also inside them, which is where surface rendering techniques fall short. Volume rendering techniques make it possible to explore the inner structures of volumetric data and allow visual representation of transparent and complex datasets.

The advantage of volume rendering algorithms is that there is no need to determine surfaces in advance which removes the need for the segmentation step and a polygon representation, instead they render every voxel in the volume raster directly, without explicit conversion to geometric primitives. Volume visualization methods can achieve soft surfaces by integrating the contribution from the voxels in entire volume towards final rendered image [5]. Therefore, this method is applicable even in cases where dataset can't be faithfully described by polygons or has a lot of small details that can't be lost. It is also applicable in cases when geometric surfaces for dataset are unavailable or too cost-ineffective to generate.

IV. VOLUMETRIC RAY CASTING

Volumetric ray casting, also commonly referred to as Volume ray tracing, is the most common technique for achieving image-based volume rendering.

The 'ray casting' part of the name comes from the similarity with the traditional ray casting in that they both define rays that represent beam from a camera's viewpoint and only consider these primary rays without spawning the secondary, reflected and refracted, ones. The difference is that the ray casting interacts only with surface data and stops as soon as it encounters an intersection between a ray and a surface in a scene while volume ray casting does not stop the computation at the surface level but passes through the object, sampling it along the way [6].

This technique shows the best results in materials where light goes through the object. Examples would include smoke, marble, and skin. When light hits one of these surfaces, a fraction of the light is reflected (ray casting only deals with this case) and a fraction is scattered into the object. Ray marching is a technique to compute the scattered light. Image projection is performed by simulating the absorption of light along the ray path to the eye and calculating light intensity when arriving at the camera. The advantage of this technique lies in the fact that it does not directly calculate intersections with the objects in the scene, so no mesh data of the objects needs to be provided.

Volumetric ray marching employs simplified Emission-Absorption Optical Model, taking into consideration, as per its name, emission and absorption while ignoring light scattering as it is too computationally complex while not affecting quality too much. This means it only deals with

primary rays, unlike some more common graphic rendering methods, no secondary rays like reflection or refraction or shadows are considered. The Emission-Absorption Optical Model places two assumptions: the volume is assumed to consist of particles that emit light. The volume is assumed to consist of black particles that absorb light [16]. Meaning the portion of light that passes through a particle will be absorbed. Volume rendering integral is an equation that computes light intensity at a point s along the ray, taking into consideration light attenuation.

$$I(s) = I(s_0)e^{-\tau(s_0,s)} + \int_{s_0}^{s} q(s_1)e^{-\tau(s_1,s)}\,ds_1 \qquad (1)$$

Where parts of the equations are:

- $I(s_0)$ is initial light intensity at point s_0, s_0 being a starting point
- $e^{-\tau(s_0,s)}$ is intensity reduction from s_0 to s
- $I(s_1)$ is light intensity (emission) at point s_1, s_1 being a point on the ray further down than point s_0
- $e^{-\tau(s_1,s)}$is intensity reduction from s_1 to s
- τ is an optical depth measure, measuring degree of absorption of light
- $\int_{s_0}^{s} q(s_1)e^{-\tau(s_1,s)}\,ds_1$ gathers contributions of all points on the ray between s_0 and s_1

There is no closed form solution to this integral, but it can be numerically approximated with a discrete sum of volume samples along the ray [7].

Each point in the volume is considered to emit and absorb light, and thus contributes to the final color and opacity. There are multiple compositing schemes for deciding the output color. The most common one is Volume compositing or accumulation. For the emission-absorption model, the accumulated color and opacity are computed according to a following set of equations:

$$C_{out} = C_{in} + (1 - \alpha_{in})\alpha C \qquad (2)$$

$$\alpha_{out} = \alpha_{in} + (1 - \alpha_{in})\alpha \qquad (3)$$

These equations are known as Front-to-Back Compositing Equations since we go from the front of the volume to the back, collecting color on the way. C_{out} is the resulting accumulated color over n voxels seen from the front of the volume [8]. Variable α_{out} is the opacity of the result. Marching from beginning of the ray towards the end accumulating color and opacity using these equations is repeated until bounding box is exited, or some other stopping criteria is defined e.g., predefined maximal opacity or number of steps is exceeded. Another scheme that is going to be mentioned in continuation is Maximum Intensity Projection. It simply takes color of the highest density point along the ray.

V. MARCHING CUBES ALGORITHM

The Marching cubes algorithm is a computer graphics algorithm for creating polygonal representation of surfaces with constant density – isosurfaces. It was first presented in the paper *Marching cubes: A high resolution 3D surface construction algorithm* [9]. The algorithm processes the 3D data in a scan-line order and calculates triangle vertices using linear interpolation. The density of the desired surface is specified as an input from the user.

The Marching cubes algorithm belongs to the surface rendering algorithms. Its advantage is that it works on lookup tables so it's relatively high speed. The algorithm works on a divide-and-conquer principle, it divides the problem of creating surface to the small cells, it focuses on each cell and applies to it one iteration of the algorithm. The first step of the algorithm would be splitting the space into a uniform grid of cells, in three dimensions that would be cubes. Secondly, the isovalue or a density of a desired isosurface is defined, in this case the beam density. Next step is assigning a Boolean value to each voxel according to its value in a relation to the designated isovalue – value one if the voxel's value is greater than the isovalue, meaning that the voxel is inside the isosurface or lies on the surface and value zero if it is lesser than the isovalue and the voxel is outside of the isosurface.

For each cube eight pixels that make up cube vertices are taken into polygon evaluation. Four vertices are pixels of one slice and the other four belong to adjacent slice. The surface intersects cubes at edges where one vertex is outside the surface and another one is inside, calculating the exact intersection location with interpolation in later steps of the algorithm. There are eight vertices in each cube and two possible states for each one, equating to $2^8 = 256$ ways a surface can intersect the cube [10]. These 256 possible cases of surface-edge intersections are enumerated and can be searched found in the lookup table. The states given to each voxel in the previous step are now composed into an 8-bit index, where each bit corresponds to a Boolean value assigned to the vertex. That index is used to fetch data from the lookup table. Algorithm iterates over all cubes, adding triangles to a mesh, and the final mesh is the union of all added triangles. The smaller the cubes, the smaller the mesh triangles will be, making approximation more closely match the target function.

VI. BEAM VISUALIZATION

Visualizing the beam is the most important part of the project. The visualization was realized in Unity with the ray marching technique using 3D textures and shaders. The shaders were written in High-Level Shader Language (HLSL). Results are three-dimensional even though inputs are two-dimensional images.

High-level shading language is a programming language used for writing programmable shaders in DirectX. It is very similar to the Cg shader language and will be used to write the shaders in the next steps. The Shader that will draw the ultrasonic beam implements volume rendering using ray marching algorithm and as a result gives a rendered volume representing the ultrasonic beam. Programmable parts of the graphics pipeline are vertex shader and fragment shader [15]. The vertex shader processes and performs transformations on individual vertices. Fragment shader is where GPU calculates the final RGB color for every pixel. The implementation is done in Unlit shader type as the beam doesn't need to be affected by lightning models.

Most important task for the vertex shader is transforming vertices' coordinates to the clip space so that they can be used by rasterizer and be projected on the screen. The vertex shader receives object space coordinates as its inputs, the conversion from object space to the clip space is done by putting vertices through a sequence of transformation matrices. To make things easier, all the transformation matrices can be combined into one matrix by the name Model View Projection matrix (MVP). In Unity there is a built-in helper function UnityObjectToClipPos(float3 pos) that does exactly that, it takes a position point and directly transforms it from an object space to the camera's clip space in homogeneous coordinates [11].

Fragment stage is where the ray marching will be done. There is a ray created through each pixel, pixel's local position is the ray origin, and the ray direction is the vector from camera to the pixel. Direction vector is calculated by using helper function ObjSpaceViewDir() and inverting it, since it returns opposite direction, from pixel to camera. Using this ray, the unit cube in which volume is drawn is intersected. The points of intersection are helpful for calculating marching step size. Distance between the start and the end point of intersection is divided by the defined maximum step size. The max step size limit is there because of performance concerns, if the number of steps is too great visualization would not run smoothly in real-time.

Finally, the ray marching is executed in the loop that runs for a number of steps. In each step the density is read from the B&W 3D texture and the color from the colored 3D texture (or from the transfer function) using adjusted local vertex position.

These 3D textures are a special Unity Asset type, they are bitmap images that contain information in three dimensions as opposed to the standard two. Texture coordinates are placed on a unit cube and texture data is accessed by three-dimensional texture coordinates. This kind of data type is fit for storing and manipulating volumetric data. The 3D texture containing volumetric data was built from a sorted sequence of 2D images through a script and is of the input properties for the shader. What is done with color calculation in the next step depends on the technique chosen.

In the compositing technique color is sent to the method where accumulation formulas are used for collecting and interpolating colors through the steps. If the pixel density does not satisfy threshold the pixel color is not taken into calculation. Compositing can be terminated early if the currently accumulated opacity is so high it makes contributions of future samples insignificant.

In Maximum intensity projection (MIP) technique the color of a voxel with maximum intensity along the ray is returned after checking all the pixels along the same ray.

In both cases the final pixel color is returned. The color of the beam can be chosen to either be read from a 3D texture or from a defined transfer function. The case of using textures was implemented in a way that a color for the voxel is read from the texture in the ray marching algorithm, while when using transfer function, the density is first calculated using ray marching and then at the end, based on that value the color is read from the transfer function.

A. Transfer function

Transfer functions define mapping from voxel density to RGB color and opacity based on user defined values. Using transfer functions provides a better overview of the data. Transfer functions can be used to highlight important parts and omit unnecessary ones thus better emphasizing nuances inside the volume. Transfer function takes the scalar value, in this case density and relates it to a corresponding color and transparency, for example if it is desirable to highlight higher density parts higher opacity should be assigned to those densities [15]. Deciding how a transfer function should look is not an easy task, there is no fixed guide to perfect function because of variations from dataset to dataset.

In Unity transfer function can be created by using a 2D texture with height of one, this is enough since needed information, color and opacity, can be stored in one pixel. Texture color is read and set based on input gradient stored in form of a Unity's Gradient class.

B. Collision

For detecting collisions Unity's default colliders were not close enough of an approximation for the beam, so the mesh had to be reconstructed from the images. The reconstruction was accomplished with the Marching cubes algorithm. The script inputs are a stack of B&W images and a minimum density value defined by the user, that value is going to be considered as the threshold in the script.

In the preprocessing step when assigning labels for each pixel the slices were traversed line by line, in each one the first pixel with equal or higher density than defined was assigned a label one and all subsequent ones until the pixel below the threshold. After labeling the pixels were used in mesh creation using the Marching cubes algorithm with interpolation. The resulting triangles were decided using the case table and together they formed the resulting beam mesh. The mesh data is then saved in an instance of Unity's Mesh class.

The script is set up by attaching it to the desired GameObject that must have a Mesh Renderer and a Mesh Filter component. After mesh creation the Mesh collider can be added to the beam, Unity automatically adapts the Mesh collider to the selected Mesh - in this case the beam.

Beam Collision script listens for the collision events and reacts by calling function that changes material color on a defect when it is under collision to a color that can be seen easily. After the beam collider is no longer in contact with the defect the material color is reverted to the original one. OnCollisionEnter() and OnCollisionExit() are physics engine specific functions that are triggered on update frames when the collision is detected and when the contact between colliders has stopped.

C. User interface

The User Interface is divided into three sections, which can be seen in Fig.2. The first section, on the left, consists of settings regarding beam visualization. Most important ones are Focal depth and Refraction angle, these settings

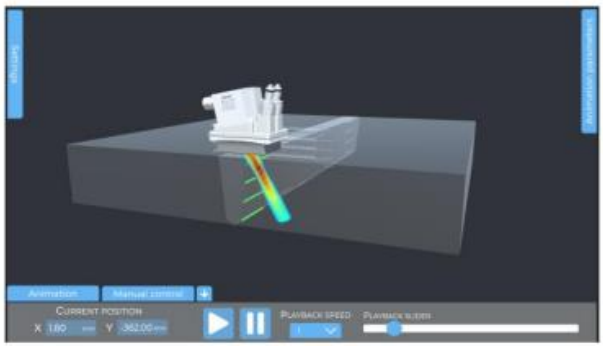


Fig. 2. Highlighted defects result.

can be chosen from their drop-down lists, each combination of the two maps to the one set of images and by extent textures that will be used for visualization. For output color there are Max Intensity Projection and Compositing options. For Compositing the choice of current Z-slice is enabled, by using the slider the user can select any cross-section on z-axis within the beam for a more detailed view. The panel on the right is tasked with handling the input of the auto-movement path parameters. The user can fine tune the course of simulation by inputting the values for the desired simulation output in the suitable fields. Once all the parameters are confirmed automatic simulation will be ready to be played.

The middle section, at the bottom of the screen, controls the movement of the probe model. Animation tab contains buttons for playing and pausing and the slider for rewinding the animation. Playback speed can be slowed down or accelerated. In the manual control tab user can either input exact probe position or use two sliders for selecting horizontal line position and sliding UT probe by a selected increment along it. The components used in the process of building the interface were extended from the Quantum UI package [12] which was downloaded from the Unity Asset store web page. Another custom package that was used during UI development for smoother UI transitions is LeanTween [13].

VII. RESULTS

In this chapter, the results of the implementation are presented and discussed. The comparison between beam visualization using two different technique types and two color acquisition methods will be shown. Since results are visualization, the described results are mostly observations. Numerical approximation of the result accuracy can be explored in the future work.

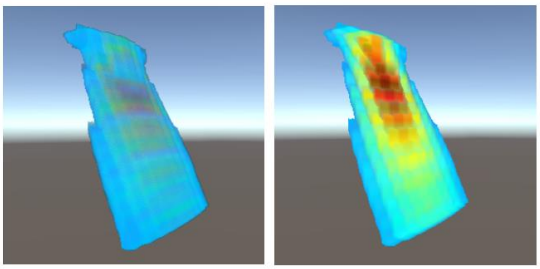


Fig. 3. Comparison between Compositing(left) and MIP (right) using color from texture and a cutoff density = 0.3 at a 30-degree refraction angle.

As previously explained, the beam can be rendered with two techniques: Maximum intensity projection and Compositing. On Fig.3. when using Max projection, the highest density can be clearly seen from all of the angles (an effect more visible in movement than on images) while with Compositing we can see color accumulation effect, most of the beam is in shades of blue as it is the color most present in the photos, so it contributes the most, but there are also some traces of red in the middle.

Max intensity projection results can be improved upon by adding transparency. With this feature enabled voxels with lesser density values will be more transparent than the ones with the higher ones. This makes important parts of the beam stand out more. The result is outer blue rays being less visible than the middle ones in yellow-red range. Transparency effect is achieved by setting alpha color value to depend on voxel density, in this case it is just multiplied by a parameter in the shader.

To get better insight of inner structure there are two options: removing parts outside of desired density segments or slicing off the parts of the volume that are not needed. Extracting segment from the beam is done by selecting minimum and maximum density value, all the other parts that are not included in this range are going to be transparent as can be seen on Fig.4. Cutting the volume along one of the axes makes it possible to select a slice or a part of the volume that is of interest. For example, on the z-axis if maximum and minimum are set, everything in front of the minimum and behind the maximum z-value will be removed.

Coloring the volume after calculating densities can be done in different ways. The most common coloring method for volume rendering is using transfer function, since it is easily adjustable and can define both color and transparency. Another advantage of the transfer function is that we do not need additional color texture, saving time and space. The advantage of reading color from texture is that everything is colored exactly like in the original dataset, there is no approximation and thus it is more similar to the source data. Also, it saves the time that would otherwise be needed for defining perfect transfer function for a particular dataset. Even though the B&W texture used for density calculation was lower resolution the transfer function gave very smooth results, which is not the case with a color from texture one where it caused a bit of a blocky look to colors. If the input images were of better

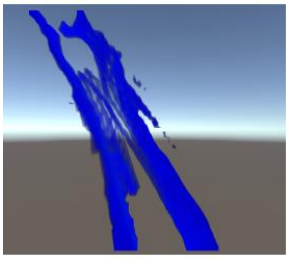


Fig. 4. Result beam with an alpha value between 0.07 and 0.115, only volume with density inside that range is visible.

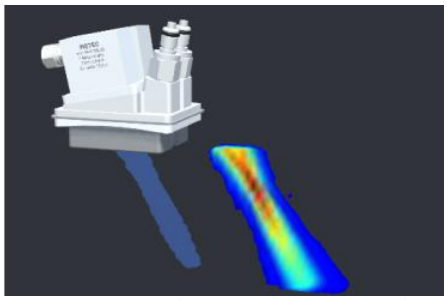


Fig. 5. Mesh reconstruction for the density threshold of 0.3 and the original visualized beam

resolution the result would quite probably be significantly better and more similar to the transfer function result.

The focus depth change from 30 to 60 mm causes the rendered result to consist of a more intense and larger surface of red area in comparison with the 30 mm one, this difference shows how the ray with depth change shifts focus to a deeper depth, around 60 mm deep from probe position. The refraction angle, as the name says, changes the angle of the probe and in extension angle in which rays enter the block. The probe can be imagined as if it was positioned in the top left corner.

The marching cubes algorithm result is a mesh that approximates the beam. On the Fig.5. the comparison between the generated mesh and the beam using the density threshold of 0.3 can be seen. The mesh is lightly shaded on the image for the comparison, but it can be disabled from rendering in simulation if desired.

VIII. Future work

Future work could improve upon offering more scanning path options. Possibility to simulate scanning of complex surfaces like the curved surfaces or the surfaces with obstacles would make this program applicable for use in a wider number of cases since real objects are not regularly shaped blocks and commonly have welds and other obstacles on trajectory. Implementing this feature would be more complex than it is for flat surfaces. Additional complications to path calculation would present obstacles as they would prevent the probe from moving freely. Another small issue is that the results with current images come out a little pixelated, the cause is that the current images were generated in dimensions of 43 x 31 pixels to save time for a course of development. If the simulation images can be regenerated in a bigger resolution, with their quality improved so would the quality results, the final visualization would be clearer and more accurate.

The results could be evaluated numerically, to calculate how close the simulation represents the real-life beam and detection.

IX. Conclusion

The possibilities and advantages of using volume rendering in displaying datasets have been explored by applying described techniques and algorithms to a practical problem of visualizing ultrasonic beam for quality inspection. The achieved simulation can be used as a helping guide in minimizing the number of redundant scans in search of defects inside object structure during quality control phase. The implementation was done using Unity and writing custom shaders, results achieved showcased one possibility of use of volume rendering in the quality inspection field. The biggest limitation of the final application is no support for complex trajectories and that could be a potential course for future work.

Acknowledgments

I would like to extend my gratitude to INETEC for supplying the materials, 3D models, image sequences and providing mentoring help in the creation of this paper.

References

[1] L. J. Bond. “Fundamentals of Ultrasonic Inspection[1]” In: Nondestructive Evaluation of Materials. ASM International, Aug. 2018.,URL:https://www.asminternational.org/documents/10192/22533690/05511G_SampleArticle.pdf/a8c3979c-3b35-f304-68f2-151271f2e3b8.

[2] Robert A. Drebin, Loren Carpenter, and Pat Hanrahan. “Volume Rendering”. In: SIGGRAPH Comput. Graph. 22.4 (June 1988), pp. 65–74. ISSN: 0097-8930. DOI: 10.1145/378456.378484. URL: https://doi.org/10.1145/ 378456.378484.

[3] Andreas G. Schreyer and Simon K. Warfield. “Surface Rendering”. In: 3D Image Processing: Techniques and Clinical Applicatio Ed. by Davide Caramella and Carlo Bartolozzi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 31–34. URL: https://doi.org/10.1007/978-3-642-59438-0_4.

[4] Heavy.AI. Volume Rendering. URL: https://www.heavy.ai/technical-glossary/volume-rendering (visited on 06/2022)

[5] Multiple authors. Volume Rendering. URL: https://www.sciencedirect. com/topics/computer-science/volume-rendering (visited on 06/2022)

[6] ETH Zurich, “Direct Volume Rendering “, https://cgl.ethz.ch/teaching/former/scivis_07/Notes/stuff/StuttgartCourse/VIS-Modules-06-Direct_Volume_Rendering.pdf

[7] Tino Weinkauf. Introduction to Visualization and Computer Graphics. 2015. URL: https://www.kth.se/social/files/565e35dff27654457fb84363/08_VolumeRendering.pdf

[8] M. Ikits, J. Kinss, A.Lefohn, C. Hansen, “GPU Gems, Chapter 39.Volume Rendering Techniques”, https://developer.nvidia.com/gpugems/gpugems/part-vi-beyond-triangles/chapter-39-volume-rendering-techniques

[9] William Lorensen and Harvey Cline. “Marching Cubes: A High Resolution 3D Surface Construction Algorithm”. In: ACM SIGGRAPH Computer Graphics 21 (Aug. 1987), pp. 163–. DOI: 10.1145/37401.37422.

[10] Paul Bourke. Polygonising a scalar field. URL: http://paulbourke.net/ geometry/polygonise/ (visited on 06/2022).

[11] Unity Technologies. Unity User Manual. URL: https://docs.unity3d.com/Manual/ (visited on 06/2022)

[12] Quantum Tek. Unity Asset Store package - Quantum UI. URL: https : / / assetstore.unity.com/packages/tools/gui/quantum-ui-162077 (visited on 06/2022)

[13] Dented Pixel. Lean Tween. URL: https://assetstore.unity.com/packages/tools/animation/leantween-3595.

[14] L. Radisson, “Raytracing Studies“ https://observablehq.com/@makio135/raytracing-studies

[15] Randima Fernando and Mark J. Kilgard. “Chapter 1. Introduction, Chapter 4. Transformations”. In: The Cg Tutorial: The Definitive Guide to Programmable Real-Time Graphics. Addison-Wesley Professional, 2003. URL: https : / / developer.download.nvidia.com/CgTutorial/cg_tutorial_frontmatter.html.

[16] Tino Weinkauf. Introduction to Visualization and Computer Graphics. 2015. URL: https://www.kth.se/social/files/565e35dff27654457fb84363/08_VolumeRendering.pdf

Istraživanje izazova i problema suradnje dizajnera i razvojnih programera tijekom procesa razvoja digitalnog proizvoda

A. Ilić*, M. Čarapina* i V. Uglješić*
* Tehničko veleučilište u Zagrebu / Informatičko-računarski odjel, Zagreb, Hrvatska
ana.ilic@tvz.hr, vesna.ugljesic@tvz.hr, mia.carapina@tvz.hr

Radom se istražuje problematika procesa razvoja aplikacija od skica ekrana (engl. wireframe) do proizvoda, s fokusom na prijenos informacija između dizajnera i programera. U praksi čest je slučaj da se razvijena programska podrška u prvim fazama razlikuje od skice koju je osmislio dizajner ili dizajnerski tim. Istraživanje detektira uzroke zbog kojih dolazi do spomenutih odstupanja osvrćući se na iskustva dizajnera i programera.

Ključne riječi – UX, UI, digitalni proizvod, razvoj, suradnja

I. UVOD

Dizajneri i programeri često nailaze na probleme prilikom zajedničke suradnje, ali za stvaranje uspješnog proizvoda i optimalnog korisničkog iskustva vrlo je bitna dobra i organizirana suradnja između ove dvije strane [1]. Nerijetko se dešava da je upravo taj odnos zanemaren i to rezultira problemima u razvoju digitalnog proizvoda.

Dizajn i programski razvoj digitalnog proizvoda imaju zajednički cilj, a to je uspješan krajnji proizvod i stoga je potrebno tijek rada prilagoditi željama i potrebama obje strane kako bi taj cilj bio ostvaren [2]. Ponekad se čini kao da dizajneri i programeri pričaju dva različita jezika te često imaju različite stavove, dolaze iz različitih okolina i funkcioniraju na različite načine. Mogući razlog koji dovodi do izazova u stvaranju digitalnog proizvoda leži u pretpostavci da su dizajneri uglavnom fokusirani na korisničko iskustvo i dizajn korisničkog sučelja, a programeri na funkcionalnosti, tehničke mogućnosti i ograničenja.

Ovim radom istražuju se izazovi i problemi suradnje dizajnera i programera tijekom procesa razvoja digitalnog proizvoda, a temeljem prikupljenih podataka i rezultata predlažu se strategije koje bi pridonijele procesu stvaranja kvalitetnijih digitalnih proizvoda.

II. METODOLOGIJA ISTRAŽIVANJA

U siječnju 2023. godine provedeno je anketno istraživanje putem platforme Google Forms s ciljem prikupljanja iskustava, mišljenja i stavova dizajnera i razvojnih programera tijekom procesa razvoja programske podrške, a s fokusom na fazu predaje (engl. handoff) dizajnerskog rješenja razvojnim programerima te izrade inicijalne verzije programskog rješenja. Anketni upitnik korišten u istraživanju sastojao se od pet pitanja pri čemu su sva pitanja, osim prvog, bila otvorenog tipa kako bi ispitanici mogli detaljno opisati svoja iskustva i stavove. Odgovorom na prvo pitanje sudionici su mogli odrediti opredjeljuju li se kao dizajneri ili razvojni programeri nakon čega je uslijedilo grananje na dva potpitanja ovisno o opredjeljenju ispitanika.

Pitanja koja su postavljena dizajnerima u sklopu ankete bila su:

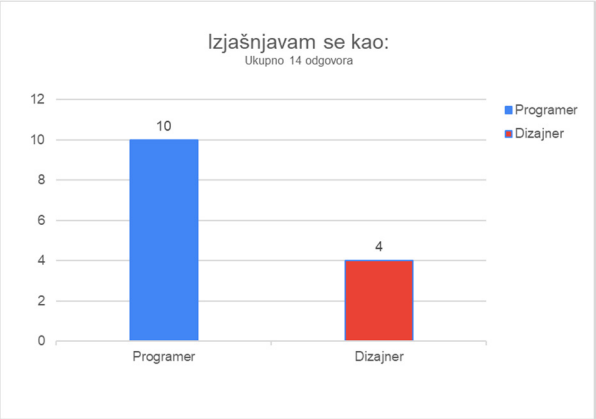
1. Je li Vam se dešavalo da se u handoff fazi prva verzija programskog rješenja razlikuje od predanog dizajnerskog rješenja (Vašeg ili tima u kojem ste radili)? Ako je odgovor DA, što smatrate da su razlozi koji dovode do toga?
2. Što bi htjeli poručiti razvojnim programerima kako bi taj proces bio uspješniji? (Pitanje za dizajnere)

Pitanja koja su postavljena razvojnim programerima u sklopu ankete bila su:

3. Je li Vam se dešavalo da ste na prvu verziju programskog rješenja dobivali povratnu informaciju od dizajnera da izvedeno programsko rješenje ne odgovara predanom dizajnu? Ako je odgovor DA, što smatrate da su razlozi koji dovode do toga?
4. Što bi htjeli poručiti dizajnerima kako bi taj proces bio uspješniji? (Pitanje za programere)

III. REZULTATI ISTRAŽIVANJA I ANALIZA

U anketnom upitniku sudjelovalo je ukupno 14 ispitanika (N=14) od čega 10 razvojnih programera i 4 dizajnera koji su zaposlenici hrvatskih tvrtki te se profesionalno bave razvojem mobilnih i web aplikacija. Važno je napomenuti da je anketa proslijeđena u četiri tvrtke (Infinum, Netgen, Artum Design i Bien Studio). Budući da je broj ispitanika relativno malen, na temelju provedenog istraživanja nije moguće dati globalne i pouzdane zaključke no dobiveni podaci mogu poslužiti kao koristan uvid u problematiku.



Slika 1. Zanimanje ispitanika.

Analiza odgovora na pitanje "Je li Vam se dešavalo da se u handoff fazi prva verzija programskog rješenja razlikuje od predanog dizajnerskog rješenja (Vašeg ili tima u kojem ste radili)?" pokazuje da su svi anketirani dizajneri prilikom izrade digitalnog proizvoda iskusili da se prva verzija programskog rješenja razlikuje od predanog dizajnerskog rješenja. Kao glavni razlog navode nedostatak komunikacije i suradnje od samog početka procesa te ističu važnost komunikacije s razvojnim programerima. Neki od navedenih razloga zašto dolazi do odstupanja od prvobitnog dizajnerskog rješenja, po mišljenju dizajnera, su brzina u razvoju, nedostatak poznavanja procesa dizajna, nedostatak jasnih smjernica te tehnička i vremenska ograničenja i izmjene u dizajnu koje nisu raspravljene.



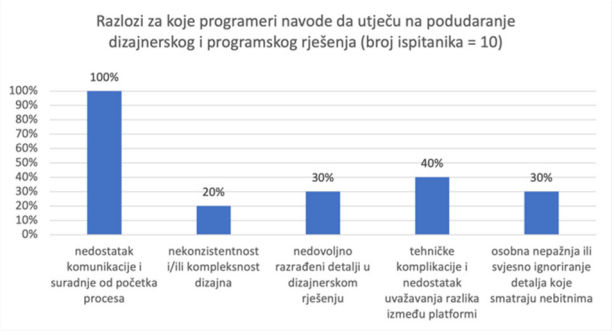
Slika 2. Razlozi za koje dizajneri navode da utječu na podudaranje dizajnerskog i programskog rješenja.

Za što uspješniju suradnju, dizajneri predlažu otvorenu i redovitu komunikaciju od samog početka. Navode da je potrebno uključiti glavnog dizajnera i voditelja projekta u redovne sastanke s razvojnim programerima kako bi se kroz cijeli proces razvoja pratilo kako dizajnersko rješenje postaje živi proizvod te kako bi se brže rješavale zapreke na putu. Neki od kvalitativnih odgovora su:

"Handover ne može biti jedan sastanak – Osim inicijalnog sastanka u kojem developeri i dizajneri detaljno prolaze kroz dizajn, potrebno je uključiti glavnog dizajnera u redovne sastanke developmenta i PM-a kako bi se kroz cijeli proces developmenta pratilo kako dizajnersko rješenje postaje živi proizvod te kako bi se brže rješavale zapreke po putu."

"Sudjelovati u procesu (kickoff meeting, wireframes, UI, imati nekakve checkpoint sastanke) od samog početka. Bolje proučiti tipografiju i tipografsku skalu koja se koristi."

Iz perspektive razvojnih programera, 90% ispitanika ističe kako se dešavalo da prime povratnu informaciju od dizajnera da programsko rješenje ne odgovara predanom dizajnu. Kao glavne razloge navode nekonzistentnost i kompleksnost u dizajnu, nedovoljno razrađene detalje u dizajnerskom rješenju, tehničke komplikacije i nedostatak uvažavanja razlika između platformi, pa čak i osobnu nepažnju ili namjerno ignoriranje nekih detalja koje smatraju nebitnima. Kao i dizajneri, programeri nedostatak komunikacije navode kao jedan od glavnih razloga zašto dolazi do problema u pretvorbi dizajnerskog u programsko rješenje. Ističu da se to najčešće dešavalo kada ne bi sudjelovali u procesu od inicijalne faze osmišljavanja dizajnerskog rješenja. Programeri smatraju da dizajneri često ne poznaju standard tehnologije za koju rade dizajn te da zbog toga dolazi do problema u implementaciji grafičkih elemenata. Navode i da je bitno detaljno razjasniti sve aspekte dizajnerskog rješenja prije kretanja u razvoj te da su komunikacija i suradnja od samog početka projekta ključne za razvoj kvalitetnog proizvoda.



Slika 3. Razlozi za koje programeri navode da utječu na podudaranje dizajnerskog i programskog rješenja.

Neki od prijedloga razvojnih programera za poboljšanje suradnje i razvojnog procesa su:

"Komunikacija, komunikacija i još komunikacije. U razvoj se ne bi smjelo ići dok se čitko i jasno ne razumije problem i svima bude jasno kako izvesti rješenje. Ako će biti eventualnih dorada ili se smatra da se u početnom koraku nešto krivo komuniciralo, npr. novo dizajn. rješenje zahtjeva složenu manipulaciju s animacijama i sl. tada se isto mora jasno komunicirati. Ujedno vidjeti da se prijedlozi dizajnerskog rješenja što više baziraju na core UI-u, što se jako rijetko ili gotovo nikada ne koristi (npr. upitno je koliko je potrebno izmišljati novi switch button ili cell separator)."

"Kako bi proces bio kvalitetniji, dizajneri bi trebali poznavati određene standarde vezano za tehnologiju za koju rade dizajn. Kao primjer navodim Flutter, a za Flutter postoji Material 3 standardizacija koju kad bi dizajneri poznavali i pratili bilo bi svima lakše. Jedan glavni primjer je coloring schema u Material 3, ako dizajner poznaje te standarde cijeli dizajn što se tiče coloringa će biti drugačije složen."

IV. DISKUSIJA

Istraživanje je pokazalo da je jedan od najčešćih izazova za dizajnere činjenica da naknadnim uključivanjem razvojnih programera u projekt moraju promijeniti mnoge stvari zbog tehničkih ograničenja ili nekih drugih čimbenika koji ograničavaju razvoj npr. nedostatak vremena. Da bi se izbjegle ovakve situacije predlaže se uključiti razvojne programere što ranije u proces dizajna, upoznati ih s idejom proizvoda i predstaviti osmišljeni način rada i interakcije korisnika s proizvodom. Kontinuirana komunikacija i jasna slika o trenutnom stanju u razvojnom procesu ključni su za uspjeh [3].

Kao i svaki odnos, suradnja dizajnera i programera se odvija u oba smjera. Rezultati istraživanja pokazuju da je bitno uključiti dizajnere u razvoj ranih verzija programske podrške i tražiti povratne informacije o implementiranom dizajnu. Zajednička vizija, međusobno razumijevanje i kvalitetna komunikacija od samog početka razvoja proizvoda ključ su izrade kvalitetnog digitalnog proizvoda [4].

Najranije verzije skica ekrana (engl. wireframe), čak i one na papiru, dobar su početak za dobivanje povratnih informacija od programera. Zahvaljujući skicama programeri mogu steći dojam o tome kako je sučelje organizirano, što im pomaže u planiranju implementacije dizajna i strukturiranju koda [2]. Što ranijim uvidom u skice programeri mogu primijetiti moguće probleme tijekom razvoja.

Mnogi dizajneri i programeri ne komuniciraju tijekom procesa dizajna nego se susreću tek na predaji dizajnerskog rješenja razvojnim programerima, odnosno u fazi kada se smatra da je dizajn već „gotov“. Tada se često javlja potreba da se takav gotovi prijedlog dizajnerskog rješenja mijenja, najčešće zbog tehničkih ograničenja u razvoju koja dizajneri nisu predvidjeli. Loše pripremljena specifikacija dizajna i dokumentacija te nedostatak međusobne komunikacije siguran su recept za ovakav ishod stvari. Postizanje ravnoteže između vizualne i tekstualne dokumentacije iznimno je važno. Za uspješno prenošenje dizajna potrebno je stvarati prototipove i prezentirati ih programerima što češće. Prototip na interaktivan način prikazuje planirano korisničko iskustvo [5]. Programeri tako dobivaju jasan uvid u planirane interakcije i animacije, iskustvo i cjelokupan doživljaj proizvoda te ih uspješnije mogu implementirati.

Nadalje, dizajneri i programeri posjeduju različite vještine, znanja i navike te pristupaju rješavanju problema na različite načine. Dizajneri su usmjereni na korisnike, a programere se često dovodi u situaciju da „samo“ pretvaraju dizajn u programsku podršku što može dovesti do problema. Nerijetko se dešava da programeri ne razumiju dovoljno radni proces dizajnera, i obrnuto. Ograničeno razumijevanje može dovesti do nesuglasica i nesporazuma. Za nesmetanu suradnju programeri bi trebali razumjeti osnove alata koje dizajneri koriste, njihove načine rada i način izražavanja (terminologiju i žargon), kako bi mogli surađivati s njima. Također, dizajneri bi trebali biti upoznati s procesom rada programera, njihovom terminologijom i alatima koje koriste. Obje strane bi trebale biti otvorene za učenje i razumijevanje međusobne perspektive [2]. Komunikacija je ključna za uspjeh projekta i treba biti redovita i otvorena [2]. Dizajneri i programeri trebaju imati redovne sastanke i razmjenjivati informacije o napredovanju projekta. Kreiranjem zajedničkog jezika i standarda, dizajneri i programeri mogu razumjeti jedni druge i raditi zajedno na stvaranju kvalitetnog proizvoda [6].

Na kraju, dizajneri i programeri mogu imati različite vizije završnog proizvoda, što dovodi do neslaganja i potrebe za naknadnim izmjenama. Za prevladavanje ovog problema bitno je od početka uspostaviti jasne ciljeve dizajna i razvoja te zajedno raditi na pronalaženju načina za njihovo postizanje.

V. ZAKLJUČAK

Kako bi se maksimalno iskoristili pozitivni učinci zajedničke suradnje bitno je održavati redovitu i konstruktivnu komunikaciju te razviti uzajamno razumijevanje radnih procesa, navika, terminologije i žargona. Kvalitetna suradnja od najranijih početaka procesa rezultira boljim razvojem i boljim konačnim proizvodom te čini radno okruženje produktivnim i pozitivnim.

S obzirom na mali broj ispitanika koji su sudjelovali u istraživanju nije moguće dati globalne i pouzdane zaključke no dobiveni podaci mogu poslužiti kao koristan uvid u problematiku.

REFERENCE

[1] Ungar, J., & White, J. (2008). Agile User Centered Design: Enter the Design Studio - a Case Study. In Extended Abstracts on Human Factors in Computing Systems (CHI '08) (pp. 2167–2178)

[2] Jones, A.J. (2019). Factors for successful Agile collaboration between UX designers and software developers in a complex organisation. (Doctoral dissertation). University of Chester, United Kingdom

[3] Caballero, Leydi & Moreno, Ana & Seffah, Ahmed. (2016). How Agile Developers Integrate User-Centered Design Into Their Processes: A Literature Review. International Journal of Software Engineering and Knowledge Engineering. 26. 1175-1201.

[4] Ambler, Scott. (2008). Tailoring Usability into Agile Software Development Projects 10.1007/978-1-84628-941-5_4.

[5] Johanna Kollmann, (2008). Designing the User Experience in an agile context. Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Life Sciences, University College London, 2008.

[6] Brown, Judith & Lindgaard, Gitte & Biddle, Robert. (2011). Collaborative Events and Shared Artefacts: Agile Interaction Designers and Developers Working Toward Common Aims. Proceedings - 2011 Agile Conference, Agile 2011. 10.1109/AGILE.2011.45.

www.dspace.com



Move mobility forward so that it can move everyone tomorrow.

... and come to dSPACE:
dspace.com/go/career-croatia

dSPACE

Ericsson Nikola Tesla



Pretvaramo kreativne ideje u globalnu stvarnost



U središtu Vaših tehnoloških potreba!

STORM Grupa je na hrvatskom i regionalnom tržištu prepoznata kao vodeći sistem integrator. Grupa ima 9 tvrtki članica koje pokrivaju različite segmente IT industrije s jednim ciljem - korisnicima ponuditi zaokruženu i potpunu uslugu po principu „ključ u ruke“!

Mi smo stručnjaci za:

- **Mrežna i infrastrukturna rješenja**
- **Usluge podatkovnog centra**
- **Sigurnosna rješenja**
- **Aplikativna i poslovna rješenja**

Više informacija na www.stormgrupa.hr

STORM GRUPA

INFODOM

Napredak pokretan inovacijama

Iskustvo temeljeno na znanju

InfoDom je vodeći ponuđač fleksibilnih digitalnih platformi za javni i privatni sektor koji osigurava postizanje poslovne izvrsnosti i ekonomske održivosti temeljenih na najboljim svjetskim i domaćim standardima. S više od dva desetljeća iskustva, mrežom renomiranih ICT partnera, s odjelom razvoja vlastitih proizvoda i rješenja osigurava ključne kompetencije za buduća tržišna kretanja kojima postižu poslovnu agilnost u okvirima digitalne transformacije.

Sredstvima EU osiguravaju provedbu digitalne transformacije

InfoDom kontinuirano ulaže u razvoj novih proizvoda i usluga te u istraživanja, u sklopu čega su realizirali nekoliko EU financiranih projekata (Razvoj inovativne platforme za digitalnu transformaciju poduzeća, Razvoj novog proizvoda za samoprocjenu regulatorne usklađenosti, procjenu zrelosti i kontroling za područje održive proizvodnje i prerade hrane te CEKOM – Centar kompetencija za digitalnu transformaciju prehrambene industrije na ruralnim područjima).

Pogled u budućnost

Strateški razvoj InfoDoma temelji se na iskustvima iz prethodnih razdoblja s jedne strane, i snažnom razumijevanju budućih tehnoloških i poslovnih trendova s druge strane, fokusirajući se na pet područja razvoja: komercijalizaciju proizvoda, usmjerenost na poslovnu izvrsnost, strateška partnerstva, održivost indeksa rasta te stvaranju inkubatora znanja.

Referenti projekti – 250 korisničkih iskustava

U dosadašnjoj praksi InfoDom se ističe referencama iz područja javnog sektora i državne uprave te realnog sektora.


BIRAM VIŠE.
telemach
210+ TV kanala | 24.000+ naslova u Video klubu
Sports Mode | EON Kids
EON
0800 555 500
telemach.hr

PROZA Station

Napredno upravljanje u oblaku za transformatorske stanice i pristupne podatkovne centre

Digitalna redundancija, veća efikasnost i niži troškovi



info.digital@koncar.hr

KONČAR
DIGITAL

Energija koja pokreće Hrvatsku

HEP

Više od struje



sponsors

general sponsor

HEP

golden sponsor

KONČAR

silver sponsor

telemach

bronze sponsors

INFODOM

STORM GRUPA

dSPACE

Ericsson Nikola Tesla

sponsors

SIEMENS energy

AD NET

TEH mjerne tehnologije d.o.o.

ALTPRO RESEARCH · DEVELOPMENT · PRODUCTION

MIPRO
Jadranski trg 1/II, HR-51000 Rijeka, Croatia

phone: +385 51 423 984
mipro@mipro.hr • www.mipro.hr